Available online at www.sciencedirect.com
ScienceDirect
Intelligence 36 (2008) 29 – 34
INTELLIGENCE

The stability of general intelligence from early adulthood to middle-age

Lars Larsen [a,*], Peter Hartmann [b], Helmuth Nyborg [c]

[a] Center for Geropsychology, Department of Psychology, University of Aarhus, Denmark
[b] Dansk psykologisk Forlag, Copenhagen, Denmark
[c] IDRU, Department of Psychology, University of Aarhus, Denmark

Received 30 August 2005; received in revised form 29 December 2006; accepted 4 January 2007
Available online 8 February 2007

Abstract

Early cross-sectional studies suggested that cognitive functions begin to decline in young adulthood, whereas the first longitudinal studies suggested that they are mainly stable in adulthood. A number of more contemporary longitudinal studies support the stability hypothesis. However, drop out effects have the consequence that most longitudinal studies end up with relatively few subjects.

In the present study we determined absolute as well as differential stability in general intelligence *g*, and in verbal and arithmetic abilities, longitudinally for 4000+ adult male veterans drawn from the Vietnam Experience Study (VES). The subjects were given five cognitive tests in their early adulthood. Approximately 18 years later, 14 cognitive tests were administered. Two tests, one verbal and one arithmetic, were administered on both occasions. A Principal Axis Factor analysis was conducted separately on the tests from first and second testing in order to extract both a "g_{young}" and a "g_{old}" general intelligence factor. g_{young} was then correlated with g_{old} to determine the differential stability of *g*. The absolute scores from the recurrent tests were correlated to determine the differential stability and compared using an ordinary *t*-test in order to estimate the absolute stability.

The differential stability coefficients were: 0.85 for *g*; 0.79 for arithmetic; and 0.82 for verbal ability. With respect to absolute stability of the specific tests, we found a significant increase in verbal score (mean scores; 107.16, 116.52), but no change in arithmetic score. Problems associated with different concepts of stability, level of analysis and potential practice effects were discussed.

Keywords: General intelligence; Differential stability; Absolute stability

1. Introduction

Questions of stability and change are obviously central to the scientific study of adult psychological development. In particular, with respect to cognitive abilities, these questions have generated considerable interest as well as controversy (Jensen, 1980).

Whether age-related stability or change is best explained through a single general ability factor, *g*, or through different types of abilities that follow different developmental trajectories is of great consequences for theories of general intelligence (Jensen, 1998; Spearman, 1927) as well as for theories of multiple intelligence (Cattell, 1971; Horn, 1970; Horn & Cattell, 1966, 1967; Thurstone, 1938).

* Corresponding author. Department of Psychology, Jens Chr. Skous vej 4, DK-8000 Aarhus C, Denmark. Tel.: +45 89424900.
E-mail address: larsl@psy.au.dk (L. Larsen).

0160-2896/$ - see front matter
doi:10.1016/j.intell.2007.01.001

The study of age-related stability in intelligence is not without problems, however. How do we best define stability, the proper level of analysis and type of empirical design?

The distinction between differential and absolute stability is obviously very important (Caspi & Bem, 1990). These kinds of stability are independent of each other and each contributes to the general picture of stability versus change. Differential stability thus sheds light on the relative change of subjects within a population, whereas the absolute change informs us of whether the group as unity gains or loses cognitive ability over time.

Moreover, choice of level of analysis may affect the outcome, since stability at the highest order factor level, *g*, might camouflage larger changes at the lower order group factor level and, in particular, at the specific factor level. It is, in other words, very important to distinguish between various levels of analysis, as cognitive change might be more important and pronounced at some than other levels.

Finally, generation effects may confound results, as they can easily be mistaken for age-related change in cross-sectional studies. These effects can be controlled by implementation of longitudinal designs. On the other hand, longitudinal studies tend to overestimate stability and growth because of practice and drop out effects (Salthouse, 1992; Siegler & Botwinick, 1979). The latter have the consequence that the majority of longitudinal studies covering longer stretches of time encompass relatively few subjects.

Research on adult psychological development was insufficient during the first half of the 20th century, because developmental psychologists focused primarily on early development (Schaie, 2000). However, the recruitment needs of the military during World War I did motivate researchers to construct mental tests suitable for adults, so that they could direct large numbers of young and middle-aged draftees into different military functions (Yerkes, 1921). The construction of psychometric assessment tools, such as the Army Alpha, provided the foundation for later empirical investigations of age differences in adulthood, such as the ones conducted by Jones and Conrad (1933). The early cross-sectional studies suggested that cognitive functions decline between young adulthood and middle-age.

Some years later Kuhlen (1940) pointed out that cultural change could be mistaken for age-related change when interpreting results from cross-sectional studies. During the 1950s and 60s researchers became increasingly aware of the necessity of studying the same individual over longer periods of time, in order to control for possible differences in the characteristics of different generations, the so-called cohort effect (Schaie, 1965). This marked the beginning of a major paradigmatic change in emphasis, from cross-sectional studies of adult development to longitudinal studies (Schaie, 2000). An early example of such longitudinal investigations illustrates the point well.

In 1950 Owens conducted the first longitudinal study of the development of mental abilities in adulthood (Owens, 1953; Owens & Clampitt, 1952), and retested 127 males, using the Army Alpha test, Form 6, with which these males had been tested when entering the Iowa State College as freshmen during the Winter Quarter in 1919. The Army Alpha, Form 6, consists of 8 subtests: Following Directions, Arithmetical Problems, Practical Judgement (common sense), Synonym–Antonym (verbal opposites), Disarranged Sentences, Number Series Completion, Verbal Analogies and Information. Individual differences in cognitive functions tended to remain stable over 30 years from early adulthood to approximately age fifty, with the exception of verbal analogies (significant increase) and Disarranged Sentences (significant decrease). In absolute terms Owens found no significant decrease on any subtest, but a significant increase on both the Total score and four of the subtests, namely; Practical Judgement, Synonym–Antonym, Disarranged Sentences and Information. Owens explained the differences in outcome between his cross-sectional and longitudinal studies with differences in length of education, which favoured the younger groups in cross-sectional studies and introduced a confounding cohort effect.

It is beyond the scope of this article to present a thorough review of the extensive longitudinal research that followed, so it will have to suffice to mention that several more recent studies have confirmed that individual differences in measures of mental ability are relatively stable in adulthood (Arbuckle, Maag, Pushkar, & Chaikelson, 1998; Deary, Whalley, Lemmon, Crawford, & Starr, 2000; Eichorn, Hunt, & Honzik, 1981; Hertzog & Schaie, 1986; Plassmann et al., 1995; Schwartzman, Gold, Andres, Arbuckle, & Chaikelson, 1987).

With respect to absolute change, results are somewhat more ambiguous. The Owens study shows an increase mainly in verbal skills and no significant increase on subtests such as Following Directions, Arithmetical Problems and Number Series Completion (Owens, 1953). Results from the Intergenerational study (Eichorn et al., 1981) show an increase in both verbal and non-verbal test scores. Results from the Concordia study (Arbuckle et al., 1998) on the other hand show an increase in one verbal subtest (Vocabulary) and a decrease in another (Verbal Analogies), while scores on three non-verbal subtests (Picture Completion, Picture Anomalities, Paper Formboard) decreased. Arbuckle et al. do however mention, that the lower stability of these non-verbal subtests from the 1942 standardization of the M test, which was used, may be due to the fact the Picture Completion and Picture Anomalities subtests had lower reliability than the other subtests (Blair, 1959), and further that the visual stimuli of these two subtests require fine visual discriminations and therefore can be affected by loss of visual acuity (Arbuckle et al., 1998, p. 672).

In the present study we determined absolute as well as differential stability in general intelligence *g*, and in verbal and arithmetic abilities, longitudinally over a period of almost two decades, for 4000+ adult male subjects.

2. Method

2.1. Subjects

Adult male subjects (N=4321–4385) were drawn from the Vietnam Experience Study (VES: Centers for Disease Control, 1988, 1989). VES was conducted in order to assess possible long-term effects of military service in Vietnam. Subjects, about half of them doing service in Vietnam and half serving elsewhere, and the selection of tests relevant for cognitive studies have been described in details elsewhere (Larsen, 2003; Nyborg & Jensen, 2000, 2001), so only a summary will be provided here.

The VES population is fairly representative for the US population with respect to education, income, occupation and race, but subjects scoring below the 10th percentile in the pre-induction cognitive aptitude test were excluded, in accordance with a US Congress mandate. This obviously truncates the lower-end tail of the ability distribution. The mean age at first testing in 1967–1971 was 19.92 (1.72) years and at second testing in 1985/86 38.35 (2.52) years, resulting in a mean testing interval of 17.90 (1.86) years.

2.2. The cognitive ability tests

The subjects were given five cognitive tests in their early adulthood (t_1) at induction into the military (test number 13, 15, 17–19). Later, at the second testing (t_2) at middle-age, 14 cognitive tests were administered. The description below of the 19 cognitive tests used in this analysis is adapted from Nyborg and Jensen (2000):

1. Grooved Pegboard Test (GPT) (Right hand): A measure of manual dexterity and fine motor speed. The speed score is the reciprocal of the number of seconds, taken to place a set of pegs in a grooved hole with right hand as quickly as possible (scores reversed in order to get positive correlations).
2. Grooved Pegboard Test (GPT) (Left hand): As Grooved Pegboard Test (GPT) (Right hand), but with the left hand (scores reversed).
3. Paced Auditory Serial Addition Test (PASAT): A measure of mental control, mental speed and computational and attentional abilities. The subject mentally adds a sequence of numbers in rapid succession. Score is the total number of correct responses.
4. Rey–Osterrieth Complex Figure Drawing (CFD) (Direct copy): A measure of visual–spatial ability and memory. The subject reproduces a complex spatial figure while the figure is in full view.
5. CFD (Immediate recall): The figure is reproduced right after the figure has been removed from view.
6. CFD (Delayed recall): The figure is reproduced 20 min after the figure has been removed from view, with other activities intervening.
7. Wechsler Adult Intelligence Scale, Revised (WAIS-R): General information scale. The score is scaled.
8. Wechsler Adult Intelligence Scale, Revised (WAIS-R): Block design scale. The score is scaled.
9. Word List Generation Test (WLGT): A measure of verbal fluency. The subject generates as many words as possible within 60 s that begin with the three letters: F, A, and S. The score is the total number of words generated.
10. Wisconsin Card Sort Test (WCST): A measure of concept-formation, problem-solving and set-switching abilities and use of feedback in decision-making. The score is the ratio of correct responses to countable responses.
11. Wide Range Achievement Test (WRAT): Measures the ability to read aloud a list of single words (un-timed test). Total raw score.
12. California Verbal Learning Test (CVLT): A measure of verbal learning and memory. Subjects recall a list of 16 words over five repeated learning trials. The score is the total correct over 5 trials.

13. Army Classification Battery (ACB): Verbal test, a measure of verbal reasoning. Administered at time of induction.
14. Army Classification Battery (ACB): Verbal test, a measure of verbal reasoning. Administered in 1985–86.
15. Army Classification Battery (ACB): Arithmetic reasoning test. Administered at time of induction.
16. Army Classification Battery (ACB): Arithmetic reasoning test. Administered in 1985–86.
17. Pattern Analysis Test (PAT): A visual–spatial measure of pattern recognition. Administered at time of induction.
18. General Information Test (GIT): Administered at time of induction.
19. Armed Forces Qualification Test (AFQT): A general aptitude battery. Total score on four subtests (Word knowledge, Paragraph comprehension, Arithmetic reasoning, Mathematics knowledge). Administered at time of induction. Two of the administered cognitive ability tests were the same at t_1 and t_2, namely the ACB (verbal) and ACB (arithmetic) subtests of the Army Classification Battery (tests 13–14 and 15–16). The remaining 15 tests were administered either at t_1 or t_2. A Principal Axis Factoring (PAF) of all 19 tests indicated that the two recurrent tests had identical factor loadings, across time (t_1 and t_2), on the first extracted unrotated factor (PAF1). This indicates that the recurrent tests were equally well defined by *g* across time.

3. Analyses

Due to the low number of recurrent tests, only a few specific and simple analyses could be conducted.

3.1. Estimation of differential stability

A Principal Axis Factor analysis (PAF) was conducted separately on the tests from t_1 and t_2 in order to extract a "g_{young}" general intelligence factor from the five early cognitive tests and a "g_{old}" from the fourteen cognitive tests administered at middle-age. g_{young} was then correlated with g_{old} to determine the differential stability of *g*.

The absolute scores obtained from the recurrent tests, ACB (verbal) and ACB (arithmetic) were correlated to determine the differential stability of these specific tests (containing *g* variance as well as specific variance).

The contribution of specific tests at t_1 to the differential stability of "g_{old}" was determined by correlating residual scores from the initial five tests at t_1 (after extracting g_{young}) with "g_{old}", in order to determine any independent effect of the five initial tests to later *g*.

The residual scores from the five initial tests at t_1 (after extracting g_{young}) were correlated with the residual scores of ACB (verbal) at t_2 and ACB (arithmetic) at t_2 (after extracting g_{old}) in order to determine any independent effect of the five initial tests to later ACB (verbal) and ACB (arithmetic) at t_2.

3.2. Estimation of absolute stability

The absolute scores obtained from the recurrent test ACB (verbal) and ACB (arithmetic) were compared using an ordinary *t*-test (repeated measure) to compare raw scores at first (t_1) and second (t_2) testing in order to estimate the absolute stability of specific test scores (containing *g* variance as well as specific variance). The two recurrent tests showed practically identical factor loadings across time, both when conducting a joint factor analysis of all the 19 tests from t_1 and t_2 and when conducting separate factor analysis of the test taken at t_1 and t_2 and comparing the factor loadings for the two recurrent tests. This ensures the structural stability of these recurrent tests, and enables comparability of total scores in relation to *g*.

The fact that only two tests were recurrent made it meaningless to try and obtain a measure for absolute stability in *g* based on the recurrent tests. To get there, at least three recurrent tests would be needed for a valid factor analysis, since no factor can be validly estimated with less than three indicators. Moreover, we could have instigated an anchoring procedure for g_{young} and g_{old} by using the recurrent tests' absolute scores as reference points, but with only two recurrent tests, this anchoring procedure would be based on a simple averaging of the two tests, and this would most likely not add any important knowledge besides the simple comparison of the specific tests. Obviously, we would have preferred to be able to separate the absolute stability of *g* from the absolute stability of the specific tests.

4. Results

The differential stability coefficients were as follows: 0.85 for *g* ($N=4321$); 0.79 for ACB (arithmetic) ($N=4385$); and 0.82 for ACB (verbal) ($N=4384$). All coefficients were statistically highly significant (i.e. $p<0.0001$).

With respect to the contribution of the five initial tests to g_{old} while controlling for g_{young} it was found that no correlation exceeded $r=|0.11|$. Naturally, this reached significance due to the large number of subjects, but the factual contribution of the specific abilities to later g was deemed unimportant.

As for the initial five tests' contribution to later ACB (arithmetic) and ACB (verbal) while controlling for g_{young} and g_{old} it was found that the only notable correlations were seen between $PAT_{residuals}$ and $AFQT_{residuals}$ which correlated $r=-0.29$ and $r=-0.23$ with ACB (verbal) $t_{2\text{-residuals}}$. As could be expected, a correlation was also found between ACB (arithmetic) $t_{1\text{-residuals}}$ and ACB (arithmetic) $t_{2\text{-residuals}}$ of 0.36, and furthermore between ACB (verbal) $t_{1\text{-residuals}}$ and ACB (verbal) $t_{2\text{-residuals}}$ of $r=0.44$. All correlations were statistically highly significant ($p<0.0001$).

With respect to absolute stability of the specific tests, we found a significant increase in ACB (verbal) score from t_1 to t_2, but no change in ACB (arithmetic).

5. Discussion

The distinction between differential and absolute stability is important. The differential stability coefficients for general intelligence g, and for the specific arithmetic and verbal skills were all large and highly significant. On the other hand, the absolute stability for arithmetic and verbal skills differed markedly by a gain of roughly 1/3 SD for the latter and no changes in arithmetic skills. In terms of differential stability, our results are in accordance with a large number of previously mentioned longitudinal studies, finding support for the stability hypothesis. In terms of absolute stability, the tendencies of our results resemble the ones found in the study conducted by Owens (1953), namely higher scores on subtests requiring verbal skills and no significant changes in scores on subtests requiring arithmetic skills.

The present study highlights the importance of distinguishing between the stability of higher and lower order factors. Since general intelligence g is a distillate of the common source of individual differences in all mental tests (Jensen, 1998, p. 74), if only looking at g for stability we may actually fail to spot the specific sources of change or stability in test behavior over time. In other words, although g is without questioning the best measure of general intellectual ability, it is too general to detect potentially important changes at the level of lower order cognitive ability factors. The stability of the two specific tests (of about .8 or 64%) may be roughly explained by 70–80% stability of g and 20–30% stability in the specific tests, based on the relationship between the recurrent test before and after g has been partialled out. We speculate that the observed changes in verbal skills over time reflect training-related optimisation of genotypic dispositions.

A problem in many longitudinal studies of cognitive abilities is that practice effects may confound results (Ferrer, Salthouse, Stewart, & Schwartz, 2004; Salthouse, 1992). Practice effects can be controlled when using refined sequential designs (Baltes, Reese, & Nesselroade, 1977, Schaie & Willis, 2002).

Closer inspection of the present database revealed that 77–78 subjects apparently did not take ACB-tests of arithmetic and verbal skills at the time of their induction (t_1), but they certainly took them at the second testing time (t_2). This would enable us to monitor possible practice effects. Unfortunately we could not definitively rule out the possibility that the subjects who appeared to have been tested only once at t_2 in fact also took the tests at the time of induction, but that the results somehow got lost. The archival nature of the data did not allow us to dig further into this possibility. Assuming for a moment that they were actually tested only at t_2, the post-hoc analyses in the present study indicated no practice effects, as subjects tested only once obtained a Verbal mean score of 116.49 (SD 21.34), and an Arithmetic mean score of 102.06 (SD 22.95), whereas subjects tested twice obtained a Verbal mean score of 116.52 (SD 23.07) and an Arithmetic mean score of 104.61 (SD 24.43) (Table 1). None of the differences come close to statistical significance, but due to the obvious limitations mentioned above we cannot use the results to support a certain conclusion. It should also be kept in mind that a practice effect has to be rather strong to show up when test and retest are conducted on average 18 years apart (Schaie, 1996; Zelinsky & Burnight, 1997).

Overall, our findings provide support for the outcome of many other longitudinal studies, suggesting that general intelligence g shows high differential stability from early adulthood to middle-age. In fact, g measured in early adulthood predicts this very ability later in life with a precision that equals the reliability of the tests.

Arithmetic ability did not change in this study — in absolute terms — but verbal ability appeared to increase over time. As mentioned previously, Arbuckle, Maag, Pushkar, and Chaikelson (1998) also observed an

Table 1
Test for changes in Verbal and arithmetic test–retest raw scores

t-test for dependent samples	Mean	SD	N	*p*
Verbal 1	107.16	22.26		
Verbal 2	116.52	23.07	4384	0.000
Arithmetic 1	104.43	22.01		
Arithmetic 2	104.61	24.63	4385	0.446

increase in one verbal subtest (Vocabulary), but further found that Verbal Analogies decreased.

Finally, the observation that g_{young} is an excellent predictor of later general high-level ability, did not say anything about some very real absolute instability at the lower levels of specific ability factors.

Acknowledgement

We are grateful to The Centers for Disease Control in Atlanta for providing the archival VES data.

References

Arbuckle, T. Y., Maag, U., Pushkar, D., & Chaikelson, J. (1998). Individual differences in trajectories of intellectual development over 45 years of adulthood. *Psychology and Aging*, *13*(4), 663−675.

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). Life-span developmental psychology: Introduction to research methods. Monterey, CA: Brooks/Cole.

Blair, W.R.N. (1959). *Normative data for the revised examination "M": Annex 1 to HQC 3077–2 (M3B1)*. Unpublished manuscript, Department of National Defence, Canada, Army.

Caspi, A., & Bem, D. J. (1990). Personality continuity and change across the life course. In L. A. Pervin (Ed.), *Handbook of personality research* (pp. 549−575). New York: The Guildford Press.

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

Centers for Disease Control. (1988). Health status of Vietnam veterans. *Journal of the American Medical Association*, *259*, 2701−2719.

Centers for Disease Control. (1989). Health status of Vietnam veterans: Vol. IV. Psychological and neuropsychological evaluation. Atlanta, Georgia: Center for Environmental Health and Injury Control.

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: Follow-up of the 1932 Scottish mental survey. *Intelligence*, *28*(1), 49−55.

Eichorn, D. H., Hunt, J. V., & Honzik, M. P. (1981). Experience, personality, and IQ: Adolescence to middle age. In D. H. Eichorn, J. A. Clausen, N. Haan, M. P. Honzik, & P. H. Mussen (Eds.), *Present and past in middle life*. New York: Academic Press.

Ferrer, E., Salthouse, T. A., Stewart, W. F., & Schwartz, B. S. (2004). Modelling age and retest processes in longitudinal studies of cognitive abilities. *Psychology and Aging*, *19*(2), 243−259.

Hertzog, C., & Schaie, K. W. (1986). Stability and change in adult intelligence: 1. Analysis of longitudinal covariance structures. *Psychology and Aging*, *1*(2), 159−171.

Horn, J. L. (1970). Organization of data on life-span development of human abilities. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology*. New York: Academic Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, *57*, 253−270.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 107−129.

Jensen, A. R. (1980). *Bias in mental testing*. London: Methuen.

Jensen, A. R. (1998). The g factor: The science of mental ability. Westport, CT: Praeger.

Jones, H. E., & Conrad, H. S. (1933). The growth and decline of intelligence: A study of a homogenous group between the ages of ten and sixty. *Genetic Psychology Monographs*, *13*, 223−298.

Kuhlen, R. G. (1940). Social change; a neglected factor in psychological studies of the life span. *School & Society*, *52*, 14−16.

Larsen, L. (2003). Testosterone as a factor in psychological and behavioral traits. *Psykologisk ph.d.-skriftserie*, *Vol. 1* (pp. 83−86). Risskov: Psykologisk Institut, Århus Universitet.

Nyborg, H., & Jensen, A. R. (2000). Black–white differences on various psychometric tests: Spearman's hypothesis tested on American armed services veterans. *Personality and Individual Differences*, *28*, 593−599.

Nyborg, H., & Jensen, A. R. (2001). Occupation and income related to psychometric *g*. *Intelligence*, *29*, 45−55.

Owens, W. A. (1953). Age and mental abilities: A longitudinal study. *Genetic Psychology Monographs*, *48*, 3−54.

Owens, W. A., & Clampitt, R. R. (1952). The effects of a thirty year age increment upon individual and trait differences in intelligence. *Proceedings of the Iowa Academy of Sciences, Des Moines, IA*, *Vol. 59* (pp. 385−388).

Plassmann, B. L., Welsh, K. A., Helms, M., Brandt, J., Page, W. F., & Breitner, J. C. S. (1995). Intelligence and education as predictors of cognitive state in late life: A 50-year follow-up. *Neurology*, *45*, 1446−1450.

Salthouse, T. (1992). Mechanisms of age-cognition relations in adulthood. Hillsdale, N.J.: Lawrence Erlbaum.

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, *64*, 92−107.

Schaie, K. W. (1996). Intellectual development in adulthood. Cambridge, England: Cambridge University Press.

Schaie, K. W. (2000). The impact of longitudinal studies on understanding development from young adulthood to old age. *Journal of Behavioral Development*, *24*(3), 257−266.

Schaie, W. K., & Willis, J. L. (2002). Research methodology in adult development and aging. In K. W. Schaie & S. L. Willis (Eds.), *Adult development and aging* (pp. 109−129). New Jersey: Prentice Hall.

Schwartzman, A. E., Gold, D., Andres, D., Arbuckle, T. Y., & Chaikelson, J. (1987). Stability of intelligence — A 40 year follow-up. *Canadian Journal of Psychology*, *41*, 244−256.

Siegler, I. C., & Botwinick, J. (1979). A long-term longitudinal study of intellectual ability of older adults: The matter of selective subject attrition. *Journal of Gerontology*, *34*, 242−245.

Spearman, C. (1927). The abilities of man: Their nature and measurement. New York: Macmillan.

Thurstone, L. L. (1938). Primary mental abilities. Chicago: University of Chicago Press.

Yerkes, R. M. (Ed.). (1921). Psychological examining in the United States Army. *Memoirs of the National Academy of Sciences*, *Vol. 15* (pp. 1−890).

Zelinsky, E. M., & Burnight, K. P. (1997). Sixteen-year longitudinal and time lap changes in memory and cognition in older adults. *Psychology and Aging*, *12*, 503−513.