



# Secular rise in economically valuable personality traits

Markus Jokela<sup>a,1</sup>, Tuomas Pekkarinen<sup>b,c,1</sup>, Matti Sarvimäki<sup>b,c,1,2</sup>, Marko Tervio<sup>b,1</sup>, and Roope Uusitalo<sup>d,1</sup>

<sup>a</sup>Institute of Behavioural Sciences, University of Helsinki, 00014 Helsinki, Finland; <sup>b</sup>Department of Economics, Aalto University School of Business, 00100 Helsinki, Finland; <sup>c</sup>VATT Institute for Economic Research, 00100 Helsinki, Finland; and <sup>d</sup>School of Business and Economics, University of Jyväskylä, 40014 Jyväskylä, Finland

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved May 8, 2017 (received for review June 23, 2016)

Although trends in many physical characteristics and cognitive capabilities of modern humans are well-documented, less is known about how personality traits have evolved over time. We analyze data from a standardized personality test administered to 79% of Finnish men born between 1962 and 1976 ( $n = 419,523$ ) and find steady increases in personality traits that predict higher income in later life. The magnitudes of these trends are similar to the simultaneous increase in cognitive abilities, at 0.2–0.6 SD during the 15-y window. When anchored to earnings, the change in personality traits amounts to a 12% increase. Both personality and cognitive ability have consistent associations with family background, but the trends are similar across groups defined by parental income, parental education, number of siblings, and rural/urban status. Nevertheless, much of the trends in test scores can be attributed to changes in the family background composition, namely 33% for personality and 64% for cognitive ability. These composition effects are mostly due to improvements in parents' education. We conclude that there is a "Flynn effect" for personality that mirrors the original Flynn effect for cognitive ability in magnitude and practical significance but is less driven by compositional changes in family background.

personality traits | cognitive ability | cohort effects | earnings | Flynn effect

There are many well-documented trends in average physical characteristics and cognitive capabilities of modern humans. Average height and body mass index have been on the rise around the world (1–4). Average IQ scores have increased at a rate of 0.2 SD per decade since the 1950s (5). In this study, we document similar trends in economically valuable personality traits of young adult males, as measured by a standardized test.

Recent findings in economics and psychology show that personality traits, especially conscientiousness and neuroticism, are important predictors of outcomes such as education and income in various populations. The predictive power of personality tests can be higher or lower than that of IQ depending on the measures used (6–8). Although most studies have reported contemporaneous correlations, there is evidence that traits measured at adolescence predict educational attainment and adult income (9–13). Recent studies also show that employment growth has been strong in occupations that require high levels of social skills (14, 15).

Previous evidence on trends in personality traits has been constrained by a lack of high-quality data on representative samples of successive cohorts of the same source population. Comparisons of cross-sectional studies of US college students have shown positive trends over time in traits such as extraversion and narcissism (16–18). However, students who participate in surveys are known to differ systematically from those who do not participate in characteristics such as academic achievement and vocational interests (19–21). Moreover, the selectivity of college admissions has changed over time, which has changed the composition of college student populations by socioeconomic backgrounds (22, 23). There are some studies where the same personality test was given to different cohorts of the same source population at the same age (24–27), but generalizing their findings to wider populations is problematic due to self-selection of survey respondents (19, 20). On the other hand, researchers have used large and representative data on high school seniors in the United States. However, most items in this dataset measure

social attitudes and personal values, and researchers have had to construct proxy measures for personality from a small number of items. Results have been mixed; some argue that personality traits have remained stable (28), whereas others claim to find increases in individualistic traits (29, 30).

Our data come from the Finnish Defense Forces (FDF), which has tested all military conscripts since 1982. Finnish men are drafted to military service in the year they turn 18, and most start their service at age 19 or 20. Both cognitive ability and personality tests are taken in the second week of military service in standardized group-administered conditions. Due to the comprehensive conscription system that grants relatively few exceptions, these data cover 79% of the population of Finnish men born between 1962 and 1976 ( $n = 419,523$ ). We also have test data for three additional cohorts born between 1977 and 1979 who took the personality test at the local draft board. However, these test results may not be directly comparable with earlier cohorts due to differences in the testing environment. The test score data have been linked with information on later life income and demographic background variables derived from administrative registers and population censuses. We present the data in more detail in *Materials and Methods*.

In comparison with earlier work, our test score data have both strengths and weaknesses. The main strength is that we observe a large and stable fraction of Finnish men over birth cohorts and that the test items remained unchanged during the period we examine. This facilitates the interpretation of changes in test results across cohorts. The most serious weakness of our data is that it does not include women. We also do not have test results for those men who chose to do the civilian service or were exempted from service due to medical reasons.

Another important limitation of the FDF personality test is that its scales do not directly correspond to standard personality scales and it has not been validated in a peer-reviewed journal.

## Significance

The secular rise in intelligence across birth cohorts is one of the most widely documented facts in psychology. This finding is important because intelligence is a key predictor of many outcomes such as education, occupation, and income. Although noncognitive skills may be equally important, there is little evidence on the long-term trends in noncognitive skills due to lack of data on consistently measured noncognitive skills of representative populations of successive cohorts. Using test score data based on an unchanged test taken by the population of Finnish military conscripts, we find steady positive trends in personality traits that are associated with high income. These trends are similar in magnitude and economic importance to the simultaneous rise in intelligence.

Author contributions: M.J., T.P., M.S., M.T., and R.U. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>1</sup>M.J., T.P., M.S., M.T., and R.U. contributed equally to this work.

<sup>2</sup>To whom correspondence should be addressed. Email: matti.sarvimaki@aalto.fi.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1609994114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1609994114/-DCSupplemental).

The FDF test measures eight traits (see legend of Fig. 1A). We conducted an online test using a short version of the test to see how these scales relate to the widely used Five-Factor Model (FFM) (see *SI Appendix* for details). The results from our convenience sample ( $n = 231$ ) suggest that the FDF scales capture three of the FFM scales (extraversion, conscientiousness, and neuroticism) but not agreeableness and openness.

## Results

**Cohort Trends in Test Scores.** Fig. 1A shows the evolution of average scores for each of the eight personality traits in our data, measured in SDs of the earliest birth cohort and centered at its mean. All but one of the traits exhibit a clear upward trend. The increase is largest for self-confidence, sociability, and leadership motivation, where averages for the 1976 cohort are about 0.6 SD above the average of the 1962 cohort. Average scores for activity-energy and achievement striving increase about 0.4 SD, whereas deliberation and dutifulness increase about 0.2 SD. The only trait without a clear trend is masculinity. In *SI Appendix*, we show that these trends are unlikely to be driven by changes in selection out of military service, in age at test, or in the validity of test responses. Personality test scores have a structural break after the change in test administration, and there is no consistent trend for the three postchange cohorts.

To put the magnitude of the trends in context, Fig. 1B shows the changes in cognitive test scores over the same period. Average scores for all three subtests exhibit secular increases of similar magnitude as seen for personality traits, varying from 0.2 SD for verbal reasoning to 0.6 SD for visuospatial reasoning. General cognitive ability, defined as the sum of cognitive subscores, increased at a rate of 0.018 SD per year, which is in line with previous evidence for positive trends in IQ scores across many countries (5), also known as the “Flynn effect.” Cognitive scores also show the end of the Flynn effect, which has been dated at around the 1970s birth cohorts in Finnish (31), Norwegian (32), and Danish conscript data (33).

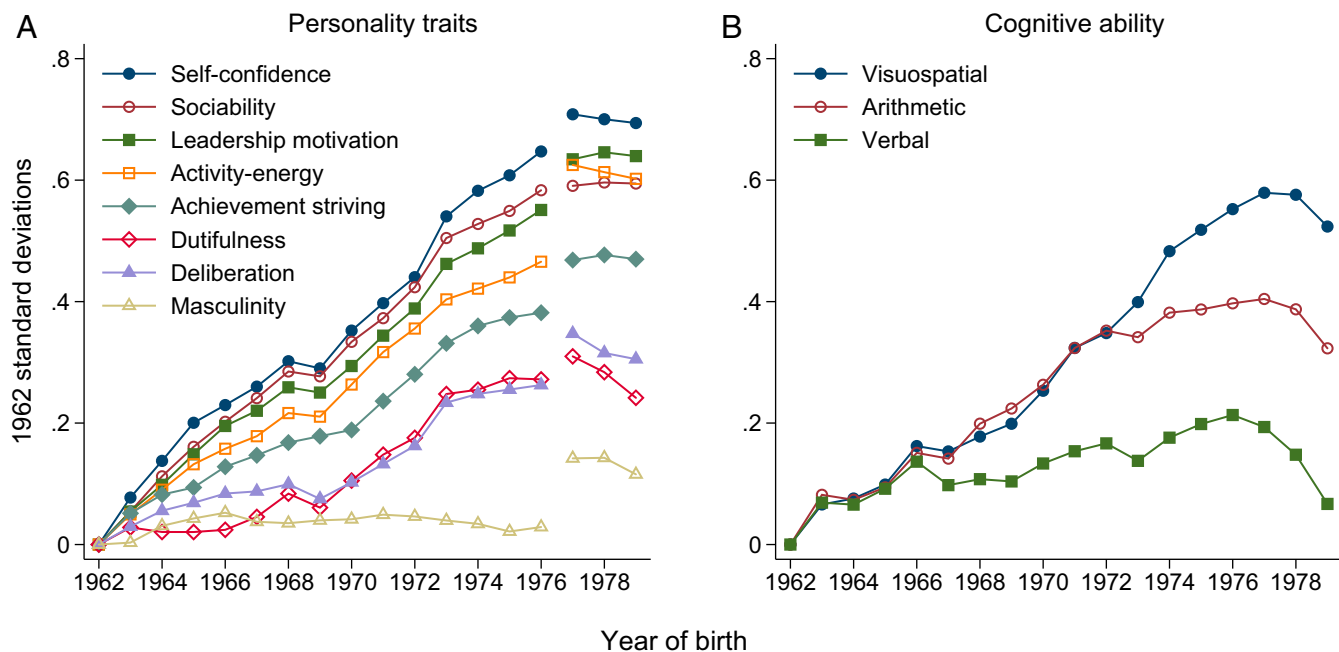
**Predictive Validity.** A natural concern in interpreting the rise in any test scores is that later cohorts may have become more motivated or more adept at test-taking without any actual trends

in underlying traits. We are unable to measure changes in motivation or in the ability “to game” the test, nor are we able to link test scores to trait-typical behavior (e.g., whether individuals with high sociability scores were highly sociable in real life). However, we can assess the predictive validity of test scores for income in later life. Fig. 2A plots the rank correlation of each personality trait with earnings at age 30 (the latest age at which we observe earnings for all cohorts). With the exception of masculinity—the only trait without a clear trend in Fig. 1A—all traits show a persistent and strong positive association with earnings, with rank correlations of about 0.1–0.2.

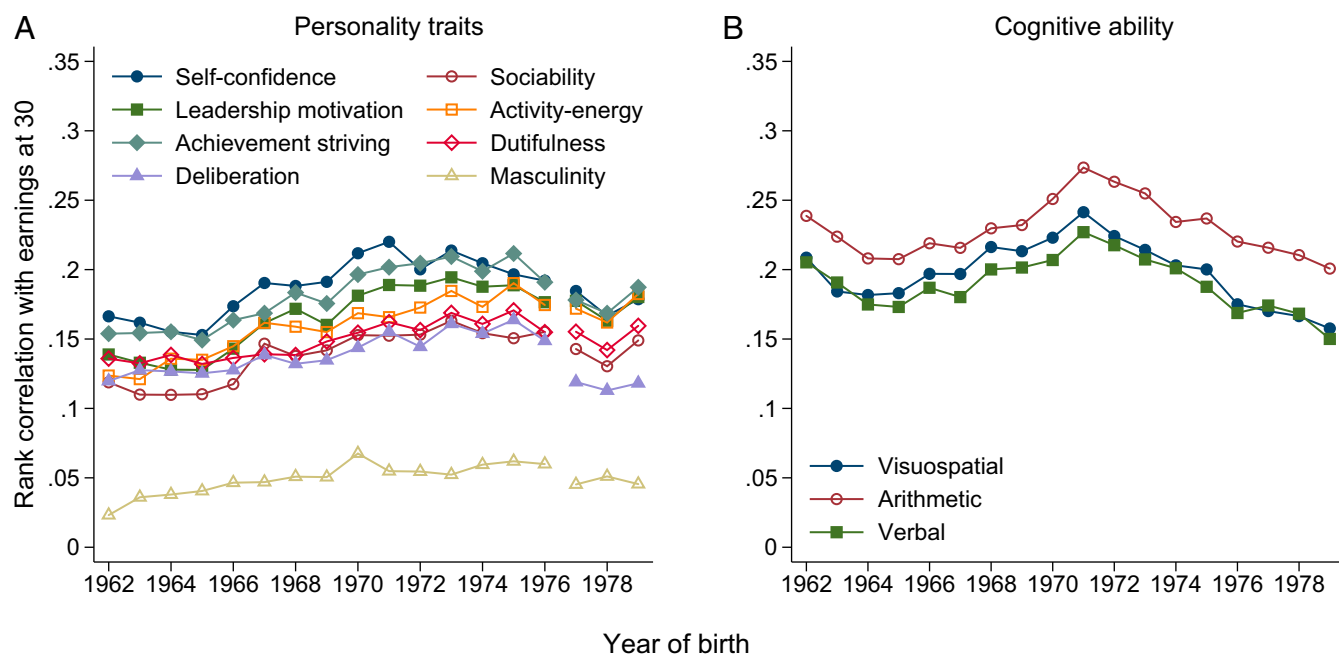
**Stochastic Dominance.** Test scores are ordinal measures of underlying traits, and treating them as if measured on an interval scale can result in misleading interpretations. In particular, conclusions may depend on arbitrary scaling decisions (34–36). *SI Appendix*, Figs. S2–S5 show that, with the exception of masculinity, the shifts took place across the entire distributions of test scores. That is, distributions of scores of the later cohorts dominate the distributions of earlier cohorts in the sense of first-order stochastic dominance. Thus, our conclusion about positive trends in personality test scores is robust to any monotonic transformation of the raw test scores.

**Anchoring Test Scores to Earnings.** To obtain a quantitative interpretation for the trends, we convert the test scores to interval scale by anchoring them to later-life earnings. We use average annual earnings at age 30–34, which we observe up to the 1976 birth cohort and which has been shown to be a good proxy for lifetime income (37). We regress these earnings on all personality test scores and use the resulting estimates to predict earnings for each combination of test scores (*SI Appendix*, Table S2). This predicted earnings measure is our anchored personality test score; cognitive test scores are anchored similarly. In addition to summarizing the test scores on a one-dimensional interval scale, this approach also provides an economic context for our results.

Fig. 3 depicts the means of the anchored test scores across birth cohorts. The trends are very similar for personality and cognitive ability, showing an increase of about €2500 between the 1962 and



**Fig. 1.** Average scores for measures of (A) personality traits and (B) cognitive ability by birth year for native-born military conscripts in Finland. All scores are depicted in base year SDs, with base year means normalized at zero. The break in personality test scores reflects a change in test administration.



**Fig. 2.** The relation of earnings and (A) personality traits and (B) cognitive ability by birth cohort, measured as the within-cohort rank correlation between the test score and annual earnings at age 30. *SI Appendix, Fig. S1* shows the same relations for average annual earnings at age 30–34, which is a better measure of lifetime earnings but not observed for the last three cohorts. The break in personality test scores reflects a change in test administration.

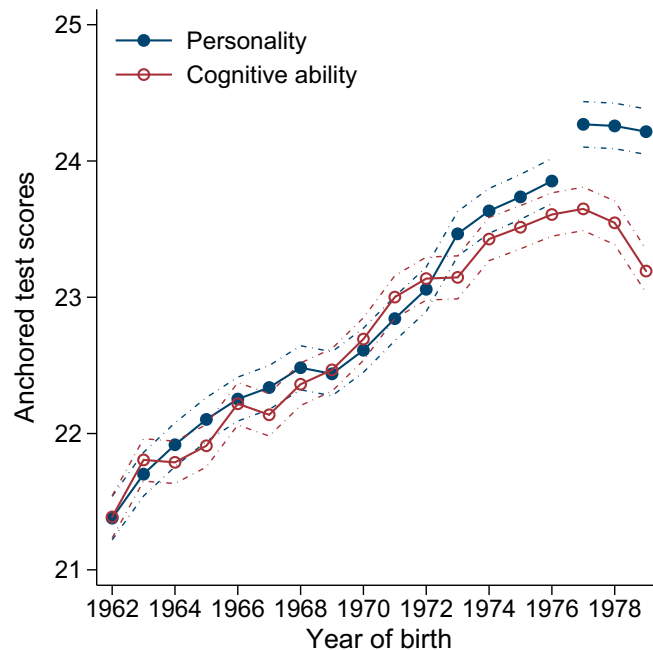
1976 cohorts for personality and €2200 for cognitive ability. Put differently, based on a time-invariant model for the relation of test scores and earnings, the increase in personality test scores predicts about 12% higher earnings for the 1976 cohort than for the 1962 cohort; based on cognitive test scores, the predicted increase is 10%. In *SI Appendix, Fig. S7*, we show that these trends are very similar when using alternative income measures. Furthermore, anchoring the test scores to completed education yields trends that are qualitatively similar to income-related anchorings, despite the very different scale of measurement.

Although personality traits are correlated with each other and with cognitive abilities, both have independent power in predicting earnings (*SI Appendix, Table S2*). Trends in personality are similar across levels of cognitive ability (*SI Appendix, Fig. S17*). An anchoring regression where both sets of test scores are controlled at the same time results in smaller magnitudes for both trends, namely 7% for personality and 8% for cognitive scores (*SI Appendix, Fig. S6*). For a different approach, *SI Appendix, Table S5* reports an exploratory factor analysis, which shows that cognitive abilities and personality traits load into distinct factors. Confirmatory factor analysis (*SI Appendix, Fig. S13*) indicates that personality and cognitive factors are both related to earnings, but the correlation between them is only 0.41.

**Measurement Error.** The trends reported in Fig. 3 may understate cohort trends due to measurement error in individual test scores. We investigate the impact of measurement error in *SI Appendix*. Using brothers' test scores as instrumental variables (IVs) results in about a 9 percentage point higher increase in anchored personality test scores (*SI Appendix, Fig. S11*). For cognitive scores, the same approach yields a 4 percentage point higher increase. However, these instruments are not without problems, as able brothers may be directly helpful for one's earnings (38). Thus, we view the ordinary least squares (OLS)-based trends as conservative and the IV estimates as more likely to be upwards biased; a structural equation model suggests trends in between the two but much closer to the former (*SI Appendix, Table S9*).

**Personality and Background Variables.** Fig. 4 plots the trend in the anchored personality test scores separately by levels of background

variables. It reveals a stable regularity between family background and personality across birth cohorts. Anchored scores are positively correlated with parental income, parental education, and urban childhood environment among all cohorts, whereas their association with the number of siblings is negative. Fig. 4 also shows that the positive trend is visible in every demographic subgroup. For



**Fig. 3.** Average of anchored test scores by birth cohort, with anchoring to average annual earnings at age 30–34 (in 1,000s of 2010 Euros) using the 1962–76 birth cohorts for estimating the prediction model. Dashed lines depict 95% confidence intervals. The break in personality test scores reflects a change in test administration.

example, the test scores of men with parents in the bottom quintile of the income distribution in the last cohort were nearing the level that was seen for those with parents in the top quintile in the earliest cohort in our data.

Trends in personality coincide with trends in background characteristics that are correlated with personality. There have been decreases in family size and increases in parental income, parental education, and urbanization (*SI Appendix, Table S11*). All of these background variables have been evolving to the direction that predicts higher levels of personality traits that are in turn known to predict higher incomes (39, 40). It is therefore natural to ask to which extent the observed cohort trends in personality can be explained as merely reflecting changes in the composition of the population by background characteristics.

The similarity of trends across backgrounds already suggests that changes in parental education, family size, and urbanization cannot fully explain the change in personality traits. Nevertheless, changes in backgrounds explain a sizeable fraction of the trends. We decompose the changes in test scores using the reweighting procedure proposed in ref. 41 (details in *SI Appendix*). In effect, we ask what the distribution of test scores would be if the relationship between background variables and test scores stayed the same as it was in 1962 but the distribution of background variables was the same as it was in 1976.

Table 1 reports the results of this exercise. We find that 33% of the increase in the anchored personality test score can be attributed to changes in background characteristics. The traits most affected by changing backgrounds are achievement striving and dutifulness, for which over 40% of the increase can be

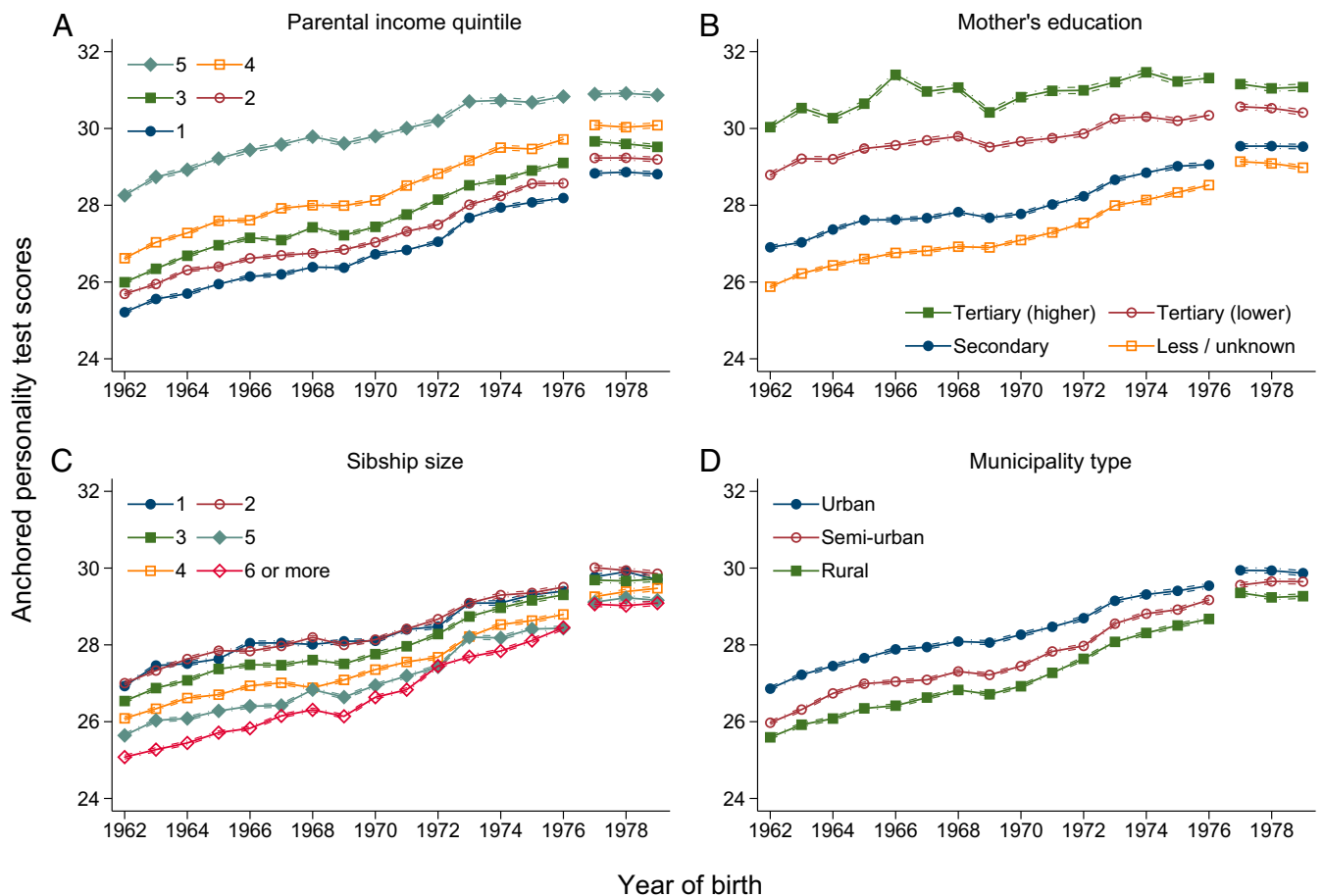
attributed to a composition effect. For other personality traits (besides masculinity, which has no clear trend), the composition effect accounts for 14–34% of the increase.

Beneficial trends in background characteristics are more important for explaining changes in cognitive ability than in personality, with 64% of the increase in anchored cognitive test scores attributable to the change in composition. For verbal scores, the change in backgrounds predicts an even larger increase than was actually observed.

Only one previous study has presented evidence on the role of demographic changes behind the Flynn effect. An analysis with sibship size as the only background variable was conducted in Norway (42). There a comparison between birth cohorts of 1938–40 and 1974–85 found that 35% of the increase in verbal scores and 13% of the increase in visuospatial scores can be attributed to the decrease in sibship size (there was no increase in arithmetic scores). If we use sibship size as the only background variable, we explain only 18% of the increase in cognitive ability. *SI Appendix, Table S13* reports a similar analysis one background variable at a time; it shows that vastly improved parental education levels are the main driver of composition effects.

## Conclusions

We find a Flynn effect for personality—that is, a secular rise in personality traits that are associated with higher earnings. The fact that the trend is positive is clear from the way distributions of test scores shift up across birth cohorts. Various methods of quantifying the economic importance of these changes all point toward the trend in personality being similar in magnitude and



**Fig. 4.** Evolution of economically valuable personality traits across birth cohorts by (A) parental income quintile, (B) mother's education level, (C) sibship size, and (D) urbanization of birth place. Test scores are anchored to earnings at age 30–34. *SI Appendix, Fig. S18* shows that the broad picture is similar when this analysis is repeated one trait at a time. Dashed lines depict 95% confidence intervals. The break in personality test scores reflects a change in test administration.

**Table 1. Cohort trends and demographic backgrounds**

Variable	Change between 1962 and 1976 cohorts		Share predicted, %
	Observed	Predicted	
<b>Personality</b>			
Self-confidence	0.65	0.16	25
Sociability	0.58	0.15	26
Leadership motivation	0.55	0.19	34
Activity-energy	0.47	0.09	20
Achievement striving	0.38	0.17	44
Dutifulness	0.27	0.11	41
Deliberation	0.26	0.04	14
Masculinity	0.03	0.00	-15
All (anchored)	0.57	0.19	33
<b>Cognitive ability</b>			
Visuospatial	0.55	0.25	45
Arithmetic	0.40	0.26	65
Verbal	0.21	0.25	119
All (anchored)	0.44	0.28	64

“Observed” is the actual difference in means between the birth cohorts, and “predicted” is the mean of predicted values for this difference, based on age at test, parental income, mother’s and father’s levels of education, sib-ship size, and rural/urban status, using the model estimated for the 1962 cohort. All variables were measured in 1962 SDs. Bootstrapped SEs are below 0.007 for all observed and below 0.015 for all predicted means.

economic importance to the rise in cognitive abilities. The trends in personality are also similar across levels of cognitive ability and across demographic subgroups.

Our results on traits related to extraversion (i.e., sociability and activity–energy) are consistent with studies reporting increasing levels of extraversion (16, 24–26). Our findings for conscientiousness-related traits are in agreement with findings from freshman psychology students at the University of Amsterdam between 1982 and 2007 (25) and from the Baltimore Longitudinal Study of Aging between 1989 and 2004 (43). We also found increasing levels of self-confidence. This trend is in contrast to findings from the Monitoring the Future study (28) but is in agreement with cross-temporal meta-analysis of US college students (17). A positive trend has been reported for narcissism at least in the United States (18). We cannot distinguish self-confidence associated with narcissism from self-esteem; we can only see that this measure of self-confidence predicts high earnings for the person himself.

Growing evidence suggests that the Flynn effect has ended and may have reversed in Western Europe (32, 33, 44–46). The last three birth cohorts in our data coincide with the peak in cognitive test scores in Finland (31). There is no clear trend for personality scores between these cohorts, which suggests that the end of the Flynn effect could also be reflected in personality traits. However, the data on these three birth cohorts are not fully comparable with our main data, and thus, it is not possible to make strong conclusions from them.

The causes of the Flynn effect are still unclear (5), and our data do not reveal the ultimate cause of the cohort trends in personality either. Of course, we cannot distinguish between birth year and year of test as causal factors behind the trends. However, we can rule out trends in personality traits being mere reflections of changes in broadly defined socioeconomic backgrounds. Nevertheless, trends in background variables are indeed favorable and explain about two-thirds of the rise in cognitive ability and one-third of the trends in personality.

## Materials and Methods

**Psychological Testing in the FDF.** FDF has tested all conscripts with a battery of psychological tests since 1955. Initially the test consisted of only a cognitive test that measured reading skills, mathematical skills, and logical reasoning

skills. In 1982, the FDF introduced a personality test that measures eight personality traits. Test results are one of the criteria used in selecting conscripts to officer training.

The validity of the test and its predictive power for successful military service have been evaluated in several internal reports of the FDF. The results of these (mainly unpublished) studies have been summarized and the test procedure described in detail in ref. 47. Only those who enter service take the tests; those who are exempted (e.g., on prior health grounds) and those who choose to do nonmilitary service do not take the test. Test results of professional military officers were retracted by the FDF.

**Administration of the Test.** Both the cognitive test and the personality test are administered in the second week of military service. The tests are organized in standardized group-administered conditions at all FDF units. Between 1995 and 2000, the personality test was administered already at the call-up, on average 18 mo before entering the service. The purpose was to use the test scores in placement of conscripts already before they started their service. However, the results were not widely used for this purpose, and the FDF was concerned that test conditions at local draft boards were not sufficiently standardized. In 2001, the FDF reverted to testing conscripts at the start of service (47). The cognitive test has always been administered in the military service.

The test is a 2-h paper-and-pencil test where conscripts are asked to choose a correct alternative from a list (cognitive ability test) or whether they agree or disagree with statements (personality test). Completed answer sheets are sent to the Finnish Defense Research Agency for optical scanning. The test leaflets were unchanged from 1982 to 2000 but have not been released by the FDF. In 2001, the personality test was revised, and both the content and the results of the new test remain classified.

*SI Appendix, Table S1* reports means and SDs for each test score by cohort, and *SI Appendix, Figs. S2–S5* show the full distributions of the raw scores for both personality and cognitive test. Observed scores vary over the entire range of possible values. The distributions of cognitive test scores are roughly normal but those of personality test scores less so. Ceiling effects may cause attenuation of trends for measures of self-confidence and sociability.

**Content of the Personality Test.** The test contains between 18 and 33 items for each of the eight personality traits. Altogether there are 218 statements with a response scale of yes/no. The scores are formed by summing up the number of statements to which a person agrees (or, in case of reverse-coded statements, disagrees with). We observe the raw scores but not individual items. Internal reliability varies between 0.6 and 0.9 by trait; average Cronbach alpha is 0.75 (47).

Self-confidence measures the person’s self-esteem and beliefs about his abilities (32 items; e.g., whether the person feels to be as good and able as others and can meet other people’s expectations). Sociability measures the person’s level of gregariousness and preference for socializing with others (33 items; e.g., whether the person likes to host parties and not withdraw from social events). Leadership motivation measures how much the person prefers to take charge in groups and influence other people; it includes 30 items. Activity–energy measures how much the person exerts physical effort in everyday activities and how quickly the person prefers to execute activities (28 items; e.g., whether the person tends to work fast and vigorously and prefers fast-paced work). Achievement striving, dutifulness, and deliberation all represent personality traits that are related to the higher order personality factor conscientiousness. Achievement striving measures how strongly the person wants to perform well and achieve important life goals (24 items; e.g., whether the person is prepared to make personal sacrifices to achieve success). Dutifulness measures how closely the person follows social norms and considers them to be important (18 items; e.g., whether the person would return money if given back too much change at a store). Deliberation measures how much the person prefers to think ahead and plan things before acting (26 items; e.g., whether the person prefers to spend money carefully). Masculinity measures the person’s occupational and recreational interests that are traditionally considered as masculine (27 items; e.g., whether the person would like to work as a construction manager).

The FDF questionnaire also includes questions about mental health and questions assessing the validity of the answers. These include four mental health subscales from the Minnesota Multiphasic Personality Inventory (MMPI) but not other measures of normal personality. Of these variables we use only the lie score, which measures socially desirable responding—that is, attempts to give an overly favorable impression of one’s conduct. *SI Appendix, Table S12* shows that trends in test scores cannot be attributed to changes in response validity as measured by the lie score.

**Content of the Cognitive Ability Test.** Cognitive ability is measured with subtests of verbal, arithmetic, and visuospatial reasoning. Each subtest is composed of 40 multiple-choice questions in order of increasing difficulty. The test–retest reliabilities of the subtests vary between 0.76 and 0.88 (47). Verbal reasoning involves choosing synonyms or antonyms of a given word, selecting a word that belongs to the same category as a given word pair, choosing which word on a list does not belong in the group, and choosing similar relationships between two word pairs. Arithmetic reasoning involves completing a series of numbers that follow a certain pattern, solving short verbal problems, computing simple arithmetic operations, and choosing similar relationships between two pairs of numbers. The visuospatial reasoning task is a set of matrices containing a pattern problem with one removed part, and the participant needs to decide which of the given alternative figures completes the matrix; it is similar to Raven's Progressive Matrices (48).

**Register Data.** We use register data on the Finnish population compiled by Statistics Finland to obtain adult outcomes and background variables. These data provide information on basic demographics, family situation, living conditions, educational attainment, labor market status, and earnings of all Finnish residents. This information was linked to test scores by Statistics Finland using personal identification numbers and deidentified before being made available to researchers.

Income data are from the Finnish Tax Authority. We measure earnings as the average annual earnings during ages 30–34, where “earnings” is the sum of labor market income and entrepreneurial income; we do not drop zeros. We

deflate all values to 2010 Euros using the Statistics Finland CPI. In *SI Appendix*, we also use alternative income measures derived from the same data.

Information about the identity of parents and brothers comes from the Finnish Population Register. Childhood municipality of residence comes from the Population Censuses of 1970, 1975, and 1980. We define childhood municipality as the municipality of residence in the first census after the year of birth. We drop those who are not observed at that point as they are likely to be foreign-born. We use Statistics Finland's Statistical Grouping of Municipalities to divide municipalities into urban, semiurban, and rural. We define sibship size as the number of children with the same biological mother.

Data on educational attainment are from the Register of Completed Education and Degrees maintained by Statistics Finland. These data contain information on the highest educational qualification that the individual has obtained and the date at which the individual received the qualification. We use it to obtain parents' level of education and the eventual level of education for the conscripts.

Permission to use the register data was approved by Statistics Finland (license TK-53-228-14) and by FDF (AJ23378). Personal data were processed following the regulations in Personal Data Act 523/1999 and the guidelines of Finnish Advisory Board on Research Integrity. The use of administrative data in scientific research does not require explicit consent from the subjects in Finland.

**ACKNOWLEDGMENTS.** We thank Kai Nyman and Kari Laitinen at the FDF for help in assisting with access to data and interpreting test scores and Annalaina Kotilainen for excellent research assistance. T.P., M.S., and R.U. were supported by Strategic Research Council at the Academy of Finland Grants 293445 and 303686. M.T. was supported by European Research Council Grant ERC-240970.

1. Hatton TJ, Bray BE (2010) Long run trends in the heights of European men, 19th–20th centuries. *Econ Hum Biol* 8:405–413.
2. Fogel RW (2004) *The Escape from Hunger and Premature Death, 1700–2100: Europe, America, and the Third World* (Cambridge Univ Press, Cambridge, UK).
3. Floud R, Fogel RW, Harris B, Hong SC (2011) *The Changing Body: Health, Nutrition, and Human Development in the Western World Since 1700* (Cambridge Univ Press, Cambridge, UK).
4. Finucane MM, et al.; Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index) (2011) National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* 377:557–567.
5. Pietschnig J, Voracek M (2015) One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspect Psychol Sci* 10:282–306.
6. Borghans L, Duckworth AL, Heckman JJ, ter Weel B (2008) The economics and psychology of personality traits. *J Hum Resour* 43:972–1059.
7. Almlund M, Duckworth AL, Heckman JJ, Kautz T (2011) Personality psychology and economics. *Handbook of the Economics of Education*, eds Hanushek EA, Machin S, Woessmann L (Elsevier, Amsterdam), Vol 4, pp 1–181.
8. Borghans L, Golsteyn BHH, Heckman JJ, Humphries JE (2016) What grades and achievement tests measure. *Proc Natl Acad Sci USA* 113:13354–13359.
9. Goldberg LR, Sweeney D, Merenda PF, Hughes JE (1998) Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. *Pers Individ Dif* 24:393–403.
10. Van Eijck K, de Graaf PM (2004) The Big Five at school: The impact of personality on educational attainment. *Netherlands Journal of Social Sciences* 40:24–40.
11. Heckman JJ, Stixrud J, Urzua S (2006) The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *J Labor Econ* 24:411–482.
12. Mueller G, Plug E (2006) Estimating the effect of personality on male and female earnings. *Ind Labor Relat Rev* 60:3–22.
13. Lindqvist E, Vestman R (2011) The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *Am Econ J Appl Econ* 3:101–128.
14. Deming DJ (2015) The Growing Importance of Social Skills in the Labour Market, National Bureau of Economic Research Working paper no. 21473 (National Bureau of Economic Research, Cambridge, MA). Available at [www.nber.org/papers/w21473](http://www.nber.org/papers/w21473).
15. Weinberger CJ (2011) The increasing complementarity between cognitive and social skills. *Rev Econ Stat* 96:849–861.
16. Twenge JM (2001) Birth cohort changes in extraversion: A cross-temporal meta-analysis, 1966–1993. *Pers Individ Dif* 30:735–748.
17. Gentile B, Twenge JM, Campbell WK (2010) Birth cohort differences in self-esteem, 1988–2008: A cross-temporal meta-analysis. *Rev Gen Psychol* 14:261–268.
18. Twenge JM, Konrath S, Foster JD, Campbell WK, Bushman BJ (2008) Egos inflating over time: A cross-temporal meta-analysis of the Narcissistic Personality Inventory. *J Pers* 76:875–902, discussion 903–928.
19. Peterson RA (2001) On the use of college students in social science research: Insights from a second-order meta-analysis. *J Consum Res* 28:450–461.
20. Porter SR, Whitcomb ME (2005) Non-response in student surveys: The role of demographics, engagement and personality. *Res Higher Educ* 46:127–152.
21. Porter SR, Umbach PD (2006) Student survey response rates across institutions: Why do they vary? *Res Higher Educ* 47:229–247.
22. Bailey MJ, Dynarski SM (2011) Gains and Gaps: Changing Inequality in US College Entry and Completion, National Bureau of Economic Research Working paper no. 17633 (National Bureau of Economic Research, Cambridge, MA). Available at [www.nber.org/papers/w17633](http://www.nber.org/papers/w17633).
23. Heckman JJ (2000) Policies to foster human capital. *Res Econ* 54:3–56.
24. Mroczek DK, Spiro A, 3rd (2003) Modeling intraindividual change in personality traits: Findings from the normative aging study. *J Gerontol B Psychol Sci Soc Sci* 58:153–165.
25. Smits IAM, Dolan CV, Vorst HC, Wicherts JM, Timmerman ME (2011) Cohort differences in Big Five personality factors over a period of 25 years. *J Pers Soc Psychol* 100:1124–1138.
26. Billstedt E, et al. (2013) Secular changes in personality: Study on 75-year-olds examined in 1976–1977 and 2005–2006. *Int J Geriatr Psychiatry* 28:298–304.
27. Kawamoto T, Endo T (2015) Personality change in adolescence: Results from a Japanese sample. *J Res Pers* 57:32–42.
28. Trzesniewski KH, Donnellan MB (2010) Rethinking “generation me” a study of cohort effects from 1976–2006. *Perspect Psychol Sci* 5:58–75.
29. Twenge JM, Campbell WK (2008) Increases in positive self-views among high school students: Birth-cohort changes in anticipated performance, self-satisfaction, self-liking, and self-competence. *Psychol Sci* 19:1082–1086.
30. Twenge JM, Campbell WK (2010) Birth cohort differences in the monitoring the future dataset and elsewhere: Further evidence for generation me—Commentary on Trzesniewski & Donnellan (2010). *Perspect Psychol Sci* 5:81–88.
31. Koivunen S (2007) Suomalaismiesten kognitiivisen kykyprofiilin muutokset 1988–2001: Flynnin efektiä suomalaisessa aineistossa? MSc Thesis (University of Jyväskylä, Jyväskylä, Finland). Available at [urn.fi/URN:NBN:fi:juyu-2007274](http://urn.fi/URN:NBN:fi:juyu-2007274).
32. Sundet JM, Barlaug DG, Torjussen TM (2004) The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence* 32:349–362.
33. Teasdale TW, Owen DR (2005) A long-term rise and recent decline in intelligence test performance: The Flynn Effect in reverse. *Pers Individ Dif* 39:837–843.
34. Bond TN, Lang K (2013) The evolution of the Black-White test score gap in Grades K–3: The fragility of results. *Rev Econ Stat* 95:1468–1479.
35. Cunha F, Heckman J, Schennach S (2010) Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78:883–931.
36. Jacob B, Rothstein J (2016) The measurement of student ability in modern assessment systems. *J Econ Perspect* 30:85–108.
37. Böhlmark A, Lindquist M (2006) Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden. *J Labor Econ* 24:879–896.
38. Grönqvist E, Öckert B, Vlachos J (August 3, 2016) The intergenerational transmission of cognitive and non-cognitive abilities. *J Hum Resour*, 10.3368/jhr.52.4.0115-6882R1.
39. Blanden J, Gregg P, Macmillan L (2007) Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *Econ J (Oxf)* 117:C43–C60.
40. Groves MO (2005) Personality and the intergenerational transmission of economic status. *Unequal Chances: Family Background and Economic Success*, eds Bowles S, Gintis H, Osborne Groves M (Princeton Univ Press, Princeton, NJ), pp 208–231.
41. DiNardo J, Fortin NM, Lemieux T (1996) Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica* 64:1001–1044.
42. Sundet JM, Ingrid B, Tams K (2008) The Flynn effect is partly caused by changing fertility patterns. *Intelligence* 36:183–191.
43. Terracciano A, McCrae RR, Brant LJ, Costa PT, Jr (2005) Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychol Aging* 20:493–506.
44. Shayer M, Ginsburg D, Coe R (2007) Thirty years on—A large anti-Flynn effect? The Piagetian test Volume & Heaviness norms 1975–2003. *Br J Educ Psychol* 77:25–41.
45. Flynn JR (2009) Requiem for nutrition as the cause of IQ gains: Raven's gains in Britain 1938–2008. *Econ Hum Biol* 7:18–27.
46. Woodley M, Meisenberg G (2013) In the Netherlands the anti-Flynn effect is a Jensen effect. *Pers Individ Dif* 54:871–876.
47. Nyman K (2007) *Varusmiesten johtajavalintojen luotettavuus*, Publication Series 1 (National Defense University, Department of Behavioral Sciences, Helsinki, Finland).
48. Raven J, Raven JC, Court JH (2000) *Manual for Raven's Progressive Matrices and Vocabulary Scales* (Oxford Psychologists Press, Oxford).