**EURAMERICA** Vol. 33, No. 4 (December 2003), 757-800


# Has Verbal Ability Declined in America?

Min-Hsiung Huang
Institute of European and American Studies, Academia Sinica
e-mail: mhhuang@sinica.edu.tw

## Abstract

Using data from the General Social Survey (GSS), 1974-1996, researchers have debated whether the observed intercohort decline represents an actual intercohort decline in verbal ability in the U.S. population. Some researchers speculate that the observed intercohort decline in verbal ability in GSS data is spurious since previous literature has shown that verbal ability increases over the life span and peaks in old age. Other researchers, however, maintain that aging effects on verbal ability are not large enough to explain the steep decline in verbal ability observed in GSS data. By identifying some of the

Received July 2, 2003; accepted September 9, 2003; last revised October 6, 2003.
Proofreaders: Hui-Wen Ko, Hsin-Wen Fan

* This research was supported by Academia Sinica through the Learning 2000 Project. The opinions expressed herein are those of the author, not those of Academia Sinica or any other funding agencies. An earlier version of this paper was presented at the 2001 American Sociological Association annual meeting in Anaheim, California, USA. Data from the General Social Survey is available from the Interuniversity Consortium for Political and Social Research, University of Michigan, Ann Arbor, Michigan, or at http://www.icpsr.umich.edu:8080/GSS/homepage.htm. I greatly appreciate the comments of Duane F. Alwin, Norval D. Glenn, Walter R. Gove, Tony Tam (譚康榮), the members of the *EurAmerica* Editorial Board, and two anonymous reviewers. I also thank Yu-Ling Lin (林毓玲) for research assistance and Audrey Chen (陳慧文) for English editing. The author alone is responsible for any inadequacies remaining in this manuscript. Direct correspondence to Min-Hsiung Huang (黃敏雄), Academia Sinica, Institute of European and American Studies, Nankang, Taipei, 11529, Taiwan (tel: 886-2-3789-7246).

inadequacies of earlier research and considering a strategy that takes changes in the selectivity of education over time into account, this study attempts to reconcile some of these disparate findings. The findings from this study did not support an age interpretation. Intercohort decline in verbal ability began with persons born in late 1940s and ended with persons born in early 1960s, independent of aging effects on vocabulary acquisition. The intercohort decline in GSS vocabulary test scores resembles the widespread test score decline observed in the U.S. during the 1960s and 1970s in its onset, end, and magnitude.

**Key Words:** vocabulary test scores, test score decline, verbal ability, cohort analysis, aging

## I. Introduction

One of the most important educational goals in the U.S. is to help every child become a successful reader. United States President George W. Bush's "Reading First Program," which authorized $900 million in federal funds for fiscal 2002, is based on a belief that early literacy is the key to children's entire education and should be a top priority for parents and educational policy makers. Despite a concerted effort made previously by Democratic and Republican administrations at the federal and the state government levels, the goal for universal literacy among school children is far from being achieved. As of 1998, only 31 percent of American fourth-grade students demonstrated proficiency for reading texts appropriate for fourth-graders, while more than a third (38 percent) were unable to reach the basic level (Loomis & Bourque, 2001: 56). For the poor and minority children in cities, the percentage of students who lacked basic reading skills was almost two times greater than the national average.

In an international comparison of reading literacy for 15-year-old students in 2000, the United States ranked near the middle among 32 countries and 18 percent of American students were below the level of capable of basic reading tasks (The OECD Program for International Student Assessment, 2001). Therefore, the U.S. government is greatly concerned that a significant number of their students may not have the necessary literacy skills to benefit sufficiently from their learning opportunities to be successful in the labor force after they complete school.

Reading skills are not only important for personal development, but changes in the average level of verbal ability in a nation may also have significant economic consequences (Bishop, 1989). Students' level of academic achievement is often regarded as a direct measure of a nation's human capital which affects its economic growth (Lee & Lee, 1995). Therefore, to what extent have academic test scores changed over time is an important research question. To address this straightforward question,

however, is less than simple. Researchers have debated over the generalizability and validity of the "Great Test Score Decline" observed in the U.S. during the 1960s and 1970s: 1) did it pertain to all subject areas, 2) was it a national problem, 3) was it large enough to be of serious educational concern, and 4) was it just an artifact of specific tests (Bracey, 1991; Koretz, 1986, 1992; Rotberg, 1990; Stedman & Kaestle, 1991)?

There has been a great deal of methodological debate on this topic. Even when researchers use the same data set to report trends in test scores, the results are inconsistent. For example, using nationally-representative data from the General Social Survey (GSS), researchers have presented contradictory results (Alwin, 1991; Glenn, 1994; Wilson & Gove, 1999a, 1999b; Glenn, 1999; Alwin & McCammon, 1999, 2001). The essence of the debate centered around whether the observed long-term intercohort decline in vocabulary test scores, beginning with individuals born in the period around 1916, was spurious. Some speculated that the observed intercohort decline in verbal ability was, in fact, a function of the growth in verbal ability as people age (Wilson & Gove, 1999a, 1999b). Others (Glenn, 1999; Alwin & McCammon, 1999), however, maintained that intercohort decline in verbal ability was not due to aging effects. The inconsistencies in the findings stem from two methodological problems. First, the fact that year of birth (cohort) is a linear function of survey year (period) and age makes it a formidable task to disentangle aging, period, and birth cohort effects.[1] Second, for a given level of schooling, persons born in recent years tend to be less selective than persons born in earlier years, as the years of education that the average person obtains increases over time. Therefore, the absence of a control for changes in the selectivity of education over

[1] Since year of birth is the subtraction of year of survey and age, one cannot separate age, period, and cohort effects. Furthermore, all three variables cannot be in a regression model as independent variables at the same time as the third variable will not vary when holding any two of these three variables constant (Firebaugh, 1997).

time prohibits a correct investigation of intercohort trend in education-adjusted vocabulary test scores.

It is unfortunate that our understanding on trends in verbal scores is limited by these methodological difficulties that prevent researchers from a better use of the GSS data. This is because the GSS has some advantages over other data sets which are used to report trends in test scores. These advantages include (1) a nationally-representative sample, (2) a household sample (instead of a school sample) which includes people who are not in school, (3) a series of replicating cross-sectional surveys since 1974 for adults age 18 and over, (4) a wide age range, (5) a wide birth year range, and (6) a set of vocabulary test items which have not been changed since 1974. The disadvantage of using the GSS data for trend analysis in test scores, however, is that there are only 10 test items in the GSS vocabulary test. Because of this disadvantage, there is reason for caution in the interpretation of GSS vocabulary test score trends.

In presenting intercohort changes in vocabulary test scores, I considered the possibility that selectivity of education has changed over time. I also attempted to reconcile some of the inconsistencies in previous research. This analysis reports several findings. First, previous studies have presented flawed intercohort trends and age trends in GSS vocabulary test scores because changes in educational selectivity were not taken into account. Second, there was an intercohort decline in GSS vocabulary test scores — a decline similar to the wide-spread test score decline of the 1960s and 1970s with respect to its onset, end, and magnitude. Third, vocabulary increased before individuals reach their late-thirties and declined after their late-sixties. There was no significant age-related increase between the late-thirties and the early-sixties — a result that is consistent with Schaie's (1996) findings based on longitudinal data. Finally, contrary to suggestions by Wilson and Gove (1999a, 1999b), the intercohort decline in GSS vocabulary test scores was not primarily a result of age-related increases in vocabulary over the life course.

## II. Aging versus Cohort Explanations: A Brief Review

Using data from the GSS, 1974-1990, Alwin (1991) reported an intercohort decline in education-adjusted mean verbal scores starting with individuals born in the period around 1916. According to his analysis, the effects of aging partially explained the downcast level of vocabulary knowledge in the oldest cohorts—those who were age 79 or older in 1990. Except for these oldest cohorts, Alwin (1991) concluded that the effects of aging on vocabulary test scores were probably insignificant. Such a conclusion was based on the assumption that there were no period effects on verbal scores, and that a linear age variable contributes little to explaining the variance of vocabulary scores when cohort and amount of schooling are already controlled. Glenn (1994) confirmed Alwin's (1991) finding of a long-term intercohort decline in GSS vocabulary test scores. With respect to aging effects, Glenn (1994) maintained that vocabulary scores, after adjusting for cohort and education, tended to decline, rather than increase, after young adulthood.

In the GSS, 1974-1998, the correlation between age and birth year is 0.9. One of the strategies for overcoming the age-period-cohort identification problem is to depend on theoretical guidance or "side information" that suggests the relative plausibility of alternative interpretations of observed trends. For example, Wilson and Gove (1999a) argued that the observed intercohort decline in the GSS education-adjusted vocabulary scores reported by Alwin (1991) and by Glenn (1994) was spurious since verbal ability increases over the life span and peaks in old age.

The essence of Wilson and Gove's (1999a, 1999b) critique was five-fold. First, the finding of declining verbal achievement was not consistent with "the Flynn effect," which notes that massive IQ gains over time in many countries is due to the increasing years of schooling worldwide (Wilson & Gove, 1999a:

253). For example, Flynn (1998) discovered that whites gained 25 IQ points from 1918 to 1995 on the Wechsler-Binet tests in the U.S. Second, Wilson and Gove (1999a: 255-256) argued that the observed intercohort decline in education-adjusted vocabulary scores in the GSS was spurious since an education selection effect led to an unfair comparison of a more selective group (persons born in earlier years) to a less selective group (persons born in recent years) for a given level of schooling. Third, the spurious intercohort decline in vocabulary scores was the mirror image of true age-related increases in verbal ability over the course of a lifetime. Education-adjusted mean vocabulary scores increased with age and peaked at about age 65, when a significant decline followed (Wilson & Gove, 1999a: 259). Therefore, the age trend in GSS vocabulary test scores was consistent with the literature on cognitive development, which states that crystallized intelligence, such as that measured by the GSS vocabulary test, peaks and starts to decline in old age (Cattell, 1963, 1971). Empirical evidence of these age-related patterns in verbal ability were consistent with those presented in Schaie's (1996) analysis using data from the Seattle Longitudinal Study. Fourth, after controlling for the effects of education (though not before), Wilson and Gove discovered that year of survey was negatively correlated with verbal scores in all age categories (18-34, 34-50, 50-66, and 66-81). Although survey year and birth year were highly correlated ($r > 0.82$) in each age category, the negative correlation between survey year and verbal score was not the result of an intercohort decline in GSS vocabulary test scores, but was solely a negative period effect since schooling had become less selective over time (Wilson & Gove, 1999a: 264). Wilson and Gove also hypothesized the negative period effect could have been due to word obsolescence, a decrease in reading, changes in the ethnic composition or age structure of the populations, social disorganization, and the way in which the GSS was administered. Finally, Wilson and Gove noted that there had not been an intercohort decline in vocabulary acquisition. Instead of a decline, there might has been a modest

intercohort increase in average vocabulary size since words used in the GSS vocabulary tests have become somewhat less frequently used today than when the test was constructed (Wilson & Gove, 1999b: 301).

In responding to Wilson and Gove's (1999a) age interpretation of the observed intercohort trend in GSS vocabulary test scores, Glenn (1999) used a different strategy to display mean vocabulary by age. Glenn traced four birth cohorts (1920-1929, 1930-1939, 1940-1949, and 1950-1954) as they grew older through the GSS years, from 1974 to 1996. Glenn found that the GSS data showed no age-related increases in vocabulary scores within cohorts after young adulthood. Alwin and McCammon (1999), on the other hand, created several four-year cohort categories and three-year age categories. In each category of cohort or age, the correlation between age and birth year was largely reduced, but the year of survey became highly correlated with both age and cohort ($r > 0.98$). Alwin and McCammon examined aging effects on vocabulary test scores within each cohort category and examined cohort effects on vocabulary test scores within each age category. By assuming an absence of period effects, Alwin and McCammon concluded that aging effects were not large enough to explain the larger intercohort patterns in GSS vocabulary test scores.

Therefore, while Wilson and Gove promote an explanation based on aging, Glenn (1999) and Alwin and McCammon (1999) support a cohort explanation for the intercohort changes in GSS vocabulary scores. None of them, however, have provided compelling findings to support their positions, partly because changes in the selectivity of education over time were not taken into account. Wilson and Gove correctly argued that changes in the selectivity of education generated a distorted intercohort trend in education-adjusted vocabulary test scores. However, with respect to education-adjusted vocabulary test scores in the GSS data, if changes in the selectivity of education bring about a spurious intercohort decline, they will also produce a false age trend which, Wilson and Gove suggest, is in line with literature

that "consistently indicates that verbal abilities, especially vocabulary knowledge, continue to improve over the life course, and that significant decline often does not begin until age 60 or later" (Wilson & Gove, 1999a: 258). Because of the absence of a control for temporal changes in the selectivity of education, Wilson and Gove (1999a, 1999b) overestimated the effects of age on vocabulary test scores throughout their analyses. For the same reason, cohort effects were not accurately measured by Glenn or Alwin and McCammon. Without controlling for changes in the selectivity of education, one cannot adequately evaluate the relative plausibility of aging versus cohort interpretations of the observed intercohort differences in education-adjusted vocabulary test scores.

## III. Data

I used data from the GSS, 1974-1998. The GSS is a series of replicating cross-sectional surveys for adults age 18 and over. For each survey year, there is a sample size of about 1,500 respondents. Since 1975, the GSS samples have been designed to give every household an equal probability of inclusion in the sample. In order for the GSS samples to be representative of the U.S. population on the individual level, and to take into account the over-sampling of blacks in 1982 and 1987, I applied the relevant statistical weights (Davis & Smith, 1992).

A ten-item multiple-choice vocabulary test was administered to respondents in the GSS survey years of 1974, 1976, 1978, 1982, 1984, 1987-1991, 1993-1994, 1996, and 1998. For each word item, GSS respondents were asked to choose the one word out of five possible matches that came closest in meaning to the stimulus word (first presented and in capital letters). Hauser and Huang (1997) provided sample vocabulary test items which are reproduced in Table 1. In the GSS, the vocabulary test score variable represents the total number of correct word items.

In the analyses, persons who were at least 24 years old were included since many younger people are still in school. In addition,

Table 1 Sample Vocabulary Test Items

| | | | | | |
|---|---|---|---|---|---|
| a. LIFT | 1. sort out | 2. raise | 3. value | 4. enjoy | 5. fancy |
| b. CONCERN | 1. see clearly | 2. engage | 3. furnish | 4. disturb | 5. have to do with |
| c. BROADEN | 1. efface | 2. make level | 3. elapse | 4. embroider | 5. widen |
| d. BLUNT | 1. dull | 2. drowsy | 3. deaf | 4. doubtful | 5. ugly |
| e. ACCUSTOM | 1. disappoint | 2. customary | 3. encounter | 4. get used | 5. business |
| f. CHIRRUP | 1. aspen | 2. joyful | 3. capsize | 4. chirp | 5. incite |
| g. EDIBLE | 1. auspicious | 2. eligible | 3. fit to eat | 4. sagacious | 5. able to speak |
| h. CLOISTERED | 1. miniature | 2. bunched | 3. arched | 4. malady | 5. secluded |
| i. TACTILITY | 1. tangibility | 2. grace | 3. subtlety | 4. extensibility | 5. manageableness |
| j. SEDULOUS | 1. muddled | 2. sluggish | 3. stupid | 4. assiduous | 5. corrupting |

Source: Miner (1957: 53).

younger respondents in the GSS may also be less representative of the younger population (Alwin & McCammon, 1999). In some cases of replicating Wilson and Gove's analyses, however, individuals 18 to 23 years of age were included for sample consistency. I included only individuals who are native-born not only because non-natives score significantly lower on the GSS vocabulary test, but also because there was a significant relationship between nativity and cohort membership (Alwin & McCammon, 1999). Finally, I selected valid cases for respondents' years of education completed and the number of adults in the household. These case selections resulted in a total of 16,155 cases. When individuals 18 to 23 years of age were included, the sample size was 17,803.

## IV. Methods

### A. Distribution of the GSS Vocabulary Test Scores

Because the distribution of GSS vocabulary test scores was censored on the right-hand side, the OLS model tended to underestimate the effects of key independent variables (e.g., age, birth year, and years of education). In addition, there were ceiling or floor effects on the number of correct answers in other subpopulations (e.g., individuals in high-scoring or low-scoring birth cohorts). For this reason, the analyses of the vocabulary test score data were based on a two-sided Tobit specification, which compensated for censoring at both ends of the distribution on the assumption that the true distribution of test scores is Gaussian (Maddala, 1983; Amemiya, 1984). The use of the Tobit specification, relative to the use of the OLS specification, did not make a significant difference to the results.

### B. Changes in the Selectivity of Education over Time

Schooling shapes knowledge of vocabulary and selects for verbal ability as well. The relationship between years of schooling

completed and measured verbal ability among adults reflects (1) the effects of the educational experience on vocabulary acquisition, and (2) selectivity of schooling on the basis of pre-existing characteristics that affect verbal ability such as innate ability and socioeconomic background. As the mean level of schooling years completed in the population increases over time, the standard of elite education increases as well. For example, an average high school graduate born in 1920 is more highly selected than is an average high school graduate born in 1970. Since the selectivity of education changes over time, controlling for respondents' years of schooling completed for persons of different birth cohorts is likely to introduce a "pseudocontrol" which may distort the results (Lieberson, 1978, 1985).

To deal more precisely with the effects of educational selectivity and their changes over time, a data set which includes not only measures of adult verbal ability but also measures of preadult ability across several cohorts was needed. However, to the author's knowledge, no longitudinal national data sets exist at the present time with the required measures of preadult ability across several birth cohorts which cover a wide range in time. Given what is available in the GSS, some indirect evidence sheds some light on the topic. For example, for a given level of schooling, respondents' relative educational standing within cohorts changed across cohorts as the average level of educational attainment in the general population increased over time. Therefore, intercohort changes in respondents' relative educational standing within cohorts may be considered a proxy measure for intercohort changes in the selectivity of education.

To measure respondents' relative educational standing within cohorts, I created 16 birth cohorts which included the 1885-1900 cohort, a set of 14 five-year cohorts from 1901 to 1970, and the 1971-1974 cohort. For each birth cohort, I assigned a score to every respondent according to his/her cumulative percentile in the percentage distribution for respondents' years of schooling completed, as shown in Table 2. For example, respondents born between 1901 and 1905 and with 12 years of schooling have a

Table 2 Cumulative Percentage Distribution for Respondents' Years of Education by Birth Cohort: General Social Survey (GSS), 1974-1998

| | Birth Cohort | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1885-1900 | 1901-1905 | 1906-1910 | 1911-1915 | 1916-1920 | 1921-1925 | 1926-1930 | 1931-1935 | 1936-1940 | 1941-1945 | 1946-1950 | 1951-1955 | 1956-1960 | 1961-1965 | 1966-1970 | 1971-1974 |
| No schooling | 2.9 | 0.3 | 0.2 | 0.0 | 0.1 | 0.2 | 0.4 | 0.0 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| 1 year | 3.3 | 0.3 | 0.2 | 0.1 | 0.3 | 0.3 | 0.6 | 0.1 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| 2 years | 3.7 | 1.3 | 0.4 | 0.5 | 0.6 | 0.5 | 0.7 | 0.4 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.2 | 0.0 |
| 3 years | 5.9 | 2.5 | 1.5 | 2.0 | 1.2 | 1.1 | 1.0 | 0.9 | 0.2 | 0.1 | 0.0 | 0.1 | 0.1 | 0.0 | 0.2 | 0.0 |
| 4 years | 7.7 | 3.8 | 3.6 | 3.1 | 2.7 | 1.9 | 1.8 | 1.1 | 0.5 | 0.2 | 0.1 | 0.2 | 0.3 | 0.0 | 0.2 | 0.0 |
| 5 years | 13.2 | 5.7 | 5.2 | 4.5 | 4.2 | 3.3 | 2.5 | 1.4 | 0.8 | 0.5 | 0.1 | 0.2 | 0.5 | 0.0 | 0.2 | 0.0 |
| 6 years | 16.9 | 8.5 | 8.6 | 8.4 | 7.3 | 6.0 | 4.0 | 2.7 | 1.3 | 1.1 | 0.7 | 0.3 | 0.5 | 0.0 | 0.2 | 0.0 |
| 7 years | 23.8 | 14.5 | 14.6 | 12.9 | 11.5 | 8.3 | 6.2 | 4.7 | 3.0 | 2.2 | 1.3 | 1.1 | 0.7 | 0.4 | 0.2 | 0.0 |
| 8 years | 52.0 | 38.2 | 37.0 | 28.0 | 22.1 | 16.4 | 13.5 | 9.5 | 7.2 | 4.6 | 2.6 | 2.3 | 1.6 | 1.2 | 0.9 | 2.4 |
| 9 years | 56.8 | 44.2 | 42.6 | 34.0 | 26.2 | 20.3 | 19.4 | 14.7 | 11.8 | 7.5 | 4.7 | 4.2 | 3.7 | 3.7 | 2.9 | 4.8 |
| 10 years | 64.8 | 53.3 | 50.7 | 40.8 | 33.1 | 26.1 | 28.4 | 20.6 | 18.0 | 13.4 | 8.4 | 7.9 | 7.2 | 7.2 | 5.6 | 8.9 |
| 11 years | 68.9 | 58.7 | 56.0 | 46.6 | 40.1 | 33.6 | 34.5 | 26.5 | 23.5 | 17.8 | 12.0 | 11.9 | 11.3 | 12.7 | 10.6 | 10.7 |
| 12 years | 80.6 | 76.3 | 80.6 | 75.7 | 71.6 | 70.5 | 67.8 | 64.4 | 61.0 | 53.3 | 42.8 | 45.6 | 44.6 | 44.1 | 38.4 | 43.5 |
| 13 years | 82.8 | 81.4 | 86.8 | 80.1 | 76.9 | 76.6 | 73.5 | 70.9 | 68.7 | 62.2 | 52.3 | 54.5 | 54.9 | 54.7 | 49.8 | 53.6 |
| 14 years | 87.2 | 87.1 | 91.0 | 86.7 | 84.0 | 86.3 | 81.1 | 78.2 | 76.2 | 72.4 | 64.2 | 68.4 | 67.6 | 69.2 | 63.3 | 62.5 |
| 15 years | 90.5 | 88.0 | 92.7 | 89.7 | 86.8 | 88.5 | 83.5 | 80.9 | 79.7 | 76.2 | 69.2 | 74.0 | 72.0 | 74.3 | 68.9 | 72.0 |
| 16 years | 96.7 | 95.6 | 96.5 | 94.7 | 93.6 | 94.8 | 90.9 | 90.9 | 89.3 | 88.2 | 85.9 | 88.4 | 88.6 | 89.6 | 89.8 | 89.9 |
| 17 years | 97.8 | 97.2 | 97.7 | 96.3 | 95.6 | 96.3 | 94.0 | 93.1 | 92.0 | 91.7 | 90.2 | 91.8 | 92.6 | 94.0 | 94.2 | 97.0 |
| 18 years | 98.2 | 97.8 | 98.7 | 98.4 | 98.1 | 97.9 | 96.9 | 96.4 | 96.1 | 95.2 | 94.9 | 96.3 | 96.7 | 97.3 | 96.8 | 99.4 |
| 19 years | 98.5 | 98.7 | 99.2 | 99.3 | 98.6 | 98.6 | 98.0 | 97.8 | 97.7 | 97.5 | 96.6 | 97.9 | 98.0 | 98.8 | 98.5 | 100 |
| 20 years | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

cumulative percentile of 76.3, based on the cumulative percentage distribution of respondents' years of education completed for respondents born in that cohort. Therefore, these respondents were assigned a score of 76.3. Having also 12 years of education, respondents born between 1966 and 1970, however, have a much lower score of 38.4. To control for changes in the selectivity of education over time, I control for these scores, which indicate respondents' relative educational standing within cohorts. For another example, according Table 2, an elementary school graduate with eight years of schooling, born between 1885 and 1900, has an educational standing comparable to that of a high school graduate born between 1941 and 1945. In this analysis, I treat these two persons as having the same "competitive advantage" with respect to their pre-existing characteristics, such as inherent cognitive ability and socioeconomic background, factors known to affect verbal ability.

A similar strategy, a ridit transformation of respondents' years of schooling completed, was used by Lieberson (1978: 960) to consider the possibility that education was a pseudocontrol in examining the income differences found between Southern-born Blacks living in the North and Northern-born Blacks. In some status attainment models such as the vacancy competition model, one's standing in educational attainment relative to others of the same cohort is often regarded as more important than one's absolute years of schooling completed in obtaining a high-paying job (Sørensen, 1979; Thurow, 1975; Sakamoto & Powers, 1995; Liu & Sakamoto, 2002). This is because one's relative level of educational attainment is regarded as a better measure of "competitive advantage."

A correlation coefficient of 0.522 between these percentile scores and vocabulary test scores was nearly equal to a correlation coefficient of 0.517 between years of education completed and vocabulary test scores. This indicates that relative standing within a cohort in educational attainment predicted verbal scores almost as well as the number of years of schooling completed. Respondents'

years of schooling completed and respondents' relative schooling standing within cohorts had a correlation of 0.89. This high correlation of 0.89 between absolute and relative levels of educational attainment was partially driven by a limitation in the measure of respondents' relative level of educational attainment. The limitation was that those who had extremely low or extremely high levels of educational attainment changed very little in their relative educational standing over time. For instance, the correlation coefficient was reduced to 0.77 when the sample was limited to those with nine to 15 years of schooling. For individuals with 11 to 13 years of schooling, the correlation between the absolute and the relative levels of schooling was as low as 0.53.

For each birth cohort in the GSS, 1974-1998, Figure 1 displays the percentage of respondents not completing high school, completing high school, completing some college, and completing a bachelor's degree or beyond. The amount of schooling completed increased significantly for individuals born before the late 1940s. For individuals born between the late 1940s and the early 1970s, the increase in the mean level of schooling completed was less significant. This implies that changes in educational selectivity on vocabulary test scores were largely confined to cohorts born before 1946.

In this paper, I address four research questions. (1) Has there been an intercohort decline in GSS vocabulary test scores? (2) What is the age trend in GSS vocabulary test scores? (3) Is the observed intercohort decline in vocabulary test scores attributable to either an aging effect or a period effect? (4) Does the intercohort decline in vocabulary test scores resemble the well-known test score decline of the 1960s and 1970s (Koretz, 1986) in its onset, end, and magnitude?

## V. Results

### A. Intercohort Trends in Vocabulary Test Scores

Figure 2 displays three intercohort trends in vocabulary test

Figure 1
Percentage of Completed Educational Degree by Birth Cohort
(GSS, 1974-1998)

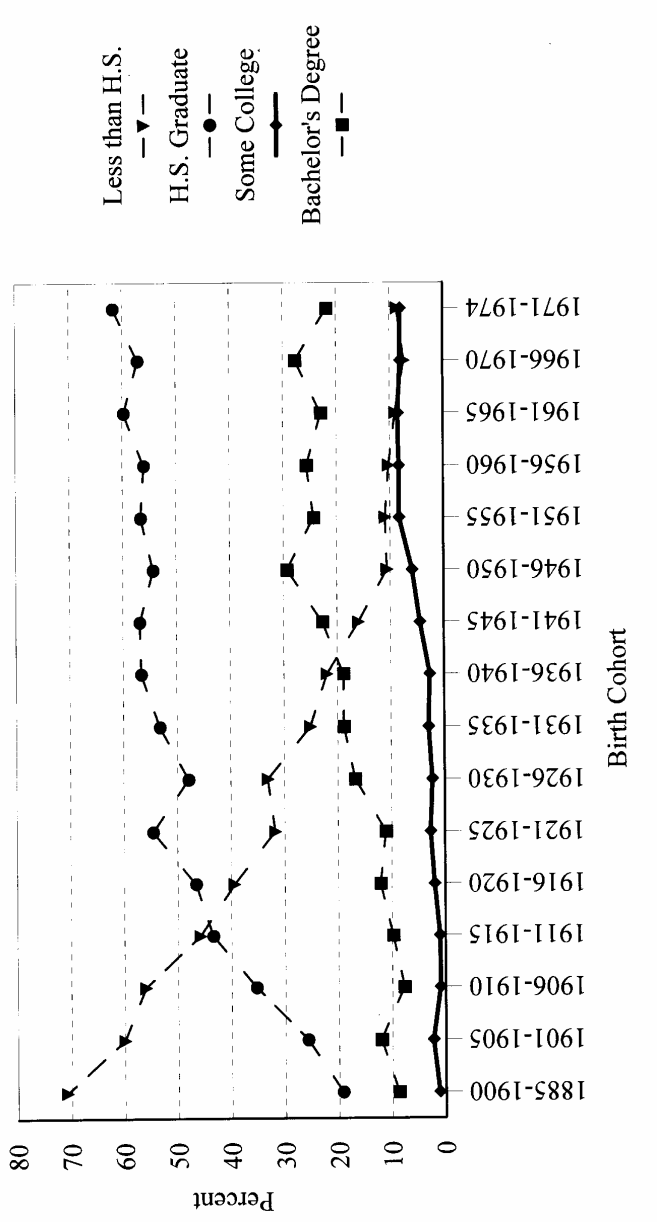
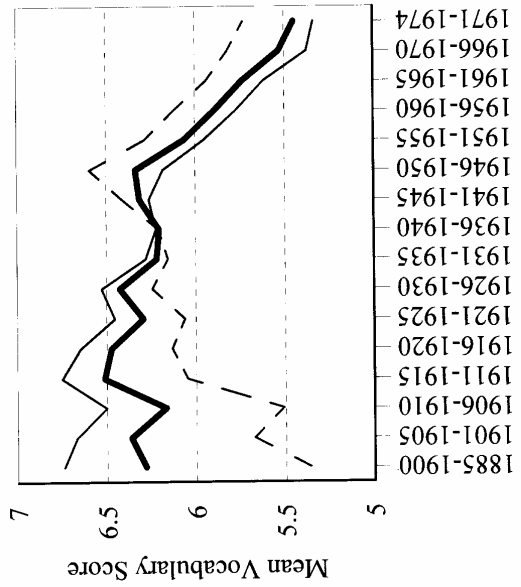


Figure 2
Mean Vocabulary Scores by Birth Cohort, Ages 24-89
(GSS, 1974-1998)




Note: The vocabulary scores range from zero to ten.

scores based on a baseline model, an education model, and a standing-adjusted model. The baseline model used included birth cohort only:

$$E[y] = \alpha + \sum_{1}^{I} \beta_i x_i \quad (1)$$

where y was the number of correct vocabulary test scores items, α was the intercept, the $x_i$ were a set of 15 dummy variables representing 16 birth cohorts (the 1946-1950 cohort was the reference group), and the $\beta_i$ were cohort effects. Table 3 reports the age range within each birth cohort. In the education model, a continuous variable for respondents' years of education was added to the baseline model. In the standing-adjusted model, respondents' relative educational standing within cohorts was added to the education model.[2]

As displayed in Figure 2, the education model revealed a long-term intercohort decline in vocabulary test scores, a decline consistent to that reported by Alwin (1991). The decline began with the 1911-1915 birth cohort and continued to the most recent cohort, persons born in 1971-1974. This long-term intercohort decline in vocabulary test scores was recognized by Glenn (1999) and Alwin and McCammon (1999), though interpreted as spurious by Wilson and Gove (1999a, 1999b). Wilson and Gove (1999a, 1999b) came to this conclusion because (1) changes in educational selectivity over time were not taken into account; (2) the intercohort decline manifest in the GSS vocabulary scores stood out as an anomaly, because it was inconsistent with other studies whose results suggested a relatively modest increase in verbal IQ (Flynn 1984a, 1984b, 1987, 1994, 1998); and (3) the observed intercohort decline in education-adjusted vocabulary test scores was, in fact, a mirror image of genuine age-related accumulations in vocabulary knowledge.

The standing-adjusted model, unlike the education model,

[2] For these three models, Table A1 in the appendix shows the regression estimates.

Table 3
Age Range by Birth Cohort and Range of Birth Years by Age Category:
GSS, 1974-1998

| Birth Cohort | N | Age Range | | Age Category | N | Range of Birth Year | |
|---|---|---|---|---|---|---|---|
| | | Minimum | Maximum | | | Minimum | Maximum |
| 1885-1900 | 273 | 74 | 89 | 84-89 | 253 | 1885 | 1914 |
| 1901-1905 | 317 | 69 | 89 | 80-83 | 277 | 1891 | 1918 |
| 1906-1910 | 521 | 64 | 89 | 76-79 | 463 | 1895 | 1922 |
| 1911-1915 | 738 | 59 | 87 | 72-75 | 639 | 1899 | 1926 |
| 1916-1920 | 951 | 54 | 82 | 68-71 | 739 | 1903 | 1930 |
| 1921-1925 | 1,117 | 49 | 77 | 64-67 | 766 | 1907 | 1934 |
| 1926-1930 | 1,071 | 44 | 72 | 60-63 | 871 | 1911 | 1938 |
| 1931-1935 | 1,026 | 39 | 67 | 56-59 | 891 | 1915 | 1942 |
| 1936-1940 | 1,200 | 34 | 62 | 52-55 | 985 | 1919 | 1946 |
| 1941-1945 | 1,573 | 29 | 57 | 48-51 | 1,070 | 1923 | 1950 |
| 1946-1950 | 1,964 | 24 | 52 | 44-47 | 1,170 | 1927 | 1954 |
| 1951-1955 | 1,831 | 24 | 47 | 40-43 | 1,389 | 1931 | 1958 |
| 1956-1960 | 1,672 | 24 | 42 | 36-39 | 1,560 | 1935 | 1962 |
| 1961-1965 | 1,147 | 24 | 37 | 32-35 | 1,701 | 1939 | 1966 |
| 1966-1970 | 586 | 24 | 32 | 28-31 | 1,716 | 1943 | 1970 |
| 1971-1974 | 168 | 24 | 27 | 24-27 | 1,665 | 1947 | 1974 |

takes into consideration the fact that selectivity of education changes over time. As displayed in Figure 2, results from the standing-adjusted model suggest that a systematic intercohort decline in the GSS vocabulary scores started much later, beginning with persons born in the period around 1946 to 1950. The intercohort decline manifest in the GSS vocabulary scores does not stand out as an anomaly in the presence of the "Flynn effect." The decline in test scores during the 1960s and 1970s is well-known. As Koretz (1986) suggested:

> Although not all indicators of educational achievement showed large declines over the past two decades, the great majority did, leaving no question that the decline was real and not an artifact of specific tests. The decline was widespread, appearing among many types of students, on many types of tests, in many subject areas, and in all parts of the nation. Moreover, in many instances, the decline was large enough to be of serious educational concern. (p. 31)

According to Koretz, the test score decline started with individuals born in 1946 for the Scholastic Aptitude Test (SAT), in 1949 for American College Testing Program (ACT) tests, and in 1951 for the Iowa Tests of Educational Development (ITED) Grade 12 and the Minnesota Scholastic Aptitude Test administered to high school juniors. Therefore, the intercohort decline in vocabulary test scores closely resembled the test score decline in the 1960s and 1970s at its onset. The test score decline ended with persons born in 1962 for the SAT, in 1958 for the ACT, and in 1962 for ITED Grade 12. From the present analysis, however, Figure 2 shows that the intercohort decline in GSS vocabulary test scores, after controlling for respondents' relative educational standing within cohorts, ended with persons born in 1970-1974. This is because those born later are younger, and younger people tend to have lower verbal scores. When age was properly controlled, as will be shown later, the vocabulary test scores intercohort decline ended at earlier cohorts. The resemblance between the intercohort decline in vocabulary test scores and the

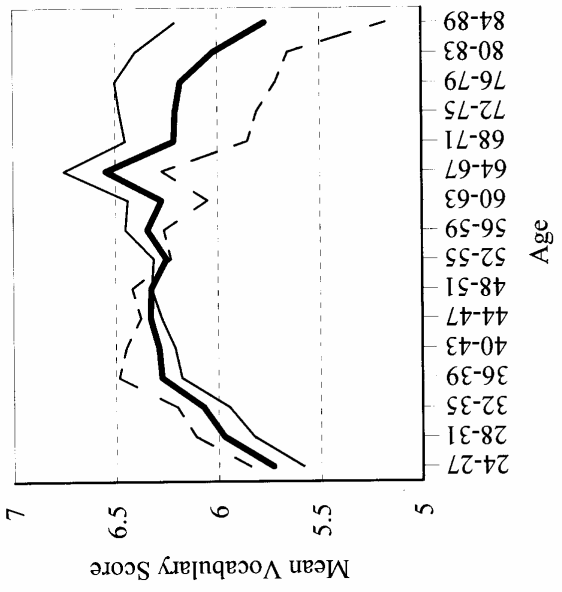
widespread test score decline during the 1960s and 1970s will be discussed separately.

## B. Age Trends in Vocabulary Test Scores

To display age trends in vocabulary test scores, the same strategy used to display intercohort trends in vocabulary test scores was used. I regressed vocabulary test scores on a set of 15 dummy variables representing 16 age categories (the 28-31 age category was the reference group), with and without controlling for the amount of schooling and respondents' relative educational standing within cohorts. Table 3 reports the range of birth years for each age category. Based on three sets of regressions, Figure 3 displays three age trends in vocabulary test scores.[3] After controlling for the amount of schooling completed, vocabulary knowledge increased with age until individuals reached their mid-sixties. This is the age trend that Wilson and Gove (1999a) believed to be intrinsic and consistent with the literature on cognitive psychology. But after controlling for respondents' relative educational standing within cohorts, as shown in Figure 3, aging appeared to increase one's vocabulary knowledge before the late-thirties. Between one's late-thirties and early sixties, there was no significant increase in vocabulary knowledge. The rise in the mid-sixties was probably attributable to the fact that persons still surviving tend to have higher verbal ability. That is, those who died a few years before they reach their mid-sixties were likely to have lower verbal ability than those who survived. The drop in verbal ability after the mid-sixties, on the other hand, resulted from the biological effects of aging. Without taking into account changes in the selectivity of education over time, Wilson and Gove overestimated aging effects during middle age.

[3] The regression estimates for these three sets of regressions are presented in Table A2 of the appendix.

Figure 3
Mean Vocabulary Scores by Age Category, Ages 24-89
(GSS, 1974-1998)




Note: The vocabulary scores range from zero to ten.

This is further demonstrated in Table 4. Table 4 shows the effects of age independent of period effects. Table 4 was modeled on Wilson and Gove's Table 2 (1999a: 262). As such, the extent to which Wilson and Gove's (1999a) findings were affected by controlling for changes in the selectivity of education over time could be examined. As with Wilson and Gove's table, four age categories (24-35, 36-50, 51-65 & 66-89) were created to minimize the correlation between age and the year of survey.[4] In Table 4, there are four regression models for each age category. Model 1 shows that the effects of aging on vocabulary test scores were positive for ages 24 to 35 and negative for ages 66 to 89. Controlling for the year of survey additionally, as shown in Model 2, did not change the effects of aging since age and survey year were not significantly correlated. Wilson and Gove (1999a) suggested that an intrinsic age trend in vocabulary test scores should appear when vocabulary test scores are regressed on age and years of education. Indeed, Model 3 shows that aging effects were positive for age categories 24 to 35, 36 to 50, and 51 to 65, but negative for the oldest age category, 66-89. However, after holding changes in the selectivity of education constant by controlling for respondents' relative educational standing within cohorts, as shown in Model 4, the effects of aging become insignificant between the ages of 36 and 50 and between 51 and 65. Therefore, the regression results presented in Table 4 are consistent with the standing-adjusted age trend displayed in Figure 3. The results also suggest that the standing-adjusted age trend displayed in Figure 3 was not significantly affected by a period effect because aging effects in Table 4 were independent of period effects.

[4] For the age categories, 24-35, 36-50, 51-65, and 66-89, respectively, the correlations between age and survey year were 0.079, -0.004, -0.018, and 0.086; the correlations between age and birth year were -0.36, -0.5, -0.5, and -0.57; and the correlations between birth year and year of survey were 0.9, 0.87, 0.88, and 0.77.

Table 4
Unstandardized Coefficients from the Regression of Vocabulary Test Scores on Age, Survey Year, Education, and Relative Educational Standing within Cohorts by Age: GSS, 1974-1998

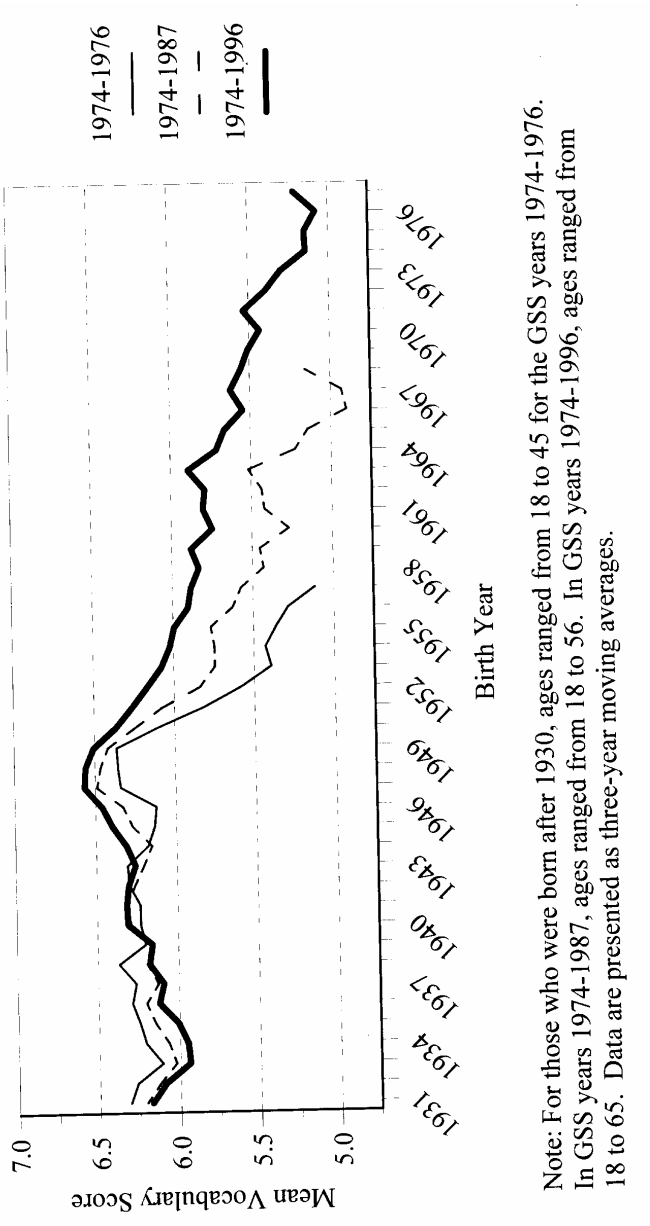
| Independent Variables | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| *Ages 24 to 35 (N=5,082)* | | | | |
| Age | 0.043*** | 0.046*** | 0.047*** | 0.045*** |
| Survey Year | ---- | -0.015*** | -0.029*** | -0.027*** |
| Education | ---- | ---- | 0.408*** | 0.250*** |
| Educational Standing | ---- | ---- | ---- | 0.015*** |
| $R^2$ | 0.006 | 0.009 | 0.268 | 0.272 |
| *Ages 36 to 50 (N=4,924)* | | | | |
| Age | -0.005 | -0.004 | 0.013* | 0.007 |
| Survey Year | ---- | 0.015*** | -0.023*** | -0.012** |
| Education | ---- | ---- | 0.411*** | 0.237*** |
| Educational Standing | ---- | ---- | ---- | 0.019*** |
| $R^2$ | 0.000 | 0.003 | 0.306 | 0.314 |
| *Ages 51 to 65 (N=3,399)* | | | | |
| Age | -0.014 | -0.014 | 0.018* | 0.008 |
| Survey Year | ---- | 0.003 | -0.030*** | -0.022*** |
| Education | ---- | ---- | 0.378*** | 0.232*** |
| Educational Standing | ---- | ---- | ---- | 0.017*** |
| $R^2$ | 0.001 | 0.001 | 0.281 | 0.288 |
| *Ages 66 to 89 (N=2,750)* | | | | |
| Age | -0.041*** | -0.043*** | -0.018** | -0.026*** |
| Survey Year | ---- | 0.018** | -0.021*** | -0.014** |
| Education | ---- | ---- | 0.362*** | 0.259*** |
| Educational Standing | ---- | ---- | ---- | 0.013** |
| $R^2$ | 0.011 | 0.014 | 0.288 | 0.291 |

Note: The regression models are based on a two-sided Tobit Specification which takes the censored distribution of the vocabulary test scores into account. $R^2$ statistics are based on OLS models.
$^*p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$

Wilson and Gove (1999a, 1999b) also presented three misleading figures to support their thesis of lifelong acquisition in vocabulary until old age. In each figure, Wilson and Gove showed mean vocabulary score by birth year for respondents born between 1945 and 1976 for the three sets of selected GSS survey years. In the first two figures (Wilson & Gove, 1999a, Figures 3-4, pp. 259-260), these three sets of selected GSS survey years were: 1974-1976, 1974-1987, and 1974-1996. In the last figure (Wilson & Gove, 1999b, Figure 2, p. 292), these three sets of GSS survey years were 1974-1978, 1984-1988, and 1993-1996. By adding data from later survey years, Wilson and Gove included older respondents from the same cohorts. For respondents from the same birth cohorts, Wilson and Gove found that older respondents in later survey years had higher vocabulary scores than younger respondents in earlier survey years. Therefore, Wilson and Gove suggested that the patterns indicated in these figures were consistent with the literature that vocabulary knowledge continues to increase over the life course until age 60 or later (Wilson & Gove, 1999a: 260; Wilson & Gove, 1999b: 291).

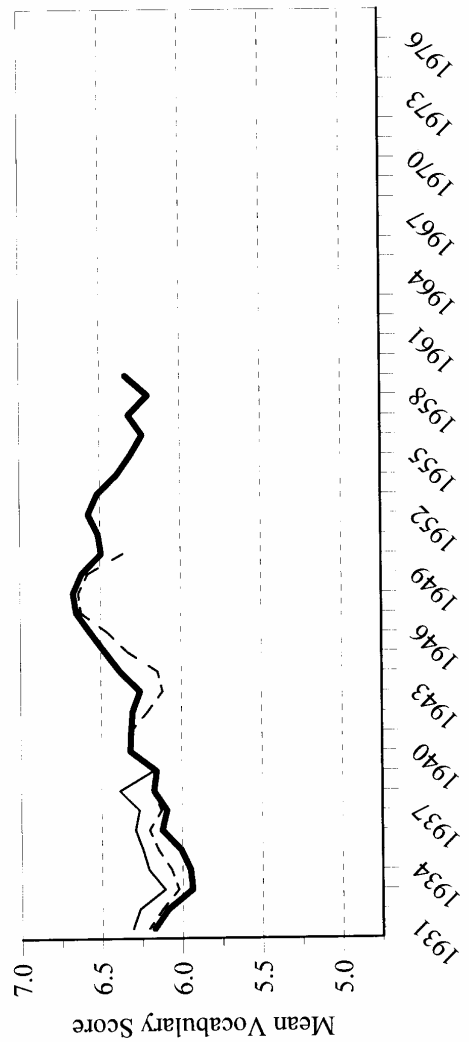
In fact, Wilson and Gove's figures did not support a long-term positive aging effect which continues until persons reach their sixties. The observed significant age-related increases in vocabulary scores in Wilson and Gove's figures were largely dominated by vocabulary increases between age 18 and age 36. If Wilson and Gove were to exclude respondents who were younger than 36 years old, they would have found no significant age-related increases in vocabulary scores. For example, modeled on Wilson and Gove's (1999a) Figure 3, I created Figure 4 which presents mean unadjusted vocabulary score by birth year for adults ages 18 to 65 who were born after 1930, for three sets of selected GSS survey years: 1974-1976, 1974-1987, and 1974-1996. While the data in Wilson and Gove's (1999a) Figure 3 were restricted to respondents ages 18 to 51 who were born after 1944, Figure 4 of the present study extended the trend lines by including respondents ages 18 to 65 who were born after 1930. As with Wilson and Gove, no sampling weights were used. For

Figure 4
Mean Unadjusted Vocabulary Score by Birth Year for Selected GSS Survey Years
Unweighted, Ages 18-65, Adults Born after 1930




Note: For those who were born after 1930, ages ranged from 18 to 45 for the GSS years 1974-1976. In GSS years 1974-1987, ages ranged from 18 to 56. In GSS years 1974-1996, ages ranged from 18 to 65. Data are presented as three-year moving averages.

Figure 5
Mean Unadjusted Vocabulary Score by Birth Year for Selected GSS Survey Years
Unweighted, Ages 36-65, Adults Born after 1930




Note: For those who were born after 1930, ages ranged from 36 to 45 for the GSS years 1974-1976. In GSS years 1974-1987, ages ranged from 36 to 56. In GSS years 1974-1996, ages ranged from 36 to 65. Data are presented as three-year moving averages.

respondents born after 1944, the results shown in the present study's Figure 4 were consistent with those presented in Wilson and Gove's (1999a) Figure 3, which indicated that cohorts' mean scores were higher for older respondents in later survey years than for younger respondents in earlier survey years. However, when respondents younger than 36 years old were excluded, Figure 5 of the present study indicated that the age-related increases in vocabulary scores were no longer significant. Thus, it was inaccurate for Wilson and Gove to argue that their figures support a long-term positive aging effect which continues until one's late-sixties. The results shown in the present study's Figures 4 and 5 indicate, again, that aging increased one's vocabulary knowledge before the late-thirties. After one's late-thirties, there were no age-related increases in vocabulary. A flat trend during middle age, between one's late-thirties to early sixties is, in fact, consistent with the literature on adult cognitive development. For example, Schaie (1996: 107-136) used longitudinal data from the Seattle Longitudinal Study and found that verbal ability did not change significantly during middle age.

## C. Are the Cohort Effects Spurious Due to Aging?

As shown in Figure 2, the intercohort decline in GSS verbal scores, after controlling for years of education and changes in educational selection over time, began with the 1946-1950 birth cohort and was sustained through the cohorts of the early 1970s. One may suspect that (1) the uninterrupted vocabulary test scores decline for these birth cohorts may be primarily attributable to the fact that more recent cohorts were also younger and therefore had lower scores, or (2) the intercohort decline in vocabulary scores was mainly due to a period effect. When birth cohorts were limited to persons born between 1946 and 1974, age, birth year, and year of survey correlated only moderately with each other.[5] Hence, it

[5] The correlation between age and year of survey was 0.48. It was -0.55 for age and birth year and 0.47 for year of survey and birth year.

was possible to examine whether the verbal score decline for cohorts born after 1950 was primarily due to an aging effect or a period effect. To address these two questions, the following specification in the cohort model was used:

$$E[y] = \alpha + \sum_{1}^{I} \beta_i w_i + \delta_1 z_1 \qquad (2)$$

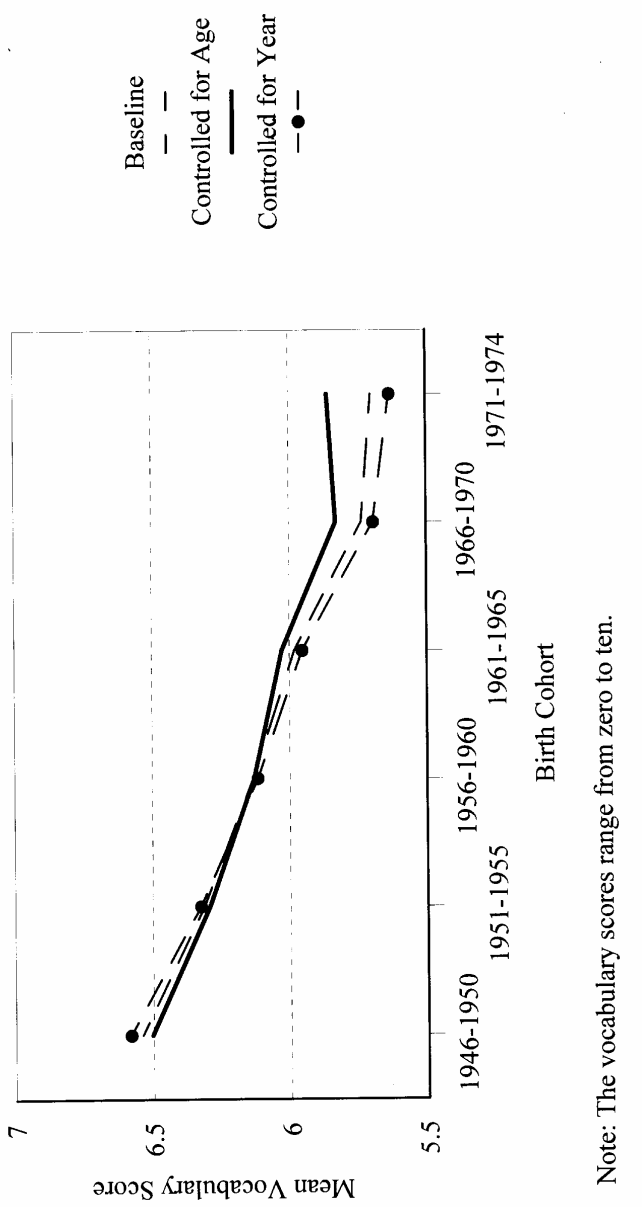
where y was the number of correct vocabulary test scores items, $\alpha$ was the intercept, the $w_i$ were dummy variables representing six birth cohorts (1946-1950, 1951-1955, 1956-1960, 1961-1965, 1966-1970, 1971-1974, where the 1946-1950 cohort was the omitted category), the $\beta_i$ were cohort effects, $z_1$ was a continuous variable for years of schooling completed, and the $\delta_1$ represented the effects of schooling. In this model, respondents' relative educational standing within cohorts was not included as an independent variable since changes in the selectivity of education were not significant for cohorts born after 1945, see Figure 1.

To examine whether aging accounted for the intercohort decline, a set of age dummy variables representing seven age categories (24-27, 28-31, 32-35, 36-39, 40-43, 44-47, 48-52, where the 24-27 age category was the omitted group) were added to the cohort model. To examine whether intercohort decline was due to a period effect, a set of one-year survey year dummy variables (1996 was the omitted category) was added to the cohort model.

Figure 6 presents the results.[6] Considerably lower than what was suggested by Wilson and Gove (1999a, 1999b), aging accounted for only about 23 percent of the intercohort decline in vocabulary test scores for persons born between 1946 and 1974. For a shorter time range, persons born between 1946 and 1970,

[6] The regression estimates for these three sets of regressions are presented in Table A3 of the appendix.

Figure 6
Education-Controlled Mean Vocabulary Test Scores by Birth Cohort, Ages 24-52
(GSS, 1974-1998)




Note: The vocabulary scores range from zero to ten.

aging accounted for only 16 percent of the intercohort decline. After controlling for age, the end of the intercohort decline took place in the 1966-1970 cohort.

I controlled for year of survey to see whether the intercohort decline was due to a period effect. However, as shown in Figure 6, the intercohort decline in vocabulary test scores persisted after controlling for year of survey.

## D. Some Characteristics of the Intercohort Decline in Vocabulary Test Scores

To what extent does the intercohort decline in vocabulary test scores resemble the well-known test score decline of the 1960s and 1970s? Because interchort trends displayed in Figures 2 and 6 were based on 5-year birth cohort aggregates, one cannot be certain as to which year the onset and the end of the decline took place. As reported earlier, the onset of the intercohort decline in vocabulary test scores took place at the 1946-1950 birth cohort. Because the year-specific (birth year) mean vocabulary test scores were the highest for 411 persons born in 1948, the vocabulary test scores intercohort decline approximately resembles the test score decline of the 1960s and 1970s at its onset.

With respect to the end of the decline in test scores during the 1960s and 1970s, Koretz (1986) suggested that it ended with persons born in the late 1950s (for students in lower grades) or early 1960s (for students in higher grades). The age-controlled intercohort trend displayed in Figure 6, which is based on 5-year birth cohort aggregates, suggests that the decline ended with persons born in the 1966-1970 cohort. It is possible, however, that the true end year of the decline occurred earlier. For example, the year-specific (birth year) mean vocabulary test scores was as low as 5.68 for the 243 persons born in 1961.

According to the age-controlled interchort trend displayed in Figure 6, the magnitude of the intercohort decline in vocabulary test scores from the 1946-1950 birth cohort to the 1966-1970 birth cohort was 0.36 standard deviation units, based on the

standard deviation of vocabulary test scores in the 1966-1970 birth cohort. Based on the standard deviation of vocabulary test scores in the 1946-1950 cohort, the magnitude of the decline was 0.33 standard deviation units. The size of the decline in GSS vocabulary test scores was smaller than the decline of 0.48 standard deviations for the SAT-Verbal and a decline of 0.4 standard deviations for the ITED-Reading Grade 12. The larger decline found for the SAT-Verbal was attributable to sample selection and changes in the composition of test-takers. Koretz (1986) concluded that the magnitude of the test score decline averaged about 0.3 standard deviations across subjects and tests over the entire period of the test score decline. Given that verbal tests evidenced a larger decline than tests of other subjects, the size of the intercohort decline in vocabulary test scores is consistent with the average size of declines found in other test data sets. Therefore, the intercohort decline in GSS vocabulary test scores, after controlling for absolute and relative level of educational attainment and aging effects, was similar to the well-known test score decline of the 1960s and 1970s in its onset, end, and magnitude.

## VI. Conclusions

Controlling for respondents' years of schooling completed, Alwin (1991) presented a long-term intercohort decline in GSS vocabulary test scores beginning with persons born early in the twentieth century. Alwin's (1991) model assumes that persons born in recent years are as selective as persons born in earlier years with respect to the same level of educational attainment. This is questionable because the average years of education completed has changed significantly for individuals born in the first half of the twentieth century. Furthermore, the plausibility of such a long-term intercohort decline in verbal ability is challenged by the fact that it is not found in other data sets.

In this study, I considered the possibility that the selectivity of education may change in correspondence to changes in the average years of schooling completed over time. After controlling for

temporal changes in the selectivity of education, the intercohort decline started much later, beginning with individuals born between 1946 and 1950. Therefore, the onset of the decline is consistent with that of the widespread test score decline of the 1960s and 1970s. After aging effects were controlled additionally, the intercohort decline in GSS vocabulary test scores resembled the massive test score decline of the 1960s and 1970s in its onset, end, and magnitude.

In previous research, the focus of the debate was on whether the observed intercohort decline in GSS vocabulary test scores was actually an increase in vocabulary as people age. Results of this study suggest that (1) vocabulary increases before individuals reach their late-thirties and declines after they reach their late-sixties and there is no significant age-related increase in vocabulary between the late-thirties and the early-sixties; (2) the intercohort decline in vocabulary test scores is not primarily due to age-related increases in vocabulary over the life course; and (3) the observed significant age-related increases in vocabulary scores in Wilson and Gove's figures (1999a, Figures 3-4, pp. 259-260; 1999b, Figure 2, p. 292) were not caused by a long-term positive aging effect that continues into the late sixties, but appear to be attributable to vocabulary increases between the ages of 18 and 36. In sum, the results of this study weaken the age interpretation.

The intercohort decline in verbal ability may not be spurious as interpreted by some earlier researchers. Alternative explanations for the decline may exist. For example, Hayes, Wolfer, and Wolfe (1996) reported that schoolbook publishers after World War II reduced their use of the more rare words to the extent that current sixth, seventh, and eighth grade schoolbooks were simpler than fifth grade schoolbooks used before World War II. These simplified schoolbooks were used by baby boomers (persons born between 1947 and 1961) and subsequent cohorts. Therefore, Hayes and his colleagues hypothesized that the declining mean SAT-verbal scores since 1963 may be partly attributed to the widespread decline in the difficulty of schoolbooks in elementary, middle, and high schools after World War II.

One should also take note that there were only ten test items in the GSS vocabulary test. Trends in GSS vocabulary test scores could be biased if some words in the test have become more difficult over time, independent of other changes in verbal ability in the general population.

## Appendix

Table A1
Effects of Birth Cohort on Vocabulary Test Scores (with and without controlling for respondents' years of education and relative educational standing within cohorts): GSS, 1974-1998

| | Baseline | | | Education | | | Standing-Adjusted | | |
|---|---|---|---|---|---|---|---|---|---|
| | *b* | *SE* | *p* | *b* | *SE* | *P* | *b* | *SE* | *p* |
| Intercept | 6.591 | 0.051 | 0.000 | 0.867 | 0.084 | 0.000 | 1.966 | 0.136 | 0.000 |
| Birth year | | | | | | | | | |
| 1885-1900 | -1.229 | 0.159 | 0.000 | 0.562 | 0.136 | 0.000 | -0.049 | 0.148 | 0.742 |
| 1901-1905 | -0.919 | 0.146 | 0.000 | 0.491 | 0.125 | 0.000 | 0.032 | 0.132 | 0.81 |
| 1906-1910 | -1.072 | 0.117 | 0.000 | 0.331 | 0.101 | 0.000 | -0.155 | 0.111 | 0.164 |
| 1911-1915 | -0.544 | 0.101 | 0.000 | 0.572 | 0.087 | 0.000 | 0.182 | 0.094 | 0.054 |
| 1916-1920 | -0.456 | 0.091 | 0.000 | 0.469 | 0.078 | 0.000 | 0.143 | 0.084 | 0.088 |
| 1921-1925 | -0.533 | 0.084 | 0.000 | 0.273 | 0.072 | 0.000 | -0.041 | 0.078 | 0.598 |
| 1926-1930 | -0.361 | 0.084 | 0.000 | 0.338 | 0.072 | 0.000 | 0.083 | 0.076 | 0.273 |
| 1931-1935 | -0.442 | 0.085 | 0.000 | 0.096 | 0.072 | 0.184 | -0.118 | 0.075 | 0.116 |
| 1936-1940 | -0.386 | 0.081 | 0.000 | 0.043 | 0.069 | 0.534 | -0.128 | 0.071 | 0.069 |
| 1941-1945 | -0.185 | 0.076 | 0.015 | 0.079 | 0.064 | 0.218 | -0.022 | 0.065 | 0.737 |
| 1946-1950 | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1951-1955 | -0.307 | 0.074 | 0.000 | -0.228 | 0.063 | 0.000 | -0.269 | 0.062 | 0.000 |

(to be continued)

Table A1 (continued)
Effects of Birth Cohort on Vocabulary Test Scores (with and without controlling for respondents' years of education and relative educational standing within cohorts): GSS, 1974-1998

| | Baseline | | | Education | | | Standing-Adjusted | | |
|---|---|---|---|---|---|---|---|---|---|
| | *b* | *SE* | *p* | *b* | *SE* | *P* | *b* | *SE* | *p* |
| 1956-1960 | -0.475 | 0.076 | 0.000 | -0.410 | 0.064 | 0.000 | -0.444 | 0.064 | 0.000 |
| 1961-1965 | -0.652 | 0.084 | 0.000 | -0.560 | 0.071 | 0.000 | -0.596 | 0.071 | 0.000 |
| 1966-1970 | -0.768 | 0.107 | 0.000 | -0.800 | 0.091 | 0.000 | -0.799 | 0.090 | 0.000 |
| 1971-1974 | -0.864 | 0.180 | 0.000 | -0.843 | 0.152 | 0.000 | -0.887 | 0.151 | 0.000 |
| Years of Schooling | --- | --- | --- | 0.417 | 0.005 | 0.000 | 0.268 | 0.016 | 0.000 |
| Educational Standing | --- | --- | --- | --- | --- | --- | 0.016 | 0.002 | 0.000 |

Note: The regression models are based on a two-sided Tobit specification which takes the censored distribution of the vocabulary test scores into account. The baseline model has a log-likelihood ratio of -35,728. It is -33080 for the education model and -33,029 for the standing-adjusted model.

Table A2
Effects of Aging on Vocabulary Test Scores (with and without controlling for respondents' years of education and relative educational standing within cohorts): GSS, 1974-1998

| | Baseline | | | Education | | | Standing-Adjusted | | |
|---|---|---|---|---|---|---|---|---|---|
| | *b* | *SE* | *P* | *b* | *SE* | *P* | *b* | *SE* | *p* |
| Intercept | 6.110 | 0.055 | 0.000 | 0.670 | 0.085 | 0.000 | 1.853 | 0.123 | 0.000 |
| Age | | | | | | | | | |
| 84-89 | -0.930 | 0.174 | 0.000 | 0.385 | 0.149 | 0.010 | -0.195 | 0.155 | 0.208 |
| 80-83 | -0.435 | 0.166 | 0.009 | 0.590 | 0.142 | 0.000 | 0.068 | 0.146 | 0.642 |
| 76-79 | -0.387 | 0.131 | 0.003 | 0.682 | 0.112 | 0.000 | 0.219 | 0.117 | 0.061 |
| 72-75 | -0.301 | 0.110 | 0.006 | 0.654 | 0.094 | 0.000 | 0.237 | 0.099 | 0.016 |
| 68-71 | -0.250 | 0.103 | 0.015 | 0.630 | 0.089 | 0.000 | 0.249 | 0.093 | 0.007 |
| 64-67 | 0.169 | 0.100 | 0.093 | 0.932 | 0.086 | 0.000 | 0.584 | 0.090 | 0.000 |
| 60-63 | -0.056 | 0.095 | 0.557 | 0.618 | 0.081 | 0.000 | 0.309 | 0.084 | 0.000 |
| 56-59 | 0.149 | 0.092 | 0.107 | 0.621 | 0.079 | 0.000 | 0.370 | 0.081 | 0.000 |
| 52-55 | 0.112 | 0.088 | 0.204 | 0.481 | 0.075 | 0.000 | 0.275 | 0.076 | 0.000 |
| 48-51 | 0.304 | 0.086 | 0.000 | 0.495 | 0.073 | 0.000 | 0.353 | 0.073 | 0.000 |
| 44-47 | 0.265 | 0.084 | 0.002 | 0.448 | 0.072 | 0.000 | 0.358 | 0.071 | 0.000 |
| 40-43 | 0.341 | 0.082 | 0.000 | 0.385 | 0.069 | 0.000 | 0.319 | 0.069 | 0.000 |
| 36-39 | 0.375 | 0.081 | 0.000 | 0.353 | 0.069 | 0.000 | 0.305 | 0.068 | 0.000 |
| 32-35 | 0.096 | 0.079 | 0.223 | 0.126 | 0.067 | 0.060 | 0.107 | 0.066 | 0.109 |
| 28-31 | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 24-27 | -0.267 | 0.078 | 0.001 | -0.241 | 0.066 | 0.000 | -0.240 | 0.066 | 0.000 |
| Years of Schooling | --- | --- | --- | 0.405 | 0.005 | 0.000 | 0.235 | 0.014 | 0.000 |
| Educational Standing | --- | --- | --- | --- | --- | --- | 0.019 | 0.001 | 0.000 |

Note: The regression models are based on a two-sided Tobit specification which takes the censored distribution of the vocabulary test scores into account. The baseline model has a log-likelihood ratio of -35,729. It is -33,144 for the education model and -33,059 for the standing-adjusted model.

Table A3
Effects of Birth Cohort on Vocabulary Test Scores (with and without controlling for age and year of survey): GSS, 1974-1998

| | Baseline | | | Baseline+Age | | | Baseline+Year | | |
|---|---|---|---|---|---|---|---|---|---|
| | *b* | *SE* | *P* | *B* | *SE* | *p* | *b* | *SE* | *p* |
| Intercept | 0.659 | 0.121 | 0.000 | 0.501 | 0.128 | 0.000 | 0.781 | 0.140 | 0.000 |
| Birth year | | | | | | | | | |
| 1946-1950 | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1951-1955 | -0.223 | 0.059 | 0.000 | -0.208 | 0.062 | 0.001 | -0.254 | 0.061 | 0.000 |
| 1956-1960 | -0.405 | 0.060 | 0.000 | -0.372 | 0.066 | 0.000 | -0.466 | 0.064 | 0.000 |
| 1961-1965 | -0.552 | 0.067 | 0.000 | -0.469 | 0.074 | 0.000 | -0.625 | 0.072 | 0.000 |
| 1966-1970 | -0.796 | 0.085 | 0.000 | -0.663 | 0.093 | 0.000 | -0.885 | 0.092 | 0.000 |
| 1971-1974 | -0.837 | 0.143 | 0.000 | -0.637 | 0.152 | 0.000 | -0.950 | 0.151 | 0.000 |
| R's Education | 0.432 | 0.008 | 0.000 | 0.429 | 0.008 | 0.000 | 0.430 | 0.008 | 0.000 |
| Age | | | | | | | | | |
| 48-52 | | | | 0.290 | 0.129 | 0.025 | | | |
| 44-47 | | | | 0.242 | 0.101 | 0.016 | | | |
| 40-43 | | | | 0.197 | 0.082 | 0.016 | | | |
| 36-39 | | | | 0.319 | 0.075 | 0.000 | | | |
| 32-35 | | | | 0.176 | 0.069 | 0.011 | | | |
| 28-31 | | | | 0.152 | 0.065 | 0.019 | | | |
| 24-27 | | | | --- | --- | --- | | | |
| Survey Year | | | | | | | | | |
| 1974 | | | | | | | -0.137 | 0.147 | 0.351 |
| 1976 | | | | | | | -0.467 | 0.129 | 0.000 |

(to be continued)

Table A3 (continued)
Effects of Birth Cohort on Vocabulary Test Scores (with and without controlling for age and year of survey): GSS, 1974-1998

| | Baseline | | | Baseline+Age | | | Baseline+Year | | |
|---|---|---|---|---|---|---|---|---|---|
| | *b* | *SE* | *P* | *B* | *SE* | *p* | *b* | *SE* | *p* |
| 1978 | | | | | | | -0.113 | 0.117 | 0.335 |
| 1982 | | | | | | | -0.055 | 0.100 | 0.584 |
| 1984 | | | | | | | -0.076 | 0.103 | 0.461 |
| 1987 | | | | | | | -0.104 | 0.092 | 0.259 |
| 1988 | | | | | | | -0.101 | 0.110 | 0.362 |
| 1989 | | | | | | | -0.005 | 0.106 | 0.962 |
| 1990 | | | | | | | -0.079 | 0.110 | 0.469 |
| 1991 | | | | | | | 0.067 | 0.104 | 0.520 |
| 1993 | | | | | | | 0.003 | 0.099 | 0.978 |
| 1994 | | | | | | | -0.049 | 0.084 | 0.555 |
| 1996 | | | | | | | --- | --- | --- |
| 1998 | | | | | | | 0.038 | 0.088 | 0.669 |

Note: The regression models are based on a two-sided Tobit specification which takes the censored distribution of the vocabulary test scores into account. The baseline model has a log-likelihood ratio of -14,713. It is -14,704 for the baseline+age model and -14,704 for the baseline+year model.

## References

Alwin, D. F. (1991). Family of origin and cohort differences in verbal ability. *American Sociological Review*, 56: 625-638.

Alwin, D. F., & McCammon, R. J. (1999). Aging versus cohort interpretations of intercohort differences in GSS vocabulary scores. *American Sociological Review*, 64: 272-286.

Alwin, D. F., & McCammon, R. J. (2001). Aging, cohorts, and verbal ability. *Journal of Gerontology*, 56B: S151-S161.

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics*, 24: 3-61.

Bishop, J. H. (1989). Is the test score decline responsible for the productivity growth decline? *American Economic Review,* 79, 1: 178-197.

Bracey, G. (1991). Why can't they be like we were? *Phi Delta Kappan*, 73, 2: 104-117.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54: 1-22.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Davis, J. A., & Smith T. W. (1992). *The NORC General Social Survey: A user's guide*. Newbury Park, CA: Sage Publications.

Firebaugh, G. (1997). *Analyzing repeated surveys*. Thousand Oaks, CA: Sage Publications.

Flynn, J. R. (1984a). IQ gains and the Binet decrements. *Journal of Educational Measurement,* 21: 283-290.

Flynn, J. R. (1984b). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin,* 95: 29-51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin,* 101: 171-191.

Flynn, J. R. (1994). IQ gains over time. In R. J. Sternberg, S. J. Ceci, J. Horn, E. Hunt, J. P. Matarazzo, & S. Scarr (Eds.), *Encyclopedia of human intelligence* (pp. 617-623). New York: Macmillan.

Flynn, J. R. (1998). IQ gains over time: Toward finding the causes. In J. Neisser (Ed.), *The rising curve: Long term gains in IQ and related measures* (pp. 25-66). Washington, DC: American Psychological Association.

Glenn, N. D. (1994). Television watching, newspaper reading, and cohort differences in verbal ability. *Sociology of Education,* 67: 216-230.

Glenn, N. D. (1999). Further discussion of the evidence for an intercohort decline in education-adjusted vocabulary. *American Sociological Review,* 64: 267-271.

Hauser, R. M., & Huang, M. H. (1997). Verbal ability and socioeconomic success: A trend analysis. *Social Science Research,* 26: 331-376.

Hayes, D., Wolfer, L., & Wolfe, M. (1996). Schoolbook simplification and its relation to the decline in SAT-Verbal scores. *American Educational Research Journal,* 33: 489-508.

Koretz, D. M. (1986). *Trends in educational achievement* (Congress of the United States, Congressional Budget Office). Washington, DC: U.S. Government Printing Office.

Koretz, D. (1992). What happened to test scores and why? *Educational Measurements: Issues and Practice,* 11, 4: 7-11.

Lee, D. W., & Lee, T. H. (1995). Human capital and economic growth: Tests based on the international evaluation of educational achievement. *Economic Letters,* 47: 219-225.

Lieberson, S. (1978). A reconsideration of the income differences found between migrants and northern-born blacks. *American Journal of Sociology,* 83: 940-966.

Lieberson, S. (1985). *Making it count: The improvement of social research and theory*. Berkeley and Los Angeles, CA: University of California Press.

Liu, J., & Sakamoto A. (2002). The role of schooling in Taiwan's labor market: Human capital, screening, or credentialism? *Taiwanese Journal of Sociology,* 29: 1-56.

Lommis, S. C., & Bourque, M. L. (Eds.). (2001). *National assessment of educational progress achievement levels 1992-1998 for reading.* Washington, DC: National Assessment Governing Board.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics.* New York: Cambridge University Press.

Miner, J. B. (1957). *Intelligence in the United States: A survey.* New York: Springer.

Rotberg, I. (1990). I never promised you first place. *Phi Delta Kappan,* 72, 4: 296-303.

Sakamoto, A., & Powers, D. A. (1995). Education and the dual labor market for Japanese men. *American Sociological Review,* 60: 222-246.

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle longitudinal study*. New York: Cambridge University Press.

Sørensen, A, B. (1979). A model and a metric for the analysis of the intra-generational status attainment process. *American Journal of Sociology,* 85, 2: 361-384.

Stedman, L. C., & Kaestle C. F. (1991). The great test score decline: A closer look. In C. F. Kaestle, L. C. Stedman, K. Tinsley, & W. V. Trollinger, Jr. (Eds.), *Literacy in the United States: Readers and reading since 1880* (pp. 129-145). New York: Yale University Press.

The OECD Program for International Student Assessment. (2001). *Knowledge and skills for life: First results from PISA 2000*. OECD.

Thurow, L. C. (1975). *Generating inequality: Mechanisms of*

*distribution in the U.S. economy*. New York: Basic Books.
Wilson, J. A., & Gove, W. R. (1999a). The intercohort decline in verbal ability: Does it exist? *American Sociological Review,* 64: 253-266.
Wilson, J. A., & Gove, W. R. (1999b). The age-period-cohort conundrum and verbal ability: Empirical relationships and their interpretation. *American Sociological Review,* 64: 287-302.

# 語文能力在美國有下降的趨勢嗎？

黃敏雄

## 摘　要

最近一些研究者使用美國社會變遷調查資料 (General Social Survey, 1974-1996)，爭議語文能力是否有隨出生年逐年下降的趨勢。有一些研究者指出，過去所觀察到語文能力隨出生年逐年下降的趨勢是假象，此乃導因於語文能力會隨年齡逐年增加。另有學者認為，年齡對語文能力的影響有限，不足以解釋整個下降趨勢。本文首先釐清過去研究在分析方法上的爭議，進而控制相同教育程度在不同世代所具有的相對位階。結果發現，年齡並不是解釋語文能力下降的主因。語文能力在美國的下降現象始於一九四〇年代末期出生的世代，止於一九六〇年代早期出生的世代，此趨勢及其變化幅度與一九六〇及一九七〇年代美國各州所發現的學習測驗分數下降情況相似。

**關鍵字：**字彙測驗分數、測驗分數下降情況、語文能力、出生年分析、年齡