Rising Verbal Intelligence Scores: Implications for Research and Clinical Practice

Bob Uttl and Cory L. Van Alstine
Oregon State University

Evidence suggests that scores on various intelligence tests have been rising at a fast rate. To find out whether performance on the Wechsler Adult Intelligence Scale (WAIS) Vocabulary subtest has also been rising, the authors searched major psychology journals for investigations involving healthy younger and older adult participants and collected the reported WAIS Vocabulary scores. The meta-analysis shows that WAIS Vocabulary scores have been rising at the rate of 0.117/year (corresponding to 1.52 IQ points/decade) for younger adults and 0.367/year (corresponding to 4.79 IQ points/decade) for older adults.

Mounting evidence suggests that raw scores on various intelligence tests have been rising at a fast rate over the last 7 decades (Flynn, 1984, 1987, 1994, 1999; Jensen, 1998; Lippmann, 1976; Neisser, 1998). When raw score gains for 25 tests in 15 countries are standardized ($M = 100$, $SD = 15$) and expressed as IQ score increases, the IQ gains per decade ($\Delta IQ$/decade) range from 1.8 to 12.5 IQ points, with an overall mean of 5.0 IQ points (Jensen, 1998). The IQ gains are largest on culture-reduced, nonverbal tests such as the Raven’s Progressive Matrices Test ($\Delta IQ$/decade = 5.69) and the Performance scale of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955, 1981, 1997; $\Delta IQ$/decade = 7.8), smaller on verbal tests such as the Verbal scale of the WAIS ($\Delta IQ$/decade = 4.2), and small or nil on scholastic achievement tests (Jensen, 1998; see also Flynn, 1998, 1999). Finally, Flynn (1999) also observed that IQ gains were small or nil on Arithmetic, Information, and Vocabulary subtests of the WAIS in English-speaking nations but substantial on Vocabulary tests in German-speaking nations.

Rising intelligence scores have important implications for both research and clinical practice. In clinical practice, the rising intelligence scores lead clinicians to conclude that their clients are more highly functioning relative to their peers than they really are. As a concrete example, Flynn (1985) explained that by using outdated norms in the Wechsler Intelligence Scale for Children manual, the number of children classified as “mentally retarded” dropped from 8.8 million in 1947 to 2.6 million in 1972 in the United States alone. In research, intelligence tests are commonly used to describe the intellectual level of research participants, to establish that the sample is representative of the general population and comparable with other samples. Because of rising intelligence scores, however, researchers frequently find that their samples appear to be special, drawn from elite intellectual groups rather than from the general population. To illustrate, if IQ scores are rising at 5 points/decade, a representative sample drawn today from the U.S. population would score about 112 IQ points on the revised WAIS (WAIS–R) because it was standardized between 1976 and 1980, or about 24 years ago.

The single most frequently used test to establish a level of verbal intellectual functioning is the Vocabulary subtest of the WAIS (Wechsler, 1955) and its newer revisions (WAIS–R and WAIS–III; Wechsler, 1981, 1997). The WAIS Vocabulary subtest consists of 40 words. An examinee is presented with 1 word at a time and asked to define each word’s meaning. The examinee’s responses receive 0, 1, or 2 points, depending on how well he or she defines the word, allowing a range of scores from 0 to 80. The Vocabulary subtest is quick to administer, correlates highly (.91–.95) with the Verbal scale of the WAIS, and comes with extensive normative data (Wechsler, 1955, 1981, 1997). For these reasons, it is frequently used as a quick index of intellectual functioning in clinical practice (e.g., premorbid functioning) and is the most frequently used test to establish the intellectual level and representativeness of research samples.

The claim (Flynn, 1984, 1998) that intelligence scores are rising for some IQ tests but not for other tests—specifically the Vocabulary subtests—suggests that users of the Vocabulary subtests have nothing to worry about and that they can safely ignore rising IQ scores when interpreting examinees’ Vocabulary subtest scores. However, such a conclusion is not warranted. First, Flynn (1987) noted that the available data regarding IQ increases on the Vocabulary subtest of the WAIS in English speaking countries are scant. The only available U.S. data are described in the WAIS–R manual (Wechsler, 1981, as cited in Flynn, 1984). The manual includes Vocabulary data on only 72 examinees, ages 35–44 years, who were tested with both WAIS–R and WAIS. The reported data show that the scaled scores increased by 1.8 points (scaled scores were standardized to $SD = 3.0$) in 26 years, or 3.5 IQ points/decade (Wechsler, 1955, p. 47). Second, Flynn’s (1987) analyses showed

Bob Uttl and Cory L. Van Alstine, Department of Psychology, Oregon State University.
Cory L. Van Alstine is now at the Department of Psychology, Pacific University.
A preliminary report of this work was presented at the annual meeting of the Northwest Cognition and Memory Society, June 2000, Bellingham, Washington. We thank John G. Bacon II and Carolyn R. Pilkenton-Taylor for assisting with the project and Amy Siegenthaler for insightful comments on the manuscript.
Correspondence concerning this article should be addressed to Bob Uttl, who is now at the Institute of Psychology, University of Tsukuba, Tsukuba City, Ibaraki 305-8572, Japan. E-mail: bob.uttl@human.tsukuba.ac.jp

Copyright 2003 by the American Psychological Association, Inc.
that vocabulary scores are rising substantially in other countries, especially Germany, suggesting that relatively smaller increases in vocabulary scores in the United States may be an artifact of limited available data. Third, most of the data sets examined for rising IQ scores were based on children and young adults (i.e., school-age children, college students, military service draftees) and, therefore, it is unclear whether increases in Vocabulary scores would be larger if the data for more mature, older adults were examined (Flynn, 1987; Jensen, 1998).

The possibility that Verbal intelligence scores may be rising at a steeper rate for older adults is indicated by several key facts. First, the average number of years a person spends in school has risen substantially during the last 4–5 decades (Newburrow & Curry, 2000; see also other U.S. Census Bureau statistical reports available at www.census.gov). The same evidence of rising educational attainment appears in the successive standardization samples that were used for the WAIS: WAIS (standardized according to U.S. 1950 Census; Wechsler, 1955), WAIS–R (standardized according to U.S. 1970 Census; Wechsler, 1981), and WAIS–III (standardized according to U.S. 1995 Census; Wechsler, 1997).

Second, in cross-sectional studies, older adults score higher on Vocabulary than younger adults (Schaie, 1996; Uttl, 2002; Uttl & Graf, 1997; Wechsler, 1997). The performance advantage is especially apparent in studies that match educational attainment of younger and older adults (e.g., Uttl, 2002), but it is also seen in the latest population-based cross-sectional studies (Schaie, 1996) including the large standardization sample of the WAIS–III based on 2,450 individuals (Wechsler, 1997).

Third, performance on Vocabulary tests of the successive standardization samples of the WAIS shows an age-related reversal from younger adults (18–24 years old) outperforming older adults (65–74 years old; age range used in a typical aging studies) on the WAIS (Wechsler, 1955), to a minimal advantage of older adults over younger adults on the WAIS–R (Wechsler, 1981), to older adults outperforming younger adults on the WAIS–III (see Figure 1; Wechsler, 1997).

Fourth, Vocabulary scores (as well as IQ scores) correlate substantially with number of years of education (e.g., Kaufman, Reynolds, & McLean, 1989; r = .45, Uttl & Pilkenton-Taylor, 2001). In combination with the rising educational attainment, this correlation suggests that younger adults’ Vocabulary scores may be rising more slowly than Vocabulary scores of older adults across the years because younger adults have not yet realized their educational potential (Flynn, 1987; Jensen, 1998).

Finally, the differential IQ gains for older versus younger adults are not without a precedent. British Raven Progressive Matrices Test data (i.e., data from the culture-reduced, nonverbal test) from two cross-sectional samples tested 50 years apart show gains between 1942 and 1992 of 20 points for ages 18–32 and larger gains of almost 30 points for ages 33–67 (Flynn, 1998).

The goal of the present study was to address the inconsistencies and evidence lacuna on rising verbal intelligence scores. We focused on a specific set of questions: (a) Are verbal intelligence scores—specifically scores on the WAIS Vocabulary subtest—rising, and what is the magnitude of any such intelligence gains? (b) Are scores on the WAIS Vocabulary subtest rising at a different rate for younger versus older adults? and (c) Can generational increases in education explain rising WAIS Vocabulary scores?

An unparalleled opportunity to investigate whether scores on the WAIS Vocabulary subtest have been rising is provided by aging studies that typically report their participants’ WAIS Vocabulary scores to establish the intellectual level of participants. If the vocabulary scores have been rising, we would expect that recent studies would report higher vocabulary raw scores than earlier studies. Because studies on aging typically include both a group of younger and older adults, they allow us to examine whether intelligence scores have been rising for both younger and older adults and whether they have been rising at different rates for the two groups. By recording the average number of years of education for each sample (when available), we were able to investigate whether any increases in the Vocabulary test scores could be attributed to parallel increases in the education level of participants. Finally, compared with its successors, the WAIS–R and WAIS III Vocabulary subtests, the WAIS Vocabulary subtest has another critical advantage: It has been in use for the longest time span. Whereas the WAIS was released in 1955 and replaced by a revised version—WAIS–R—in 1981, WAIS–R was replaced by a revision much faster, after only 16 rather than 26 years, and the WAIS–III is only several years old.

Method

To find out how WAIS Vocabulary scores have changed during the last 3.5 decades, we searched journals for studies that (a) included samples of normal healthy English-speaking adults from either Canada or the United States and (b) reported their WAIS Vocabulary raw or scaled scores (to be included, a study must have specified whether the scaled scores were corrected or uncorrected for age). For each identified study, we recorded a publication date, the number of examinees in each sample, the mean raw or scaled WAIS Vocabulary score, and the mean number of years of education (if available). If the study reported scaled scores, we used WAIS translation tables to find out the raw score equivalents.
To examine the possibility that Vocabulary scores are rising at different rates for younger and older adults, we searched journals that are likely to report data on samples of both younger and older adults: Aging, Neuro-psychology, and Cognition; Clinical Neuropsychologist; Experimental Aging Research; Journal of Abnormal Psychology; Journals of Gerontology; Neuropsychologia; and Psychology and Aging. We searched all of these journals from the first issue of 1965 or from the first available issue (i.e., the first issue of Volume 1) until the last issue of 1995. The complete data set is available from Bob Uttl or from www.alfalab.com/risingverbalintelligence.html.

Results

We located 249 samples representing 7,151 individuals whose performance on the WAIS Vocabulary subtest was reported in the selected journals from January 1965 to December 1995. To examine whether Vocabulary scores have been rising at different rates, we divided the samples into two approximately equal-sized groups: a sample of younger adults (< 50 years old) versus a sample of older adults (≥ 50 years old). The mean age of 67.8% of the younger samples fell between 18 and 24 years, and the mean age of 76.6% of the older samples fell between 65 and 74 years. Table 1 shows the number of younger and older samples, the number of younger and older examinees, and the mean age of younger and older examinees for each of the six publication periods: 1965–1969, 1970–1974, 1975–1979, 1980–1984, 1985–1989, and 1990–1995. Although the mean ages shown in Table 1 increase and decrease across the publication periods for both younger and older samples, this “bounce” may be an artifact of the arbitrary division of the data in 5-year bins. Correlations between the mean age of samples and the publication year were small (r = −.024 for younger samples and r = −.092 for older samples), and the weak relationships between the mean age and the publication year did not influence the results of the subsequent analyses.

Younger samples were drawn primarily from undergraduate students (52.1% of all samples), whereas older samples were drawn primarily from adults living in the community (68.8% of all samples) who responded to advertising for study volunteers. For other samples, the authors did not specifically specify whether or not younger adults were students, or in case of older adults, whether or not older adults were community living adults. However, we did not include any samples drawn from special populations, such as adults from assisted living homes, institutionalized adults, adults with psychiatric disorders, and so forth.

Figure 2 shows each sample mean Vocabulary raw score plotted against the publication year, for all samples, with the size of each bubble indicating the sample size. The data in Figure 2 indicate that Vocabulary raw scores were rising steadily for samples published between 1965 and 1995. A weighted least squares meta-regression model weighted by the sample size (Hedges & Olkin, 1985; Sutton, Abrams, Jones, Sheldon, & Song, 2000; Thompson, 1994) showed increases in Vocabulary raw scores of 0.289/year (95% confidence interval [CI] = 0.269, 0.309), t(7149) = 25.48, p < .0001, or 3.73 IQ points/decade (using translation tables in WAIS manual).

To determine whether Vocabulary scores are rising faster for older than younger adults, we tested the interaction between age group and publication year by including it in the weighted meta-regression model together with age group and publication year. The Age Group × Publication Year interaction term was significant, t(7147) = 13.06, p < .0001, indicating that Vocabulary raw scores were rising at different rates for older and younger adults.

The top panel of Figure 3 shows the mean Vocabulary raw scores plotted against the publication year for younger samples, and the bottom panel of Figure 3 shows the same data for older samples. Figure 3 and the corresponding analyses indicate that Vocabulary raw scores were rising faster for older samples than for younger samples. For younger adults, Vocabulary raw scores went up by 0.117/year (95% CI = 0.090, 0.144), t(3245) = 8.38, p < .0001, or about 1.52 IQ points/decade (Vocabulary = 55.34 + 0.117 * [Publication Year – 1955]), whereas for older adults, the raw scores went up at a higher rate of 0.367/year (95% CI = 0.342, 0.392), t(3904) = 28.22, p < .0001, or about 4.79 IQ points/decade (Vocabulary = 50.96 + 0.367 * [Publication Year – 1955]). Inspection of the bottom panel of Figure 3 also suggests that the IQ gains may be underestimated for older samples because of evidence of ceiling effects in the data during the 1985–1995 period.

To find out whether concomitant increases in education over the examined period could explain rising Vocabulary scores, we reanalyzed the data from the studies that reported mean education for each sample: 56 younger samples and 60 older samples. First, the meta-regression analyses weighted by the sample sizes showed smaller increases in Vocabulary scores/year of education for younger adults, 0.735/year of education (95% CI = 0.548, 0.920), t(2036) = 7.02, p < .0001, but substantial increases for older adults, 0.548/year of education (95% CI = 0.365, 0.730), t(4038) = 7.00, p < .0001.

Table 1

<table>
<thead>
<tr>
<th>Publication year</th>
<th>No. of samples</th>
<th>No. of individuals</th>
<th>Mean age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Younger</td>
<td>Older</td>
<td>Younger</td>
</tr>
<tr>
<td>1965–1969</td>
<td>17</td>
<td>16</td>
<td>268</td>
</tr>
<tr>
<td>1970–1974</td>
<td>7</td>
<td>7</td>
<td>180</td>
</tr>
<tr>
<td>1975–1979</td>
<td>15</td>
<td>15</td>
<td>287</td>
</tr>
<tr>
<td>1980–1984</td>
<td>18</td>
<td>17</td>
<td>334</td>
</tr>
<tr>
<td>1985–1989</td>
<td>31</td>
<td>36</td>
<td>689</td>
</tr>
<tr>
<td>1990–1995</td>
<td>33</td>
<td>37</td>
<td>1,487</td>
</tr>
<tr>
<td>Total</td>
<td>121</td>
<td>128</td>
<td>3,245</td>
</tr>
</tbody>
</table>
adults, 2.212/year of education (95% CI = 2.038, 2.386), t(2529) = 24.85, p < .0001.

Second, to elucidate whether increases in education over the examined time period can explain rises in Vocabulary scores, we conducted two sets of weighted meta-regression analyses, one for younger adults and one for older adults. In each set of analyses, we entered both education and publication year as predictors. For younger adults, Vocabulary scores increased by 0.890 with each year of education (95% CI = 0.704, 1.076), t(2035) = 9.33, p < .0001, and decreased by 0.225 with each publication year (95% CI = –0.276, –0.174), t(2035) = –8.66, p < .0001. For older adults, Vocabulary scores increased by 2.039 with each year of education (95% CI = 1.859, 2.219), t(2528) = 22.18, p < .0001, and increased 0.164 with each publication year (95% CI = 0.116, 0.212), t(2531) = 6.71, p < .0001—slightly less than half of the increases obtained when increases in educational attainment are not taken into account. Thus, when education is held constant, Vocabulary raw scores showed a decline for younger adults and an increase for older adults.

Discussion

We found that intelligence scores are rising on the widely used Wechsler Vocabulary test at the relatively fast rate of 0.289/year or 3.73 IQ points/decade. More importantly, we found smaller increases (0.117/year or 1.52 IQ points/decade) for young adults but larger increases (0.367/year or 4.79 IQ points/decade) for older adults. The analyses of the reduced data—the samples with available education data—indicated that, at least for older adult samples, education can explain a substantial portion of the rises in Vocabulary scores, reducing the increases in raw scores per year to approximately half, from 0.367 to 0.164/year, or from 4.79 to 2.13 IQ points/decade. Similarly, for younger adults, small rises in Vocabulary scores were reduced to decreases after controlling for education, from .117 to –0.225/year or from 1.52 IQ points/decade increase to 2.93 IQ points/decade loss. However, these latter findings are weaker, based on only about 55% of all samples (i.e., the samples with available educational data).

Our findings are consistent with and extend previous research. The finding that Vocabulary scores have been rising steadily replicates Vocabulary increases reported in the WAIS–R manual but based on a small sample of only 72 individuals (Wechsler, 1981). Our data are based on over 7,000 individuals tested with the same test items over a period of 31 years. Moreover, confidence in our findings is reinforced by noting that, by design, our data represent sampling standards of U.S. and Canadian aging research and are unlikely to be a product of chance influences afflicting small sample size studies or due to an occasional rogue investigator using selective sampling methods. Rising Vocabulary scores are in agreement with the Flynn effect observed on the multiple-choice WORDSUM vocabulary knowledge test (Huang & Hauser, 1998).
The finding that Vocabulary scores have been rising faster for older adults than for younger adults qualifies Flynn’s conclusion that increases on vocabulary tests are small or nil in English-speaking countries. Flynn based his conclusion on studies that included mostly children or young adults who might not have reached their full potential on vocabulary tests, and thus, their vocabulary test scores underestimated their verbal IQ potential (Flynn, 1987). Our data are consistent with this view; they indicate that verbal IQ achievement of older adults has been increasing steadily across the decades at a faster rate, resulting in an increasingly large gap between younger and older adults. Our data show the age-related reversal in performance on the WAIS Vocabulary subtest similar to the one observed across the successive standardizations of WAIS, WAIS–R, and WAIS–III Vocabulary tests (see Figure 2); they suggest that in 1953, younger adults outperformed older adults by 5.1 points; in 1978, older adults outperformed younger adults by 1.4 points; and in 1995, older adults outperformed younger adults by 8.1 points on the Vocabulary test. Finally, we also observed this same pattern of findings with the WAIS–R Vocabulary test (Uttl & Van Alstine, 2000), even though the WAIS–R Vocabulary subtest has not been around for as long as the WAIS. Given that our data replicate the pattern found with the standardization samples stratified to match the population characteristics, we can be confident that our findings are not an artifact of disorderly recruitment procedures or systematic shifts in recruitment procedures among other explanations.

Contrary to Flynn’s conclusions, our data suggest that scores on measures of crystallized intelligence may be rising as fast as scores on measures of fluid intelligence (Cattell, 1941; Horn & Cattell, 1966, 1967). Overall higher scores of older versus younger adults are consistent with the notion that crystallized intelligence is maintained or even increases across the adult life span. Moreover, assuming that the younger adults (~25 years old) tested a few years prior to 1965 were drawn from the same population (i.e., the same cohort) as the older adults (~65 years old) tested a few years prior to 1995, our results provide evidence that age-related increases in vocabulary scores commonly observed in cross-sectional studies of aging reflect not only cohort differences but also true intrapersonal increases in verbal knowledge.

Our findings underscore the need to interpret an individual’s or population sample’s verbal IQ scores in light of these quickly rising verbal IQ scores. Today, a normal older adult of average intelligence is expected to score about 124 IQ points on the WAIS Vocabulary test (normed between 1953 and 1954) and 112 IQ points on the WAIS–R Vocabulary test (normed between 1976 and 1980; see Uttl & Van Alstine, 2000). A normal younger adult of average intelligence is expected to score lower than an older adult, 107 IQ points on the WAIS Vocabulary subtest and 104 IQ points on the WAIS–R Vocabulary subtest, but his or her score is also expected to be higher than the score suggested by the outdated norms. As the result of these rising verbal intelligence scores, the use of the outdated Vocabulary norms will result in fewer people being classified as having a lower verbal intelligence and more people being classified as having superior verbal intelligence (Strauss, Spreen, & Hunter, 2000). In addition, this effect will be stronger for older than for younger adults.

If the Vocabulary subtest is used to establish the intellectual level of a sample drawn from a normal population, the sample will appear “highly intelligent,” drawn from “a special highly intellectually functioning population.” Moreover, the samples of young adults will appear intellectually inferior to samples of older adults, and researchers studying aging will find it difficult, if not impossible, to match younger and older adults in terms of their verbal abilities (see Uttl, 2002, for similar results with another verbal knowledge test—the North American Adult Reading Test). Our findings highlight the need (a) to periodically rerun even verbal IQ tests such as Wechsler Vocabulary subtests to counteract the effects of rising IQ scores and (b) to interpret each sample’s intellectual levels in relationship to up-to-date age-specific norms rather than in terms of numerical similarity.

What is responsible for these increases in verbal IQ scores? One possible explanation for rising IQ scores directly addressed by our study are parallel increases in educational attainment as well as in access to education over the last several decades (Flynn, 1998; Greenfield, 1998; Teasdale & Owen, 1987, 1989; Williams, 1998). When education was held statistically constant, yearly increases in verbal IQ scores of older adults were reduced to about half, thereby suggesting that a large portion of increases could be accounted for by concomitant increases in educational attainment (see Teasdale & Owen, 1987, for similar findings based on Denmark’s military service recruits). The education explanation is also consistent with smaller IQ increases observed in younger samples. In contrast to older samples who were mostly recruited from community living adults who achieved their full education potential, younger samples were drawn primarily from college undergraduates who have not yet completed their education. As the result, younger adults’ education attainment has increased only slightly over the decades and, because of the increased accessibility of education, composition of undergraduate samples has shifted from having students in the intellectual elite to having students from more diverse intellectual backgrounds. Increases in verbal intelligence scores may also be due to other factors, however, including increases in accessibility of printed, spoken, and visual media (Greenfield, 1998), environmental complexity (Schooler, 1998), nutrition (Lynn, 1998; Martorell, 1998; Sigman & Whaley, 1998), and socioeconomic status and urbanization (Flynn, 1998).

References

Flynn, J. R. (1985). Wechsler intelligence tests: Do we really have a criterion of mental retardation? American Journal of Mental Deficiency, 90, 236–244.


Received February 13, 2002
Revision received November 25, 2002
Accepted December 3, 2002

BRIEF REPORTS

Wanted: Your Old Issues!

As APA continues its efforts to digitize journal issues for the PsycARTICLES database, we are finding that older issues are increasingly unavailable in our inventory. We are turning to our long-time subscribers for assistance. If you would like to donate any back issues toward this effort (preceding 1982), please get in touch with us at journals@apa.org and specify the journal titles, volumes, and issue numbers that you would like us to take off your hands.