International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report IR-11-028

National Variation in Cognitive Life Cycle Development

Vegard Skirbekk (skirbekk@iiasa.ac.at)
Daniela Weber (weberd@iiasa.ac.at)
Valeria Bordone (bordone@iiasa.ac.at)

Approved by

Wolfgang Lutz
Leader, World Population Program

January 26, 2012 – Revised November 7, 2013

Contents

Abstract

Maintaining cognitive functioning through mid- to late-life is relevant for both the individual and societal aim of active ageing and reducing dependence in old age. This paper analyses life-cycle variation in country-level rank ordered cognitive performance over a 40-year period. For the cohort born between 1949 and 1952, we observe standardized mathematical test scores at teen age from the First International Mathematics Study (FIMS) and cognitive test performance at mid-life, based on the SHARE survey. This allows us to compare the relative country ranking in 1964 and the performance in 2004 of the same birth cohort. Our results show that those countries which had the highest scores in math tests taken by 13 years old grade level students are not the same countries that, 40 years later, have the top performing scores in cognitive tests among mid-age adults. This highlights the importance of considering country-level influences on cognitive change over the life cycle, in addition to individual characteristics. Further studies are required to explore the link between specific contextual factors and cognitive functioning before we will be able to formulate relevant policy implications from these results.

Acknowledgments

This work was supported by Starting Grant of the European Research Council, Grant Agreement 241003-COHORT and the “Wittgenstein Award” of the Austrian Science Fund (FWF): Z171-G11.

The authors are grateful to Bilal Barakat, Bill Butz, and Elke Loichinger who gave us precious comments on earlier versions of this paper. We also would like to thank the participants to the European Population Conference (June 2012) and to the Annual Scientific Meeting of the Gerontological Society of America (November 2012). We are thankful to Kathryn Platzer and Matthew Cantele for the English editing.

The English data on the 52-55 year olds were made available through the UK Data Archive (UKDA). ELSA was developed by a team of researchers based at the National Centre for Social Research, University College London and the Institute for Fiscal Studies. The data were collected by the National Centre for Social Research. The funding is provided by the National Institute of Aging in the United States, and a consortium of UK government departments co-ordinated by the Office for National Statistics. The developers and funders of ELSA and the Archive do not bear any responsibility for the analyses or interpretations presented here.

The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

This paper uses data from SHARE release 2.5.0, as of May 24th 2011. The SHARE data collection has been primarily funded by the European Commission through the 5th framework programme (project QLK6-CT-2001- 00360 in the thematic programme Quality of Life), through the 6th framework programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th framework programme (SHARE-PREP, 211909 and SHARE-LEAP, 227822). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064, IAG BSR06-11, R21 AG025169) as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

About the Authors

Vegard Skirbekk is the Leader of the Age and Cohort Change Project (ACC) with the World Population Program at the International Institute for Applied Systems Analysis (IIASA).

Daniela Weber is a Research Assistant with the World Population Program at the International Institute for Applied Systems Analysis (IIASA).

Valeria Bordone is a Research Scholar at the Vienna University of Economics and Business (WU) and at the International Institute for Applied Systems Analysis (IIASA).

Preface

This is a revision of the original IR-11-028 published January 26, 2012. This version includes three main changes in addition to some language revisions.

First, we excluded the countries England, Japan, and the USA from the analysis reported in the main text, as there are some limits in comparability. Further details on the comparability issues as well as a more comprehensive analysis including those countries can now be found in Appendix B.

A more detailed description of the data can be found in Appendix A. Information on the FIMS data is currently difficult to attain, therefore we decided to include this for interested readers.

In this version, the cohorts 1944-47 have been excluded because of comparability issues due to the time in which the FIMS was administered. In any case, the results did not significantly differ from those of the cohorts 1949-52.

Finally, the abstract has been modified according to the changes in the main text.

National Variation in Cognitive Life Cycle Development

Vegard Skirbekk
Daniela Weber
Valeria Bordone

1 Introduction

Maintaining cognitive functioning through mid- to late-life is relevant for both the individual and the societal aim of active aging and reducing dependence in old age. Several studies suggest a strong degree of age-related cognitive plasticity at the individual level: cognitive decline associated with aging could slow down by engaging in physical exercise and keeping mentally active through participation in social activities (e.g. Ball et al., 2007; Baltes, 1987; Baltes & Labouvie, 1973; Baltes, Lindenberger, & Staudinger, 2006; Hertzog, Cramer, Wilson, & Lindenberger, 2009; Kramer & Willis, 2002, 2003). There may, however, also be country-level influences on the average cognitive performance of a population such as length of compulsory schooling (Brinch & Galloway, 2012), but so far those influences have not received much attention in research literature. With a quasi-longitudinal approach, this paper sets out to compare math achievement at teen age and cognitive performance at mid-life across countries for the birth cohort 1949-1952.

Earlier evidence, both from the USA and Europe, has found considerable stability in terms of individual-level rank-ordering of cognitive functioning, i.e., a person with high cognitive performance at younger ages tends also to score high in cognitive tests later in life (e.g. Deary, Whalley, Lemmon, Crawford, & Starr, 2000; Zimprich & Mascherek, 2010). Furthermore, individuals with higher cognitive performance at young ages will have lower incidence of dementia in old age (e.g. Deary, Whiteman, Starr, Whalley, & Fox, 2004; Snowdon, Kemper, Mortimer, Greiner, Wekstein, & Markesbery, 1996); while people with lower initial ability are likely to experience relatively greater cognitive decline in old age (e.g. Bourne, Clayton, Murch, & Grant, 2006; Deary, Starr, & MacLennan, 1998; Richards, Shipley, Fuhrer, & Wadsworth, 2004). Health status, genetic traits, behavior, socio-economic status as well as the degree of cognitive engagement and education contribute to alter the path of individual aging (Comijs, van den Kommer, Minnaar, Penninx, & Deeg, 2011; Engelhardt, Buber, Skirbekk, & Prskawetz, 2010; Glymour, Kawachi, Jencks, & Berkman, 2008).

Yet, this quite strong life-cycle stability in cognition at the individual-level does not necessarily imply that a cohort´s performance over the life cycle will be stable in a cross-country comparative framework, too. Macro-level factors such as social structures, culture and lifestyle, opportunities for physical and mental activity, education systems, as well as employment and retirement patterns, could all significantly affect how well a certain cohort in a specific country maintains its cognitive skills over time. Although limited, few attempts of considering macro-level influences showed that contextual factors do affect individual cognitive functioning and, in turn, the average performance of a society – e.g., Aneshensel, Ko, Chodosh, & Wight (2011) and Lang et al. (2008) found that neighborhood socioeconomic disadvantage has a large negative impact on cognitive functioning; Booth and colleagues (2000) concluded that access to local facilities is significantly associated with being active and it can in turn inform the development of policies to promote cognitive health of older people; Jefferis, Power, & Hertzman (2002) used the 1958 British Birth Cohort Study to investigate the role of the postnatal environment on cognitive function until early adulthood, finding an overwhelming influence.

This study draws on the limitations and challenges of previous research in order to explore the life cycle variation of cognitive functioning at country-level. Understanding whether nations age differently will give insights into country-level influences on cognitive aging patterns. To do this, we propose to combine several datasets and use a quasi-longitudinal approach that allows the investigation on long-run changes in the cross-country comparison of a cohort's cognitive performance (following Lee, 2010). In the remaining of the paper, results are presented and the approach used here (as well as its drawbacks) is discussed.

2 Background: Limitations and Challenges of Previous Research

Two strands of research have mainly looked at, first, the development of students' academic improvement trajectories (e.g. Bloom, 1964; Lichten, 2004) and, second, cognitive performance over time. This latter, overcoming the limitations of small-scale local studies of the first, has used large cross-sectional samples to examine national–level aggregate students' mental growth patterns (e.g. Lee, 2010), but has not always been able to overcome confounding effects of age or grade and cohort on growth.

Country-specific case studies have increased the knowledge on academic performance either at separate stages of development (e.g. early childhood, childhood, adolescence) or at separate levels of the educational system (e.g. preschool, elementary, secondary, post-secondary) (see Lee, 2010 for a discussion of these studies), while the few studies that used an international comparative approach have looked at specific age groups (i.e. the young or the elderly) (e.g. Dewey & Prince, 2005; Gonzales et al., 2008; OECD, 2011; Skirbekk, Loichinger, & Weber, 2012). Although limited on the age range considered, there exist some empirical research that capitalized on cohort-based tracking of achievement with repeated cross-sectional or quasi–longitudinal samples (e.g. on math achievement growth patterns between US states (Coley, 2003) or countries (Lee & Fish, 2008)). Yet, to our knowledge, there is no study that has addressed cross-national variation in mental performance over different stages of life, from younger to older ages. One reason for this is that there are no international comparable data that track students´ achievement from kindergarten or elementary school through high school age, not to mention mid-life or old age.

With this study, we aim at offering a first step to overcome this gap in the literature. Adopting a quasi-longitudinal approach, we compare country performances of the 1949-1952 birth cohort when in school (around 13 years old, as we will further describe in the following) and at mid-life (i.e. at age 52-55). If the life-cycle stability found at individual level applies also to within-cohort national level, we may expect that (in addition to genetic and individual characteristics) macro-level conditions in early life, such as environment where children are born or family conditions during childhood and human capital investment at younger ages are likely to be the most important factors in determining country average cognitive performance also later in life. On the contrary, variation in the national ranking of country performances between teen age and mid-life may rather point at the effect of life-course changes.

3 Methodological Approach

In order to investigate whether the cognitive stability usually found at the individual level (i.e., performing well at younger ages goes along with scoring high at older ages) is confirmed at the country level, this work mainly follows a descriptive approach. We first create a relative rank-ordering of countries for the 1949-1952 cohort, based on the results from mathematical tests of a nationally representative sample at teen age (i.e., 13 years old grade level pupils) in 1964. Second, we create a relative rank-ordering of countries for the same cohort group, based on the performances in several cognitive tests from surveys to which a nationally representative sample participated at ages 52-55 in 2004. Third, we compare the two relative ordering to identify whether nations with good students are also top performers in terms of cognitive functioning of their growing old population forty years later.

The quasi-longitudinal approach used has the important advantage of avoiding problems deriving from nonrandom dropout, retest-practice effects and noncognitive change associated with longitudinal surveys (Thorvaldsson, Hofer, Berg, & Johansson, 2006).

4 Data and Measures

This study uses data from several cross-country comparable datasets. Data at younger ages are derived from the 1964 First International Mathematics Study (FIMS) on 13 years old grade level pupils. This determines the choice of the birth cohort under focus (i.e. 1949-1952). The FIMS survey was designed to be representative of particular age- and school year-groups in twelve countries and was conducted by the International Association for the Evaluation of Educational Achievement (IEA) (Husén, 1967a, 1967b; Wolf, 1967). It was, in general, carried out by teachers, whereas the test was conducted in the classroom. An international committee proposed to develop items to cover several categories of intellectual process. For instance, one of the questions in the study was: "There are 227 boys in a school. Every boy in the school belongs to either the music club or the sports club, and some boys belong to both clubs. The music club has 120 members, and 36 of these are members of the sports club. What is the total membership of the sports club?" For further details on the survey, we refer the reader to Appendix A. The whole questionnaire is available in Husén 1967b (Appendix II, pp. 312ff). FIMS was the first of several subsequent internationally comparable school tests (e.g. Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009; OECD, 2011), and the only one, so far, whose teenage interviewees have already reached mid-age.

The FIMS was carried out in Australia, Belgium, England, Finland, France, Germany[1], Israel, Japan, the Netherlands, Scotland, Sweden, and the United States. We are able to hypothetically follow the 1949-1952 cohort, observing a representative sample again at mid-ages in 2004 (i.e. aged 52-55), in Belgium, France, Germany, the Netherlands, and Sweden using the Survey of Health, Ageing and Retirement in Europe (SHARE) (Börsch-Supan et al., 2005)[2]. Although we do not track the same individuals tested by the FIMS survey, we observe nationally representative samples of these populations over a 40-year life span. Of particular interest is the (re-)examination at an age when age-related cognitive decline is already visible (i.e. between ages 52 and 55).

As indicators of the different domains of cognitive functioning in adulthood, which may vary differently over time (Richards, Shipley, Fuhrer, & Wadsworth, 2004), we rely on four items: *immediate recall* tests the ability to recall as many words as possible within a specific time, out of ten common words read by the interviewer; the *delayed recall* test consists in recalling, with a few minutes of delay, as many words as possible out of the ten words read out before the immediate recall task; *fluency* refers to the performance of naming as many animals as possible within one minute; *numeracy* tests the ability of performing some simple calculations based on real life situations. While recall measures range from 0 to 10, fluency has no fixed ceiling and numeracy scores from 1 to 5, indicating at which step of the numerical battery the interviewee stopped answering correctly[3]. For the exact wording of the tests, we refer the reader to the SHARE questionnaire, available on the website http://www.share-project.org/.

Based on these datasets, we examine the extent of changes in the ranking of countries that lag further behind and those matching the top performances. A description of the samples at teen age and early fifties for all participating countries, including coverage, gender composition and response rates is provided in Appendix A.

For math performance at younger ages and both recall tests at mid-life, means are our measure of the country-specific level of ability. Consequently, countries are ordered according to their mean achievement. Standard deviations (S.D.) provide information about the size of the age-specific cognitive variation within each country. For numeracy at mid-life, countries´ scores are calculated as the proportion of the

---

[1] Data were collected in the Federal Republic of Germany before reunification, hereafter referred to as Germany.

[2] Data on cognitive abilities of nationally representative samples of the non-institutionalized population aged 50+ are also collected by the English Longitudinal Study of Ageing (ELSA) in England (Marmot & Banks, 2003), the Health and Retirement Study (HRS) in the USA (National Institute on Aging, 2007) and the Japanese Study of Aging and Retirement (JSTAR) in Japan (Ichimura, Shimizutani, & Hashimoto, 2009). Although these surveys have been designed to be comparable (e.g. on http://www.rieti.go.jp/en/projects/jstar/index.html it is written that JSTAR “is designed to ensure to the maximum extent possible, comparability with preceding surveys, such as HRS, SHARE and ELSA”), they differ from each other in the inclusion and/or the formulation of some questions. For this reason, we consider them in a separate additional analysis, reported in Appendix B.

[3] Following SHARE’s calculation, the numeracy score ranges from 1 (poor numerical ability, if the interviewee has not answered any math question correctly) to 5 (excellent numerical ability, if the interviewee has answered all the math questions correctly), being 2 = fair if the interviewee does not answer the first question correctly, but gives the right answer to the math question that follows; 3 = good corresponds to a correct answer to the first question only; 4 = very good corresponds to two correct answers (at both first and second questions).

population with the top two performance levels on the five-point scale. Individual sampling weights are applied in all our calculations.

5 Results

Math achievement at teen age

Country-level means of performance and standard deviations from the FIMS data collected in 1964 - at teen age for the cohort group considered - are presented in Table 1a. As shown in the first column, the FIMS scores range from 15.3 points for Sweden up to 30.4 points for Belgium.

Table 1: a) Math test scores at 13 years old grade level in 1964; b) Mean scores and standard deviation (S.D.) of immediate recall, delayed recall, and fluency tests at 52-55 years old in 2004; and relative frequencies reaching the final question of the numerical test battery at 52-55 years old in 2004. 1949-1952 cohort. 5 countries.

Cohort 1949-1952									
	a) 13 years old grade level		b) 52-55 years old						
	FIMS	S.D.	Immediate recall	S.D.	Delayed recall	S.D.	Fluency	S.D.	Numeracy
Belgium	30.4 (1)	13.7	5.5 (3)	1.6	4.0 (4)	1.8	21.5 (3)	6.2	5.1% (4)
France	21.0 (4)	13.2	5.0 (5)	2.0	3.6 (5)	2.0	21.3 (4)	7.6	54.0% (5)
Germany	25.4 (2)	11.7	6.0 (1)	1.8	4.5 (2)	1.8	22.7 (2)	7.2	72.6% (1)
Netherlands	21.4 (3)	12.1	5.5 (3)	1.7	4.3 (3)	1.8	20.8 (5)	5.9	60.4% (3)
Sweden	15.3 (5)	10.8	5.7 (2)	1.4	4.7 (1)	1.6	25.5 (1)	7.2	64.5% (2)

Sources: a) Table 1.2 (Husén, 1967b); b) SHARE, wave 1. Authors' calculation.

The standard deviations give us additional information: Belgium registers the greatest within-country variation at teen age in 1964 (S.D. = 13.7); while Sweden is the country with the lowest variation in pupils' performance (S.D. = 10.8).

Cognitive performance at ages 52-55

The scores of the four tests of cognitive functioning undertaken in 2004 at age 52-55 are reported in Table 1b. Among the Continental European countries, in terms of both immediate and delayed recall as well as in fluency and numeracy, Germany and Sweden score the highest: on average, Germans aged 52-55 recall 6 words out of ten immediately and 4.5 with a few minutes of delay and they name almost 23 animals within 1 minute; their Swedish counterparts recall 5.7 and 4.7 words, respectively and name about 26 animals. France, with on average only 5 words recalled immediately, 3.6 words recalled delayed and 21 animals named on average within 1 minute, is the country with the largest within-country variation in memory tasks among the 52-55 year olds (S.D. = 2 in recall tests and = 7.6 in fluency). The smallest standard deviation in recall measures is registered in Sweden (S.D. immediate recall = 1.4; S.D. delayed recall = 1.6).

Among Germans aged 52-55, about 73% reached the last question of the numerical battery. In France, only 54% reached the final stage of the numerical test.

Cross-country comparisons over time

Of interest for this study is the relative performance of the countries concerning the 1949-1952 cohort at the two points in time, i.e., the changes between the country ranking in 1964 among 13 years old grade level students and in 2004 among the 52-55 year olds. Our results show that the countries which had the top performing results in math tests undertaken by 13 years old grade level students are not the same countries that, 40 years later, have the top performing mid-age adults in cognitive tests.

The most striking result of the analysis is the positive shift in the relative ordering of Sweden, in spite of its weak performance at younger ages. On the contrary, Belgium shows a relative decline in the positions. Germany confirms its position among the top scores from teen age (in the mathematical school test) to mid-life (in the recall, fluency and numeracy cognitive tests).

6 Discussion

Understanding how cognitive functioning is maintained from the teens to the fifties is of great importance, not only for the individual, but also for national prosperity in aging economies (Engelhardt et al., 2010; OECD, 2011; Schmidt & Hunter, 2004). This study identifies reordering in the national-level ranking of cognitive performance for the cohort group born in 1949-1952 over time.

One implication of this study is that more weight should be given to national-level influences to understand life-cycle variation in cognitive functioning: the average country performance can represent more than the sum of individual behaviors. The result of this study may have high potential in translating into policy messages. However, it has to be carefully interpreted and be considered as a starting point for further data collection and research. A change in the relative performance of the countries at the two points in time could be due to either a relative improvement in the cognitive performance of the cohorts in the countries we observe performing better at mid-ages (e.g. Sweden) or to a poorer preservation of the cognitive abilities over the life course of the countries performing better at younger ages than later in life (e.g. Belgium). Moreover, we acknowledge two main limitations of this study. First, the difference between school tests at younger ages and cognitive tests of adults may be argued to affect the estimated changes in the rank-ordering. Second, the relative rank-order findings may derive from the fact that a different subsample of individuals is retained at the later time point. However, we believe that the selection of the mid-aged sample is rather homogeneous among the countries considered (i.e. similar selection by causes of death, gender-specific mortality, and health-specific problem incidence).

Despite the need to improve on these aspects, our results show that about 40 years after the first tests, changes can be observed in the ranking position of the countries considered. This hints to the importance of expanding the research on cross-country differences in cognitive aging trajectories, with more investigations of national-level factors next to individual-level analysis. In contrast to individual-level findings which report a high degree of temporal stability, the outcome of this study is in line with the findings of cohort-based estimates of school achievement at different grades which show considerable changes in the international rank-ordering of a cohort with age – for instance, Lee (2010) showed that relative international performance of cohorts of US pupils varies strongly between school levels. While general forms and characteristics of national aggregate growth curves over the academic years are known relatively well, we add knowledge on a longer perspective (from teen to mid-age) in a cross-country comparative perspective.

A number of national-level factors, such as social structures, environment, culture, opportunities for physical and mental activities, education systems, employment and retirement patterns, but also attitudes towards aging, are likely to contribute in explaining the between-country variation in cognitive performance over time (e.g. Gutchess & Indeck, 2009). However, identifying the specific mechanisms at work goes beyond the goal of this study.

The availability of longer panel data from childhood to old adulthood will allow future research to investigate the specific impact of national-level factors on cognitive development. A long-term assessment of changes in the growth trajectories between sequential cohorts will also allow more direct policy-oriented conclusions.

7 References

Aneshensel, C. S., Ko, M. J., Chodosh, J., & Wight, R. G. (2011). The urban neighborhood and cognitive functioning in late middle age. *Journal of Health and Social Behavior, 52*(2), 163-179.

Ball, K., Edwards, J. D., & Ross, L. A. (2007). The impact of speed of processing training on cognitive and everyday functions. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 62B* (Special Issue I), 19-31.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On dynamics between growth and decline. *Developmental Psychology, 23*(5), 611-626.

Baltes, P. B., & Labouvie, G. V. (1973). Adult development of intellectual performance: Description, explanation, and modification. In C. Eisdorfer, & M. P. Lawton (Eds.), *The psychology of adult development and aging* (pp. 157-219). Washington, DC: American Psychological Association.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (2006). Lifespan theory in developmental psychology. In W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology* (pp. 569-664). New York: Wiley.

Bloom, B. (1964). *Stability and change in human characteristics*. New York: John Wiley.

Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., & Weber, G. (Eds.). (2005). *Health, ageing and retirement in Europe - First Results from the Survey of Health, Ageing and Retirement in Europe*. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

Booth, M. L., Owen, N., Bauman, A., Clavisi, O., & Leslie, E. (2000). Social–cognitive and perceived environment influences associated with physical activity in older australians. *Preventive Medicine, 31*(1), 15-22.

Bourne, C., Clayton, C., Murch, A., & Grant, J. (2006). Cognitive impairment and behavioural difficulties in patients with Huntington's disease. *Nursing Standard, 20*(35), 41-44.

Brinch, C., & Galloway, T. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Science, 109(2),* 425-430.

Coley, R. J. (2003). *Growth in school revisited: Achievement gains from the fourth to the eighth grade* (ETS Policy Information Report). Princeton, NJ: ETS.

Comijs, H. C., van den Kommer, T. N., Minnaar, R. W. M., Penninx, B. W. J. H., & Deeg, D. J. H. (2011). Accumulated and differential effects of life events on cognitive decline in older persons: Depending on depression, baseline cognition, or ApoE ε4 status? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66B*(suppl 1), i111-i120.

Deary, I. J., Starr, J. M., & MacLennan, W. J. (1998). Is age kinder to the initially more able? Differential aging of verbal ability in the HOPE study. *Intelligence, 26*(4), 357-375.

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.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology, 86*(1), 130-147.

Dewey, M. E., & Prince, M. J. (2005) in Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., & Weber, G. (Eds.). (2005). *Health, ageing and retirement in Europe - First Results from the Survey of Health, Ageing and Retirement in Europe*. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

Engelhardt, H., Buber, I., Skirbekk, V., & Prskawetz, A. (2010). Social involvement, behavioural risks and cognitive functioning among older people. *Ageing and Society, 30*(5), 779–809.

Glymour, M. M., Kawachi, I., Jencks, C. S., & Berkman, L. F. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology and Community Health, 62*(6), 532-537.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences.

Gutchess, A. H., & Indeck, A. (2009). Cultural influences on memory. *Progress in Brain Research, 178*(1), 37-50.

Health and Retirement Study, (HRS, wave 7) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, (2004).

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2009). Enrichment effects on adult cognitive development. *Psychological Science in the Public Interest, 9*(1), 1-65.

Husén, T. (Ed.). (1967a). *A comparison of twelve countries: International study of achievement in mathematics* (Vol. 1). Uppsala: Almqvist & Wiksells Boktryckeri AB.

Husén, T. (Ed.). (1967b). *A comparison of twelve countries: International study of achievement in mathematics* (Vol. 2). Uppsala: Almqvist & Wiksells Boktryckeri AB.

Ichimura, H., Shimizutani, S., & Hashimoto, H. (2009). JSTAR First Results 2009 Report.Discussion papers 09047, Research Institute of Economy, Trade and Industry (RIETI).

Jefferis, B. J. M. H., Power, C., & Hertzman, C. (2002). Birth weight, childhood socioeconomic environment, and cognitive development in the 1958 British birth cohort study. *British Medical Journal, 325*(7359), 305.

JSTAR datasets, produced by the Research Institute of Economy, Trade and Industry (RIETI), Hitotsubashi University, and the University of Tokyo, are distributed by RIETI in Tokyo, Japan.

Kramer, A. F., & Willis, S. L. (2002). Enhancing the cognitive vitality of older adults. *Current Directions in Psychological Science, 11*(5), 173-177.

Kramer, A. F., & Willis S. L. (2003). Cognitive plasticity and aging. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 267-302). San Diego, CA: Academic Press.

Lang, I. A., Llewellyn, D. J., Langa, K. M., Wallace, R. B., Huppert, F. A., & Melzer D. (2008) Neighborhood deprivation, individual socioeconomic status, and cognitive function in older people: Analyses from the English Longitudinal Study of Ageing. *Journal of the American Geriatrics Society, 56*(2), 191-198.

Lee, J. (2010). Tripartite growth trajectories of reading and math achievement: Tracking national academic progress at primary, middle, and high school levels. *American Educational Research Journal,47*(4), 800-832.

Lee, J., & Fish, R. (2008). Are Asian schools more effective than American schools? International and interstate comparisons of school effects. Paper presented at the annual meeting of the American Educational Research Association (AERA), New York.

Lichten, W. (2004). On the law of intelligence. *Developmental Review, 24*(3), 252-288.

Marmot, M., & Banks, J. (2003). *Health, wealth and lifestyles of the older population in England : The 2002 English Longitudinal Study of Ageing*. London: Institute for Fiscal Studies.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 Assessment frameworks*. Chestnut Hill, MA: Boston College, TIMSS & PIRLS International Study Center

National Institute on Aging. (2007). *Growing older in America: The Health and Retirement Study*. Washington, DC: National Institutes of Health.

OECD. (2011), *PISA 2009 Results: Students on line*. Paris: OECD Publishing.

Richards, M., Shipley, B., Fuhrer, R., & Wadsworth, M. E. J. (2004). Cognitive ability in childhood and cognitive decline in mid-life: Longitudinal birth cohort study. *British Medical Journal, 328*(7439), 552-554.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*(1), 162-173.

Snowdon, D. A., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. Findings from the Nun study. *Journal of the American Medical Association, 275*(7), 528-532.

Skirbekk, V., Loichinger, E., & Weber, D. (2012). Variation in cognitive functioning as a refined approach to comparing aging across countries. *Proceedings of the National Academy of Sciences, 109*(3), 770-774.

Thorvaldsson, V., Hofer, S. M., Berg, S., & Johansson, B. (2006). Effects of repeated testing in a longitudinal age-homogeneous study of cognitive aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*(6), P348-P354.

Wolf, R. M. (1967). *Data Bank Manual. Phase I: International Study of Achievement in Mathematics: A Comparison of Twelve Countries*. Uppsala: Almqvist & Wiksells Boktryckeri AB.

Zimprich, D., & Mascherek, A. (2010). Five views of a secret: Does cognition change during middle adulthood? *European Journal of Ageing, 7*(3), 135-146.

8 Appendices

8.1 Appendix A

Description of datasets

FIMS. The First International Mathematics Study allows investigation on exactly parallel lines of salient futures of the school system in participating countries and the role of mathematics. Table A.1 shows sample size, proportion by gender, mean age and number of pupils tested (per 100,000 pop.) in the five countries considered here and the additional three countries considered in Appendix B.

Table A.1. Summary statistics of FIMS data.

1964	Cohort 1949-1952 13 years old grade level			
	Sample size	Males (%)	Mean age	Number of pupils tested per 100,000 of pop.
Belgium	2,645	65	14.0	29
England	3,089	52	14.3	7
France	3,449	54	13.6	7
Germany	4,475	48	13.7	8
Japan	2,050	51	13.4	2
Netherlands	1,443	56	13.1	12
Sweden	2,828	51	13.7	37
USA	6,544	51	14.0	3

Sources: Tables 14.3B and 14.4 (Husén, 1967a).

Regulations specifying when children should at the latest begin and a minimum age at which the pupils must attain compulsory schooling are summarized in Table A.2.

Table A.2. Age regulation, by country. Figures relate to the time of testing (May, 1964).

Country	Mandatory age of entry to school (in years)	Median age of entry to school (in years)	Median age of entry to school (months)	Age when compulsory ceases (in years)
England	5	5	2	15
Belgium	6	6	2	14
France	6	6	3	16
Germany	6	6	9	15
Netherlands	6	6	5	14
Japan	6	6	6	14
USA	6	5	9	16
Sweden	7	7	1	15

Sources: Table 13.1 (Husén, 1967a).

The populations tested in the participating countries can be seen in relation to their own education structure, as shown in Table A.3.

Table A.3. Location of population tested in the respective school systems.

Country	Location
Belgium	5 e. (2 e A 3 in Enseignement Technique)
England	3rd Form
France	5 e. (CFE in ecole primaire)
Germany	7. Klasse (Schulleistungsjahr)
Japan	Ni-nen. 2nd Grade
Netherlands	6 e. in primary schools
Sweden	Arskurs 7
USA	8th Grade

Sources: Table 13.4 (Husén, 1967a).

The percentage of the school timetable devoted to the study of mathematics is relevant when studying attainments in mathematics (Table A.4): a between-country comparison of the percentage of the curricula devoted to the study of mathematics at around age 13 in 1964 does not show significant differences.

Table A.4. Percentage of the timetable devoted to mathematics.

Country	Percent Mathematics
Belgium	11 (Academic)-14 (Vocational)-12 (Others)
England	14
France	13 (Lycee sauf classes term.)-15 (CEG)-18 (Ecole Primaire)
Germany	12
Japan	13
Netherlands	11
Sweden	11
USA	14

Sources: Table 13.7 (Husén, 1967a).

The following describes the surveys used to gain information on cognitive performances of the 1949-1952 cohort in 2004. Table A.5 reports the individual response rates for the 2004 samples, by country; Table A.6 includes sample size, proportion by gender and mean age in the eight countries included in the comparative analysis.

ELSA. The English Longitudinal Study of Ageing investigates the lives of people in England. The study covers a broad range of topics such as health, economic situation, and quality of life (Marmot & Banks, 2003).

HRS. The Health and Retirement Survey is a large-scale longitudinal project that studies the transitions that individuals undergo toward the end of their working lives and in the following years in the USA. The HRS provides a body of multidisciplinary data that addresses the challenges and opportunities of aging (National Institute on Aging, 2007).

JSTAR. The Japanese Study of Aging and Retirement includes information on economic, social, and health conditions in Japan (Ichimura, Shimizutani, & Hashimoto, 2009). It was conducted by the Research Institute of Economy, Trade and Industry (RIETI), Hitotsubashi University, and the University of Tokyo, Japan for the first time in 2007. Therefore, Japanese samples in our study were observed in 2007, at the age of 55–58 (i.e. they are 3 years older than the samples in the other countries considered).

SHARE. The Survey of Health, Ageing and Retirement in Europe is a multidisciplinary and cross-national panel database of micro data on health, socioeconomic status as well as social and family networks of more than 45,000 individuals in fifteen European countries (Börsch-Supan et al., 2005).

Table A.5. Individual response rates for the 2004 samples.

Country	Individual Response Rate (%)
Belgium	90.5
England	67.0
France	93.3
Germany	86.2
Japan	60.0
Netherlands	87.8
Sweden	84.6
USA	88.6

Sources: http://www.share-project.org; http://hrsonline.isr.umich.edu/sitedocs/sampleresponse.pdf; http://www.ifs.org.uk/elsa/report03/ch1.pdf. Note: Individual response rate for Germany refers to the current Federal Republic of Germany.

Table A.6. Summary statistics on 1949-1952 cohort.

2004	Cohort 1949-1952 (52-55 years old)		
	Sample size	Males (%)	Mean age
Belgium	484	52.40	54.44
England	1207	49.56	53.89
France	279	55.12	53.47
Germany	329	51.62	53.46
Japan	530	48.87	56.60
Netherlands	355	48.65	53.41
Sweden	331	50.07	53.44
USA	1751	53.28	53.03

Sources: ELSA, wave 2 for England; HRS, wave 7 for the USA; JSTAR, wave 1 for Japan; and SHARE, wave 1 for European countries. Authors' calculation. Note: Japanese data are from 2007 and refer to people aged 55-58 belonging to the 1949-1952 birth cohort.

8.2 Appendix B

An extension of the study to England, Japan and the USA

For the sake of completeness, we have decided to include the analysis also on ELSA for England, JSTAR for Japan and HRS for the USA, taking into account the possible limitation of their samples when commenting on the results. The following characteristics of these surveys made us choosing not to keep these countries in the main part of the study.

First, ELSA and HRS differ from SHARE in not including the numeracy test and in the time allowed to answer the recall questions: SHARE allows one minute, while ELSA and HRS allow two minutes.

Second, in order to avoid period effects, we have to consider all the surveys on adults at the same point in time as the first wave of SHARE (i.e., wave 2 for ELSA and wave 7 for HRS). However, JSTAR has been conducted only in 2007, therefore catching the 1949-1952 cohort at ages 55-58, i.e. three years later than in the other countries considered.

Third, interviews in ELSA may be affected by re-test effect. As the respondents considered are undertaking the cognitive tests for the second time in wave 2, their performance may be better than at first exposure to the test. It has to be noted that the HRS sample, despite deriving from wave 7, is not affected by retest effects as the 1949-1952 cohort participated to HRS for the first time in 2004 (the year we observe them)[4].

Table B.1 shows the results of the analysis on eight countries. If we look at pupils' performance (Table B.1a), we see Japanese teenagers leading the ranking of the eight countries considered, with a mathematical test score of 31.2. England is at the mid of the ranking and the USA almost at the bottom. Moreover, England and Japan register the greatest within-country variation at teen ages in 1964 (their standard deviations are 18.5 and 16.9, respectively).

Once we compare country average performance in cognitive tests at mid-life (Table B.1b), it emerges that English scores are higher than those from the other countries in both immediate and delayed recall tests: mid-aged English on average recall above 6 words out of 10 immediately and 5 words a few minutes after they had been read. This result may however be inflated by possible retest effects that affect performances when undertaking the same test more than once. The relative improvement of the US sample in the recall tests may be partly affected by a higher availability of time that respondents had to answer the tests, as mentioned above. It is however interesting to notice that, based on data from HRS 2006[5], 95% of respondents of the 1949-1952 cohort group answered the immediate recall question within 60 seconds (i.e. the same time allowed by SHARE). We acknowledge, however, that respondents who took between 1 and 2 minutes to respond repeated a few more words, on average, than respondents who completed their task within 60 seconds. The Japanese sample belonging to the 1949-1952 cohort, despite being 3 years older than the others in 2004, performs as comparable to the USA in delayed recall and almost as the Dutch sample in the numerical test. However, an additional challenge in the comparability of the Japanese results refers to the fact that the 1964 test scores for all the countries (including Japan) as well as 2004 indicators for all other countries, but Japan, are

---

[4] Exceptions, i.e. individuals interviewed as partners in previous waves and then considered as respondents if they turn 50 years old before wave 7, are not included.

[5] We tested the time used to answer by respondents of the HRS 2006 because in the previous 2004 wave the time used by the interviewees to answer the recall tasks was not reported.

calculated by using population weights on the samples. However, JSTAR does not provide weights for its dataset.

Table B.1. a) Math test scores at 13 years old grade level in 1964; b) Mean scores and standard deviation (S.D.) of immediate recall, delayed recall and fluency tests at 52-55 years old in 2004; and relative frequencies reaching the final question of the numerical test battery at 52-55 years old in 2004. 1949-1952 cohort. 8 countries.

Cohort 1949-1952									
	a) 13 years old grade level		b) 52-55 years old						
	FIMS	S.D.	Immediate recall	S.D.	Delayed recall	S.D.	Fluency	S.D.	Numeracy
Belgium	30.4 (2)	13.7	5.5 (5)	1.6	4.0 (7)	1.8	21.5 (4)	6.2	55.1% (5)
England	23.8 (4)	18.5	6.3 (1)	1.6	5.2 (1)	1.8	22.3 (3)	6.4	
France	21.0 (6)	13.2	5.0 (8)	2.0	3.6 (8)	2.0	21.3 (5)	7.6	54.0% (6)
Germany	25.4 (3)	11.7	6.0 (2)	1.8	4.5 (5)	1.8	22.7 (2)	7.2	72.6% (1)
Japan	31.2 (1)	16.9	5.4 (7)	1.6	4.9 (2)	2.0			59.3% (4)
Netherlands	21.4 (5)	12.1	5.5 (5)	1.7	4.3 (6)	1.8	20.8 (6)	5.9	60.4% (3)
Sweden	15.3 (8)	10.8	5.7 (4)	1.4	4.7 (4)	1.6	25.5 (1)	7.2	64.5% (2)
USA	17.8 (7)	13.3	5.9 (3)	1.5	4.9 (2)	1.8			

Sources: a) Table 1.2 (Husén, 1967b); b) ELSA, wave 2 for England; HRS, wave 7 for the USA; JSTAR, wave 1 for Japan; and SHARE, wave 1 for European countries. Authors' calculation. Note: Japanese data are from 2007 and refer to the age 55-58 belonging to the 1949-1952 birth cohort.