

Cognition across the lifespan: Aging and gender differences

Nichols, E. S.^{a,b*}, Wild, C. J.^b, Owen, A. M.^{b,c,d}, & Soddu, A.^{a,b}

^aThe Department of Physics and Astronomy, the University of Western Ontario, London, Ontario, Canada

^bBrain and Mind Institute, the University of Western Ontario, London, Ontario, Canada

^cThe Department of Physiology and Pharmacology, the University of Western Ontario, London, Ontario, Canada

^dThe Department of Psychology, the University of Western Ontario, London, Ontario, Canada

* Corresponding author

E-mail: enicho4@uwo.ca (ESN)

21 Abstract

22 Maintaining cognitive health across the lifespan has been the focus of a multi-billion-dollar industry.
 23 In order to guide treatment and interventions, a clear understanding of the way that proficiency in
 24 different cognitive domains develops and declines across the lifespan is necessary. Additionally, there
 25 are gender differences in a range of other factors, such as anxiety and substance use, that are also
 26 known to affect cognition, although the scale of this interaction is unknown. Our objective was to
 27 assess differences in cognitive function across the lifespan in men and women in a large,
 28 representative sample. Over 45,000 individuals were tested on 12 cognitive tasks. Segmented
 29 regression was used to model the trajectory of three cognitive domains: short-term memory, verbal
 30 abilities, and reasoning. Each domain showed a unique trajectory, suggesting that not all cognitive
 31 abilities develop and decline in the same way. Gender differences were found in all three domains;
 32 however, after controlling for socio-demographic factors, these differences were greatly reduced or
 33 disappeared. These results suggest that the trajectory of cognition across the lifespan differs for men
 34 and women, but is greatly influenced by environmental factors. We discuss these findings within a
 35 framework that describes gender differences in cognition as likely guided by a complex interplay
 36 between biology and environment.

38 Introduction

39 By 2020, roughly 22% of the world's population will be over 65, a total of approximately 1.7
 40 billion people [1]. The consequences of our aging population are many, including an increasing focus
 41 on maintaining cognitive health; more so than ever before, individuals are seeking ways to keep their

42 minds sharp. This new interest in maintaining cognitive health and reversing, or stalling, normal
43 cognitive decline has led to the creation of billion-dollar industry, promoting products as wide-ranging
44 as “brain training” software and pharmaceutical interventions such as nootropics. Yet, in order to be
45 able to evaluate these approaches as potential tools and treatments, it is important that we first have
46 a clear understanding of how cognition changes across the lifespan in average, healthy individuals.
47 Additionally, because of the often-cited cognitive differences between women and men [2–5], it is
48 important to characterize cognition in each population; if gender differences in cognitive abilities do
49 exist, then men and women may respond differently to cognitive aging interventions.

50 In healthy individuals, cognitive abilities develop rapidly throughout childhood. By six, most
51 children have developed some degree of inhibitory control, verbal fluency, and task switching [6–8].
52 By twelve, they are able to plan and organize, and use conceptual strategies and reasoning [9]. In
53 adolescence, these abilities continue to develop, with most teenagers having good attentional
54 control, verbal fluency, and processing speed [6]. By 18, executive function is thought to be mature
55 [10], although research suggests that some processes continue to develop in early adulthood [11].
56 Young adulthood is where most researchers agree that cognitive abilities peak; however there is large
57 variability within this period across different cognitive functions [6,11]. Mid to late adulthood is then
58 characterized by a slow decline in most cognitive abilities [8,12], and while it can be problematic, this
59 decline is considered part of healthy aging.

60 Differences in cognitive abilities between men and women are less clear; although several
61 gender disparities in cognitive abilities appear to exist, recent studies have found these differences to
62 be mediated by underlying factors related to gender, rather than being inherent to gender itself. For
63 example, during childhood, girls are thought to develop faster than boys in verbal fluency and

information processing [2,3], and boys are thought to develop faster than girls in spatial reasoning, working memory, and number processing [4,13]. However, Tzuriel and Egozi [14] found that gender differences in mental rotation disappeared after visuospatial training. Similarly, Krinzinger and colleagues [15] found that number processing advantages in boys were mediated by attitudes toward mathematics, and similar results have been found in young adults [16]. Differences in verbal processing have been less clear, with some suggesting that they are due to variability in instruction and strategy [17,18], and others suggesting a hormonal link [19,20]. Reports of gender differences in age-related cognitive decline are largely thought to be the result of cohort effects [21–23], although others have found gender-specific links to brain-derived neurotrophic factor [24] and brain metabolic activity [25]. Realistically, the truth likely lies somewhere in between, with a multifaceted interaction of biology and environment [25,26].

Finally, there are a number of sociodemographic factors known to affect cognition. For example, socioeconomic status (SES) and its relationship with cognitive abilities has been widely studied; it is generally agreed that higher SES predicts better performance on cognitive tasks [27,28]. Additionally, anxiety, depression, and substance abuse also have known detrimental effects on cognition, with higher levels of all three being associated with poorer cognitive outcomes [29–31]. Such sociodemographic factors also interact with gender; women tend to experience higher levels of anxiety [32] and depression [33], while men tend to experience higher levels of substance abuse [34], although women may be more at risk specifically for alcohol abuse (35, but see 36). Thus, there is a complex interaction of age, gender, and other sociodemographic variables that must be considered when studying cognitive abilities across the lifespan.

The internet provides a unique opportunity for examining cognition across the lifespan in the general population on a huge scale, allowing data to be sampled from participants from a broad range of SES, geographical, and educational backgrounds. There is a growing body of evidence to suggest that online cognitive assessment is as reliable as in-person pen-and-paper testing [11,37], addressing concerns that data collected in this manner may not be valid. Leveraging the power of the internet provides us with a cross-sectional snapshot of both demographics and cognition from a larger and more diverse sample than would be possible to collect in the laboratory.

The first goal of the present study was to characterize cognitive abilities across the lifespan, ranging from adolescence to late adulthood. Specifically, we sought to address whether differences exist between cognitive domains; do memory and reasoning show the same pattern, or do they peak at different ages? Do reasoning and verbal abilities show the same rate of decline, or does one remain resilient to aging more so than the other? The second goal was to examine whether these age effects differed between genders, and what factors may influence these differences. Specifically, we sought to address whether differences exist in some cognitive domains and not others, whether men and women peak at the same age, and whether they decline at the same rate. Further, we explored the demographic and social factors that affect the genders differently, and whether controlling for these differences affects the observed pattern of cognitive abilities across the lifespan. Based on smaller studies using more limited time windows, we predicted that gender differences would manifest with men outperforming women in memory and reasoning, but with women outperforming men in verbal abilities, and that the pattern of these abilities would show an increase up to early adulthood, and a slow decline into mid and late adulthood. By taking into account studies of the effects of mental

health and sociodemographic variables on cognition, we also predicted that controlling for these factors would eliminate gender differences in cognitive abilities.

Materials and methods

Participants

All data for this study were collected with the CBS (www.cambridgebrainsciences.com) online platform, which has previously been used for other large-scale studies of cognition [38,39]. A total of 65,994 participants registered and completed all portions of the study, with a final number of 45,779 being included after removing outliers and rows with missing data (described in the data preprocessing section below). A summary of the final sample's demographics is included in Table 1, broken down by gender. All participants gave informed consent, and ethics approval was obtained through the Cambridge Psychology Research Ethics Committee (2010.62).

Table 1. Comparison of demographic variables across women and men.

Measure	Mean (SD) or Percentage		$t(df)$, W , $\chi^2(df, N)^1$	p	Effect size ²
	Women	Men			
<i>N</i>	13,444	32,335			
Age (years)	28.08 (11.01)	28.22 (10.30)	-1.31(23,696)	0.220	-0.01
Sleep (hours per night)	7.09 (1.64)	6.93 (1.64)	9.51(25,241)	< .001	0.10
Alcohol (units per week)	1.72 (1.85)	1.78 (2.00)	-3.10 (27,022)	.003	-0.03
Caffeine (units per day)	3.24 (4.69)	4.22 (5.53)	-19.36 (29,414)	< .001	-0.19
Cigarettes (per day)	1.48 (4.54)	1.71 (5.06)	-4.68 (27,830)	< .001	-0.05
Highest education completed			2.17e ⁸	.989	6.27e ⁻⁵
Some high school	10.20%	9.80%			
High School	8.40%	10.60%			
Some post-secondary	27.80%	30.50%			
Post-secondary degree	27.20%	27.70%			
Professional degree	26.40%	21.40%			

Level of employment			2.21e ⁸	.003	0.01
No answer	4.10%	2.70%			
Unemployed	10.30%	12.10%			
Full time student	28.70%	25.40%			
Employed and student	15.50%	11.90%			
Employed part time	10.00%	6.20%			
Employed full time	31.40%	41.80%			
Exercise			2.18e ⁸	.974	5.35e ⁴
Never	10.60%	11.00%			
Infrequently	37.80%	34.70%			
Weekly	19.90%	18.20%			
Several times a week	25.30%	28.30%			
Every day	6.50%	7.90%			
Depressive feelings			2.36e ⁸	< .001	0.07
No answer	1.10%	1.10%			
Never	10.40%	17.50%			
Occasionally	56.80%	54.00%			
Quite often	21.40%	18.20%			
Nearly every day	7.30%	6.40%			
All the time	3.10%	2.80%			
Anxiety			2.58e ⁸	< .001	0.15
No answer	1.30%	1.50%			
Never	13.00%	24.00%			
Occasionally	48.80%	49.90%			
Quite often	21.20%	14.80%			
Nearly every day	10.80%	6.50%			
All the time	5.00%	3.20%			
Video games			2.42e ⁹	< .001	0.09
Never	38.00%	17.10%			
Monthly	26.00%	20.30%			
Weekly	21.40%	29.40%			
Daily	14.60%	33.20%			
Siblings			2.19e ⁸	.111	0.01
Only child	12.40%	10.60%			
Youngest	30.40%	31.00%			
Middle	16.40%	17.40%			
Oldest	40.80%	41.00%			
Religiosity			2.34e ⁸	< .001	0.06
Atheist	31.90%	42.40%			
Agnostic	31.50%	31.50%			

Religious lapsed	19.50%	15.10%			
Religious practicing	13.10%	8.10%			
Very religious	3.90%	2.90%			
Political leaning			68.67(2, N = 45,779)	< .001	0.04
Liberal	47.20%	44.90%			
Middle	45.20%	45.10%			
Conservative	7.60%	10.00%			
Tech savvy			2693.40(1, N = 45,779)	< .001	0.24
Yes	64.00%	89.10%			
No	36.00%	10.90%			

¹Welch's *t*-test used to compare numeric variables, Wilcoxon Rank-Sum used to compare ordinal variables, and χ^2 used to compare categorical variables

²Effect sizes used were Cohen's *d* for *t*-tests, *r* for Wilcoxon Rank-Sum tests, and Cramer's *V* for χ^2 tests

Materials

Sociodemographic questionnaire

The sociodemographic questionnaire included questions about the individual's age and gender, lifestyle such as exercise, substance use, and sleep, mental health such as depressive symptoms and anxiety, and other demographics such as education, employment, and level of technical savviness. Demographics included in the present study are listed in Table 1. Demographic information collected but not of interest here included country of birth, hours slept the night before completing the study, and favourite type of music. The full demographic questionnaire is available in Appendix 1.

Cognitive battery

Prior to filling in the questionnaire, participants completed the 12 tests in the CBS test battery. Test order was fixed across participants. Detailed descriptions of the tests can be found in the Supplementary Material of Hampshire et al. [31], but in brief they are: (1) 'Monkey Ladder'

(visuospatial working memory); (2) 'Grammatical Reasoning' (verbal reasoning); (3) 'Double Trouble' (a modified Stroop task); (4) 'Odd One Out' (deductive reasoning); (5) 'Spatial Span' (short-term memory); (6) 'Rotations' (mental rotation); (7) 'Feature Match' (feature-based attention and concentration); (8) 'Digit Span' (verbal working memory); (9) 'Spatial Planning' (planning and executive function); (10) 'Paired Associates' (shape-location associative memory); (11) 'Interlocking Polygons' (visuospatial processing); and (12) 'Token Search' (working memory and strategy).

Factor analysis

Imaging studies have underscored the fact that there is rarely a one-to-one mapping between cognitive functions and the brain areas, or networks, that underpin them. One approach to this issue is to examine the complex statistical relationships between performance on any one cognitive task (or group of tasks), and changes in brain activity to reveal how one is related to the other. In order to do this most effectively, large amounts of data need to be included because of the natural variance in cognitive performance (and brain activity) across tests and across individuals. In the age of computerized internet testing and so-called 'big data', this problem becomes much easier to solve. Hampshire et al. [31] collected data on the 12 CBS tasks from 45,000 participants. These data were then subjected to a factor analysis and 3 discrete factors relating to overall cognitive performance were identified. Each one of these factors represents an independent cognitive function that is best described by a combination of performance on multiple tests, something that no single test can assess, and were labeled as encapsulating aspects of short-term memory, reasoning, and verbal abilities, respectively. This technique allows an individual's performance to be compared to a very large normative database in terms of these descriptive factors rather than performance on a single test.

Here, the same 12 tests were used to create three “composite” scores reflecting performance based on the factor analysis by Hampshire and colleagues, and the three composite scores were calculated as follows. First, the 12 individual test scores were normalized ($M = 0.0$, $SD = 1.0$). Then, the three cognitive domain scores were calculated using the formula $Y = X(Ar^+)^T$, where Y is the $N \times 3$ matrix of domain scores, X is the $N \times 12$ matrix of test z-scores, Ar is the 12×3 matrix of varimax-rotated principal component weights (i.e., factor loadings) from Hampshire et al. All 12 tests contributed to each domain score, as determined by their component weights. Because scores were demeaned, a domain score of 0.0 is the mean of the population that was used to derive the loadings. Thus, a score above zero indicates that someone is above average.

Data preprocessing

Only data from the participants who completed all questionnaire items and all 12 tests were included in analysis. 65,994 participants met these requirements. Data were then cleaned to remove impossible and improbable questionnaire responses, removing 5,732 participants. Examples of improbable responses include smoking over 60 cigarettes per day, sleeping more than 17 hours the night before, or consuming more than 50 alcoholic drinks per day. Test scores were then filtered for outliers in two passes: scores greater than six standard deviations were assumed to be technical errors and were first removed, eliminating 7,298 participants. Then, scores greater than four standard deviations from the recalculated mean were identified, assumed to be performance outliers, and removed, eliminating 7,157 participants. Finally, individuals younger than 12 and older than 69 were removed because of low numbers outside of this age range, eliminating 28 participants. 45,779 participants were included in the final analysis.

Statistical analyses

Data were analyzed in R (version 3.5.2, R Core Team, 2018) and RStudio (version 1.1.463).

Specific packages included: ‘Segmented’ [40] for computing regressions with breakpoints, ‘MatchIt’ [41] for matching samples on demographic variables, ‘parallel’ for parallel computing, and ‘boot’ [42] for calculating confidence intervals. Figs were produced using ‘ggplot2’ [43].

To examine the differences in demographic variables between genders, three different tests were used: Welch’s *t*-tests for continuous variables, Wilcoxon Rank Sum tests for ordinal variables, and chi-square tests for categorical variables. *P*-values were corrected for multiple comparisons using a false discovery rate and were considered significant at $p < .05$. Effect size was calculated using the appropriate measures for each test: Cohen’s *d* for *t*-tests, *r* for Wilcoxon Rank Sum tests, and Cramer’s *V* for chi-square tests. Results are included in Table 1. Measures of skew and kurtosis indicated that domain scores were normally distributed, and histograms are shown in Fig 1.

Fig 1. Histograms of domain scores by gender. Dashed lines indicate mean.

Segmented linear regression models were constructed to predict each of the 3 domain scores from participants’ reported age and were estimated using maximum likelihood estimation. Segmented regression was used to fit a model in which there is a change in the linear relationship – such as a “peak” that indicates a transition from increasing to decreasing performance with age – without imposing a pre-determined shape (e.g., quadratic or cubic) through adding one or more piecewise linear relationships [40,44]. The value of the independent variable (i.e., age) at which this change occurs is referred to as a breakpoint. The relationship between cognitive performance and age was modeled separately for each gender.

198 The segmented regression technique used here requires that the number of breakpoints, and
199 (optionally) initial estimates of their locations, are provided. To determine the number of these points
200 in each score, we fit each segmented regression model multiple times with one or more breakpoints
201 and selected the model with the lowest Bayesian Information Criterion (BIC)[40,45]. The number of
202 breakpoints was estimated separately for each domain score and gender. The algorithm converged on
203 consistent breakpoint locations regardless of whether initial estimates were provided (from visual
204 inspection of local regression curves, shown in Fig S1), or not. To confirm that a model with one or
205 more breakpoints predicted the data better than a linear model, the Davies' test [46] was used to
206 determine whether there was a statistically significant change in slope. The estimated breakpoint
207 location was taken as the age at which there was peak performance in all regression models except
208 for two cases. First, in men's verbal scores, in which there were two breakpoints and the breakpoint
209 with the highest score was used as peak age. Second, in women's reasoning scores, in which the
210 highest score was at the lower boundary of our age range. Slopes of the increasing and decreasing
211 segments, as well as the middle segment for men's verbal scores, were obtained using the 'slope'
212 function of the 'segmented' package, and 95% confidence intervals (CIs) were calculated for peak age,
213 score at peak age, and all slopes.

214 Differences in these parameters between men and women were analyzed by bootstrapping
215 with 10,000 replications the difference of the estimated parameter values from models that were
216 separately estimated for men and women; that is, the sex-by- peak age, peak score, and slope
217 interactions were evaluated using randomized bootstrapping, because estimation of breakpoint
218 parameters in our segmented regression did not allow for interaction with other variables. On each
219 bootstrap iteration, a random sample of 13,444 men were selected in order to match the female

sample size. To determine whether these values differed significantly between genders, the lower and upper 2.5% quantiles of the bootstrapped difference values were produced; if these bounds included zero, then it could be interpreted as no significant difference between the genders.

In segmented models where multiple breakpoints were deemed a better solution than a single point as determined using BIC, the increasing or decreasing portion of the curve (i.e., the data to the left or right of the “peak”) was characterized by two increasing or decreasing linear segments with different slopes (as can be seen in Fig 2C, women’s reasoning scores). In order to compare slopes between the genders in these cases, bootstrapping was conducted by fitting the segmented model, then calculating the average slope to the left (in the case of men’s verbal scores) or right (in the case of women’s reasoning scores) of the peak. The rest of the bootstrapping parameters were kept the same as described above, with 95% confidence intervals of the difference values being used to detect significant differences between genders.

Post-hoc analyses

Given the differences in demographic variables between the genders, a second set of analyses were run in samples matched across genders for all demographic variables included in Table 1. These follow-up analyses were performed because, although the genders do realistically differ on measures such as anxiety and sleep, such factors are known to affect cognition and may contribute variance to the domain scores. Because it is difficult to account for the strong collinearity between gender and our demographic variables, matching the samples on these variables allowed us to examine gender differences when controlling for differences in socio-demographic variables. Descriptive information for these two new samples is summarized in Table S1, and histograms of domain scores are shown in

241 Fig S2. Local regression curves are shown in Fig S3. The same set of analyses were performed as
242 outlined in the section above.

243 Results

244 Demographics

245 As reported in Table 1, a total of 13,444 women and 32,335 men completed the relevant
246 demographic questionnaire items and all 12 cognitive tests. Women and men differed on several
247 demographic factors, but not for age, education, exercise, and number of siblings (all p s > 0.05). While
248 all significant p -values were $\leq .003$, the largest effect sizes were seen in hours of sleep (Cohen's d =
249 0.10), units of caffeine per day (Cohen's d = -0.19), anxiety level (Wilcoxon's r = 0.15), and technical
250 savviness (Cramer's V = 0.24).

251 Cognitive domain scores – unmatched samples

252 Short-term memory

253 STM scores for each gender were submitted to segmented regression with age as the sole
254 predictor, entered as a continuous variable. In both cases, the breakpoint corresponded to a peak;
255 that is, a transition from increasing to decreasing performance with increasing age (Fig 2A). Results
256 are reported in Table 2, and slopes with 95% CIs bounds that did not include zero were interpreted as
257 a significant effect of age.

258

259 **Fig 2. Regression lines for STM, Verbal, and Reasoning scores across the lifespan, ranging from 12 to**
260 **69 years of age.** 95% simultaneous confidence bands are shown in translucent colour around the line,
261 and 95% confidence intervals for peak age are shown in translucent rectangles.

262

263 **Table 2: Segmented regression parameter estimates for age, from regression models estimated for**
264 **each composite score, for models estimated with N = 45,779**

Score	Gender	Coef	SE	<i>t</i>	<i>p</i>
STM	Women	0.03	0.01	3.83	< .001
	ΔAge	-0.05			
	Men	0.04	0.01	6.48	< .001
	ΔAge	-0.07			
Verbal	Women	0.04	0.01	8.16	< .001
	ΔAge	-0.05			
	Men	0.10	0.01	8.44	< .001
	ΔAge1	-0.09			
	ΔAge2	-0.02			
Reasoning	Women	-0.01	0.001	-9.62	< .001
	ΔAge	-0.01			
	Men	0.003	0.004	0.73	.468
	ΔAge	-0.03			

265 Note: *p*-values for delta parameters measured by Davies' test

266

267 A model with one breakpoint was found to best estimate women's memory scores. The peak
268 in women's STM scores occurred at age 20.47 [95% CI = 19.39, 21.55], with a score of 0.02 [95% CI = -
269 0.01, 0.05]. The slopes of the segments to the left and right of the breakpoint were 0.03 [95% CI =
270 0.02, 0.05] and -0.02 [95% CI = -0.02, -0.02], respectively, indicating that age was a significant
271 predictor of STM performance in these age ranges; specifically, increasing age was associated with
272 increasing scores up to the age of 20 years, after which it was associated with decreasing
273 performance. Davies' test for a change in slope was significant ($p < .001$), indicating that the linear
274 relationship changed at the breakpoint, as can be seen in Fig 2A.

275 Men's memory scores were also best estimated by a segmented model with one breakpoint.
276 The peak in men's STM score occurred at age 20.48 [95% CI = 19.85, 21.12], with a score of 0.30 [95%
277 CI = 0.29, 0.32]. Slope of the increasing segment was 0.04 [95% CI = 0.03, 0.05], and slope of the

decreasing segment was -0.024 [95% CI = -0.03, -0.02], showing a significant effect of age on STM score in men. The change in slope was significant, as measured by the Davies' test ($p < .001$).

In order to determine whether women and men differed in peak age, peak score, or increasing and decreasing slopes, 95% quantiles of bootstrapped difference values were calculated for each parameter. As can be seen in Table 3, there was no significant difference in the age at which women and men peaked in STM performance. However, men reached a significantly higher overall score than women at their peak ages, a difference of 0.28 standard deviations. When comparing how STM scores increased leading up to peak age and how quickly they declined afterward, women and men did not differ significantly.

Table 3: Comparisons between genders on key measures of cognitive performance over the lifetime

Score	Measure	Women [95% CI]	Men [95% CI]	Difference [95% CI]
STM	Peak age	20.471 [19.388, 21.553]	20.482 [19.845, 21.119]	-0.011 [-4.696, 3.441]
	Peak score	0.021 [-0.007, 0.049]	0.304 [0.286, 0.323]	-0.283 [-0.331, -0.219]
	Increasing slope	0.032 [0.015, 0.048]	0.042 [0.029, 0.054]	0.010 [-0.071, 0.036]
	Decreasing slope	-0.023 [-0.025, -0.021]	-0.024 [-0.025, -0.023]	0.001 [-0.002, 0.005]
Verbal	Peak age	23.206 [21.996, 24.418]	39.202 [35.986, 42.428]	-15.996 [-26.362, -3.858]
	Peak score	0.067 [0.033, 0.101]	0.116 [0.074, 0.157]	-0.049 [-0.145, -0.002]
	Increasing slope	0.042 [0.032, 0.052]	0.014 [0.007, 0.027] ^a	0.028 [0.012, 0.176]
	Decreasing slope	-0.006 [-0.008, -0.004]	-0.013 [-0.017, -0.009]	0.007 [-0.001, 0.019]
Reasoning	Peak age	12	23.512 [22.248, 24.777]	-11.512 [-16.956, -4.221]
	Peak score	0.208 [0.168, 0.249]	0.196 [0.163, 0.228]	0.012 [-0.136, 0.046]
	Increasing slope	—	0.003 [-0.004, 0.009]	—
	Decreasing slope	-0.019 [-0.021, -0.018] ^a	-0.027 [-0.029, -0.026]	0.008 [0.004, 0.012]

Note: Values are missing for women's reasoning increasing slope as both segments were negative

^a Combined slope across two segments is reported. Slopes of the individual segments are reported in-text.

292

293 **Verbal abilities**

294 Results of segmented regression of verbal scores are also summarized in Table 2. A model with
 295 one breakpoint was again found to best estimate women's verbal scores. The peak in women's verbal
 296 scores occurred at age 23.21 [95% CI = 22.00, 24.42] with a score of 0.07 [95% CI = 0.03, 0.10], as can
 297 be seen in Fig 2B. Slope of the increasing segment was 0.04 [95% CI = 0.03, 0.05], and slope of the
 298 decreasing segment was -0.006 [95% CI = -0.008, -0.004], showing a significant relationship between
 299 age and verbal abilities. Davies' test for a change in slope was significant ($p < .001$), indicating that the
 300 linear relationship changed at the breakpoint.

301 Men's verbal scores were best estimated by a segmented model with two breakpoints. As can
 302 be seen in Fig 2B, men first had a breakpoint at age 18.85, at which point the rate at which scores
 303 were increasing, slowed. The peak in men's verbal score occurred at age 39.20 [95% CI = 35.99,
 304 42.42], with a score of 0.12 [95% CI = 0.07, 0.16]. Slope of the initial increasing segment was 0.10
 305 [95% CI = 0.07, 0.12], the slope of the second increasing segment was 0.005 [95% CI = 0.002, 0.007]
 306 and slope of the decreasing segment was -0.01 [95% CI = -0.02, -0.009], indicating a significant
 307 relationship between age and verbal abilities in all three sections. The change in slope was significant,
 308 as measured by the Davies' test ($p < .001$).

309 As summarized in Table 3, men reached a peak in verbal abilities at a significantly later age
 310 than women. Men also had significantly higher scores at peak age, with a difference of 0.05 standard
 311 deviations. When comparing how scores increased up to peak age, women's scores improved at a
 312 faster rate than men's, however there was no difference when comparing the rate of decline from
 313 peak age to age 69.

Reasoning

A model with one breakpoint was again found to best estimate women's reasoning scores. However, this breakpoint occurred at age 38.24 years, and indicated a transition from a gradual to steeper decline: scores declined with a slope of -0.014 [95% CI = -0.017, -0.011] from age 12 to age 38.24, at which point the negative slope increased to -0.029 [95% CI = -0.034, -0.023]. Davies' test for a change in slope was significant ($p < .001$), indicating that the linear relationship changed. As can be seen in Fig 2C, the highest predicted scores for women occurred at age 12 with a score of 0.21 [95% CI = 0.12, 0.25]. However, because this is the cut-off age of our sample, it is not possible to determine whether this is indeed a true peak, or if scores are higher at earlier ages.

Men's reasoning scores were best estimated by a segmented model with one breakpoint. The breakpoint in men's reasoning score occurred at age 23.51 (95% CI = 22.25, 24.78), with a score of 0.20 [95% CI = 0.16, 0.23]. The change in slope was significant, as measured by the Davies' test ($p < .001$), however the slope of the initial segment was 0.002 [95% CI = -0.004, 0.010], and slope of the decreasing segment was -0.027 [95% CI = -0.029, -0.026], indicating that only the second segment showed a significant effect of age. Similar to women, this suggests that we did not capture a developmental increase in reasoning abilities within the current sample, and it is possible that the true peak occurs earlier than age 12.

Because we do not have a reliable measure of peak age in either gender, we compared between genders the age at which reasoning scores began to decline. In this sample, women began to decline in reasoning abilities significantly earlier than men, however reasoning scores at that age did not differ between genders (Table 3). Because women did not show an increase in reasoning scores

within our age range, we could not compare men and women on this measure. However, when comparing how scores declined after peak age, men declined significantly faster than women.

Cognitive domain scores – matched samples

In the matched samples, the general pattern of results was similar to the unmatched samples in both STM and reasoning. However, the previously found gender differences in verbal abilities disappeared in the matched sample. Segmented regression lines for the matched sample are shown in Fig 3.

Fig 3. Regression lines for STM, Verbal, and Reasoning scores across the lifespan, ranging from 12 to 69 years of age, in a socio-demographically matched sample. 95% simultaneous confidence bands are shown in translucent colour around the line, and 95% confidence intervals for peak age are shown in translucent rectangles.

Short-term memory

Results of the segmented regression for STM scores of both genders in the socio-demographically matched sample are reported in Table 4. Both women and men again showed a significant change in slope as measured by the Davies' test ($p < .001$ for both genders). As can be seen in Table 5, after matching women and men on sociodemographic variables, no significant differences were found in the age at which women and men peaked in STM, nor in the slopes of the increase and decrease in scores surrounding peak age. However, men still reached a higher overall score than women at their peak ages by a standard deviation of 0.21.

Table 4: Segmented regression parameter estimates for age, from regression models estimated for each composite score in a demographically matched sample

Score	Gender	Coef	SE	<i>t</i>	<i>p</i>
-------	--------	------	----	----------	----------

STM	Women	0.04	0.01	4.10	< .001
	ΔAge	-0.06			
	Men	0.05	0.01	3.61	< .001
	ΔAge	-0.07			
Verbal	Women	0.05	0.01	7.58	< .001
	ΔAge1	-0.13			
	ΔAge2	-0.03			
	Men	0.14	0.03	5.36	< .001
	ΔAge1	-0.13			
	ΔAge2	-0.02			
Reasoning	Women	-0.01	0.001	-8.83	< .001
	ΔAge	-0.02			
	Men	0.01	0.01	1.10	.272
	ΔAge	-0.04			

Table 5: Comparisons between genders matched on socio-demographic variables

Score	Measure	Women [95% CI]	Men [95% CI]	Difference [95% CI]
STM	Peak age	20.423 [19.365, 21.482]	19.651 [18.609, 20.693]	0.772 [-2.089, 4.319]
	Peak score	0.046 [-0.009, 0.101]	0.259 [0.187, 0.330]	-0.213 [-2.63, -0.159]
	Increase	0.036 [0.019, 0.053]	0.049 [0.022, 0.075]	-0.013 [-0.132, 0.028]
	Decrease	-0.023 [-0.025, -0.022]	-0.025 [-0.027, -0.023]	0.002 [-0.001, 0.005]
Verbal	Peak age	24.889 [22.263, 27.515]	28.423 [25.330, 31.517]	-3.534 [-20.490, 6.098]
	Peak score	0.071 [0.033, 0.108]	0.104 [0.050, 0.158]	-0.033 [-0.091, 0.019]
	Increase	0.035 [0.016, 0.048] ^a	0.022 [0.006, 0.045] ^a	0.013 [-0.012, 0.036]
	Decrease	-0.006 [-0.008, -0.003]	-0.008 [-0.011, -0.005]	0.002 [-0.003, 0.014]
Reasoning	Peak age	12	19.623 [17.804, 21.542]	-7.623 [-12.816, -2.226]
	Peak score	0.223 [0.187, 0.271]	0.131 [0.060, 0.201]	0.092 [-0.047, 0.151]
	Increase	–	0.015 [-0.012, 0.041]	–
	Decrease	-0.020 [-0.021, -0.018] ^a	-0.025 [-0.027, -0.023]	0.005 [0.003, 0.008]

^a Combined slope across two segments is reported.

364

365 **Verbal abilities**

366 Both women and men again showed a significant change in slope as measured by the Davies'
367 test ($p < .001$ in all tests). There were no significant differences in the age at which women and men
368 peaked in verbal abilities, scores at peak age, nor in the slopes of the increase and decrease in scores
369 surrounding peak age (Table 5).

370 **Reasoning**

371 Reasoning scores in our sample of women began to decrease at a significantly earlier age than
372 men, however scores at that age did not differ between genders. While we did not capture an
373 increase in reasoning abilities in either gender in our sample, reasoning scores decreased significantly
374 faster in men than women (Table 5).

375 **Discussion**

376 After creating three cognitive domain scores from the 12 CBS tests based on their underlying
377 factor structure, we were able to replicate previous findings suggesting that not all cognitive domains
378 develop and decline in the same way. Specifically, STM increased rapidly from age 12 to the early 20s,
379 at which point it decreased at a steady rate until age 69, the upper limit of our sample's age range.
380 Verbal abilities also peaked in early adulthood, while reasoning did not show a clear peak in scores,
381 instead being characterized by either a decline from age 12, or a plateau followed by a decline. These
382 results were consistent with previous studies showing that cognition is not a unitary concept, and
383 different cognitive abilities have separable developmental trajectories [11,12]. However, they extend
384 the results of those studies in several important ways:

385 First, when examining the progression of STM, verbal abilities, and reasoning in men and
386 women separately, all three cognitive domains also showed unique gender differences. Although men
387 and women peaked in STM performance at the same age, men reached a slightly higher score than
388 women. In verbal abilities, women reached their peak faster and earlier, but men again reached
389 higher scores. While women's reasoning began to decline earlier than men's, men declined at a faster
390 rate. These results extend what is known from previous gender research. For example, there is
391 evidence that men lose grey matter volume more rapidly with age than women, especially in fronto-
392 temporal regions [47]; this in turn may lead to a faster rate of decline in cognitive function, fitting the
393 pattern we see here in the reasoning domain. In contrast, women are thought to have better verbal
394 processing than men; however we see the opposite here, with men reaching a higher peak score than
395 women. One possible explanation for this discrepancy could be the age at which verbal abilities are
396 tested. For example, Burton and colleagues (19) tested a sample of university students, which is
397 common in Psychology research. Looking at the pattern of verbal abilities in men and women in the
398 current unmatched sample, women seem to outperform men at age 23, which, if we were to only
399 examine individuals around this age, may lead to the erroneous conclusion that women have superior
400 verbal abilities. Similarly, men are frequently reported to be better at mental rotation than women
401 [19], a test included in our reasoning domain. In our sample, we found that peak reasoning scores did
402 not differ between genders, but women declined much earlier than men. Again, comparing genders
403 within a limited age range would have led to the erroneous conclusion that men outperform women
404 in this domain, when in reality it is a difference in trajectory of reasoning abilities. The present results
405 underline the need to take the progression of cognitive abilities across the lifespan into account when
406 studying gender differences.

407 While these results are presumed to be reflective of the cognitive performance of the
408 population, they are complicated by the differences in socio-demographic factors. Several of the
409 factors showing the largest differences, such as sleep and anxiety, have known effects on cognitive
410 function [38], making it difficult to determine what is driving the observed gender differences. Further
411 complicating the matter is that because these socio-demographic factors are gender-dependent, it is
412 not possible to include them in the model due to issues with multicollinearity. By matching men and
413 women on these factors, however, we were able to limit their effect on the data as much as possible
414 and the results show that this greatly reduced the differences in cognitive performance and aging. Of
415 course, there are numerous demographic factors that we did not control for, and it is impossible to
416 truly capture all of them. Additionally, there are some socio-demographic differences that may have
417 biological underpinnings. For example, depression is more prevalent in women, and this may be in
418 part due to the presence of sex-specific forms such as premenstrual dysphoric disorder [48]. It is
419 therefore difficult to disentangle the environment from biological sex differences, however
420 accounting for these differences, regardless of their origin, is necessary for describing gender
421 differences in cognition alone.

422 As noted above, controlling for gender-specific differences in socio-demographic factors
423 greatly reduced the differences in cognitive performance and aging. While men still reached a higher
424 peak score than women, the difference between peak scores decreased from .28 SDs to .21 SDs.
425 Notably, all differences in verbal abilities disappeared, with women appearing more like men in the
426 nature of their progress, having no differences in peak age, score, or the rate of improvement or
427 decline. However, although the gender gap was smaller (or absent) in the matched sample, this does
428 not mean that differences in the unmatched sample should be ignored. While they may not

necessarily be inherent to biology, environmental influences are a part of life, and they do drive gender differences in cognitive abilities. Thus, it is reasonable to conclude that gender differences in cognitive abilities, based on biological sex alone, are minimal; however, there are notable effects of environmental factors that in turn drive gender differences in cognition.

One large area of disparity that remained even when controlling for environmental factors was with respect to the age at which reasoning abilities began to decline. Women declined significantly earlier than men, even when controlling for demographic factors. We were also not able to capture a reliable measure of the age at which reasoning abilities peak in either gender. In women, in both the full and matched samples, scores declined from 12 years of age, onward. This could be because 12 is the age at which women's reasoning abilities do indeed peak. However, it is also possible that women peak earlier, but due to lack of data we were unable to determine the true peak from the current sample. Similarly, both unmatched and matched samples of men showed a plateau in reasoning scores, rather than an increase, until the point at which they began to decline. There are several possible explanations here. First, it is possible that men do peak in early adulthood, somewhere between 18 and 24 years of age, but the increase in reasoning abilities was not captured due to too small a sample size or noisy data. Second, they could follow a similar trajectory to women, with a slow decline before a steeper one, again not captured due to a lack of data. Because our sample of men was very large (over 32,000 in the unmatched sample), it is unlikely that either of these options are the case. Third, this plateau could be a true peak in reasoning, lasting several years, before beginning to decline. Previous research does suggest that reasoning abilities are relatively mature by age 12 [6,9], and another large-scale study has shown that by age 18, reasoning abilities have begun to decline [12]. Thus, although it is not possible to confirm that decline begins around age 12 in the

451 current sample of women, the data follow a pattern that fits previous research and supports this
452 claim.

453 The results presented here offer some insight into how to tailor interventions for cognitive
454 decline appropriately for each gender. For example, women are known to experience more anxiety
455 than men [32], a fact reflected in the current sample. Anxiety is known to correlate negatively with
456 working memory [49]. Thus, to improve working memory, or protect against its decline, therapies
457 should perhaps focus on reducing anxiety in everyone, with a targeted focus on women. Another
458 example is substance abuse, which is more prevalent in men [34]. Because substance abuse
459 negatively affects cognition [30], especially with respect to aging [50], a focused campaign aimed to
460 reduce drug and alcohol consumption in men may yield a slowing in cognitive decline at the male
461 population level. These gender-focused interventions can be combined with other treatments known
462 to provide protection from cognitive decline, such as frequent exercise [51] for a well-rounded
463 defence against cognitive aging.

464 **Conclusions**

465 By examining a sample of over 45,000 individuals, ranging from 12 to 69 years of age, we
466 showed how different cognitive abilities vary across the lifespan. Each domain had a unique
467 relationship with age, demonstrating that not all cognitive processes follow the same pattern.
468 Importantly, we found differences in the way women and men cognitively age, and showed that these
469 disparities are reduced when controlling for socio-demographics such as sleep and anxiety.
470 Nevertheless, some gender differences remained, supporting the notion that gender differences in
471 cognition are likely guided by a complex interplay of both biology and environment.

472

References

1. United Nations Department of Economic and Social Affairs. World Population Prospects 2019: Highlights. 2019.
2. Anderson P, Anderson V, Northam E, Taylor HG. Standardization of the Contingency Naming Test (CNT) for school-age children: a measure of reactive flexibility. Clin Neuropsychol Assess. 2000;1: 247–273. Available: [http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Standardization+of+the+Contingency+NamingTest+\(CNT\)+for+school+aged+children:+A+measure+of+reactive+flexibility#0](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Standardization+of+the+Contingency+NamingTest+(CNT)+for+school+aged+children:+A+measure+of+reactive+flexibility#0)
3. Karapetsas AB, Vlachos FM. Sex and Handedness in Development of Visuomotor Skills. Percept Mot Skills. 1997;85: 131–140. doi:10.2466/pms.85.5.131-140
4. Krikorian R, Bartok JA. Developmental Data for the Porteus Maze Test. Clin Neuropsychol. 1998;12: 305–310. doi:10.1076/clin.12.3.305.1984
5. Feng J, Spence I, Pratt J. Playing an Action Video Game in Reduces Gender Differences Spatial Cognition. Psychol Sci. 2007;18: 850–855.
6. Anderson P. Assessment and Development of Executive Function (EF) During Childhood. Child Neuropsychol. 2002;8: 71–82. doi:10.1076/chin.8.2.71.8724
7. Rizeq J, Flora DB, Toplak ME. Changing relations among cognitive abilities across development: implications for measurement and research. Clin Neuropsychol. 2017;31: 1353–1374. doi:10.1080/13854046.2017.1317034
8. Diamond A. Executive Function. Annu Rev Psychol. 2013;64: 135–168. doi:10.1016/B978-0-12-385157-4.01147-7
9. Anderson V, Anderson P, Northam E, Jacobs R, Catroppa C. Developmental Neuropsychology

- 495 Differential Development of Attention and Executive Functions in 3- to 12-Year-Old Finnish
496 Children. *Dev Neuropsychol.* 2001;20: 385–406. doi:10.1207/S15326942DN2001
- 497 10. Lee K, Bull R, Ho R. Developmental Changes in Executive Functioning. *Child Dev.* 2013;84: 1933–
498 1953. doi:10.1111/cdev.12096
- 499 11. Hartshorne JK, Germine LT. When Does Cognitive Functioning Peak? The Asynchronous Rise
500 and Fall of Different Cognitive Abilities Across the Life Span. *Psychol Sci.* 2015;26: 433–443.
501 doi:10.1177/0956797614567339
- 502 12. Salthouse TA. When does age-related cognitive decline begin? *Neurobiol Aging.* 2009;30: 507–
503 514. doi:10.1016/j.neurobiolaging.2008.09.023
- 504 13. Vecchi T, Girelli L. Gender differences in visuo-spatial processing: The importance of
505 distinguishing between passive storage and active manipulation. *Acta Psychol (Amst).* 1998;99:
506 1–16. doi:10.1016/S0001-6918(97)00052-8
- 507 14. Tzuriel D, Egozi G. Gender Differences in Spatial Ability of Young Children : The Effects of
508 Training and Processing Strategies. *Child Dev.* 2010;81: 1417–1430.
- 509 15. Krinzinger H, Wood G, Willmes K. What accounts for individual and gender differences in the
510 multi-digit number processing of primary school children? *Zeitschrift fur Psychol / J Psychol.*
511 2012;220: 78–89. doi:10.1027/2151-2604/a000099
- 512 16. Sokolowski HM, Hawes Z, Lyons IM. What explains sex differences in math anxiety? A closer
513 look at the role of spatial processing. *Cognition.* 2019;182: 193–212.
514 doi:10.1016/j.cognition.2018.10.005
- 515 17. Scheuringer A, Wittig R, Pletzer B. Sex differences in verbal fluency: the role of strategies and
516 instructions. *Cogn Process.* 2017;18: 407–417. doi:10.1007/s10339-017-0801-1

- 517 18. Scheuringer A, Pletzer B. Sex differences and menstrual cycle dependent changes in cognitive
518 strategies during spatial navigation and verbal fluency. *Front Psychol.* 2017;8: 1–12.
519 doi:10.3389/fpsyg.2017.00381
- 520 19. Burton LA, Henninger D, Hafetz J. Gender differences in relations of mental rotation, verbal
521 fluency, and SAT scores to finger length ratios as hormonal indexes. *Dev Neuropsychol.*
522 2005;28: 493–505. doi:10.1207/s15326942dn2801_3
- 523 20. Griksiene R, Ruksenas O. Effects of hormonal contraceptives on mental rotation and verbal
524 fluency. *Psychoneuroendocrinology.* 2011;36: 1239–1248. doi:10.1016/j.psyneuen.2011.03.001
- 525 21. Wu Y, Zhang D, Pang Z, Oksuzyan A, Jiang W, Wang S, et al. Gender-specific patterns in age-
526 related decline in general health among Danish and Chinese: A cross-national comparative
527 study. *Geriatr Gerontol Int.* 2012;12: 431–439. doi:10.1111/j.1447-0594.2011.00784.x
- 528 22. Cornelis MC, Wang Y, Holland T, Agarwal P, Weintraub S, Morris MC. Age and cognitive decline
529 in the UK Biobank. *PLoS One.* 2019;14: 1–16. doi:10.1371/journal.pone.0213948
- 530 23. Lipnicki DM, Crawford JD, Dutta R, Thalamuthu A, Kochan NA, Andrews G, et al. Age-related
531 cognitive decline and associations with sex, education and apolipoprotein E genotype across
532 ethnocultural groups and geographic regions: a collaborative cohort study. *PLoS Med.* 2017;14.
533 24. Laing KR, Mitchell D, Wersching H, Czira ME, Berger K, Baune BT. Brain-derived neurotrophic
534 factor (BDNF) gene: A gender-specific role in cognitive function during normal cognitive aging
535 of the MEMO-Study? *Age (Omaha).* 2012;34: 1011–1022. doi:10.1007/s11357-011-9275-8
- 536 25. Malpetti M, Ballarini T, Presotto L, Garibotto V, Tettamanti M, Perani D. Gender differences in
537 healthy aging and Alzheimer’s Dementia: A 18F-FDG-PET study of brain and cognitive reserve.
538 *Hum Brain Mapp.* 2017;38: 4212–4227. doi:10.1002/hbm.23659

- 539 26. Miller DI, Halpern DF. The new science of cognitive sex differences. Trends Cogn Sci. 2014;18:
540 37–45. doi:10.1016/j.tics.2013.10.011
- 541 27. Blums A, Belsky J, Grimm K, Chen Z. Building Links Between Early Socioeconomic Status,
542 Cognitive Ability, and Math and Science Achievement. J Cogn Dev. 2017;18: 16–40.
543 doi:10.1080/15248372.2016.1228652
- 544 28. Lubinski D. Cognitive epidemiology: With emphasis on untangling cognitive ability and
545 socioeconomic status. Intelligence. 2009;37: 625–633. doi:10.1016/j.intell.2009.09.001
- 546 29. Zaremba D, Schulze Kalthoff I, Förster K, Redlich R, Grotegerd D, Leehr EJ, et al. The effects of
547 processing speed on memory impairment in patients with major depressive disorder. Prog
548 Neuro-Psychopharmacology Biol Psychiatry. 2019;92: 494–500.
549 doi:10.1016/j.pnpbp.2019.02.015
- 550 30. Crego A, Holguín SR, Parada M, Mota N, Corral M, Cadaveira F. Binge drinking affects
551 attentional and visual working memory processing in young university students. Alcohol Clin
552 Exp Res. 2009;33: 1870–1879. doi:10.1111/j.1530-0277.2009.01025.x
- 553 31. Hampshire A, Highfield RR, Parkin BL, Owen AM. Fractionating Human Intelligence. Neuron.
554 2012;76: 1225–1237. doi:10.1016/j.neuron.2012.06.022
- 555 32. McLean CP, Asnaani A, Litz BT, Hofmann SG. Gender differences in anxiety disorders:
556 Prevalence, course of illness, comorbidity and burden of illness. J Psychiatr Res. 2011;45: 1027–
557 1035. doi:10.1016/j.jpsychires.2011.03.006
- 558 33. Parker G, Brotchie H. Gender differences in depression. Int Rev Psychiatry. 2010;22: 429–436.
559 doi:10.3109/09540261.2010.492391
- 560 34. Compton W, Thomas YF, Stinson FS, Grant BF. Prevalence, Correlates, Disability, and

- 561 Comorbidity of DSM-IV Drug Abuse and Dependence in the United States. Arch Gen Psychiatry.
562 2007;64: 566–576. Available:
563 [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2707116&tool=pmcentrez&rende](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2707116&tool=pmcentrez&rendetype=abstract)
564 [rtype=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2707116&tool=pmcentrez&rendetype=abstract)
- 565 35. Grant BF, Chou SP, Saha TD, Pickering RP, Kerridge BT, Ruan WJ, et al. Prevalence of 12-month
566 alcohol use, high-risk drinking, and DSM-IV alcohol use disorder in the United States, 2001-
567 2002 to 2012-2013: Results from the National Epidemiologic Survey on Alcohol and Related
568 Conditions. JAMA Psychiatry. 2017;74: 911–923. doi:10.1001/jamapsychiatry.2017.2161
- 569 36. Bratberg GH, Wilsnack SC, Wilsnack R, Håvås Haugland S, Krokstad S, Sund ER, et al. Gender
570 differences and gender convergence in alcohol use over the past three decades (1984-2008),
571 the HUNT Study, Norway. BMC Public Health. 2016;16: 1–12. doi:10.1186/s12889-016-3384-3
- 572 37. Wesnes KA, Brooker H, Ballard C, Mccambridge L, Stenton R, Corbett A. Utility, reliability,
573 sensitivity and validity of an online test system designed to monitor changes in cognitive
574 function in clinical trials. Int J Geriatr Psychiatry. 2017. doi:10.1002/gps.4659
- 575 38. Wild CJ, Nichols ES, Battista ME, Stojanoski B, Owen AM. Dissociable effect of self-reported
576 daily sleep duration on high-level cognitive abilities. Sleep. 2018;in press: 1–11.
577 doi:10.1093/sleep/zsy182
- 578 39. Owen AM, Hampshire A, Grahn JA, Stenton R, Dajani S, Burns AS, et al. Putting brain training to
579 the test. Nature. 2010;465: 775–8. doi:10.1038/nature09042
- 580 40. Muggeo VMR. segmented: An R package to fit regression models with broken-line
581 relationships. R news. 2008;8: 20–25. doi:<http://dx.doi.org/10.1192/bjp.195.1.A6>
- 582 41. Ho DE, Imai K, King G, Stuart EA. MatchIt : Nonparametric Preprocessing for. J Stat Softw.

- 2011;42: 1–28. Available: <http://www.jstatsoft.org/v42/i08>
42. Canty A, Ripley B. boot: Bootstrap R(S-Plus) Functions. R Packag version 13-20. 2019.
43. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2016.
44. Muggeo VMR. Estimating regression models with unknown break-points. Stat Med. 2003;22: 3055–71. doi:10.1002/sim.1545
45. Tiwari RC, Cronin KA, Davis W, Feuer EJ, Yu B, Chib S. Bayesian model selection for join point regression with application to age-adjusted cancer rates. J R Stat Soc Ser C Appl Stat. 2005;54: 919–939. doi:10.1111/j.1467-9876.2005.00518.x
46. Davies RB. Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternatives Published by : Oxford University Press on behalf of Biometrika Trust Stable URL : <http://www.jstor.org/stable/2336019> Your use of the JSTOR archive indica. Biometrika. 2002;89: 484–489.
47. Kryspin-Exner I, Lamplmayr E, Felnhöfer A. Geropsychology: The gender gap in human aging-a mini-review. Gerontology. 2011;57: 539–548. doi:10.1159/000323154
48. Albert PR. Why is depression more prevalent in women? J Psychiatry Neurosci. 2015;40: 219–221. doi:10.1503/jpn.150205
49. Moran TP. Anxiety and Working Memory Capacity: A Meta-Analysis and Narrative Review. Psychol Bull. 2016;142. doi:10.1037/bul0000051.supp
50. Woods AJ, Porges EC, Bryant VE, Seider T, Gongvatana A, Kahler CW, et al. Current Heavy Alcohol Consumption is Associated with Greater Cognitive Impairment in Older Adults. Alcohol Clin Exp Res. 2016;40: 2435–2444. doi:10.1111/acer.13211
51. Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, et al. Exercise training

605 increases size of hippocampus and improves memory. Proc Natl Acad Sci. 2011;108: 3017–
606 3022. doi:10.1073/pnas.1015950108

607

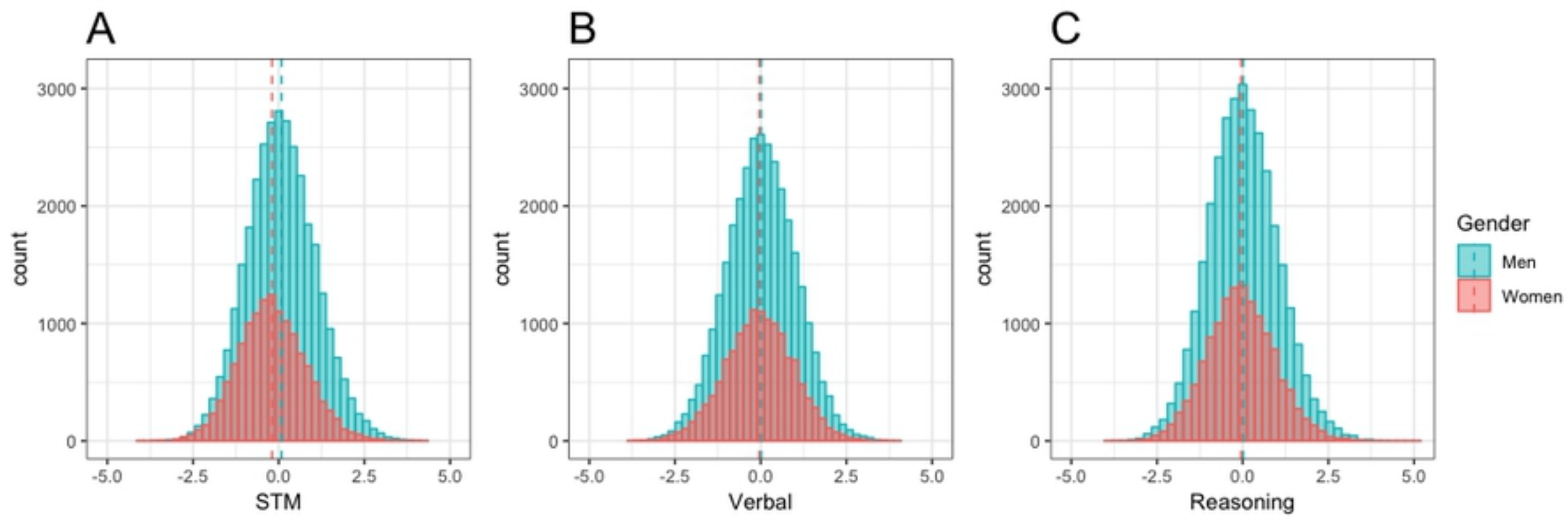
608 **Supporting information**

609 **S1 Fig. Local regression curves for STM, Verbal, and Reasoning scores across the lifespan, ranging**
610 **from 12 to 69 years of age.** 95% confidence bands are shown in translucent colour around the line.

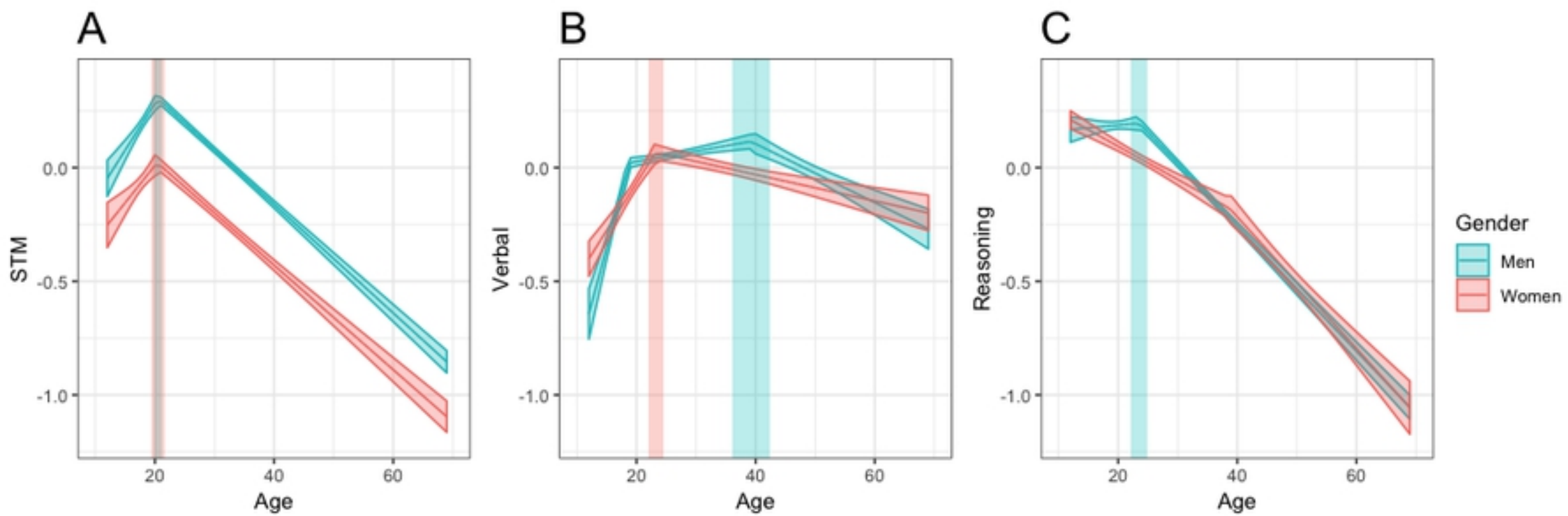
611 **S2 Fig. Histograms of cognitive domain scores by gender, in a socio-demographically matched**
612 **sample.** Dashed lines indicate mean.

613 **S3 Fig. Local regression curves for STM, Verbal, and Reasoning scores across the lifespan, ranging**
614 **from 12 to 69 years of age, in a socio-demographically matched sample.** 95% simultaneous
615 confidence bands are shown in translucent colour around the line.

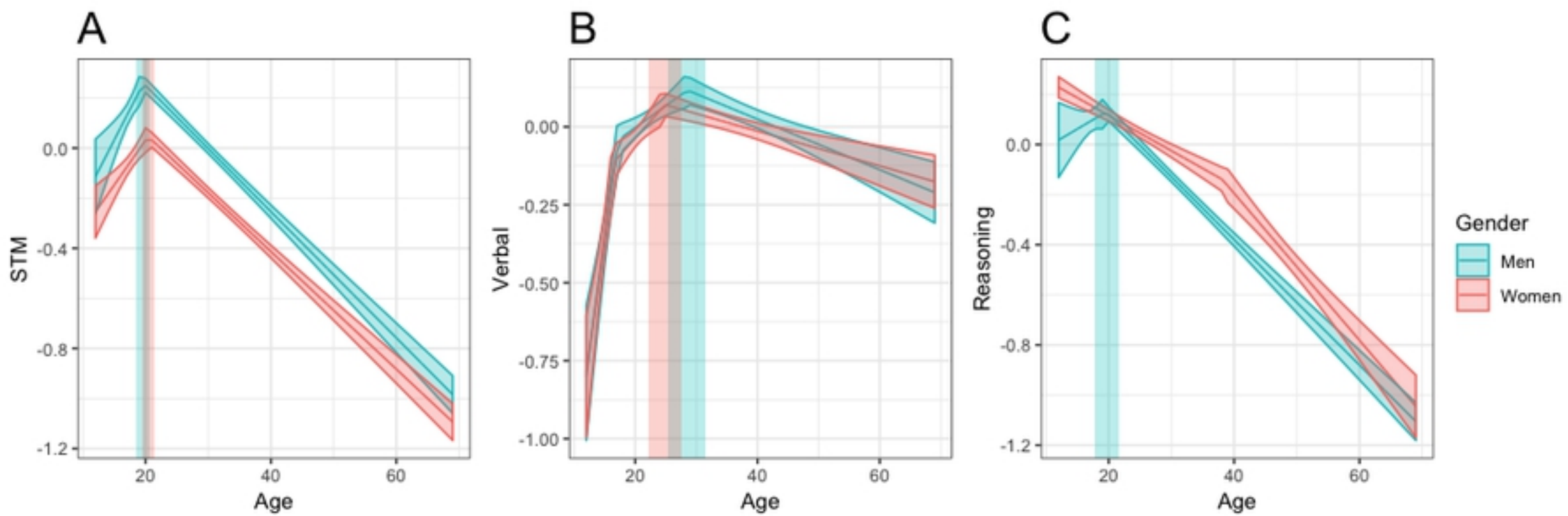
616 **S1 Table. Comparison of demographic variables across women and men in a matched sample.**



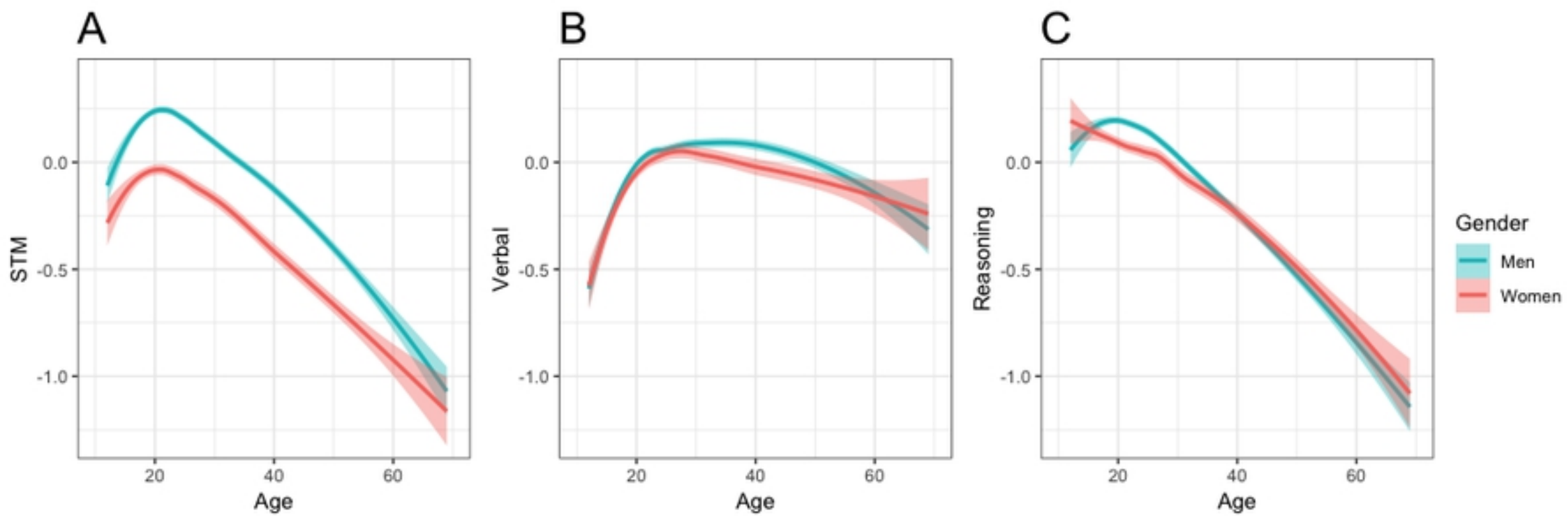
Figure



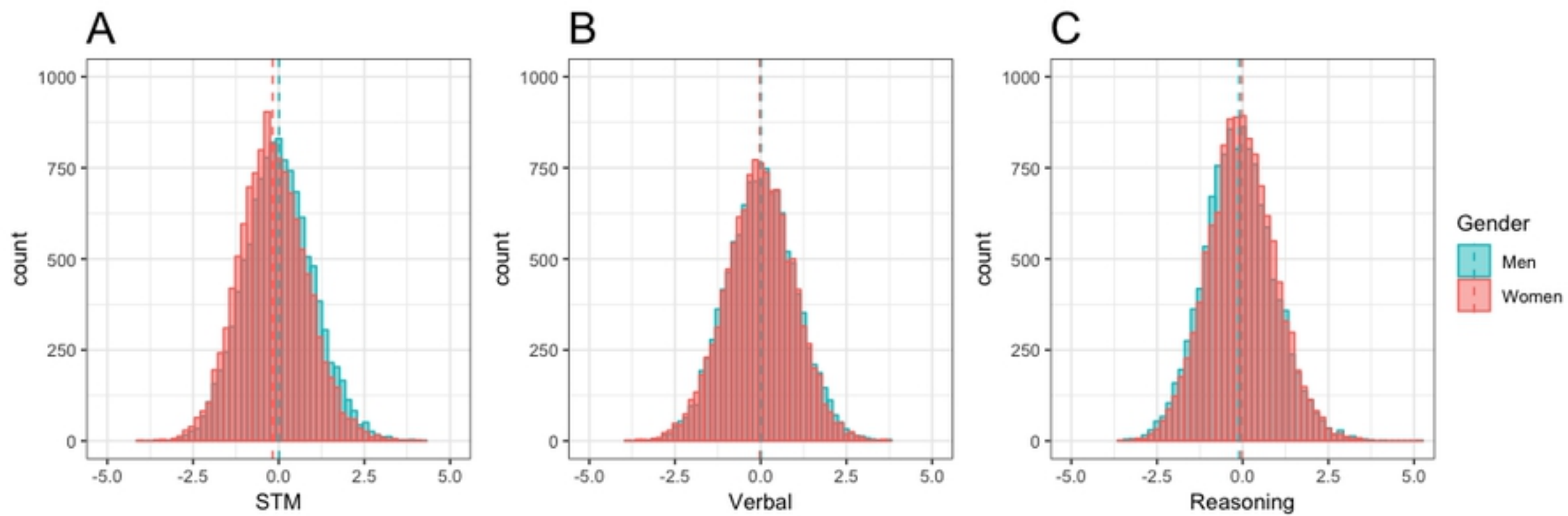
Figure



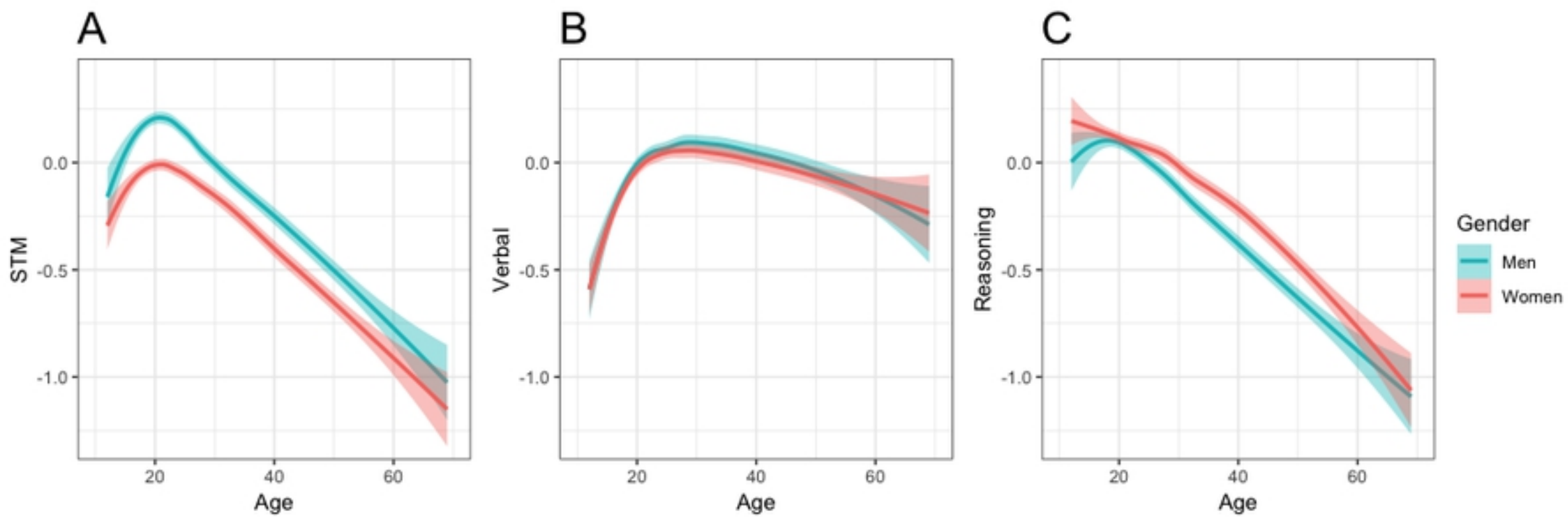
Figure



Figure



Figure



Figure