What can we Expect of Cognition after 70? Cognitive Decline, Stability, and Gain between 70 and 79 Years in Three Swedish Birth Cohorts

Peter Karlsson1,2, Valgeir Thorvaldsson1, Ingmar Skoog3, Johan Skoog1 and Boo Johansson1

1Department of Psychology, University of Gothenburg, Gothenburg, Sweden
2School of Health and Welfare, Halmstad University, Halmstad, Sweden
3Institute of Neuroscience and Physiology, Department of Psychiatry and Neurochemistry, University of Gothenburg, Gothenburg, Sweden

Abstract

There is substantial heterogeneity in cognitive aging trajectories. Although the number of individuals showing cognitive decline increases with age, significant proportions remain stable or show gains. The aim of this study was to investigate birth cohort differences regarding proportions of individuals showing cognitive decline, stability, or gain between ages 70 and 79. We analyzed longitudinal data from three Swedish birth cohorts from the Gothenburg H70 Birth Cohort Studies, born 1901-02, 1906-07 and 1930, measured on four cognitive measures (i.e., verbal meaning, figure identification, logical reasoning and spatial ability). We used the standard error of measurement (SEM = \(\sqrt{1-r_{xx}}\)) at baseline (i.e., 70 years of age) to categorize participants as showing cognitive decline (if scores decreased by >1 SEM between ages 70 and 79), cognitive stability (if change was ≤1SEM), or cognitive gain (if increasing by >1 SEM).

Analyses, using \(\chi^2\)-tests, showed that a substantial proportion remained stable over the 9-year period. Our findings also indicated significant cohort differences for all four cognitive measures, with weak to moderate effect sizes. Interestingly earlier born cohorts contained a higher proportion of participants showing cognitive gain and a smaller proportion showing cognitive decline, compared to later born cohorts. Possible explanations for these cohort differences may relate to cohort differences in cognitive reserve, selective survival, and psychometric properties of the cognitive tests. These findings have implications for instance regarding the debate concerning postponing retirement age, and the use of cognitive tests in evaluations of for example work capability.

Keywords: Aging; Cohort differences; Heterogeneity in cognitive aging

Introduction

Aging is stereotypically often associated with universal cognitive decline [1]. A large body of research, however, demonstrates a substantial heterogeneity in cognitive aging trajectories [2-8] and Wilson et al., maintain that cognitive change in old age reflects factors that are specific to the individual rather than inevitable developmental processes [9]. Although the number of individuals showing cognitive decline increases with age, a significant proportion remain stable or may even show gains [1]. Schaie categorized participants of several age groups (i.e., 25-32, 32-39, 39-46, 46-53, 53-60, 60-67, 74-74, 81-88 years) from the Seattle longitudinal study as showing either cognitive decline, stability or gain, over a seven-year period [10]. Even though the proportion of participants showing cognitive gain or stability decreased in older age groups, and proportion showing decline increased, a significant proportion (≥55%) evinced cognitive stability and gain in all age groups, including the oldest.

Barnes et al., reported that 9% of the participants were cognitive maintainers, 58% minor decliners and 33% major decliners [2]. Yaffe et al., reported that thirty percent of their participants were categorized as maintainers, 53% as minor decliners and 16% as major decliners [8]. Cognitive aging is also a reflection of the historical context and the numerous life-course influences. In that context it is not surprising that many studies have shown that later born cohorts tend, on average, to perform better on several cognitive abilities which is often referred to as Flynn effects [11-18]. But, do these birth cohort differences manifest themselves also regarding proportions of individuals showing cognitive decline, stability, and gain? To the best of our knowledge there are no previous studies investigating possible birth cohort differences in heterogeneity of aging (i.e., cohort differences in the proportions of individuals showing cognitive decline, stability, or gain). Most studies investigating cohort differences in cognitive aging have addressed this question in more refined analyses of mean trends, rather than focusing on the straightforward question typically asked by older people themselves about the likelihood of stability and change in cognition in later life. Although mean trends analyses are informative and important, it is also important to focus on the heterogeneity of cognitive aging trajectories.

The aim of this study was to investigate proportions of individuals showing cognitive decline, stability, or gain between ages 70 and 79 years. A further aim was to study possible birth cohort differences in these proportions. For this purpose we analyzed data from three...
Swedish birth cohorts from the Gothenburg H70 Birth Cohort Studies, Sweden, born 1901-02, 1906-07 and 1930, with measurements at the same ages, 70 and 79, on four cognitive measures (i.e., verbal meaning, figure identification, logical reasoning and spatial ability). Using data from these three cohorts, we have previously reported findings indicating a steeper cognitive decline in later born cohorts [19,20]. For this reason our hypothesis is that there will be a higher proportion of decliners in later born cohorts.

**Design and Methods**

**Participants**

Participants for all three cohorts were identified and systematically selected for the Gothenburg H70 Birth Cohort Studies, through the Swedish Revenue Office Register. The design of the H70 study has been described in more detail elsewhere [21,22]. The H70 study was approved by the Ethics Committee of University of Gothenburg, and informed consent was obtained from all participants (Table 1).

**Table 1:** Sample characteristics in the H70 study stratified by birth cohort, gender, and measurement occasions.

<table>
<thead>
<tr>
<th>Cognitive Test</th>
<th>N</th>
<th>Women (%)</th>
<th>Men (%)</th>
<th>Age 70 M (SD)</th>
<th>Age 79 M (S)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verbal Meaning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 1901-02</td>
<td>175</td>
<td>70.29</td>
<td>29.71</td>
<td>17.53 (6.37)</td>
<td>17.53 (7.30)</td>
</tr>
<tr>
<td>Cohort 1906-07</td>
<td>207</td>
<td>60.87</td>
<td>39.13</td>
<td>18.85 (6.35)</td>
<td>16.64 (7.82)</td>
</tr>
<tr>
<td>Cohort 1930</td>
<td>104</td>
<td>46.15</td>
<td>53.85</td>
<td>21.48 (5.26)</td>
<td>21.36 (4.94)</td>
</tr>
<tr>
<td>Total</td>
<td>486</td>
<td>61.11</td>
<td>38.89</td>
<td>19.02 (6.29)</td>
<td>19.39 (6.65)</td>
</tr>
<tr>
<td><strong>Logical Reasoning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 1901-02</td>
<td>165</td>
<td>69.70</td>
<td>30.30</td>
<td>12.61 (4.60)</td>
<td>12.41 (4.95)</td>
</tr>
<tr>
<td>Cohort 1930</td>
<td>119</td>
<td>45.38</td>
<td>54.62</td>
<td>16.80 (4.60)</td>
<td>14.49 (5.30)</td>
</tr>
<tr>
<td>Total</td>
<td>284</td>
<td>59.51</td>
<td>40.49</td>
<td>14.24 (5.03)</td>
<td>13.93 (5.28)</td>
</tr>
<tr>
<td><strong>Spatial Ability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 1901-02</td>
<td>87</td>
<td>70.11</td>
<td>29.89</td>
<td>13.38 (6.64)</td>
<td>11.73 (7.28)</td>
</tr>
<tr>
<td>Cohort 1906-07</td>
<td>209</td>
<td>59.81</td>
<td>40.19</td>
<td>15.94 (6.90)</td>
<td>11.57 (7.43)</td>
</tr>
<tr>
<td>Cohort 1930</td>
<td>110</td>
<td>44.55</td>
<td>55.45</td>
<td>19.65 (6.83)</td>
<td>15.74 (6.41)</td>
</tr>
<tr>
<td>Total</td>
<td>406</td>
<td>57.88</td>
<td>42.12</td>
<td>16.47 (7.19)</td>
<td>13.84 (7.16)</td>
</tr>
<tr>
<td><strong>Perceptual Speed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 1901-02</td>
<td>81</td>
<td>71.60</td>
<td>28.40</td>
<td>16.45 (8.54)</td>
<td>14.91 (6.81)</td>
</tr>
<tr>
<td>Cohort 1906-07</td>
<td>206</td>
<td>59.71</td>
<td>40.29</td>
<td>19.51 (6.79)</td>
<td>14.77 (7.58)</td>
</tr>
<tr>
<td>Cohort 1930</td>
<td>127</td>
<td>45.67</td>
<td>54.33</td>
<td>25.79 (7.83)</td>
<td>21.38 (7.04)</td>
</tr>
<tr>
<td>Total</td>
<td>414</td>
<td>57.73</td>
<td>42.27</td>
<td>20.69 (8.29)</td>
<td>18.59 (7.80)</td>
</tr>
</tbody>
</table>

The cognitive measures used in the H70 are included in the Dureman and Sädes psychometric test battery [26] that was commonly used when the H70 study started in 1971 [27]. Verbal ability was measured with a Synonym test. Participants are asked to match a target word with a synonym word among five alternatives. The time limit was 7 minutes and maximum score was 30. We measured logical reasoning through a Figure Logic test. In this test participants are presented with geometrical figures, organized in rows with five figures per row. Four of the figures in each row share one common feature and one figure lacks this feature. Participants have to use logical reasoning to identify, as fast as possible, the common feature and then mark the figure that lacks this common feature. The test consists of 30 rows of figures, with an 8 minute time limit and a maximum score of 30.

Spatial ability was measured using a Swedish version of the Block Design test [28]. In this test participants are asked to organize colored wooden blocks in accordance with seven different patterns presented on cards (with increasing level of difficulty). The test has a 20 minute time limit and a maximum score of 42. Perceptual speed was measured with a Figure Identification test. Participants are asked to match, as fast as possible, a target figure with an identical figure placed in line with four distractor figures. The test has a time limit of 8 minutes and a maximum score of 60. These tests were administered in a similar manner for all three cohorts at the same ages, except that the Figure Logic test was omitted at ages 75 and 79 for the 1906-07 cohort. Therefore, data from the 1906-07 cohort, on the Figure Logic test, was omitted from the analyses. Further, for the 1901-02 cohort only half of the participants were given the perceptual speed and block design test at age 70. In order to minimize the risk of biasing influences of floor effects (due to factors such as dementia and other conditions producing a more severely compromised cognition) we omitted all participants with a score of zero (on either of the cognitive measures) at age 70, or a score of zero on both the 75- and 79-year measurements, from the analyses. In all, 60 participants were omitted from the analyses (25 from cohort 1901-02, 14 from cohort 1906-07, and 21 from cohort 1930).

**Cohort 1906-07:** This cohort was selected for inclusion in the H70 study, in the same manner as the first, in 1976-77. Twelve hundred and eighty-one participants, born between July 1, 1906- June 30, 1907 were selected and the response rate was 81% yielding another representative sample [23]. The participants in this cohort were given a number from 6-10, and participants with numbers 6 and 7 were selected for psychometric testing (n=513). For this subsample participation rate at age 70 was 75%. Attrition rate due to mortality was 26% from age 70-79, and attrition due to other reasons was 18%.

**Cohort 1930:** This cohort was selected for inclusion in the H70 study in the year 2000. Seven hundred and sixty-seven individuals born in 1930 on days 3, 6, 12, 18, 21, 24, or 30 of each month were sampled. The response rate was 71% at baseline. Comparisons of non-responders to responders yielded no significant differences regarding gender, marital status, 3-year mortality rate, inpatient psychiatric care [24] or specific diagnoses such as dementia or cardiovascular disorders [25]. Every second participant was randomly selected for psychometric testing (n=254), and the response rate for this subsample was 90% at age 70. Attrition rate due to mortality was 14% from age 70-79, and attrition due to other reasons was 31%.
Statistical analyses

We used the standard error of measurement (SEM =\sqrt{1-r_x}) at baseline (i.e., 70 years of age) to categorize participants as showing subsequent cognitive decline (if scores decreased by >1 SEM between ages 70 and 79), cognitive stability (if change was \leq 1 SEM), or cognitive gain (if increasing by >1 SEM). This was done for each of the four cognitive measures. To calculate SEM, we used the standard deviation (s), for each test, of the total sample at age 70 and the reliability coefficients (r_x) originally reported in Dureman et al., [26]. For the Synonym test reported reliability was 0.88 (test-retest), for the Figure Logic test the reliability was 0.89 (split-half), for the Block Design test 0.91 (test-retest) and for the Figure Identification test 0.96 (split-half). We used \chi^2-tests to analyze cohort differences in the proportion of participants showing cognitive decline, stability and gain. Further, we estimated effect sizes using Cramer’s V.

We also performed a sensitivity analysis to test the robustness of the results to the specific cut-off used to categorize participants as decliners, stable and gainers. Here we evaluated several alternative cut-offs. First we used SEM in the same manner as described above but setting the reliabilities to 0.80 and 0.70 when calculating SEM. Further, we used the standard deviations for the total sample at baseline, for each of the four cognitive measures, to categorize participants. Here we used both a change of >1 SD and >0.5 to categorize participants as showing decline or gain (otherwise showing stability). The results were largely the same no matter which cut-off we used. Therefore we have chosen to report only the results from the analyses using SEM based on the reliabilities reported in Dureman et al., [26].

Results

The distributions of participants categorized as showing cognitive decline, stability or gain from age 70-79 are presented in table 2, stratified by cohort and cognitive measure. As expected, a substantial proportion showed decline from age 70 to 79, most pronounced in the perceptual speed test, i.e., 60-79%. A significant proportion, however, demonstrated stability or even improvements in all tests. This pattern is especially evident in the verbal meaning test where as many as 63 to 75% showed stable or improved performance. The Chi-Square tests indicated significant cohort differences in the distributions for all four cognitive measures, with weak to moderate effect sizes. The proportion of participants with weak to moderate effect sizes (for verbal meaning \chi^2 (4) = 11.28 p = 0.024, Cramer’s V = 0.11; logical reasoning \chi^2 (2) = 19.38 p = 0.000, Cramer’s V = 0.26; spatial ability \chi^2 (4) = 9.52 p = 0.049, Cramer’s V = 0.11; for perceptual speed \chi^2 (4) = 27.99 p = 0.000, Cramer’s V = 0.18).

Even though effect sizes are quite weak, a distinctive pattern, common to all four cognitive measures is apparent as shown in figure 1 (Table 2). For all cognitive measures, earlier born cohorts contained a higher proportion of participants showing cognitive gain, and a smaller proportion of participants showing cognitive decline, compared to later born cohorts. This pattern is especially evident when comparing the 1901-02 cohort to the 1930 cohort.

Discussion and Implications

We investigated birth cohort differences regarding proportions of participants showing cognitive decline, stability and gain from age 70 to 79, in four cognitive measures (spatial ability, perceptual speed, verbal meaning and logical reasoning) using data from three population-based samples. Our main findings indicate significant cohort differences for all four cognitive measures, with weak to moderate effect sizes. For all cognitive measures, earlier born cohorts, especially the 1901-02 cohort, consists of a smaller proportion of participants showing decline, and a larger proportion of participants showing cognitive gain compared to later born cohorts (the 1930 cohort in particular). This is in line with our hypothesis.

Interestingly, our overall findings related to the question whether we should expect decline or hope for stability and even gains in the years after age 70 show that a substantial proportion in fact remained stable or showed performance gains. For the verbal meaning test, which is highly experience and education based, this was in fact the case for as many as 62 to 75 percent. The range for logical reasoning was 45 to 52 percent, for spatial ability 24 to 44 percent and for the perceptual speed test 20 to 40 percent. Thus, the proportion demonstrating stability or gain was quite substantial across tests in all three cohorts. Possible reasons for these findings could relate to cohort differences in cognitive reserve.

<table>
<thead>
<tr>
<th>Cognitive Measure</th>
<th>Cohort 1901-02</th>
<th>Cohort 1906-07</th>
<th>Cohort 1930</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Decline N (%)</td>
<td>Stability N (%)</td>
<td>Gain N (%)</td>
</tr>
<tr>
<td>Verbal Meaning</td>
<td>43 (24.6)</td>
<td>97 (55.4)</td>
<td>55 (20.0)</td>
</tr>
<tr>
<td>Logical Reasoning</td>
<td>80 (48.5)</td>
<td>28 (17.0)</td>
<td>57 (34.5)</td>
</tr>
<tr>
<td>Spatial Ability</td>
<td>49 (56.3)</td>
<td>27 (31.0)</td>
<td>11 (12.6)</td>
</tr>
<tr>
<td>Perceptual Speed</td>
<td>49 (60.5)</td>
<td>5 (6.2)</td>
<td>27 (33.3)</td>
</tr>
</tbody>
</table>

Table 2: Proportions of participants categorized as showing cognitive decline, stability and gain, stratified by birth cohort and cognitive measure.

a; Data for cohort 1906-07 on the Logical Reasoning test was not collected at ages 75 and 79 and therefore omitted from the present analyses.
To the extent that later born cohorts, on average, possess a higher cognitive reserve, as indicated by their higher level of performance (Table 1), they should, in accordance with the reserve hypothesis, evoke cognitive decline later in life (i.e., after more severe age-related brain pathology), compared to earlier-born cohorts. But, when they start to decline they are expected to decline at a faster rate [29-31]. This could possibly explain our findings, especially the higher proportion of decliners in later born cohorts.

Another possible explanation relates to cohort differences in selective survival. As life-expectancy has increased continuously in Sweden since the 19th century [32], the remaining average life-expectancy at age 70 is therefore somewhat higher in later born cohorts. This is also mirrored in the attrition rates due to mortality (with 35%, 26% and 12% for the 1901-02, 1906-07, and 1930 cohort respectively). For this reason, later born cohorts may include a larger proportion of more frail individuals, who have survived almost into their 80s but at the expense of greater manifest cognitive decline. However, this explanation would suggest a larger slope variability component (i.e., larger heterogeneity) in later born cohorts, especially the 1930 cohort, as compared with the earlier born cohorts. This was not supported by our data [19,20]. Further, studies comparing overall health and functioning indicates that the 1930 cohort is, on average, less frail than earlier birth cohorts [25,33].

An additional explanation for our findings may relate to the psychometric properties of the cognitive tests used in the H70. Gains in manifest test scores are not necessarily evidence of gains in the latent trait the tests are purported to measure. Rather, gains could also be related to practice/retest effects [34-36]. There are reports of substantial practice gains even when the repeated measurements are separated by several years [37-39]. Practice effects may also vary due to interaction between participant’s level of ability and task difficulty [40]. For example, on easy tasks the benefits of repeated testing may be greater for the less able individuals, while the more able may benefit more on difficult tasks. To the extent that the cognitive tests analyzed in this study could be considered easy, cohort differences in practice effects may provide yet another possible explanation for our findings of a relatively steeper average decline among members of the 1930 cohort. However, Thorvaldsson et al., report limited practice effects in the cognitive test battery used in the H70 [41].

Also, the tests may vary in sensitivity to detect within-person change depending on level of performance. For instance, it could be easier to detect decline in high-performing individuals and gain among low-performers. Since later born cohorts, on average, perform at a higher level compared to earlier born, this would make it more likely to detect decline in later born cohorts and gain in earlier born. The main strength of this study is that we used data from age-homogenous, representative, population-based samples, measured at the same ages with the same cognitive tests. A major limitation relates to the lower baseline response rate in the 1930-cohort (71%) compared to the 1901-02 and 1906-07 cohorts (85% and 81% respectively), as a lower response rate is likely associated with a somewhat more selective sample. Further, there exists a possibility that attrition might have compromised the representativeness of the cohorts for, partly, different reasons in the three cohorts. For instance, attrition due to mortality was higher in the earlier born cohorts compared to the 1930 cohort, while attrition due to other reasons (e.g., relocation, refusal to participate) was higher in the 1930 cohort.

Lastly, using an arbitrarily chosen cut-off to categorize participants confers the risk of obtaining spurious findings, i.e., results that are specific for the particular cut-off rather than reflecting real cohort differences. Therefore we performed a sensitivity analysis to evaluate the robustness of our results to the cut-off chosen. Here we tested several ways to categorize participants but the analyses revealed the same results, no matter the cut-off used. Thus, we are fairly confident that our findings are not spurious and only a reflection of particular cut-offs when categorizing participants. In sum, we found that substantial proportions of individuals remain cognitively stable, or even show gains, from age 70 to 79. Further, our findings indicate significant cohort differences regarding these proportions. For all four cognitive measures, later born cohorts contains a smaller proportion of participants showing cognitive gain, and a larger proportion of participants showing decline compared to earlier born cohorts. One implication of these cohort differences is that although later born cohorts evidence higher average cognitive functioning (i.e., Flynn effects) we can expect a higher proportion of individuals showing cognitive decline in later born cohorts. This in turn may have implications in the debate regarding postponing retirement (as a strategy to deal with possible economic strain imposed by population aging).

Further, our results may also have implications concerning older individuals own expectations concerning aging, but may also be of importance for the planning and funding of healthcare. Finally, our findings of cohort differences also have important implications concerning the use of cognitive testing to assist in evaluations of for instance disability, dementia status and work capability. The reported birth cohort differences needs to be taken into account in practices such as standardizations of cognitive tests, establishment of norms and cut-off values and decision-making based on cognitive evaluations.

Acknowledgement
The study was supported by a grant from FAS/Forte.

Funding

References


