Trends and determinants of the Flynn effect in cognitive functioning among older individuals in 10 European countries

Philipp Hessel,1 Jonas M Kinge,2,3 Vegard Skirbekk,4,5 Ursula M Staudinger5

ABSTRACT

Background Although cognitive performance levels in old age have increased in most countries, recent evidence documents a slowing down or even decline in cohort gains in highly developed countries. The aim of this study was to assess trends and determinants in secular cohort gains in cognitive functioning among older individuals and whether cohort gains are levelling off in most advanced countries.

Methods Data for individuals aged between 50 and 84 years from the Survey of Health, Ageing and Retirement in Europe in 10 European countries between 2004 and 2013 (n=92 739) were used to assess country and age-specific changes in immediate word recall. Multivariate random intercept models were used to assess associations between secular cohort changes in immediate word recall, initial performance levels and changes in country-level socio-demographic characteristics.

Results Performance in immediate word recall improved in all countries between 2004 and 2013 (from 4.40 to 5.08 words, P<0.05). However, secular cohort gains were significantly smaller in countries with initially higher performance levels (coeff.=−0.554, 95% CI −0.682 to −0.426). Changes in socio-demographic and health conditions, including decreases in cardiovascular disease, physical activity and educational achievement, were associated with larger secular cohort gains.

Conclusions Results may either reflect that some countries are approaching the limits of cognitive plasticity, are slowing in their progress or that societal structures have not yet been optimised to improve cognitive abilities in midlife and beyond, or a combination of these interpretations.

INTRODUCTION

Given current age-specific distributions of disease and functional capacities, population ageing represents a major challenge to health systems and the economy. While poor cognitive functioning at older ages is a particular concern, a key question is whether successive cohorts will continue to be cognitively healthier than their predecessors.1 If the age of onset of cognitive impairment can be postponed at a higher speed than increases in life expectancy, future cohorts of older individuals may be able to be productive for a longer time and require fewer care resources as current cohorts.2 In contrast, if cognitive abilities do not increase at least at the speed of increases in life expectancy, the sustainability of social security systems would be severely affected in the face of an increasing share of older individuals.

Indeed, a large body of literature that goes back to 19303 4 has documented that later-born cohorts performed better at various cognitive measures from the early 20th century until the early 21st century—a phenomena also known as the Flynn effect.5–10 Originally, the Flynn effect was used to describe such cohort improvements for cognitive functioning in young adulthood. Only recently has it been investigated with adults after age 50.7 However, to date no systematic comparative evidence exists that allows to determine whether secular cohort gains in cognitive functioning in later life exist in different countries to the same degree and what affects such changes or the lack thereof.

On the one hand, several pieces of evidence suggest that secular cohort gains in cognition are likely to continue. For example, improvements in educational attainment of later-born cohorts, less likelihood of insufficient nutrition,11 reductions in tobacco use in many countries12 and more complex and mentally stimulating work experiences have been found to increase cognitive performance along cohort lines.13 Evidence also suggests that cognitive health has improved well into the second half of life,13 including a decline in the prevalence of dementia in Europe and the USA.14 15

On the other hand, secular cohort gains in cognitive functioning may differ between countries or even level off. Thus, stalling or decreasing levels of functioning among older adults have been identified in highly developed countries such as Norway, Finland and France in recent years.16–18 This stagnation or even decline has also been reported for cognitive abilities among younger cohorts.19 Different reasons for this trend have been discussed such as the levelling off in the extension of participation in tertiary education in some countries,20 increasing levels of obesity and overweight which may negatively affecting functional capacities, such as cognitive performance through elevated risks of cardiovascular disease (CVD) and diabetes,21 or relatedly, the spreading of a sedentary lifestyle may negatively affect cognition.22 In addition, classical experimental work on the plasticity of cognitive performance has demonstrated that there are most likely biology-related limits to its modifiability which are reached more quickly in later life.23

Using longitudinal data of 10 European countries from the Survey of Health, Ageing and...
Retirement in Europe (SHARE) from 2004 to 2013, the aim of this study was to investigate trends and determinants of secular cohort changes in cognitive performance among older individuals. Specifically, we investigated whether secular cohort gains in cognitive performance between 2004 and 2013 varied systematically between countries according to their initial performance levels and whether levels and changes in socio-demographic conditions of countries might mediate such differences.

Individual-level characteristics such as higher education, healthier behaviours and labour force participation have been shown to be positively associated with cognitive functioning.24,25 According to the lifespan perspective,26 positive developments in those dimensions are prone to help making better use of the developmental cognitive potential than in individuals that are lacking such resources. Nevertheless, evidence from individual-level studies does not allow to make predictions about future trends in cognitive functioning as it does not take into account secular trends in the underlying behavioural factors, for example, trends in educational or labour force participation. A cohort perspective therefore can add to our understanding of trends and determinants of cognitive functioning as it takes into account both the association between various socio-demographic characteristics and cognitive functioning, as well as trends in the socio-demographic characteristics itself.

METHODS
Data sources
Data for this study came from SHARE.27 Data are representative of the non-institutionalised population aged ≥50 years in each participating country. For the purpose of this analysis, we used data from the first (2004/2005) and fifth wave (2013), enabling us to analyse secular cohort changes in cognitive performance over nearly a decade in the 10 countries participating in both waves (Austria, Belgium, Denmark, France, Germany, the Netherlands, Sweden, and Switzerland, Italy and Spain). In addition to SHARE, we also used data from the European Union’s statistical office (Eurostat) (see table 1). Out of a total of 97 213 age-eligible observations in waves 1 and 5, 4474 observations were omitted due to missing data on at least one of the covariates.

Measures
Cognitive functioning was measured by immediate word recall, an important predictor of cognitive impairment and dementia.28 It was measured by the immediate 10-word recall test, consisting of the words butter, arm, letter, queen, ticket, or she could remember.

To assess the extent to which other key societal trends are associated with secular cohort changes in immediate word recall, we selected covariates linked with cognitive performance such as educational levels, complex and mentally stimulating work experience, caloric intake or body mass index (BMI) and sedentary lifestyles,22 prevalence of CVD29 and overall levels of economic development.10 We included the following individual-level characteristics: (a) overweight defined as a BMI of ≥25; (b) prevalence of CVD indexed by an item that asked whether a respondent had suffered a stroke, heart attack or cerebral vascular disease; (c) engagement in physical activities, measured by asking respondents whether they ever engage in activities that require a low or moderate level of energy; (d) educational achievement in terms of highest educational level obtained; (e) economic activity rates; as well as (f) gross domestic product (GDP) per capita as a country-level indicator. These measures were calculated for 2004/2005 and 2013 in order to take into account possible changes in these covariates.

To assess trends and determinants of secular cohort changes in immediate word recall, based on individual-level data from SHARE, we calculated means by country, sex and 5-year age groups between ages 50 and 84 separately for waves 1 (2004/2005) and 5 (2013) for all variables using calibrated survey weights. To derive measures of secular cohort changes within each country, sex and 5-year age group stratum, we subtracted the means in 2013 from those in 2004/2005. For GDP per capita, we used the country-specific and year-specific values derived from Eurostat.

Statistical analysis
Changes in immediate word recall over time were compared for each country, sex and age group. Multivariate analyses of the association between change in immediate word recall between 2004/2005 and 2013, initial performance levels in 2004/2005 and covariates were conducted using linear regression models including random-intercepts for country, sex and 5-year age group strata.

The model can be formally written as

\[ Y_{csat} = \beta_0 + \beta_1 Y_{csat-1} + \beta_2 X_{sat-1} + \beta_3 (X_{sat} - X_{sat-1}) + \mu_{csa} + \epsilon_{csat} \]

where \( Y_{csa} \) represents the change in immediate word recall between time \( t-1 \) (2004/2005) and \( t \) (2013) within country \( c \), sex \( s \) and age group \( a \), \( \beta_0 \) represents the mean in immediate word recall.
recall in the population, $Y_{csat-1}$ represents immediate word recall within a country, sex and age group strata at time $t-1$ (2004/2005). $X_{csat-1}$ is a vector of country, sex and age-specific covariates, with the exception of GDP which is a country-level indicator, at time $t-1$ . 2004/2005. $X_{csat} - X_{csat-1}$ represents the change in the vector of covariates within each country, sex and age group strata between $t-1$ and $t$. The model thus controls for levels of covariates in 2004/2005 as well as their change between 2004/2005 and 2013. Finally, $\mu_{csa}$ is a random intercept for each country, sex and age group, while $\epsilon_{csa}$ is the error term.$^3$1

RESULTS

Descriptive statistics averaging across ages 50–84 and sexes are shown in table 2. Average performance in the immediate word recall test at ages 50–84 improved in all countries between 2004 and 2013 with an average increase of 0.68 words recalled. Largest improvements among occurred in France (0.98), Italy (0.92) and Spain (0.84). In contrast, average performance among men and women increased least in Denmark (0.51), Germany (0.25) and Sweden (0.46). Educational levels as well as economic activity rates and GDP increased in all countries even though to different degrees. Trends in prevalence of CVD, overweight and physical inactivity varied between countries. Online supplementary appendix table 1 shows the distribution of the performance in the word recall test for each country.

Figure 1 shows secular changes in immediate word recall between 2004/2005 and 2013 within each 5-year age group and separately for men and women, comparing the three countries with the highest average levels of immediate word recall in 2004 (Germany, Denmark and Switzerland) with the three countries having the lowest average levels (Italy, Spain and France). As expected, changes in average immediate word recall were significantly larger among countries with initially lower performance levels across all age groups ($P<0.05$). For example, between 2004/2005 and 2013 average scores in the 10-word recall test among men aged 50–54 improved by 0.81 points (from 4.18 to 4.99) in initially lowest performing countries, but only improved by 0.39 points (from 4.97 to 5.36) in initially highest performing countries. Among women, over the same period, average scores in the immediate word recall test improved by 0.51 points (from 5.24 to 5.75) in the highest performing countries and by 0.86 points (from 4.51 to 5.37) in the initially lowest performing countries. By 2013, average levels of immediate word recall among the initially lowest performing countries improved up to a level similar or above average levels in highest performing countries in 2004.

Figure 2 shows secular changes in immediate word recall between 2004 and 2013 relative to average performance in 2004. As indicated by the regression lines, within all age groups there is a negative association between average performance in 2004 and secular cohort change in cognitive functioning over time. Hence, age groups performing better in the immediate word recall test in 2004 showed smaller gains over time. We used linear multilevel models with random intercepts for country, sex and age group to assess the association between levels of immediate word recall in 2004 with secular cohort change over the period 2004–2013. As table 3 shows, average immediate word recall performance in 2004 was significantly negatively associated with secular cohort change over time. A 1-word higher performance in the 10-word recall test in 2004 was associated with a 0.554-word (95% CI −0.682 to −0.426) smaller change between 2004 and 2013.
Neither prevalence of overweight in 2004 nor changes between 2004 and 2013 were associated with changes in immediate word recall. Increases in prevalence of CVD were associated with significantly smaller cohort gains in immediate word recall. Prevalence of physical inactivity in 2004 as well as increases in inactivity over time were significantly associated with lower average changes in immediate word recall for both sexes. Although countries with a higher share of individuals aged 50–84 participating in the labour market showed higher average levels of immediate word recall, increases in the former were not significantly associated with improvements in immediate word recall. Increases in the share of individuals with secondary as well as tertiary education were significantly associated with cohort improvements in immediate word recall. There was no significant association between levels or change in GDP and changes in immediate word recall.

The association between levels of immediate word recall in 2004 and secular cohort change over the period 2004–2013 was stronger among men than among women (P<0.05). However, no significant interaction existed between sex and any of the remaining covariates.

**DISCUSSION**

Summary

The major aim of this study was to investigate population-level trends and determinants in secular cohort gains in cognition using a sample of 10 European countries. Specifically, we investigated whether secular cohort gains in immediate word recall were slowing down in initially better performing countries compared with initially worse performing countries. Our findings show that cohort gains occurred in all countries. However, these gains were significantly greater in initially lower performing countries. Furthermore, our results suggest that decreases in the prevalence of CVD and physical inactivity as well as increases in educational achievement were positively associated with secular cohort gains in cognitive functioning.

**Interpretation of findings**

In line with studies demonstrating that later-born cohorts generally show higher levels of cognitive performance than their
predecessors,\textsuperscript{7,8,32} all countries in our sample enjoyed secular cohort gains in cognitive functioning. While the key factors contributing to secular improvements in cognitive functioning among younger cohorts have been educational expansions, improved hygiene as well as reduced infectious disease prevalence in younger adulthood,\textsuperscript{7,32} this study demonstrated that increases in secondary education, reductions in CVD and physical inactivity were significantly associated with the largest relative reductions in the share of individuals with only primary education, considered one of the central drivers of the so-called Flynn effect.\textsuperscript{9}

Furthermore, cohort gains in cognitive functioning may be stalling in higher performing countries because societal structures addressing midlife and beyond may have not yet been optimised to provide for rich physical, social and intellectual stimulation to the same degree as it has been accomplished for the first decades of life. For instance, lifelong-learning regimes need more systematic development and implementation across all levels of education. Patterns of economic activity and work settings need yet to be optimised with regard to the maintenance of cognitive health. Finally, the broad implementation of volunteering activities including proper training is yet to be accomplished, and certain subpopulations continue to be not represented in volunteering settings.\textsuperscript{35}

Lastly, there are two additional mutually non-exclusive interpretations of our results that deserve consideration. The first is that we cannot rule out the possibility that cohort improvements in initially higher performing countries follow a slower pace that cannot be captured in the course of 10 years but may be observed after 20 years of societal change. Second, older cohorts in high-performing countries may indeed be approaching the biological limits of plasticity, that is, restrictions set by biological factors such as neurological and physiological constraints. Indeed, there is experimental evidence that supports this notion of limits of cognitive plasticity.\textsuperscript{36}

### Limitations

Despite several strengths of this study, a number of limitations should be acknowledged. Focusing on memory functioning as indicator of cognitive functioning may be considered a limitation as it covers only one facet of cognitive ageing, although a crucial one with regard to cognitive ageing.\textsuperscript{36} The association between latent factors of episodic memory and indicators of fluid intelligence has been found to be substantive.\textsuperscript{37} In this context, it is also important to note that a study using Estonian data showed that the Flynn effect—with the exception of numerical reasoning—pretty much generalised across different subtests of cognitive functioning.\textsuperscript{38} Furthermore, cognitive test performances may be influenced by a ‘ceiling effect’. Hence, it may be the case that individuals in the best performing countries were close to the upper boundaries of performance, that is, recalling all 10 words. However, even in the overall highest performing country Switzerland the mean number of words recalled was only 3.09, with merely around 3.5% of individuals correctly recalling 9 or 10 words. Thus, it is unlikely that

### Table 3 Multilevel model: determinants of secular cohort changes in immediate word recall at ages 50–84 in 10 European countries 2004–2013

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Change in average immediate word recall performance between 2004 and 2013 by country, sex and 5-year age groups</th>
<th>Coefficient</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate word recall: Baseline levels: 2004</td>
<td>−0.554</td>
<td>−0.682 to −0.426</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Overweight (%)</td>
<td>0.002</td>
<td>−0.004 to −0.007</td>
<td>0.539</td>
<td></td>
</tr>
<tr>
<td>CVD (%)</td>
<td>−0.011</td>
<td>−0.019 to −0.003</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Physically inactive (%)</td>
<td>−0.027</td>
<td>−0.034 to −0.020</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Economic activity rate (%)</td>
<td>0.003</td>
<td>0.001 to −0.004</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Secondary (%)</td>
<td>0.004</td>
<td>0.000 to −0.008</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Tertiary (%)</td>
<td>−0.002</td>
<td>−0.009 to −0.004</td>
<td>0.464</td>
<td></td>
</tr>
<tr>
<td>GDP per capita (log)</td>
<td>−0.295</td>
<td>−1.047 to −0.456</td>
<td>0.441</td>
<td></td>
</tr>
<tr>
<td>Change: 2004–2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight (%)</td>
<td>−0.003</td>
<td>−0.008 to −0.003</td>
<td>0.355</td>
<td></td>
</tr>
<tr>
<td>CVD (%)</td>
<td>−0.009</td>
<td>−0.018 to −0.001</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Physically inactive (%)</td>
<td>−0.023</td>
<td>−0.032 to −0.015</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Economic activity rate (%)</td>
<td>0.004</td>
<td>−0.001 to −0.009</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>Education (ref.: primary)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary (%)</td>
<td>0.006</td>
<td>0.003 to −0.009</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Tertiary (%)</td>
<td>0.006</td>
<td>0.002 to −0.010</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>GDP per capita (log)</td>
<td>0.832</td>
<td>−0.797 to −2.460</td>
<td>0.317</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>6.112</td>
<td>−1.359 to −13.583</td>
<td>0.109</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results are from a linear multilevel model including random intercepts for country, age group and sex. The dependent variable is the change in word recall between 2004/2005 and 2013 within each country, sex and age group strata. The regression coefficient for each covariate refers to the association between the latter and average changes in the number of words on the 10-word recall test over time across all strata. Results are substantially similar when using fixed effects for country, sex and age groups instead of random effects, as well as when omitting indicators of baseline levels in 2004.
Research report

Ceiling effects affected the results. Also, ‘retest’ or ‘practice’ effects (increased procedural knowledge, decreased test anxiety, shorter ‘warm-up’ phase) and selective attrition (poor performers may be more likely to drop out, which potentially could bias the results) may have compromised the findings of this panel study. In particular, initially lower performing individuals may have benefited disproportionately from practice effects. In our SHARE sample of wave 5 (2013), around 40% of individuals had already taken the immediate word recall test at a previous wave. Omitting this group from the analyses did not substantially affect the results.

CONCLUSIONS
Absolute differences in word recall among older adults across European countries have been decreasing. Our findings suggest that with increasing socioeconomic development among European countries cohort improvements in word recall are leveling off. Or at least they are slowing down in countries that have been enjoying higher levels of socioeconomic development for longer. Although we do not know yet whether the underlying factors of this trend are biological or socistructural, or possibly a combination thereof, monitoring potential limits of the plasticity of ageing is important with regard to the social security and health system implications of population ageing.

What is already known on this subject?
► Studies based on single countries or regions therein have shown improvements in cognitive functioning and reductions in prevalence of dementia among later-born cohorts.
► Recent studies in some highly developed European countries have shown that secular cohort improvements in cognitive functioning have stopped to exist.
► No study has assessed whether countries may be approaching upper boundaries of cohort improvements in cognitive ageing and what determines cohort changes at later life.

What this study adds?
► Cognitive functioning levels have improved in all countries.
► Participation in education as well as physical activity positively contributed to cohort improvements.
► Largest improvements in cognitive functioning in the last decade were observed in countries with initially lower performance levels, while improvements in initially higher performing countries are significantly smaller or close to zero.
► To prevent stalling of improvements in cognitive functioning, societal structure structures may have to be optimised to allow older individuals to maximise participation in social, economic and physical activities.

Contributors VS had the idea for the study. UMS developed the conceptual framework. JMK did the initial analysis. PH performed the final analyses of the data and compiled the figures and tables. All authors interpreted the results and contributed to writing the final version of the manuscript.

Funding This paper uses data from SHARE waves 1 and 5 (DOI: 10.6103/SHARE.w1.600, 10.6103/SHARE.w2.600, 10.6103/SHARE.w5.600). The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 SHARE (RII-CT-2006-062193, COMPARE: CITS-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: no. 211909, SHARE-LEAP: no. 227822, SHARE M4: N°216982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-1352, P01_AG050842, P01_AG08291, P01_AG12815, R21_AG025169, R11_AG-4553-01, IAG_BSKR06-11, OGHRA-04-064, HHSN27120130071C) and from various national funders. The work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 262700 and through grant “The burden of obesity in Norway: morbidity, mortality, health service use, and productivity loss” project number 250335/F20. PH was supported by a David E Bell Fellowship from the Harvard Center for Population and Development Studies.

Patient consent Detail has been removed from this case descriptions to ensure anonymity. The editors and reviewers have seen the detailed information available and are satisfied that the information backs up the case the authors are making.

Ethics approval The study uses publicly available survey data which were collected after obtaining ethical approval from national review boards and consent of participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Additional data are available on the journal’s homepage.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES
Trends and determinants of the Flynn effect in cognitive functioning among older individuals in 10 European countries
Philipp Hessel, Jonas M Kinge, Vegard Skirbekk and Ursula M Staudinger

*J Epidemiol Community Health* published online February 9, 2018

Updated information and services can be found at: [http://jech.bmj.com/content/early/2018/02/09/jech-2017-209979](http://jech.bmj.com/content/early/2018/02/09/jech-2017-209979)

These include:

**References**
This article cites 34 articles, 2 of which you can access for free at: [http://jech.bmj.com/content/early/2018/02/09/jech-2017-209979#ref-list-1](http://jech.bmj.com/content/early/2018/02/09/jech-2017-209979#ref-list-1)

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to: [http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to: [http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to: [http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)