

Molecular genetics and subjective well-being

Cornelius A. Rietveld^a, David Cesarini^{b,c}, Daniel J. Benjamin^d, Philipp D. Koellinger^a, Jan-Emmanuel De Neve^{e,f}, Henning Tiemeier^g, Magnus Johannesson^h, Patrik K. E. Magnussonⁱ, Nancy L. Pedersenⁱ, Robert F. Krueger^{j,1}, and Meike Bartels^{k,l,1,2}

^aDepartment of Applied Economics, Erasmus School of Economics, Erasmus University, 3062 PA, Rotterdam, The Netherlands; ^bDepartment of Economics, Center for Experimental Social Science, New York University, New York, NY 10012; ^cResearch Institute for Industrial Economics, SE-102 15 Stockholm, Sweden; ^dDepartment of Economics, Cornell University, Ithaca, NY 14853; ^eSchool of Public Policy, Department of Political Science, University College London, London WC1H 9QU, United Kingdom; ^fCentre for Economic Performance, London School of Economics and Political Science, London WC2A 2AE, United Kingdom; ^gDepartment of Epidemiology and Biostatistics, Erasmus Medical Center, 3000 CA, Rotterdam, The Netherlands; ^hDepartment of Economics, Stockholm School of Economics, SE-113 83 Stockholm, Sweden; ⁱDepartment of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77 Stockholm, Sweden; ^jDepartment of Psychology, University of Minnesota, Minneapolis, MN 55455; ^kDepartment of Biological Psychology, VU University Amsterdam, 1081 BT, Amsterdam, The Netherlands; and ^lEMGO+ Institute for Health and Care Research, VU University Medical Center, 1081 BT, Amsterdam, The Netherlands

Edited* by Richard A. Easterlin, University of Southern California, Los Angeles, CA, and approved April 23, 2013 (received for review December 29, 2012)

Subjective well-being (SWB) is a major topic of research across the social sciences. Twin and family studies have found that genetic factors may account for as much as 30–40% of the variance in SWB. Here, we study genetic contributions to SWB in a pooled sample of ≈11,500 unrelated, comprehensively-genotyped Swedish and Dutch individuals. We apply a recently developed method to estimate “common narrow heritability”: the fraction of variance in SWB that can be explained by the cumulative additive effects of genetic polymorphisms that are common in the population. Our estimates are 5–10% for single-question survey measures of SWB, and 12–18% after correction for measurement error in the SWB measures. Our results suggest guarded optimism about the prospects of using genetic data in SWB research because, although the common narrow heritability is not large, the polymorphisms that contribute to it could feasibly be discovered with a sufficiently large sample of individuals.

happiness | GREML | GCTA

Subjective well-being (SWB)—most commonly measured by survey questions about a respondent’s happiness or life satisfaction—is a major topic of research across the social sciences (1, 2). SWB is conceptualized to include a continuous spectrum of positive feelings and subjective life assessments (3–5). In contrast to standard economic indicators, which focus on consumption of material goods, responses to SWB survey questions additionally convey information regarding a broad range of other determinants of well-being, including physical and mental health, social relationships, leisure, and subjective states, such as emotions and mental engagement (6, 7). Because SWB measures may represent a relatively comprehensive assessment of an individual’s feelings of well-being, much research aims to understand individual differences in SWB (8). Most of the literature examines social, economic, and psychological influences on SWB (4, 9), but there has also been recent interest in understanding how genetic factors influence SWB.

To date, most of these papers on the genetics of SWB are twin or family studies (10–17). These studies draw indirect inferences about the contribution of genes to SWB by contrasting the resemblance of relatives with different degrees of environmental and genetic similarity. The literature concludes that a moderate share, typically 30–40%, of the cross-sectional variation in SWB is accounted for by variation in genes.

Recently it has become possible to directly and inexpensively assay human genetic polymorphisms, segments of DNA that differ across individuals. For medical geneticists studying health outcomes, the availability of such data has ushered in a new era, as researchers are discovering an ever-increasing number of polymorphisms related to diseases and physical traits (18).

So far, however, the few attempts to find genetic polymorphisms associated with SWB have been unsuccessful (see ref. 19 for the earliest effort we know of). One study reported an

association (20), but follow-up work on an augmented sample from the same data did not replicate the finding (21). This lack of success is not surprising, given the lessons that have emerged from genetics research across a range of medical and social-science traits. Among the central challenges for complex traits, such as height and probably even more so SWB, is that the heritability of these traits appears to be comprised of a huge number of tiny genetic effects. Consequently, large samples of individuals—several orders of magnitude larger than those used to date in gene-discovery work in the social sciences—are needed for adequate statistical power to identify specific genetic polymorphisms (22, 23).

Nevertheless, anticipating that polymorphisms related to SWB will soon be discovered, SWB researchers have expressed excitement about the transformative potential of genetic data for social-science research (21), which complements what can be learned from twin and family studies (22–24). Most directly, knowing the functions of the relevant genes could shed light on the biological pathways that matter for SWB. If a set of polymorphisms were found to be sufficiently predictive, then they could be used in social-science research as control variables. More speculatively, such polymorphisms could be used as instrumental variables (25, 26), in effect treating the Mendelian randomization that occurs at conception as a natural experiment to learn about the causal effects of SWB [which may be especially credible when used in family samples (27); for a critical perspective, see ref. 28]. Finally, the discovery of polymorphisms associated with SWB could catalyze the study of how genetic sources of individual differences are amplified or dampened by environmental factors and, conversely, how environmental effects are modulated by genetic pathways.

For evaluating the extent to which these promises of genetic data can be realized, a critical question is: How much of the variation in SWB will eventually be predictable using molecular genetic data? In this article, we provide empirical evidence on a quantity—the “common narrow heritability,” explained below—that may help calibrate reasonable expectations about the answer to this question. We also discuss the inferences that can and cannot legitimately be drawn from this estimate as well as from heritability estimates in general. For example, we scrutinize the logical

Author contributions: C.A.R., D.C., D.J.B., P.D.K., J.-E.D.N., H.T., M.J., P.K.E.M., N.L.P., R.F.K., and M.B. designed research; C.A.R., D.J.B., P.D.K., N.L.P., R.F.K., and M.B. performed research; C.A.R., D.C., D.J.B., and M.B. analyzed data; and C.A.R., D.C., D.J.B., P.D.K., J.-E.D.N., M.J., P.K.E.M., R.F.K., and M.B. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

¹R.F.K. and M.B. contributed equally to this work.

²To whom correspondence should be addressed. E-mail: m.bartels@vu.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1222171110/-DCSupplemental.

coherence of invoking estimates of a trait's heritability to draw conclusions about its responsiveness to environmental interventions.

The estimates of 30–40% mentioned above likely overstate the amount of predictive power that can be obtained from molecular genetic data for two distinct reasons. First, the numbers refer to what is known as “broad heritability,” but “narrow heritability” is more germane and is necessarily smaller. Narrow heritability is the fraction of variance that can be accounted for in aggregate by the cumulative additive effects of all genetic polymorphisms. Narrow heritability can be understood as the R^2 from a population regression of SWB on its best linear genetic predictor; that is, a predictor in which each polymorphism enters additively, and the effect of each polymorphism is constrained to be linear in the number of reference alleles. Broad heritability, which is necessarily larger, is the fraction of variance in SWB that can be explained in aggregate by all genetic factors. Broad heritability can be understood as the R^2 from a population regression of SWB on its best genetic predictor, allowing not only for linear and additive effects but also for interactions among different polymorphisms (“epistasis”) and nonlinear effects of specific polymorphisms (“dominance”). In a seminal report drawing together evidence from various twin and family comparisons, Lykken (29) proposed that for SWB (along with several other traits including personality), most—if not all—of the genetic influences stem from higher-order epistatic interactions among genetic polymorphisms. Lykken called this phenomenon the “emergence hypothesis” (for a recent and related discussion, see ref. 30). If true, then the narrow heritability of SWB is much smaller than its broad heritability. Several recent, large-scale, twin-family studies, including both twin and sibling pairs, have indeed documented evidence for the importance of both additive and nonadditive genetic effects in explaining individual differences in SWB (14, 16).

Narrow heritability is more relevant than broad heritability for evaluating the predictive power that will be attainable using molecular genetic data because most interaction effects between polymorphisms are going to be extremely challenging to pinpoint. For genetically complex traits, we are not aware of a credible method for restricting the set of hypotheses about epistatic interactions that could be postulated. The number of possible combinations of polymorphisms that could be tested is therefore staggering, and this multiple hypothesis testing, in turn, necessitates imposing extremely stringent P -value thresholds. For a given sample and P -value threshold corrected for multiple testing, detecting even two polymorphisms whose interaction explains a given fraction of variance would require a sample size several orders of magnitude larger than the sample required to detect a single polymorphism that accounts for the same fraction of variance. As the order of the interaction increases, the requisite sample size quickly outstrips the number of people on the planet.

Second, even narrow heritability is likely to overstate the fraction of variance that discoverable polymorphisms are likely to capture. Estimates of narrow heritability from twin-family studies include additive variance attributable to any polymorphism, regardless of whether the polymorphism is common or rare among individuals. However, individual polymorphisms related to SWB that are rare in the population—which may collectively contribute much of the narrow heritability—will be much more difficult to reliably detect than polymorphisms that are common in the population.

In our empirical analysis, we estimate a parameter that cannot be estimated from twin or family data and that is necessarily smaller than narrow heritability, namely common narrow heritability: the fraction of variance that can be accounted for in aggregate by the cumulative additive effects of genetic polymorphisms that are common in the population (typically defined as minor allele frequency >1%). To do so, we use a recently developed method (31, 32) called genomic-relatedness-matrix restricted

maximum likelihood, or GREML. We apply GREML to SWB, pooling data from two large datasets, TwinGene [TG; the genotyped subsample of the Swedish Twin Registry (33)] and the Rotterdam Study (RS) (34), both datasets in which dense SNP genetic data have been collected. GREML has previously been used to estimate the common narrow heritability of height (31), intelligence (35, 36), personality traits (37), several common diseases (38), schizophrenia (39), economic and political preferences (22), as well as smoking, glucose levels, and depression (40). GREML has not previously been applied to SWB.

GREML estimates a heritability parameter by examining how, across pairs of individuals, phenotypic similarity relates to genetic similarity, after controlling for observables (in our case: age, sex, 20 principal components of the variance-covariance matrix of the genotypic data, and an indicator for dataset) (*Materials and Methods*). However, unlike in twin-family studies, where expected genetic similarity (inferred from the family pedigree) is used, GREML proceeds by first estimating the realized genetic similarity between pairs of unrelated individuals using the dense SNP data. To be more precise, realized genetic similarity is estimated using the sample covariance matrix of the individuals' genotypes. Because the genotypic data contains over half a million SNPs, this matrix is estimated very precisely. Because the covariance is a linear operator, the GREML method picks up the fraction of variance explained by the linear, additive action of the SNPs (i.e., the part of narrow heritability that is a result of the measured polymorphisms). Hence, GREML does not require the assumptions—about the degree of environmental and genetic resemblance between relatives and about the specific form of genetic effects (e.g., additive or dominance)—that tend to incite controversy when twin or family data are used to estimate narrow heritability. Because the current genotyping platforms from which our dense SNP data are obtained do not measure polymorphisms that are rare but do tag most of the genetic variation that is common in the population (41), GREML, as applied to these genotypic data, yields an estimate of common narrow heritability.

The key identifying assumption in GREML is that genetic similarity is uncorrelated with similarity in uncontrolled-for environmental factors that are exogenous to genotype (as defined by ref. 42). This assumption might be violated if the sample includes members of a shared extended family, such as siblings or close cousins. Therefore, it is standard to include in the sample only one individual from each family (in our case, one twin from a pair) and to drop individuals whose estimated genetic relatedness lies outside a small interval around zero. Because there is more random variation in the realized degree of genome sharing relative to the expected degree as the expected relatedness declines (43), uncontrolled-for environmental confounding factors are less likely to drive estimates that are based on realized relatedness among individuals whose expected relatedness is negligible.

We measure SWB in the TG and RS samples using responses to the two items from the Center for Epidemiologic Studies Depression Scale positive affect subscale that are available in both studies, namely responses regarding how frequently “During the past week, I was happy” and “During the past week, I enjoyed life.” We refer to these questions, respectively, as *Happy* and *Enjoy*. Responses are elicited using a four-point Likert scale ranging from “Rarely or none of the time (less than 1 day)” to “Most or all of the time (5–7 days).” *Combined* is a composite measure of the two variables. Because a substantial majority of responses to *Happy* and *Enjoy* are either in the highest-frequency category or the second-highest category (as is common with SWB survey measures), and because the software GCTA (genome-wide complex trait analysis) (32) that we use for the GREML analysis cannot presently handle multinomial variables, we converted the responses to binary variables (for details, see *Materials and Methods*).

Results

Table 1 reports the GREML estimates. For each sample, TG and RS, we report the fraction of variance explained by the measured SNPs for each of the two questions and for the combined SWB measure. We also report estimates with the TG and RS results pooled. In the TG sample, the GREML estimate for *Happy* is 0.10 (SE 0.10); for *Enjoy*, 0.06 (SE 0.10); and for *Combined*, 0.08 (SE 0.10). The corresponding figures in RS are 0.06 (SE 0.10), 0.04 (SE 0.10), and 0.08 (SE 0.10). Fig. 1 shows the estimates and SEs for the pooled sample. These estimates are of course more precise. For two of the three SWB measures, *Happy* ($h^2_{\text{SNP}} = 0.10$; SE 0.05) and *Combined* ($h^2_{\text{SNP}} = 0.09$; SE 0.05), the estimates are statistically distinguishable from zero at the 5% level.

These estimates are attenuated by measurement error in SWB. Moreover, because our measures of SWB are based on only two questions, the attenuation is probably more severe compared with what is typically observed for lengthier personality batteries. We estimated reliability using data on 105 Screening Across the Lifespan Twin study (SALT) respondents who answered the SALT survey twice, with 2 wk between the measurement occasions. Our estimates are 0.55 (SE 0.12), 0.41 (SE 0.15), and 0.61 (SE 0.12) for *Happy*, *Enjoy*, and *Combined*, respectively. Although there are also individuals in the RS study who participated in multiple waves and for whom two or more responses are available, these responses are at least 2 y apart in time, and thus more of the change in measured SWB is likely to reflect true changes in SWB; nonetheless, for completeness, we also report the RS estimates in Table 1.

A simple adjustment for attenuation is to divide the heritability by the retest reliability. This adjustment assumes that any change in measured SWB between one survey occasion and the next is because of classic measurement error (which is uncorrelated with genotype) and not true change. Using this adjustment and the SALT retest reliabilities, we estimate that 18%, 12%, and 15% of the variance in *Happy*, *Enjoy*, and *Combined* would be accounted for by common SNPs if the SWB variables were measured without error.

In the *SI Text*, we report results of additional analyses that examine how sensitive our baseline results are to the choice of relatedness thresholds and to whether relatedness is estimated only using SNPs available on both platforms (Table S1). Reassuringly, these estimates are very close to those in our preferred specification. We also attempted to compare the GREML estimates to conventional twin-based estimates, when both are derived from the Swedish twin sample. The twin estimates, shown in Table S2, are in the same range as the GREML estimates, but unfortunately we cannot draw strong conclusions because the twin estimates are very imprecise.

Discussion

In this article we provide evidence on the common narrow heritability of SWB. We find that 5–10% of the variance in responses to single-question survey measures of SWB is accounted for by the additive effects of the SNPs measured on presently used genotyping platforms. A correction for measurement error in the SWB measures raises the point estimates to the range 12–18%.

We interpret our findings as indicating that the common narrow heritability is smaller than the typical estimates of narrow heritability from twin-family studies [although one recent, large-scale twin-family study estimated narrow heritability in the 10–20% range, as small as our estimates of narrow heritability (16)]. A caveat to this conclusion is that the twin-based heritability point estimates in our Swedish sample are actually lower than the GREML estimates, which raises the alternative possibility that our low GREML estimates are because of anomalously low “true” heritabilities in the data we happened to study, perhaps because the SWB measures that were available in our data tap into recently experienced SWB to a greater extent than do multi-item dispositional measures of SWB.

There are three reasons why we believe that our interpretation is more compelling than this alternative. First, the twin-based estimates that we report, which are only available from the Swedish sample, are sufficiently imprecise that we have little confidence in the point estimates. Moreover, with a retest reliability of ≈ 0.5 in the SWB measure, the measurement-error-adjusted 95% confidence intervals would overlap comfortably with the consensus estimates from the literature on twin-family studies. Second, we have relatively more confidence in our GREML estimates, which are similar across our Swedish and Dutch samples. Third and relatedly, our interpretation of the data also fits with the evidence regarding personality, another complex behavioral trait for which epistatic interactions have been hypothesized to be important (29). Twin-based analyses tend to produce heritability estimates for personality around 30–50% (44, 45), but a recent study finds evidence that a substantial share of the heritability of Neuroticism, Openness, and Agreeableness is due to nonadditive factors (46). Two studies applying GREML to personality traits have been published to date, with results remarkably close to those reported here. One study reports estimates of 9% and 12% for neuroticism and extraversion (37), respectively, and another report estimates in the range 4–10% for traits assessed by the Cloninger personality inventory (47).

For SWB, the gap between the common narrow heritability we estimate and the larger estimates of narrow heritability from twin-family studies may imply that some of the narrow heritability is due to rare polymorphisms. For most traits, it is not well understood to what extent rare polymorphisms with substantial

Table 1. GREML (common narrow heritability) estimates for SWB

Sample	No. SNPs	<i>Happy</i> (“...I was happy”)				<i>Enjoy</i> (“...I enjoyed life”)				<i>Combined</i>			
		<i>n</i>	h^2_{SNP}	SE	<i>P</i> value	<i>n</i>	h^2_{SNP}	SE	<i>P</i> value	<i>n</i>	h^2_{SNP}	SE	<i>P</i> value
TG	627,011	5,682	0.10	0.10	0.16	5,742	0.06	0.10	0.28	5,670	0.08	0.10	0.22
RS	533,323	5,904	0.06	0.10	0.28	5,919	0.04	0.10	0.35	5,893	0.08	0.10	0.20
Pooled	852,597	11,484	0.10	0.05	0.01	11,558	0.05	0.05	0.19	11,461	0.09	0.05	0.03
SALT retest		105	0.55	0.12		105	0.41	0.15		105	0.61	0.12	
RS retest		7,845	0.41	0.03		7,916	0.42	0.03		7,795	0.41	0.03	

This table reports GREML estimates for *Happy*, *Enjoy*, and *Combined*. RS is the three Rotterdam cohorts pooled together; Pooled combines the TG and RS samples. We estimated the matrix of genetic relatedness after omitting one twin per pair in the Swedish data and then restricted the analyses to individuals whose relatedness did not exceed 0.025. *n* is the number of individuals used in the analyses after the relatedness threshold has been applied. In all analyses we control for sex, age, age-squared, and the first 20 principal components of the variance-covariance matrix of the genotypic data. The *P* value is from a likelihood ratio test of the null hypothesis that the fraction of variance explained is equal to zero. SALT retest is the sample of respondents in the SALT study who answered the survey twice (with 1 wk between survey occasions). RS retest is the retest correlation estimated using RS participants who answered the relevant questions in at least two different survey waves. These answers are at least 2 y apart in time.



Fig. 1. This figure shows the GREML for *Happy*, *Enjoy*, and *Combined*. The error bars represent the point estimate ± 1 SD. The sample pools the three Rotterdam cohorts (RS) and the Swedish Twin Registry TwinGene sample (TG).

effects account for the heritable variation (48). Until very recently, rare polymorphisms were not measured on standard genotyping platforms, and therefore most hypotheses regarding their role are based on indirect inferences such as those we make here (49).

Common narrow heritability is the quantity of most direct interest for assessing the potential contributions of genetic data to SWB research. Nonetheless, it may also be of interest to calibrate what our GREML results imply about narrow heritability, given that our GREML estimates do not rely on the same assumptions as the twin- and family-based estimates of narrow heritability. Two well-measured and widely studied complex traits for which reasonably reliable heritability estimates are available are height and cognitive ability. The twin-based estimates tend to fall in the range 50–80% for adult intelligence (50) and 80–90% for height (51). These estimates provide an upper bound on the narrow heritabilities. Other family-member comparisons—of full biological siblings, half-siblings, and parents and their children—suggest that the narrow heritabilities are unlikely to fall below 50% for adult intelligence (52) and 60% for height (53). By comparison, the one published GREML estimate for height is 45% (31), and GREML estimates for cognitive ability have also been around 45% (35, 36). This evidence suggests that the common narrow heritability that we estimate should be adjusted upward by a factor of roughly 1.5 to recover a ballpark estimate of narrow heritability.

Although our empirical contribution in this article focuses on estimating the common narrow heritability of SWB, we also believe it is important to highlight for SWB researchers that the conclusions that can be drawn from heritability estimates are more limited than is generally understood (for a related discussion, see ref. 54). Two misconceptions in particular appear to be widespread. First, some scholars erroneously conclude that higher heritability implies less variation left over to be explained by environmental factors. As the authors of the *World Happiness Report* (8) put it, twin studies are often misleadingly understood as “estimating the extent to which happiness depends on genetically based personality differences rather than differing circumstances.” However, as Jencks (42) explained, heritability comprises any genetic variation that ultimately contributes to phenotypic variation, regardless of the pathway, and many plausible pathways are in fact mediated by environmental factors. In the terminology of econometrics, the population regression of SWB on all genes—for which heritability is the R^2 —is a reduced-form regression, but the structural equations describing the true relationship between a gene and SWB may involve many intervening environmental variables. Although some genes may affect SWB via relatively direct physical pathways—for example, by affecting baseline serotonin levels or dopamine-receptor density—it also seems likely that many genes matter for SWB through their effects on

preferences, personality, and abilities, which in turn influence individuals’ choices about friendships, marriage, fertility, and occupational choice. Consequently, some of the variance in SWB explained by genes is the same variance that is explained by these environmental factors. Because genetic effects may operate indirectly through environmental variables, the heritability of SWB does not put any bound on the proportion of variance that could be explained by the full set of relevant environmental variables.

Second, findings of higher heritability are sometimes misinterpreted as demonstrating that there is less scope for interventions to increase SWB. For example, in their seminal paper, Lykken and Tellegen (11) conclude that “trying to be happier [may be] as futile as trying to be taller.” Such a claim may or may not be true, but it is in no way implied by the finding that SWB is heritable. The conclusion is incorrect for two distinct reasons. Related to the point above, some genetic effects may be mediated by modifiable environmental variables. The very same genotype may cause a person to grow to 5 feet or 6 feet tall, depending on nutritional intake. Furthermore, as Goldberger (55) pointed out, even if heritability were 100% and the genetic effects operated entirely through mechanisms that are difficult to modify, there may still exist powerful environmental interventions that do not contribute to outcome variance in the current population. As Bang (54) emphasized in her discussion of genetics and SWB, a heritability estimate represents the fraction of variance explained by genes in a specific population at a specific point in time. Using econometric terminology, one set of explanatory variables (in this case, genes) having a high R^2 does not rule out a large coefficient on another variable (an intervention), if the latter explanatory variable varies little across individuals in the population under study. In Goldberger’s (55) example, the introduction of eyeglasses dramatically improves vision even though eyesight is highly heritable.

To summarize what can be concluded from our findings, the magnitude of common narrow heritability provides useful information regarding the potential contributions of genetic data for research on SWB. Because our estimates are lower than typical heritability estimates for SWB, our results suggest that the scope for uses of genetic data that rely on substantial predictive power—such as using a set of polymorphisms as control variables, instrumental variables, or moderators in social-science research—may be more limited than has been assumed. At the same time, the fact that our estimates of common narrow heritability are nonnegligible suggests that—even if much of the broad heritability is due to epistatic interactions—some of the SNPs measured on existing platforms have main effects on SWB. Therefore, gene-discovery efforts with a large enough sample size are likely to be successful.

Materials and Methods

Our study combines data from the Swedish Twin Registry (STR) and the Rotterdam Study. STR is a large, population-based twin registry. Between 1998 and 2002, STR administered to twins born in 1958 or earlier a survey called the Screening Across the Lifespan Twin study (33). A subsample of SALT was recently genotyped using the Illumina HumanOmniExpress BeadChip technology as part of the TwinGene project (22). We refer to these $\approx 10,000$ genotyped SALT respondents as the TG sample. TG participants are all born between 1911 and 1958.

RS is a large, population-based prospective cohort study of elderly people ongoing since 1990 in the city of Rotterdam in the Netherlands (34). Approximately 11,000 subjects in the RS have been genotyped using the Illumina 550 K and 610 K arrays. RS respondents are divided into three cohorts, which we refer to as RS-I, RS-II, and RS-III.

To minimize the expected relatedness of the individuals in our sample, we only included one twin per family in the TG sample. If only one twin from a pair had answered both survey questions, the individual with complete phenotypic data was included in the analysis. If both twins had complete phenotypic data, one of them was chosen at random. We then pooled the

resulting sample with the RS sample and used the GCTA software to estimate the pairwise relatedness between all individuals in the pooled dataset. Following convention, we restricted the sample to individuals whose pairwise relatedness did not exceed 0.025 in absolute value.

These restrictions brought the sample size to just below 6,000 individuals in each sample. Our analysis is restricted to individuals with SNP data that passed quality controls and who answered both questions. Results in TG are based on 627,011 SNPs; in RS, on 533,323 SNPs. Because of incomplete overlap in the two samples, the number of SNPs in the pooled sample is larger: 852,597.

To convert the *Happy* and *Enjoy* measures to binary variables, we coded responses as “high” if they were the highest-frequency category and “not high” otherwise. We also constructed a composite measure of the two variables, which we call *Combined*, the value of which is “high” if both *Happy* and *Enjoy* are high and “not high” otherwise. By generating the binary variables in this way, the fraction coded as high (or equivalently, not high) is made as close as possible to one half, thereby maximizing the variance and hence statistical power. The distributions of SWB measures before they were binarized are given in Tables S3 and S4. In the analyses we assume that each binary variable we observe results from the realization of an underlying, normally distributed random variable for liability falling above or below some threshold.

Table S5 reports age and sex, as well as the fraction of individuals coded as high for the three variables. Throughout, we control for sex, age, age-squared, dummies for each of the three RS cohorts (TG is the omitted category), and to guard against population stratification, the first 20 principal components of the genotype data.

ACKNOWLEDGMENTS. We thank Pascal Arp, Mila Jhamai, Michael Moorhouse, Marijn Verkerk, and Sander Bervoets for their assistance in creating the genome-wide association study database; and Martha Widger and Cody Xuereb for research assistance. This work was supported in part by the Ragnar Söderberg Foundation (E9/11); the National Science Foundation

(Early Concept Grants for Exploratory Research: “Workshop for the Formation of a Social Science Genetic Association Consortium”, SES-1064089) as supplemented by the National Institutes of Health’s (NIH) Office of Behavioral and Social Sciences Research, and the National Institute on Aging/NIH through Grants P01-AG005842, P01-AG005842-2052, P30-AG012810, and T32-AG000186-23 to the National Bureau of Economic Research. The Swedish Twin Registry is supported by the Swedish Department of Higher Education, the European Commission European Network for Genetic and Genomic Epidemiology [ENGAGE: 7th Framework Program (FP7/2007-2013)/Grant agreement HEALTH-F4-2007-201413; and GenomEUtwin: 5th Framework program “Quality of Life and Management of the Living Resources” Grant QLG2-CT-2002-01254]; the NIH (DK U01-066134); the Swedish Research Council (M-2005-1112 and 2009-2298); the Swedish Foundation for Strategic Research (ICA08-0047); the Jan Wallander and Tom Hedelius Foundation; and the Swedish Council for Working Life and Social Research. The Rotterdam Study is funded by the Erasmus Medical Center; Erasmus University, Rotterdam; the Netherlands Organization for the Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture, and Science; the Ministry for Health, Welfare, and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. The genome-wide association study database of the Rotterdam Study was funded by the Netherlands Organization for Scientific Research Investments (175.010.2005.011, 911-03-012); the Research Institute for Diseases in the Elderly (014-93-015), and the Netherlands Genomics Initiative/Netherlands Consortium for Healthy Aging Project 050-060-810. Statistical analyses were carried out on the Genetic Cluster Computer (www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003, PI: Posthuma) along with a supplement from the Dutch Brain Foundation and the Vrije Universiteit Amsterdam. D.J.B. received financial support through National Institute on Aging/NIH Grants RO1-AG040787 to the University of Michigan and T32-AG00186 to the National Bureau of Economic Research. J.-E.D.N. acknowledges financial support from the UK Department for Work and Pensions and National Institute on Aging/NIH Grant RO1-AG040640.

- Easterlin R (2003) Explaining happiness. *Proc Natl Acad Sci USA* 100(19):11176–11183.
- Kahneman D, Deaton A (2010) High income improves evaluation of life but not emotional well-being. *Proc Natl Acad Sci USA* 107(38):16489–16493.
- Diener E (2000) Subjective well-being. The science of happiness and a proposal for a national index. *Am Psychol* 55(1):34–43.
- Seligman ME, Csikszentmihalyi M (2000) Positive psychology. An introduction. *Am Psychol* 55(1):5–14.
- Vaillant GE (2012) Positive mental health: Is there a cross-cultural definition? *World Psychiatry* 11(2):93–99.
- Diener E, Seligman MEP (2004) Beyond money: Toward an economy of well-being. *Psychol Sci Public Interest* 5(1):1–31.
- Kahneman D, Krueger AB (2006) Developments in the measurement of subjective well-being. *J Econ Perspect* 20(1):3–24.
- Helliwell J, Layard R, Sachs J, eds (2012) *World Happiness Report* (The Earth Institute, New York).
- Headley B, Muffels R, Wagner GG (2010) Long-running German panel survey shows that personal and economic choices, not just genes, matter for happiness. *Proc Natl Acad Sci USA* 107(42):17922–17926.
- Harris JR, Pedersen NL, Stacey C, McClearn GE (1992) Age differences in the etiology of the relationship between life satisfaction and self-rated health. *J Aging Health* 4(3):349–368.
- Lykken DT, Tellegen A (1996) Happiness is a stochastic phenomenon. *Psych Sci* 7(3):186–189.
- Tellegen A, et al. (1988) Personality similarity in twins reared apart and together. *J Pers Soc Psychol* 54(6):1031–1039.
- Roysamb E, Harris JR, Magnus P, Vitterso J, Tambs K (2002) Subjective well-being. Sex-specific effects of genetic and environmental factors. *Pers Individ Dif* 32(2):211–223.
- Stubbe JH, Posthuma D, Boomsma DI, De Geus EJ (2005) Heritability of life satisfaction in adults: A twin-family study. *Psychol Med* 35(11):1581–1588.
- Nes RB, Roysamb E, Tambs K, Harris J-R, Reichborn-Kjennerud T (2006) Subjective well-being: Genetic and environmental contributions to stability and change. *Psychol Med* 36(7):1033–1042.
- Bartels M, Boomsma DI (2009) Born to be happy? The etiology of subjective well-being. *Behav Genet* 39(6):605–615.
- Franz CE, et al. (2012) Genetic and environmental multidimensionality of well- and ill-being in middle aged twin men. *Behav Genet* 42(4):579–591.
- Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. *Am J Hum Genet* 90(1):7–24.
- Bartels M, et al. (2010) Heritability and genome-wide linkage scan of subjective happiness. *Twin Res Hum Genet* 13(2):135–142.
- De Neve JE (2011) Functional polymorphism (5-HTTLPR) in the serotonin transporter gene is associated with subjective well-being: Evidence from a US nationally representative sample. *J Hum Genet* 56(6):456–459.
- De Neve JE, Christakis NA, Fowler JH, Frey BS (2012) Genes, economics, and happiness. *J Neuroscience Psychology Econ* 5(4):193–211.
- Benjamin DJ, et al. (2012) The genetic architecture of economic and political preferences. *Proc Natl Acad Sci USA* 109(21):8026–8031.
- Benjamin DJ, et al. (2012) The promises and pitfalls of genoconomics. *Ann Rev Econ* 4(1):627–662.
- Beauchamp JP, et al. (2011) Molecular genetics and economics. *J Econ Perspect* 25(4):57–82.
- Smith GD, Ebrahim S (2003) ‘Mendelian randomization’: Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32(1):1–22.
- Ding W, Lehrer SF, Rosenquist JN, Audrain-McGovern J (2009) The impact of poor health on academic performance: New evidence using genetic markers. *J Health Econ* 28(3):578–597.
- Fletcher JM, Lehrer SF (2011) Genetic lotteries within families. *J Health Econ* 30(4):647–659.
- Conley D (2009) The promise and challenges of incorporating genetic data into longitudinal social science surveys and research. *Biodemography Soc Biol* 55(2):238–251.
- Lykken DT (1982) Presidential address, 1981. Research with twins: The concept of emergence. *Psychophysiology* 19(4):361–373.
- Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 109(4):1193–1198.
- Yang J, et al. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42(7):565–569.
- Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet* 88(1):76–82.
- Lichtenstein P, et al. (2006) The Swedish Twin Registry in the third millennium: An update. *Twin Res Hum Genet* 9(6):875–882.
- Hofman A, et al. (2011) The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol* 26(8):657–686.
- Davies G, et al. (2011) Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol Psychiatry* 16(10):996–1005.
- Chabris CF, et al. (2012) Most reported genetic associations with general intelligence are probably false positives. *Psychol Sci* 23(11):1314–1323.
- Vinkhuyzen AAE, et al. (2012) Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Transl Psychiatry* 2(4):e102.
- Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 88(3):294–305.
- Lee SH, et al.; Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ); International Schizophrenia Consortium (ISC); Molecular Genetics of Schizophrenia Collaboration (MGS) (2012) Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* 44(3):247–250.
- Lubke GH, et al. (2012) Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biol Psychiatry* 72(8):707–709.
- Barrett JC, Cardon LR (2006) Evaluating coverage of genome-wide association studies. *Nat Genet* 38(6):659–662.
- Jencks C (1980) Heredity, environment, and public policy reconsidered. *Am Sociol Rev* 45(5):723–736.
- Hill WG, Weir BS (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet Res (Camb)* 93(1):47–64.

Supporting Information

Rietveld et al. 10.1073/pnas.1222171110

SI Text

Swedish Twin Registry TwinGene Study (TG). Between December 2010 and May 2011, 9,836 Swedish twins were genotyped by the SNP&SEQ Technology Platform, Uppsala, using the Illumina HumanOmniExpress BeadChip genotyping platform. A detailed description of the analysis to detect pedigree errors is available in ref. 1. All genetic analyses reported in the present article are based on the dataset with cryptically related respondents removed from the sample.

Rotterdam Study. Genotyping of the Rotterdam Study (RS) samples was performed in the Genetic Laboratory, Department of Internal Medicine, Erasmus Medical Center, Rotterdam (HuGeF) using the Illumina, HumanHap 550K (RS-I and RS-II) and Illumina 610K Quad (RS-III) chip. Ethnic outliers and individuals with a mismatch between sex and typed X-linked markers, excess autosomal heterozygosity, or within-sample cryptic relatedness were removed.

RS and TG. In addition to the study-specific quality-control filters, additional control filters were imposed to prevent estimating spurious heritability due to different genotyping platforms across the two samples or due to the presence of twins in TG. SNPs with minor allele frequency <0.01, missing rate >0.05, or Hardy-Weinberg equilibrium test P value < 0.001 were excluded. Individuals with less than 95% of genotyped SNPs available were dropped, as were individuals with missing phenotypic information. In all analyses, one twin per family in the TG sample was always excluded. If only one twin from a pair had answered both survey questions, the individual with complete phenotypic data was included in the analysis. If both twins had complete phenotypic data, one of them was chosen at random for inclusion. The quality-control procedures resulted in a dataset of 627,011 SNPs for 6,767 individuals in the Swedish Twin Registry (STR); 517,635 SNPs for 3,878 individuals in RS-I; 498,999 SNPs for 2,085 individuals in RS-II; and 510,681 SNPs for 2,993 individuals in RS-III.

1. Benjamin DJ, et al. (2012) The genetic architecture of economic and political preferences. *Proc Natl Acad Sci USA* 109(21):8026–8031.

Table S1. Additional GREML (common narrow heritability) estimates of subjective well-being (SWB): Only SNPs in both datasets and no-relatedness threshold

Dataset	No. SNPs	Happy ("...I was happy")				Enjoy ("...I enjoyed life")				Combined			
		n	h^2_{SNP}	SE	P value	n	h^2_{SNP}	SE	P value	n	h^2_{SNP}	SE	P value
Only SNPs in both datasets	283,289	11,439	0.11	0.05	0.01	11,512	0.03	0.05	0.31	11,415	0.08	0.05	0.05
No-relatedness threshold	852,597	15,584	0.08	0.03	0.01	15,687	0.03	0.04	0.18	15,548	0.06	0.03	0.04

This table reports genomic-relatedness-matrix restricted maximum-likelihood (GREML) estimates for *Happy*, *Enjoy*, and *Combined* for the pooled sample (combining TG and RS). We estimated the matrix of genetic relatedness after omitting one twin per pair in the Swedish data. In the first analysis, we restricted the analyses to individuals whose relatedness did not exceed 0.025 and only SNPs that were available in all datasets. In the second analysis, we included all individuals and all SNPs. n is the number of individuals used in the analyses after the relatedness threshold has been applied. In all analyses we control for sex, age, age-squared, and the first 20 principal components of the variance-covariance matrix of the genotypic data. The P value is from a likelihood ratio test of the null hypothesis that the fraction of variance explained is equal to zero.

Table S2. Variance decomposition: Twin analyses in SALT

ACE decomposition	a^2	c^2	e^2	$-2 \ln L$
Same-sex and opposite-sex DZ				
<i>Happy</i>	0.04 (0.00–0.16)	0.06 (0.00–0.08)	0.91 (0.84–0.95)	24593.30
<i>Enjoy</i>	0.07 (0.00–0.16)	0.03 (0.00–0.11)	0.90 (0.84–0.95)	25289.21
<i>Combined</i>	0.03 (0.00–0.17)	0.08 (0.00–0.13)	0.89 (0.88–0.92)	23681.61
Only same-sex pairs included				
<i>Happy</i>	0.05 (0.00–0.15)	0.04 (0.00–0.10)	0.91 (0.85–0.96)	17652.16
<i>Enjoy</i>	0.08 (0.00–0.14)	0.00 (0.00–0.06)	0.92 (0.86–0.97)	18060.22
<i>Combined</i>	0.06 (0.00–0.17)	0.05 (0.00–0.13)	0.89 (0.83–0.95)	17047.92

The upper section reports estimates of an additive genetic/common environmental/unique environmental (ACE) model with same-sex and opposite-sex dizygotic (DZ) pairs included. The genetic and environmental factors are assumed to be the same in men and women. The lower section reports estimates from an ACE model with only same-sex pairs included. All models allow the liability threshold to vary by age and sex. The variance components are constrained to be the same in men and women. The Swedish Twin Registry's Screening Across the Lifespan Twin Study (SALT) sample is restricted to twin pairs in which both members responded to the survey and such that the age range is the same as in the TG sample. The SALT individuals used in these twin analyses are a superset of the TG sample used in the GREML analyses.

Table S3. Descriptive statistics: Single-question SWB measures

SWB measure	Sample	<i>n</i>	% Rarely or none of the time (less than 1 d)	% Some or a little of the time (1–2 d)	% Occasionally or a moderate amount of time (3–4 d)	% Most or all of the time (5–7 d)
<i>Happy</i>	RS-I	3,842	7.0	7.7	15.9	69.5
	RS-II	2,075	4.5	11.0	22.3	62.2
	RS-III	2,992	2.9	9.8	20.3	67.1
	STR	6,675	5.1	11.0	44.9	39.0
	STR+RS	15,584	5.1	10.0	30.0	55.0
<i>Enjoy</i>	RS-I	3,866	6.1	7.0	13.3	73.6
	RS-II	2,080	4.2	10.3	17.9	67.6
	RS-III	2,990	2.6	8.7	15.7	72.9
	STR	6,751	2.4	3.7	27.0	66.9
	STR+RS	15,687	3.6	6.3	20.3	69.8

This table provides summary statistics for the SWB measures used in the GREML analysis.

Table S4. Descriptive statistics: Combined SWB measure

SWB measure	Sample	<i>n</i>	0 (%)	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)
<i>Combined</i>	RS-I	3,830	3.2	2.1	5.0	6.4	8.2	11.5	63.5
	RS-II	2,070	2.7	1.7	6.4	6.3	12.0	15.9	55.1
	RS-III	2,989	1.4	1.4	5.2	5.8	11.0	14.3	61.0
	STR	6,659	1.1	1.3	3.8	8.0	21.4	28.7	35.8
	STR+RS	15,548	1.9	1.6	4.7	6.9	14.9	20.0	50.0

This table provides summary statistics for the combined SWB measure used in the GREML analysis.

Table S5. Descriptive statistics: GREML and Twin Analysis samples

Sample	<i>n</i>	% Female	Mean age (SD)	No. <i>Happy</i> (%high)	No. <i>Enjoy</i> (%high)	No. <i>Combined</i> (%high)
GREML						
RS-I	3,878	58.6	72.42 (7.35)	3,842 (69.5%)	3,866 (73.6%)	3,830 (63.5%)
RS-II	2,085	54.3	64.72 (7.93)	2,075 (62.2%)	2,080 (67.5%)	2,070 (55.1%)
RS-III	2,993	56.3	57.03 (6.78)	2,992 (67.0%)	2,990 (72.9%)	2,989 (60.1%)
RS	8,956	56.8	65.48 (9.91)	8,909 (67.0%)	8,936 (72.0%)	8,889 (60.7%)
TG	6,767	52.6	67.59 (8.92)	6,675 (38.9%)	6,751 (66.9%)	6,659 (35.8%)
TG+RS	15,723	55.0	66.39 (9.35)	15,584 (55.0%)	15,687 (69.8%)	15,548 (50.0%)
SALT						
SALT MZ	5,398	53.8	73.85 (5.21)	1,952 (37.8%)	3,308 (62.9%)	1,781 (34.6%)
SALT DZ	8,950	52.8	74.16 (5.21)	3,156 (37.0%)	5,460 (62.6%)	2,856 (33.6%)
SALT OS	5,682	50.0	76.19 (7.50)	1,917 (35.6%)	3,385 (61.2%)	1,703 (31.8%)

This table provides summary statistics for the variables used in the GREML and SALT analysis. DZ, dizygotic same-sex pairs; MZ, monozygotic; OS, opposite-sex DZ pairs. The TG sample is the subset of the SALT sample that has been genotyped, and in the TG (but not SALT) sample, one twin per family has been removed. The SALT sample is restricted to twin pairs in which both members responded to the survey and such that the age range is the same as in the TG sample. No. *Happy* refers to the number of observation individuals with nonmissing data for *Happy*; analogously for No. *Enjoy* and No. *Combined*. For *Happy* and *Enjoy*, "%high" refers to the fraction of individuals who chose the highest response category for the SWB question; for *Combined*, it is the fraction coded as high for both questions.