Neuropsychologia 47 (2009) 2174–2180

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Flynn effects on sub-factors of episodic and semantic memory: Parallel gains over time and the same set of determining factors

Michael Rönnlund[a,*], Lars-Göran Nilsson[b]

[a] *Department of Psychology and Centre for Population Studies, Umeå University, S-90187 Umeå, Sweden*
[b] *Department of Psychology, Stockholm University and Stockholm Brain Institute, Sweden*

ARTICLE INFO

Article history:
Available online 13 November 2008

Keywords:
Flynn effects
Episodic memory
Semantic memory
Cohort factors

ABSTRACT

The study examined the extent to which time-related gains in cognitive performance, so-called Flynn effects, generalize across sub-factors of episodic memory (recall and recognition) and semantic memory (knowledge and fluency). We conducted time-sequential analyses of data drawn from the Betula prospective cohort study, involving four age-matched samples (35–80 years; $N = 2996$) tested on the same battery of memory tasks on either of four occasions (1989, 1995, 1999, and 2004). The results demonstrate substantial time-related improvements on recall and recognition as well as on fluency and knowledge, with a trend of larger gains on semantic as compared with episodic memory [Rönnlund, M., & Nilsson, L. -G. (2008). The magnitude, generality, and determinants of Flynn effects on forms of declarative memory: Time-sequential analyses of data from a Swedish cohort study. *Intelligence*], but highly similar gains across the sub-factors. Finally, the association with markers of environmental change was similar, with evidence that historical increases in quantity of schooling was a main driving force behind the gains, both on the episodic and semantic sub-factors. The results obtained are discussed in terms of brain regions involved.

Massive time-related IQ gains were observed in Western societies during the 20th century. The effect is referred to as the Flynn effect following large-scale analyses by Flynn (1984, 1987; for similar observations, see Lynn, 1982; Owens, 1966; Schaie, Labouvie, & Beuch, 1973; Tuddenham, 1948). With regard to magnitude of the effects, Flynn observed gains amounting to 3 IQ points per decade on the WAIS and an even higher rate of gains on measures of fluid reasoning such as versions of Ravens matrices. Subsequent research confirmed the generality of the effects by demonstrating them in a variety of other settings (e.g., Colom, Andrés-Pueyo, & Juan-Espinoza, 1998; Daley, Whaley, Sigman, & Neumann, 2003; Lynn & Hampson, 1986; Must, Must, & Raudik, 2003; Teasdale & Owen, 1989).

As noted by Hiscock (2007), Flynn effects have important practical implications. First, normative data on batteries devoted to the assessment of cognitive functions need to be updated with frequent intervals in order to ensure their population validity. Second, the presence of an upward drift in performance entails that a significant portion of the age-related variance in test scores are attributable to generational membership, or cohort, rather than maturation. In such a case, cross-sectional data covering an age-range for which the Flynn effects apply should indicate a steeper age-related decline as compared with longitudinal data, which is in line with observations from the Seattle Longitudinal Study on various intellectual abilities (e.g., Schaie, 1994, 1996), and with more recent observations based on data from the Betula Prospective Cohort Study on WAIS-R Block Design and episodic and semantic memory (Rönnlund & Nilsson, 2006; Rönnlund, Nyberg, Bäckman, & Nilsson, 2005).

Theoretically, Flynn effects have been regarded as one of the major puzzles for science to disentangle (e.g., Neisser, 1998). In particular, the presence of Flynn effects seems hard to reconcile with the widely accepted notion that IQ-test performance is highly heritable. Given that the time window during which Flynn effects have been operative is small (perhaps 100 years) an explanation involving genetic factors has been disregarded by most researchers (but see Mingroni, 2001, 2007). Thus, an account of the Flynn effect would have to resolve the paradox of how environmental factors may raise test scores by a substantial margin in spite of the fact that the cognitive measures on which the effects operate are associated with high heritability estimates at a given historical time-point.

The model by Dickens and Flynn (2001, see also Flynn, 2007) attempts to resolve the paradox by arguing that heritability estimates are boosted by gene-environment matching and so-called *multiplier effects*. For example, a child who demonstrates aptitude for mathematics is likely to be encouraged by parents and teachers to pursue his/her studies of the subject. By virtue of the feedback and additional resources available, the child's proficiency will likely

---

[*] Corresponding author.
E-mail address: michael.ronnlund@psy.umu.se (M. Rönnlund).

0028-3932/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2008.11.007

increase. In turn, the successes that follow are likely to encourage further interest. By virtue of this progressive gene-environment matching, environmental factors may have a larger influence than is suggested by heritability estimates (in which the proposed interactive effects are credited to genetic factors alone). In particular, the existence of “social multipliers” whereby a similar positive feedback loop generates a progressive increment in general skills in society is important for understanding the Flynn effect.

What needs to be specified are what constellation of environmental factors drive the Flynn effects. At this point a multitude of suggestions have been put forward, including increased test sophistication/altered response strategies (e.g. Brand, 1987), improved nutrition (Lynn, 1990), increased cognitive stimulation caused by urbanization (e.g., Schooler, 1998), changes in family structure towards smaller families and hence more provision of attention and materials for each child (Williams, 1998; cf. Zajonc, 2001), and changes in the educational system with more schooling for later-born cohorts as compared with those born earlier (e.g., Teasdale & Owen, 1987, 1989; cf. Ceci, 1991; Gustafsson, 2001). However, the extant evidence relies mainly on comparisons of similarities with regard to the presence and magnitude of secular trends in cognitive test performance and in the background factors, which is weak at best, as it could either be reflective of unrelated factors (a rise of the oil price of the same magnitude as that of success in the ability to complete number series is insufficient as an indicator of a causal relationship among the two variables), or be reflective of a third unmeasured variable.

Another unresolved issue concerns the extent to which Flynn effect generalize across cognitive ability domains. As noted by Hiscock (2007), the evidence is still biased towards measures of global intellectual functioning (e.g., WAIS full-scale IQ or IQ based on versions of Raven’s matrices), with less knowledge concerning specific cognitive abilities, or what is referred to as group factors in hierarchical models of intelligence (e.g. Carroll, 1993). As noted by Rönnlund and Nilsson (2008), a lacuna of the knowledge on secular trends in human cognitive functions concerns declarative long-term memory. In particular, evidence on secular trends on measures of encoding and retrieval of personally experienced events, or episodic memory (Tulving, 1972, 1983), has been lacking, despite the high degree of relevance of episodic memory in regard to cognitive theory, neuropsychology, and everyday functioning.

Utilizing data from the Betula prospective cohort study (Nilsson et al., 1997, 2004), Rönnlund and Nilsson (2008) therefore examined time-lag differences in episodic and semantic memory across adult age groups (35–80 years in a so-called time-sequential design (Schaie, 1965). The results revealed Flynn effects on semantic as well as episodic memory, with a numeric trend of larger gains on the former factor, gains that were about as large as for Block Design performance (Wechsler, 1981). With regard to these findings it is warranted to point out that the relatively large gains observed for semantic memory by Rönnlund and Nilsson deviate from the trend of only minor gains in performance on measures assumed to reflect crystallized ability (e.g. WAIS subtests, such as Information) observed elsewhere (see Flynn, 1999, 2007). In addition, the results by Rönnlund and Nilsson indicated that several factors may have contributed to the gains, including improvements in nutrition, historical changes in family structure (towards smaller families), and, most important, increased formal schooling for later-born as compared with earlier-born cohorts. Rather than comparing mean-level patterns in the cognitive variables with that of markers of the proposed underlying factors, the study used hierarchic regressions involving data obtained at the individual level to determine the extent to which the positive time-related variance in memory and cognition remained following control of the background factors. Critically, the inclusion of the latter (body height, sibship size, and years of formal schooling) reduced the time-related variance almost entirely (>90%), suggesting that cumulative changes in nutrition, family structure, and formal schooling are sufficient to account for Flynn effects on episodic memory, semantic memory, and on visuospatial test performance.

As demonstrated by Nyberg et al. (2003), the episodic and semantic memory systems proposed by Tulving (1972, 1983) may be further fractionated into sub-systems. This was demonstrated in confirmatory factor analyses showing that a model dividing episodic memory into recall and recognition and semantic memory into fluency and knowledge provided a better account of the covariance structure than a unitary (“declarative”) model, and than a two-factor (“episodic-semantic”) model. The finding that episodic and semantic memory is further dividable into sub-factors is interesting also from a neuroscientific point of view, given evidence that recall and recognition (e.g., Cabeza et al., 1997) and fluency and knowledge involve partly different areas of a the brain, a widely held position being that fluency and recall are more dependent on the integrity of frontal lobe and executive functions due to the increased demands on self-initiated retrieval processes, as compared with knowledge and recognition, respectively (e.g., Abrahams et al., 2000; Moscovitch, 1992). From the viewpoint of a hypothesis that Flynn effects on memory mainly reflect improved self-initiated retrieval processes one might expect gains to be larger on episodic recall and semantic fluency as compared with episodic recognition and semantic knowledge, respectively. If this held true the finding by Rönnlund and Nilsson of relatively larger gains on semantic memory as compared with gains reported by Flynn on WAIS Vocabulary and Information (see Flynn, 1999, 2007) could reflect the fact that the results were based on scores reflecting a combination of vocabulary (recognition of synonyms; potentially showing small gains) and fluency (potentially larger gains). Given these possibilities, it was deemed to be of considerable interest to examine the extent to which the Flynn effects demonstrated by Rönnlund and Nilsson (2008) generalize across semantic and episodic sub-factors. In addition we examined whether cohort-related differences within the same set of predictors as in Rönnlund and Nilsson (nutrition, family structure, and educational attainment) provide a similar account of the time-related variance in the memory sub-factors.

1. Method

The data emanated from the Betula study (Nilsson et al., 1997, 2004). The first measurement occasion (Time 1) took place in 1988–1990. On this occasion a sample of 1000 individuals, recruited by means of sampling from the population registry in Umeå, a city in Northern Sweden with about 110,000 inhabitants, participated. The sample was stratified with regard to age, with 100 individual in each of 10 age groups (35, 40, 45, 50, 55, 60, 65, 70, 75, and 80 years at date of test). The data collection required 2 years for completion. As a consequence, the 10 groups varied with regard to birth year from 1908–1910 (80-year old) to 1953–1955 (35-year old). A second, third, and fourth test occasion was undertaken 5 years (1993–1995), 10 years (1998–2000), and 15 years (2003–2005) following the first assessment. On these occasions the participants in the original sample were reassessed.

New samples involving groups that matched the original sample with regard to age (i.e., 35–80 years) were also included at Time 2, Time 3, and Time 4. Consequently, for each age level, participants in the new samples differed systematically (i.e., 5, 15, or 15 years, respectively) from those in the first sample with regard to birth year. Table 1 provides a schematic presentation of the time-sequential design (Schaie, 1965) of the study, including age, times-of-measurement, and birth year for the groups. As can be seen, the design involves a considerable range in terms of birth-cohorts (1909–1969) despite the relatively short time span (15 years) during which the data were collected.

The same means for selection (i.e., random) and the same exclusion criteria were adopted as at the first test occasion (dementia diagnoses, mental retardation, another native tongue than Swedish; see Nilsson et al., 1997, 2004 for further details). For financial reasons, the sample size was cut to half at Time 3 and Time 4, such that about 50 individuals per age cohort were assessed rather than 100, as was the case at Time 1 and Time 2.

Table 2 provides a descriptive summary of the background measures serving as potential predictors of the Flynn effects on the memory subfactors (body height, sibsize, and years of formal schooling), as a function of time of measurement, age, and sex.

Table 1
Design of the study, including age, mean time of measurement, and birth cohort of the included groups.

Age	Time of measurement			
	Time 1	Time 2	Time 3	Time 4
35	1954	1959	1964	1969
40	1949	1954	1959	1964
45	1944	1949	1954	1959
50	1939	1944	1949	1954
55	1934	1939	1944	1949
60	1929	1934	1939	1944
65	1924	1929	1934	1939
70	1919	1924	1929	1934
75	1914	1919	1924	1929
80	1909	1914	1919	1924

As is discernible from the marginal means in Table 2, there are time/cohort-related differences in the variables. More specifically, with time there is an increment in body height over age cohorts, smaller sibship sizes, and of increased educational attainment ($ps < .001$; in terms of population-standard deviation units the effects range from about 0.6 for body height to 1.5 for schooling from the 1909 to the 1969 cohorts, see Rönnlund & Nilsson, 2008). Thus, each of the variables under consideration is a candidate factor for explaining Flynn effect in the sense that each of them was subject to considerable changes over the time/cohort window targeted in the present study.

1.1. *Memory measures and data reduction*

The memory tests were 10 tests that had been administrated in exactly the same fashion across the four measurement occasions. Six tests were assumed to reflect episodic memory. Out of these, three were tests of *recall* (free recall of 16 motorically enacted verb–noun commands, free recall of 16 verb–noun sentences encoded verbally, and category-cued recall of nouns from non-enacted sentences). Another three measures were *recognition* measures. The latter set included free-choice recognition of nouns (hits-false alarms), free-choice recognition of faces (hits-false alarms), and forced-choice recognition of family names (hits) presented together with the faces at study. The semantic tests were a test of vocabulary as a single indicator of *knowledge* and three measures of word *fluency*, including tests wherein the task was to generate as many words as possible with (a) initial letter A, (b) initial M containing five letters, and (c) professions with initial letter B, respectively, in one minute (for a more detailed descriptions of the materials and procedure, see Nilsson et al., 1997).

The data were analyzed by means of confirmatory factor analyses using AMOS 7.0 (Arbuckle, 2007). The hypothesized model involved episodic and semantic memory as second-order factors. Recognition and recall were assumed to be sub-factors of episodic memory (i.e. first-order factors). Fluency was assumed to be reflected by the three word fluency measures and the vocabulary measures was in this model hypothesized to load on the semantic factor. The fit of this model was appropriate as judged from several fit indices (e.g., RMSEA, CFI, standardized RMR, see Rönnlund & Nilsson, 2006). In the next step, factor scores for the first-order episodic factors and for fluency were computed based on the factor-score regression weights provided as part of the factor analysis in AMOS. These scores were next transformed to z scores based on means and standard deviations of the entire Sample 1 at Time 1 in order to facilitate subsequent comparisons of Flynn effects for the separate ability factors. In the case of knowledge, the vocabulary scores were subjected to the same type of transformation again with S1T1 as the basis for anchoring the z scores.

2. Results

2.1. *Time/cohort-related patterns*

To appreciate the extent to which the cohort markers predicted the time-lag effects in cognitive performances, simple and hierarchic regression analyses were performed. In the simple analyses test year was regressed on each of the measures (following age and gender). Thus, the unstandardized regression coefficients indicate the annual gain on the memory factors in terms of z-score units. These were .010, .010, .015, and .013 for recall, recognition, fluency, and knowledge, respectively (all $ps < .001$). Thus, the results signal the presence of significant Flynn effects on each of the sub-factors of episodic and semantic memory.

Table 2
Descritpive summary (M; S.D.s within parantheses) of the background variables as a function of age, time of measurement (T1–T4), and gender.

Age	Variable											
	Body height (cm)				Sibsize				Education (years of schooling)			
	Time 1 (f/m)	Time 2 (f/m)	Time 3 (f/m)	Time 4 (f/m)	Time 1 (f/m)	Time 2 (f/m)	Time 3 (f/m)	Time 4 (f/m)	Time 1 (f/m)	Time 2 (f/m)	Time 3 (f/m)	Time 4 (f/m)
35	166.9/180.5 (6.8/7.3)	165.5/179.8 (7.1/6.8)	166.2/179.0 (6.2/6.7)	166.0/180.5 (4.7/6.2)	2.8/3.6 (1.5/1.8)	3.3/3.2 (1.6/1.7)	3.4/2.9 (1.6/1.4)	2.4/2.6 (0.8/0.8)	14.2/13.7 (2.6/2.6)	12.8/12.9 (2.5/2.6)	14.4/14.4 (2.0/3.4)	14.7/14.4 (3.2/3.8)
40	164.7/177.7 (5.2/7.1)	166.6/180.3 (7.1/6.3)	167.0/179.6 (7.9/6.4)	168.6/178.9 (6.1/6.7)	3.8/3.5 (2.0/1.9)	3.6/2.9 (1.5/1.5)	2.8/3.0 (1.1/1.3)	3.1/3.0 (2.3/1.4)	13.7/14.0 (3.2/3.8)	13.3/12.8 (3.1/2. 9)	13.1/14.2 (2.3/5.1)	13.6/12.4 (2.7/2.3)
45	164.4/178.6 (5.8/6.3)	166.6/177.4 (6.7/7.8)	168.5/179.7 (6.2/7.0)	165.3/178.0 (5.4/5.3)	3.7/3.1 (2.2/1.6)	3.4/3.9 (1.8/2.5)	3.6/3.0 (1.9/1.5)	3.6/2.6 (2.0/1.2)	12.6/12.9 (3.8/4.7)	13.9/13.6 (3.3/3.5)	14.1/13.1 (2.6/3.5)	14.4/14.4 (2.3/3.5)
50	163.3/175.3 (4.9/6.8)	164.5/176.7 (6.4/6.8)	162.2/179.8 (6.1/6.3)	164.9/177.7 (6.2/6.1)	3.8/4.8 (2.1/2.7)	3.7/4.0 (2.2/2.9)	3.1/3.1 (2.1/1.4)	2.9/3.5 (1.7/1.7)	10.2/10 .7 (3.4/4.2)	11.2/11.9 (3.5/4.1)	13.5/13.3 (4.2/3.2)	13.1/14.2 (2.7/3.9)
55	163.6/174.1 (6.1/7.4)	162.8/176.5 (5.3/4.7)	164.9/177.4 (7.5/6.8)	164.1/178.4 (5.1/7.6)	4.9/4.7 (2.9/2.6)	4.1/4.4 (2.6/2.5)	3.4/3.5 (1.5/1.7)	3.0/3.8 (1.4/2.2)	9.1/8.8 (3.6/2. 8)	10.8/10.2 (4.2/3.3)	11.6/12.1 (4.0/4.5)	12.7/12.0 (4.1/2.7)
60	161.4/174.9 (5.8/6.1)	164.1/174.8 (5.5/6.0)	162.8/174.5 (4.2/3.9)	163.9/175.6 (6.9/5.6)	4.5/4.8 (2.6/3.0)	3.9/4.6 (2.2/3.0)	3.3/4.9 (1.4/3.3)	2.8/3.5 (1.6/1.8)	8.9/8.8 (3.3/3.2)	9.9/9.8 (4.0/3.5)	9.9/9.8 (3.7/3.3)	12.4/11.6 (4.4/4.0)
65	161.0/175.2 (5.1/6.1)	161.2/175.3 (5.3/6.6)	161.3/176.8 (5.4/4.9)	164.1/177.7 (5.0/5.7)	5.4/5.5 (3.3/3.3)	4.2/5.1 (2.9/3.2)	4.3/4.0 (2.6/2.8)	4.1/4.6 (3.1/2.5)	7.5/9.0 (1.8/3.7)	8.1/8.5 (2.3/3.8)	7.5/8.2 (1.7/2.9)	11.6/10.2 (3.6/4.1)
70	160.3/174.7 (5.6/6.1)	160.0/174.5 (5.6/4.8)	162.5/173.9 (5.3/5.6)	160.6/174.5 (5.6/4.7)	5.4/4.9 (3.2/3.0)	4.9/5.2 (2.9/3.0)	4.7/5.4 (2.6/3.5)	4.9/4.0 (3.2/2.3)	7.6/8.8 (2.8/3.6)	7.6/8.1 (2.9/3.9)	7.9/7.3 (2.3/2.1)	10.0/9.2 (4.2/3.6)
75	158.2/174.9 (6.9/6.5)	157.9/176.5 (6.2/5.9)	158.8/176.5 (4.5/5.8)	159.5/173.3 (4.8/6.6)	6.7/6.2 (3.5/3.0)	5.5/5.1 (2.8/2.8)	4.2/4.0 (2.7/1.6)	5.0/5.0 (2.9/3.2)	7.2/7.8 (2.2/3.3)	7.8/8.7 (3.8/4.4)	6.8/8.2 (1.6/3.1)	7.6/8.2 (2.4/3.4)
80	159.4/171.1 (6.0/6.4)	157.9/171.2 (5.6/7.3)	159.2/172.5 (5.2/5.7)	159.2/174.0 (5.3/6.2)	5.2/6.5 (3.0/3.4)	6.2/6.9 (2.8/3.3)	5.3/6.0 (2.7/3.0)	4.0/5.0 (2.6/3.0)	7.5/7.1 (2.9/3.5)	6.8/7.2 (2.7/3.5)	7.3/7.6 (2.9/4.0)	9.0/8.2 (3.1/3.4)
Total	162.3/175.7 (6.3/7.0)	162.6/176.3 (6.7/6.8)	163.4/176.9 (6.7/6.3)	166.0/176.8 (6.1/6.5)	4.6/4.7 (2.9/2.9)	4.3/4.4 (2.6/2.8)	3.8/4.0 (2.2/2.5)	3.6/3.8 (2.4/2.3)	9.8/10.1 (3.9/4.2)	10.0/10.5 (4.1/4.1)	10.6/10.8 (4.6/4.5)	12.0/11.5 (4.0/4.1)

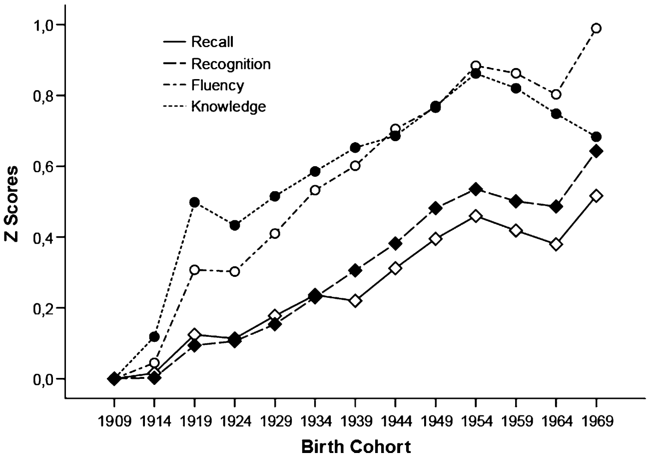

Fig. 1. Estimated cohort gradients for the episodic (recall and recognition) and semantic sub-factors (fluency and knowledge).

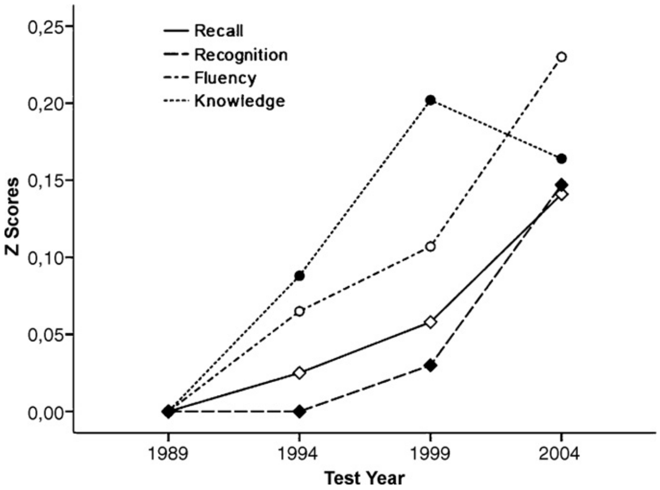

Fig. 2. Mean-level performance (*z* scores) across measurement occasions (test year) for the episodic (recall and recognition) and semantic sub-factors (fluency and knowledge).

As judged from the β-weights, the slopes of the time-related increment are highly similar across the memory sub-factors. In order to get a more fine-grained picture of the course of these Flynn effects over the full range of the included birth cohorts, the available contrasts given in Table 2 were utilized to compute cohort differences and to estimate the progressive changes in mean-level memory performance from 1909 to 1969. This was done in accord with the formula $Cd_i = \sum_j^1 (M_{ij+1} - M_{ij}))/a$, where M_{ij} is the unweighted mean for Cohort i at age j, and a is the number of observations available for each cohort; Schaie (1965). Cumulative summation of the obtained estimates across birth cohorts (i.e., $C2$–$C1$, $C3$–$C2$, . . ., $C13$–$C12$, where $C1$ = 1909 cohort and $C13$ = 1969 cohort) yields the cohort gradients depicted in Fig. 1.

As can be seen, the data in Fig. 1 confirms the impression of an upward drift in performance from the earliest-born (1909 cohort) to the birth cohort born in 1969 and a largely parallel improvement within components of second-order factors. The general pattern is of a gradual increment in performance across the memory factors, even though there is a tendency of levelling off in the improvements for cohorts born around 1950 and onwards. In the case of knowledge/vocabulary, there is even a numeric trend of slightly lower levels of performance across the subsequent birth cohorts.

As an alternative means to present these data, Fig. 2 depicts gives the mean-level performance across measurement occasions. Apart from (presumably) random fluctuations in performance over time/samples, the scaling of the gains in terms of time (rather than birth cohort) seem to indicate a trend of accelerating gain for the episodic sub-factors (relatively larger gains from T3 to T4, overall, as compared with differences between other measurement occasions). For knowledge/vocabulary a reversed trend is discernible, which mainly seems to reflect minor negative differences between the five youngest age groups assessed at Time 4 and those assessed at Time 3 (rightmost column) much in line with the pattern in Fig. 1 of a trend of stagnation, or even a minor loss, for the most recent cohorts on vocabulary.

2.2. Predictors of the Flynn effects

Having established the presence of Flynn effects on each of the studied memory factors the important issue of what caused the rising test scores remains. In order to address this issue, and, in particular, to determine whether the same or a different set of variables underlie the gains at the level of the different memory factors, regression analyses of the time-related differences in memory were conducted. In these analyses, the factor scores for each of the memory factors was the criterion and three theoretically relevant cohort markers were the predictors. The predictors were entered in four steps (blocks). The order of entry following the demographic predictors (age, sex; entered as a separate first block) was (1) body height, (2) sibsize, and (3) years of formal education. The order of entry was motivated by the hypothesized developmental sequence according to which the underlying constructs may be assumed to exert their influence. Specifically, nutritional changes (reflected by body height), were assumed to exert the earliest influence, whereas the influence with regard to education was assumed to emerge last ontogentically. In the final step, test year was entered to determine whether this variable still accounted for variance in performance beyond the influence from the foregoing set of variables.

The results are summarized in Table 3 (recall and recognition) and Table 4 (fluency and knowledge). The tables include the standardized regression weight for each of the included predictors, the increment in R^2 for each step and the total R^2 ($\sum R^2$) for the

Table 3

Summary of hierarchic regression analyses of episodic sub-factors (recall and recognition).

Block/variable	Recall				Recognition			
	β	ΔR^2	Total R^2	% SOS	β	ΔR^2	Total R^2	% SOS
1. Age	–.363*				–.356*			
Sex	–.189*	.377*	.377		–.192*	.373*	.373	
2. Body height	.064	.008*	.384	22.3	.059*	.007*	.380	21.5
3. Sibship size	–.046*	.009*	.393	54.8	–.047*	.009*	.389	54.1
4. Education (years)	.374*	.091*	.484	99.9	.379*	.094*	.483	99.9
5. Test year	–.009	.001	.485		–.002	.000	.483	

% SOS = percentage shared over simple effects.

* p <0.01.

Table 4
Summary of hierarchic regression analyses of semantic sub-factors (fluency and knowledge).

Block/variable	Fluency				Knowledge			
	β	ΔR^2	Total R^2	% SOS	β	ΔR^2	Total R^2	% SOS
1. Age	−.155*				.004			
Sex	−.193*	.221*	.221		−.106*	.097	.097	
2. Body height	.091*	.014*	.235	18.7	.066*	.011	.108	21.5
3. Sibship size	−.056*	.014*	.249	47.6	−.078*	.019	.126	26.4
4. Education (years)	.470*	.146*	.395	98.5	.487*	.155	.281	99.3
5. Test year	.007	.000	.395		−.011	.000	.281	

models. The amount of reduction of the time-related variance (% SOS, for shared over simple effect; Lindenberger & Pötter, 1998), was computed by entry of test year following each step in separate analyses and the values are presented in the last column of Tables 2 and 3.

With regard to each of the memory factors, the results show that the predictors share variance with the memory factors and that they, together, remove almost entirely the variance in memory accounted for by test year. Out of the predictors, formal schooling is assigned the largest β-value and is also the single predictor that accounts for most of the time-related variance (e.g., as judged by the % SOS) even though it was entered last among the background variables.

3. Discussion

The objectives of the study were to examine whether (1) Flynn effects generalize across sub-factors of episodic (recognition and recall) and semantic memory (knowledge and fluency) and (2) the same or different cohort factors account for the time-related gains on these memory factors.

With regard to the first issue, the results indicate that Flynn effects are highly general. More specifically, the gains were highly similar in magnitude for recall and recognition, and comparable in regard to magnitude for semantic knowledge and fluency. As such, the findings are in line with the major pattern emergent from prior studies in the sense that secular gains in cognition have been observed across a number of other cognitive domains (e.g. fluid reasoning and spatial ability). At this point, it is warranted to draw attention to the fact that over domains of cognitive abilities exceptions to this general pattern exist. For example, Schaie (1994, 1996) found a trend for numerical ability suggestive of no improvement over cohorts and even a declining trend for cohorts born after 1950. One potential reason for the strong Flynn effects on many ability factors, yet smaller effects on others, is that the Flynn effect exerts its influence via the growth of neural substrates underlying psychometric *g*, such that factors strongly related to *g* are subjected to larger changes and those depending less on *g* are associated with smaller secular gains.[1] This issue merits further attention and the extant evidence based on the so-called *method of correlated vectors* is inconclusive, with some studies indicating that the Flynn effect is mainly on the *g* factor (Colom, Juan-Espinosa, & Garcia, 2001) whereas others (e.g., Rushton, 1998) favor the view that the effects exert their influence via gains on more specific abilities and skills. Within the domain of memory functioning it further remains to be demonstrated whether the Flynn effects are restricted to declarative memory or whether they are detectable also on other memory systems.

With regard to the issue of what factors stand out as plausible predecessors of the Flynn effects, the present results indicate that improvements in nutrition, a reduction in family size, and increased formal schooling with historical time are factors which may all have contributed to the Flynn effects observed. Of particular interest at this point is that these variables accounted for virtually all of the time-related variance in performance across the memory factors (cf. also visuospatial ability, Rönnlund & Nilsson, 2008). As such, the present findings suggest that the Flynn effects primarily reflect the influence of cohort factors operative relatively early in development, unlike models (e.g., Dickens & Flynn, 2001) that appear to assume that the environmental effects are time (or period-) related (i.e. occur regardless of age or birth cohort), even though the use of time-lag comparisons are inherently ambiguous with regard to time and cohort effects and we acknowledge the possibility that time-related factors may have a role in explaining some of the secular gains in cognitive test performance. In any case the finding that environmental factors account for the time-related variations in performance seem to disfavor an alternative explanation of Flynn effects involving hybrid vigor as a main determinant (Mingroni, 2007).

Whereas the present study indicates that Flynn effects likely reflect changes in multiple factors, formal schooling stands out as a main factor. At this point it is warranted to draw attention to the fact that, both with regard to the magnitude of cognitive gains across the present set of measures and with regard to the relative influence of cohort-related variables the present findings are applicable to Sweden, and need not generalize across national and historical contexts. The present indication that our prior observation of substantial gains in semantic memory (Rönnlund & Nilsson, 2008) were not driven by inclusion of fluency measures but generalize to generalize to vocabulary, unlike the US data by Flynn (e.g., 2007; but see Uttl & van Alstine, 2003) could, for example, partly reflect variations with regard to national context.

Nevertheless, the finding that formal schooling is an important factor behind the Flynn effects, is much in line with indications elsewhere that effects of formal schooling are not restricted to measures of general knowledge and the like (reflecting crystallized intelligence) but exert a more widespread influence on cognitive performance (Ceci, 1991). As concerns specific mechanisms by which formal education could serve to enhance cognitive performance, behavioral as well as neuropsychological hypotheses have been put forward. At the behavioral level, it has been argued that differences over time in style of responding, toward more guessing and faster/more complete responding, is a key factor (Brand, 1987; Brand, Freshwater, & Dockrell, 1989), a factor that would likely be driven by educational practices. However, this factor would mainly be applicable to Flynn effects on episodic recognition (rather than recall) measures. Important to note at this point is that two of the recognition measures used at present (faces and objects) were corrected for guessing (false alarms) and that recognition of names was tested with a forced-choice procedure. Thus, a differential style of responding is not a likely explanation of the Flynn effects observed at present. In fact the finding of gains that generalize across recall

---

[1] Certainly, additional factors than the *g*-loading of a particular measure, such as a qualitative shift in educational practice, could moderate the gains on specific factors. As noted by a reviewer, the example of lack of gains in numerical ability could for example reflect the introduction of the slide rule and the calculator.

and recognition rather point to encoding factors as being important to consider. A possibility here is that increased knowledge (as reflected by the present gains over cohort on a measure of vocabulary) served to increase the general level of item-specific semantic encoding processes (e.g., verbal elaboration) over time/generations. From this viewpoint, the Flynn effects on episodic memory could be regarded to come partly as the results of more efficient semantic memory processes. Progression towards abstract reasoning with higher levels of educational attainment, possibly magnified by a shift in educational practice and thinking in general in the same direction (cf. Flynn, 2007; Schaie, Willis, & Pennak, 2005), may be a reason of similar gains in non-verbal test performance, such as on tests of non-verbal reasoning or, as demonstrated by Rönnlund and Nilsson (2008), in visuospatial performance.

At the neural level, there are findings suggesting that higher levels of education may increase reserve capacity and alter brain activity in memory tasks in a persistent fashion (Springer, McIntosh, Winocur, & Grady, 2005), including increased activity in frontal and/or temporal brain areas, activities that possibly signal enhanced item-specific encoding, much in line with the possibility that encoding factors are central behind the memory gains. A more general neurodevelopmental-schooling hypothesis has been proposed by Blair et al. (2005) in relation to the Flynn effect. According to this hypothesis, the increasing cognitive demands required in school due to the dramatic educational changes occurring during the 2000th century should have improved fluid cognition and the prefrontal cortex. Although empirical support for this hypothesis is still sparse, it should be mentioned that a few studies (e.g., Klingberg, Forssberg, & Westerberg, 2002) on fluid skill/working memory training have provided results in line with this hypothesis. Increased working memory capacity, performance on Raven's matrices and increased prefrontal and parietal activity was demonstrated after this type of training (see Klingberg, 2006; Olesen, Westerberg, & Klingberg, 2004; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005). More generally, cognitive stimulation offered by schooling and other factors, like more exposure to adult language in children, which may underlie the present finding of a relation between sibship size and Flynn effects likely promote the development of the neural substrates underlying cognitive abilities (cf. Garlick, 2002). The increased cognitive stimulation offered by prolonged schooling and other factors together with improved nutrition over cohorts (cf. Lynn, 1990; Martorell, 1998), in line with the present findings, seem to provide a full account of a growth of the underlying neural substrates underlying the Flynn effects (for evidence of larger brain sizes over cohorts, see Storfer, 1999).

More precise knowledge concerning the manifestation of Flynn effects on various cognitive measures at the neural level are currently lacking. An important issue for future research to address is the extent to which the Flynn effects observed across various types of cognitive ability measures are driven by increased neural efficiency within specific subsystems and the extent to which these gains reflect improvements in one or a few central capacities (e.g., g, executive functions) that are important across measures. The application of a sequential research design as used in the present study gathering data both at the behavioral and neural levels (e.g., by use of fMRI) for independent age-matched samples assessed on at least two different time points would seem necessary to provide definitive evidence in this regard.

In conclusion, time-sequential analyses of Swedish data from the Betula study extend the pattern of substantial time-related gains on episodic and semantic memory by demonstrating that they generalize across sub-factors of episodic as well as semantic memory. Finally, the results demonstrated that changes in multiple environmental factors, including education, family structure and nutrition likely underlie these Flynn effects.

Acknowledgements

The Betula Study is funded by the Bank of Sweden Tercentenary Foundation (1988-0082:17; J2001-0682), Swedish Council for Planning and Coordination of Research (D1988-0092, D1989-0115, D1990-0074, D1991-0258, D1992-0143, D1997-0756, D1997-1841, D1999-0739, B1999-474), Swedish Council for Research in the Humanities and Social Sciences (F377/1988-2000), the Swedish Council for Social Research (1988-1990: 88-0082, and 311/1991-2000), and the Swedish Research Council (345-2003-3883, 315-2004-6977). Also, we acknowledge the contribution by the staff in the Betula project.

References

Abrahams, S., Leigh, P. N., Harvey, A., Vythelingum, G. N., Grisé, D., & Goldstein, L. H. (2000). Verbal fluency and executive dysfunction in amyotrophic lateral sclerosis (ALS). *Neurpsychologia*, *38*, 734–747.

Arbuckle, J. L. (2007). *Amos 7.0 user's guide*. Chicago: SPSS.

Blair, C., Gamson, D., Thorne, S., & Baker, D. (2005). Rising mean IQ: Cognitive demand of mathematics education for young children, population exposure to formal schooling, and the neurobiology of the prefrontal cortex. *Intelligence*, *33*, 93–106.

Brand, C. R. (1987). Bryter still and bryter? *Nature*, *328*, 110.

Brand, C. R., Freshwater, S., & Dockrell, W. B. (1989). Has there been a "massive" rise in IQ levels in the West? Evidence from Scottish children. *Irish Journal of Psychology*, *10*, 388–393.

Cabeza, R., Kapur, S., Craik, F. I. M., McIntosch, A. R., Houle, S., & Tulving, E. (1997). Functional neuroanatomy of recall and recognition: A PET study of episodic memory. *Journal of Cognitive Neuroscience*, *9*, 254–265.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *5*, 703–722.

Colom, R., Andrés-Pueyo, A., & Juan-Espinosa, M. (1998). Generational IQ gains: Spanish data. *Personality and Individual Differences*, *25*, 927–935.

Colom, R., Juan-Espinosa, M., & Garcia, L. F. (2001). The secular increase in test scores is a 'Jensen' effect. *Personality and Individual Differences*, *30*, 553–558.

Daley, T. C., Whaley, S. E., Sigman, M., & Neumann, C. (2003). IQ on the rise: The Flynn effect in rural Kenyan children. *Psychological Science*, *14*, 215–219.

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, *108*, 346–369.

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

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

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, *54*, 5–20.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge: Cambridge University Press.

Garlick, D. (2002). Understanding the nature of the general factor of intelligence: The role of individual differences in neural plasticity as an explanatory mechanism. *Psychological Review*, *109*, 116–136.

Gustafsson, J.-E. (2001). Schooling and intelligence: Effects of track of study on level and profile of cognitive abilities. *International Education Journal*, *2*, 166–186.

Hiscock, M. (2007). The Flynn effect and its relevance to neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *29*, 514–529.

Lindenberger, U., & Pötter, U. (1998). The complex nature of unique and shared effects in hierarchical linear regression: Implications for developmental psychology. *Psychological Methods*, *3*, 218–230.

Klingberg, T. (2006). Development of a superior frontal-intraparietal network for visuo-spatial working memory. *Neuropsychologia*, *44*, 2171–2177.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *24*, 781–791.

Lynn, R., & Hampson, S. L. (1986). The rise of national intelligence: evidence from Britain, Japan and the USA. *Personality and Individual Differences*, *7*, 23–32.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature*, *297*, 222–223.

Lynn, R. (1990). The role of nutrition in secular increases in intelligence. *Personality and Individual Differences*, *11*, 273–285.

Martorell, R. (1998). Nutrition and the worldwide rise in IQ scores. In U. Neisser (Ed.), *The rising curve* (pp. 183–206). Wahington: American Psychological Association.

Mingroni, M. A. (2001). The secular rise in IQ: Giving heterosis a closer look. *Intelligence*, *32*, 65–83.

Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other trends. *Psychological Review*, *114*, 806–829.

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, *4*, 257–267.

Must, O., Must, A., & Raudik, V. (2003). The secular rise in IQs: In Estonia, the Flynn effect is not a Jensen effect. *Intelligence*, *31*, 461–471.

Neisser, U. (1998). Introduction: Rising test scores and what they mean. In U. Neisser (Ed.), *The rising curve* (pp. 3–22). Wahington: American Psychological Association.

Nilsson, L.-G., Bäckman, L., Erngrund, K., Nyberg, L., Adolsfsson, R., Bucht, G., et al. (1997). The Betula prospective cohort study: Memory, health, and aging. *Aging, Neuropsychology, and Cognition*, *4*, 1–32.

Nilsson, L.-G., Adolsfsson, R., Bäckman, L., de Frias, C. M., Molander, B., & Nyberg, L. (2004). Betula: A prosepctive cohort study on memory, health and aging. *Aging, Neuropsychology, and Cognition*, *11*, 134–148.

Nyberg, L., Maitland, S. B., Rönnlund, M., Wahlin, Å., Bäckman, L., & Nilsson, L.-G. (2003). Selective adult age differences in an age-invariant multifactor model of declarative memory. *Psychology and Aging*, *18*, 149–160.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7*, 75–79.

Owens, W. A., Jr. (1966). Age and mental ability: A second follow-up. *Journal of Educational Psychology*, *57*, 311–325.

Rönnlund, M., & Nilsson, L.-G. (2006). Adult life span patterns in WAIS-R Block Design performance: Cross-sectional versus longitudinal age gradients and relations to demographic factors. *Intelligence*, *34*, 63–78.

Rönnlund, M., & Nilsson, L.-G. (2008). The magnitude, generality, and determinants of Flynn effects on forms of declarative memory and visuospatial ability: Time-sequential analyses of data from a Swedish cohort study. *Intelligence*, *36*, 192–209.

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L.-G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Data from a population based study. *Psychology and Aging*, *20*, 3–18.

Rueda, M. R., Rothbart, McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 14931–14936.

Rushton, J. P. (1998). Secular gains in IQ not related to the g factor and inbreeding depression–unlike Black-White IQ differences: A reply to Flynn. *Personality and Individual Differences*, *26*, 217–225.

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

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist*, *49*, 304–313.

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

Schaie, K. W., Labouvie, G. V., & Beuch. (1973). Generational and cohort-specific differences in adult cognitive functioning: A fourteen-year study of independent samples. *Developmental Psychology*, *9*, 151–166.

Schaie, K. W., Willis, S. L., & Pennak, S. (2005). An historical framework for cohort differences in intelligence. *Research in Human Development*, *2*, 43–67.

Schooler, C. (1998). Environmental complexity and the Flynn effect. In U. Neisser (Ed.), *The rising curve* (pp. 67–79). Washington: American Psychological Association.

Springer, M. V., McIntosch, A. R., Winocur, G., & Grady, C. L. (2005). The relation between brain activity during memory tasks and years of education in young and older adults. *Neuropsychology*, *19*, 181–192.

Storfer, M. D. (1999). Myopia, intelligence, and the expanding human neocortex. *International Journal of Neuroscience*, *98*, 153–276.

Teasdale, T. W., & Owen, D. R. (1987). National secular trends in intelligence and education: A twenty-year cross-sectional study. *Nature*, *325*, 119–120.

Teasdale, T. W., & Owen, D. R. (1989). Continuing secular increases in intelligence and a stable prevalence of high intelligence levels. *Intelligence*, *13*, 255–262.

Tuddenham, R. D. (1948). Soldier intelligence in world wars I and II. *American Psychologist*, *3*, 54–56.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford: Clarendon Press.

Uttl, B., & van Alstine, C. L. (2003). Rising verbal intelligence: Implications for research and clinical pratice. *Psychology and Aging*, *18*, 616–621.

Wechsler, D. (1981). *Wechsler adult intelligence scale-revised*. San Antonio: The psychological corporation.

Williams, W. A. (1998). Are we raising smarter kids today? School- and home-related influences on IQ. In U. Neisser (Ed.), *The rising curve* (pp. 125–154). Washington: American Psychological Association.

Zajonc, R. B. (2001). The family dynamics of intellectual development. *American Psychologist*, *56*, 490–496.