Two Separate, Large Cohorts Reveal Potential Modifiers of Age-Associated Variation in Visual Reaction Time Performance

J.S. Talboom[1,2], M.D. De Both[1,2], M.A. Naymik[1,2], A.M. Schmidt[1,2], C.R. Lewis[1,2], W.M. Jepsen[1,2], A.K. Håberg[3], T. Rundek[4], B.E. Levin[4], S. Hoscheidt[2,5], Y. Bolla[2,5], R.D. Brinton[2,5], N.J. Schork[1,6], M. Hay[2,5], C.A. Barnes[2,5], E. Glisky[2,5], L. Ryan[2,5], and M.J. Huentelman[1,2,]

[1]The Translational Genomics Research Institute (TGen), Phoenix, Arizona 85004 USA

[2]Arizona Alzheimer's Consortium, Phoenix, Arizona 85004 USA

[3]Norwegian University of Science and Technology, Trondheim 7491 Norway

[4]University of Miami Miller School of Medicine and Evelyn F. McKnight Brain Institute, Miami, FL 33136 USA

[5]University of Arizona, Tucson, Arizona 85721 USA

[6]City of Hope National Medical Center, Duarte, CA 91010 USA

Abstract

To identify individual differences and potential factors influencing age-related cognitive decline and disease, we created MindCrowd. MindCrowd is a cross-sectional web-based assessment of simple visual (sv) reaction time (RT, index of processing speed) and paired-associate learning (PAL, index of verbal episodic memory). svRT and PAL results were combined with 22 survey questions. Analysis of MindCrowd's svRT data revealed education and reported stroke as potential modifiers of changes in processing speed and memory from younger to older ages (n_{total} = 75,666, n_{women} = 47,700, n_{men} = 27,966; ages 18-85 years old, mean $(M)_{Age}$ = 46.54, standard deviation $(SD)_{Age}$ = 18.40). To complement this work, we evaluated complex recognition reaction time (cvrRT) in the UK Biobank cohort (n_{total} = 158,249 n_{women} = 89,333 n_{men} = 68,916; ages 40-70 years old, M_{Age} = 55.81, SD_{Age} = 7.72). Similarities between the UK Biobank and MindCrowd were assessed using a subset of the MindCrowd cohort. Labeled UKBb MindCrowd (n_{total} = 39,795, n_{women} = 29,640, n_{men} = 10,155; ages 40-70 years old, M_{Age} = 56.59, SD_{Age} = 8.16), this subset was carefully selected to mirror the UK Biobank. An identical linear model (LM) was used to assess both cohorts. The LM revealed similarities between MindCrowd and the UK Biobank across most results, despite obvious cohort differences (e.g., US vs. the UK). Notable divergent findings from the UK Biobank included (1) a first-degree family history of Alzheimer's disease (FHAD) was associated with longer RTs in only. (2) Compared to being a man with more education, being a woman was associated with longer cvrRT length differences observed from younger to older ages. Divergent results from UKBb MindCrowd include more education and reported smoking. More education was associated with shorter and smoking longer cvrRTs differences observed from younger to older ages. Collected with our prior work, MindCrowd is beginning to reveal the intricate network connecting processing speed, memory, and cognition to healthy and pathological brain aging.

Reaction time (RT), an index of processing speed or efficiency in the central nervous system (CNS) [1], is an essential factor in higher cognitive function[2,3] and is profoundly affected by age [4]. In fact, of the studied demographics, age is the main factor known to influence RT[4]. Processing speed is an important limiting factor for most aspects of cognition during aging, most notably memory[5,6]. In studies where processing speed was used as a covariate, the age-related variance in various episodic memory measures was reduced or even eliminated[7,8]. Moreover, studies comparing varied factors and tests of age-related episodic memory deficit implicate age-related decline in processing speed as the main mediator[9-11]. These findings collectively suggest that RT is a useful index of age-related cognitive decline, healthy brain aging, and neurodevelopment.

RT can be operationally defined as "simple," which typically involves a non-choice reaction to a visual stimulus (svRT). RT can also be operationally defined as "complex," which involves a reaction to one or more visual stimuli after recognition (cvrRT) of correct stimuli and inhibiting incorrect stimuli [12]. svRT demonstrates variability between individuals, which is akin to paired-associate learning (PAL) and is influenced by genetic and environmental factors[13]. In addition, svRT effects are well noted across the field of neurology; for example, AD and stroke patients show slower svRT and higher inter-individual variability [14,15]. However, due to the limitations of traditional research methods, the body of work concerning RT examined only limited ranges of demographic, health, medical, and lifestyle factors in small cohorts. For example, prior work's demographics consisted of college-aged students, well-educated older adults[16-18], or athletes [19-21].

Further, with notable exceptions[22-24], many studies had few participants (e.g., n <1000) and were therefore powered to detect only variables with large effect size and to lead to spurious non-replicable findings[25]. Consequently, many RT studies had minimal ability to reveal low frequency factors or those with subtle effect sizes and conduct more sophisticated analyses (e.g., ANOVA vs. Growth Modeling) to

find interactions and moderators. Collectively, this suggests that if RT performance can inform models of disease or normative and atypical aging, we need a deeper understanding of the normal variation of RT and the genetic and environmental factors associated with RT performance.

This study aimed to characterize RT across a broad range of demographic, health, medical, and lifestyle variables commonly associated with cognitive performance and AD risk. To do this, we utilized both the MindCrowd and UK Biobank cohorts[26,27], comprising over 233 thousand combined participants, to model RT as a function of 11 or more demographic, health, medical, and lifestyle factors. These factors have been previously associated with aging and cognition [28-33]. Based on our prior work and earlier RT research[34], we hypothesized that RT, via its structure of factor association and modifications, would reveal meaningful connections to healthy brain aging.

Materials and Methods

Study Participants

MindCrowd: Overview

In January 2013, we launched our internet-based study at www.mindcrowd.org. Website visitors 18 years or older were asked to consent to our study before any data collection via an electronic consent form. As of 3-26-2020, we have had 356,674 non-duplicate or distinct visitors to the website. Of these distinct visitors, over 194,542 (54%) consented to take part. The final data set had 75,666 (54% of consented individuals) participants who completed a simple visual reaction time (svRT) and paired-associate learning (PAL) tasks and answered 22 demographic, lifestyle, and health questions. Approval for this study was obtained from the Western Institutional Review Board (WIRB study number 1129241).

MindCrowd: simple visual reaction time (svRT)

After consenting to the study and answering five demographic questions (i.e., age, biological sex, years of education, primary language, and country where they reside), participants were asked to complete a web-based svRT task. We chose svRT because it is a simple central and peripheral nervous system-dependent task influenced by intelligence and brain injury[35]. Participants were presented with a pink sphere that appeared at random intervals (between 1-10 seconds) on the screen, and they were instructed to respond as quickly as possible after the sphere appeared by pressing the enter/return key on their keyboard. Once the participant responded, the sphere disappeared until the next trial. Each participant received a total of five trials. The sphere stayed on the screen until the participant responded. The dependent variable, response time in milliseconds (msec), was recorded from the sphere's appearance on the screen to the participant's key press or screen touch.

MindCrowd: paired-associate learning (PAL)

Next, participants were presented with the PAL task. For this cognitive task, during the learning phase, participants were shown 12 word-pairs, one word-pair at a time (2s/word-pair). During the recall phase, participants were given the first word of each pair and were asked to use their keyboard to type in (i.e., recall) the missing word. This learning-recall procedure was repeated for two more trials. Before beginning the task, each participant received one practice trial consisting of three word-pairs not contained in the 12 used during the test. Word-pairs were presented in different random orders during each learning and each recall phase. The same word pairs and order of presentation were used for all participants. The dependent variable/criterion was the total number of correct word pairs entered across the three trials (i.e., 12 x 3 = 36, a perfect score).

MindCrowd: demographic, medical, health, and lifestyle questions

Upon completing the PAL task, participants were asked to fill out an additional 17 demographic and health/disease risk factor questions. These questions included: marital status, handedness, race,

ethnicity, number of daily prescription medications, a first-degree family history of dementia, and yes/no responses to the following: seizures, dizzy spells, loss of consciousness (more than 10 minutes), high blood pressure, smoking status, diabetes mellitus, heart disease, cancer, reported stroke, alcohol/drug abuse, brain disease, and memory problems). Next, participants were shown their results and provided different comparisons to other test takers based on the average scores across all participants' sex, age, and education demographics. On this same page of the site, participants were given the option to be recontacted for future research (see Supplementary Table 2 for the list of MindCrowd questions asked).

UK Biobank: study design and aims

The UK Biobank (research ethics committee approval [11/NW/0382]) is a long-term study and research resource in the United Kingdom (UK), which investigates genetic and environmental exposure to the development of the disease. The UK Biobank's stated goal is to "build a major resource that can support a diverse range of research intended to improve the prevention, diagnosis, and treatment of illness and the promotion of health throughout society." The UK Biobank began in 2006. The study is currently following about 500,000 participants in the UK, enrolled at ages 40 to 69. Initial enrollment took place from 2006 to 2010. All participants are monitored for at least 30 years after recruitment and initial assessment (i.e., termed "instance 0" by the Biobank). Potential participants were invited to visit an assessment center, where they completed a questionnaire. Participants were next interviewed about lifestyle, medical history, and nutritional habits. Lastly, vital measurements, such as weight, height, and blood pressure, were measured. The UK Biobank aims to electronically record all health-related changes and events across the entire 30-year study. Notably, this task is aided by the UK's integrated health system and corresponding electronic health record-keeping, an approach that is not possible in the United States.

UK Biobank: data procurement

All UK Biobank data were derived from Application #43036, entitled "Exploring and Accommodating Heterogeneity in Large-Scale Genetic Analyses" as a "Collaborator Project,"

UK Biobank: Complex visual recognition reaction time (cvrRT) and educational attainment.

Each participant's cvrRT was based on 12 rounds of the card-game' Snap.' Participants were shown two cards at a time with a picture on them. Participants pressed a button on a table in front of them as quickly as possible if the images cards/matched. For each of the 12 rounds, the following data were collected: the pictures shown on the cards (Index of card A, Index of card B), the number of times the participant clicked the 'snap' button, and the latency to first click of the 'snap' button. This last record of "latency to click the button" was used as the UK Biobank's criterion for regression analyses.

For Educational Attainment, the following conversions from UK Biobank (UKBb) answer codes (see http://biobank.ndph.ox.ac.uk/showcase/coding.cgi?id=100305) to MindCrowd (MC) values were made: a) "UKBb **-7** None of the above" to "MC No high school diploma," b) "UKBb **2** A levels/AS levels or equivalent" to "MC High school diploma," c) "UKBb **3** O levels/GCSEs or equivalent" to "MC High school diploma," d) "UKBb **4** CSEs or equivalent" to "MC High school diploma," e) "UKBb **5** NVQ or HND or HNC or equivalent" to "MC Some college," f) "UKBb **6** Other professional qualifications (e.g., nursing and teaching)" to" MC Some college," g) "UKBb **1** College or University degree" to "MC College degree." All UKBb participants selecting "**-3** Prefer not to answer" were removed from the final dataset before model selection and analysis. While we did our best to ensure a similar education measure across UKBb MindCrowd and the UK Biobank, we realize that there are fundamental differences between US and UK schools that we cannot control or eliminate. Table 6 lists the UK Biobank data fields used.

Data Quality Control

For the MindCrowd analysis, a final data set, including all qualifying participants up to 3-17-2020, was generated. See Supplementary Figure 9 for a flowchart detailing the following filtering steps. This dataset removed participants: a) with duplicate email addresses (only first entry kept), b) who did not complete all three rounds of the PAL test, c) whose primary language was not English, d) who was not between 18-85 years old, e) whose RT trials were above or below 1.5 x the interquartile range (IQR) and f) whose median svRT was above or below 1.5 x the IQR range of all participants of the same age (Supplementary Figure 10 details RT and IQR exclusion). Participants from either study were removed if they were missing any data (listwise deletion). Lastly, for the UKBb MindCrowd and UK Biobank analysis, participants were removed if their responses to a demographic, medical, health, and lifestyle question did not match the other study. For example, participants in the UK Biobank who responded to the “Race” question with “Prefer Not to Answer” were removed. “Prefer Not to Answer” was not a choice MindCrowd participants were given on the “Race” question. Removing these participants was done to align UKBb MindCrowd and UK Biobank cohorts as much as possible.

Statistical Methods

Statistical analyses were conducted using R[36,37]. For all analyses, multivariate linear regression was performed using the general linear model (LM) to model Median svRT or Median cvrRT (i.e., criterion or dependent variable) as a function of either 24 (MindCrowd) or 11 predictors (UK Biobank analysis). For the MindCrowd analysis, svRT was modeled as a function of PAL Performance and Age raised to the power of three (i.e., to fit and estimate nonlinear associations). Most figures were created using “*ggplot2*” bundled together as a part of the R package, “*tidyverse*[38]”. Continuous by continuous interactions (i.e., simple slopes) were estimated using the R packages "*interactions*[39]," "*sandwich*[40]," "*jtools*[41]". Categorical by categorical interactions were estimated using the R package, “*emmeans*” [42]. Adjustments for multiple comparisons were evaluated using Tukey’s method via the “*emmeans*”

package. Missing data were assessed via the "*finalfit,*[43]" "*visdat,*[44]" and "*naniar*[45]" R packages (see Supplementary Table 1 for a complete list of resources).

Model evaluation and controls

All measurements were taken from distinct samples. Model fit and violations of parametric assumptions were evaluated separately in each model. Here we evaluated different residual plots, assessing normality, homoscedasticity, outliers, residual autocorrelation, and multicollinearity. The MindCrowd LM included all 22 demographic questions, health, medical, and lifestyle questions. These questions were: Age, Biological Sex, Race, Ethnicity, Educational Attainment, Marital Status, Handedness, Number of Daily Medications, Seizures, Reported Dizziness, Loss of Consciousness, Reported Hypertension, Smoking Status, Heart Disease, Reported Stroke, Alcohol/Drug Abuse, Diabetes Mellitus, Cancer, a First-Degree Family History of Alzheimer's disease, history of brain disease, whether the test was taken on a mobile device, and the version of the MindCrowd site used. Not surprisingly, the device used to take the RT test in MindCrowd was associated with RT performance. For the UK Biobank analyses, these 11 variables included: Age, Biological Sex, Diabetes mellitus, Handedness, Reported Stroke, Reported Hypertension, Smoking Status, Reported Dizziness, Educational Attainment, and a First-Degree Family History of Alzheimer's Disease. Examination of each model's variance inflation factors (VIF) revealed no unexpected factors with a VIF > 5 (i.e., considered "highly" colinear by convention, see Supplementary Table 3).

Inclusion of polynomials and automated model selection

The MindCrowd analyses included Age and PAL Performance as first through third-degree non-orthogonal polynomials (i.e., cubic regression). This choice was based on empirical evaluations, using Bayesian information criterion (BIC) weights (a.k.a Schwarz weights)[46]. We generated, ran, and recorded BICs across seven models (i.e., base and first – sixth-degree-[nonorthogonal]polynomial). BIC

weights were calculated from raw BIC values using the "*qPCR*[47]" (v1.4-1) R for each model. The third-degree polynomial model reported the largest BIC weight, and it was 1.46E+270 times more likely to occur than the base (no polynomial) model ($_{BICHighW}$ 9.99E-01 / $_{BICLowW}$6.82E-271 = 1.46E+270) [46]. It is worth noting that a prior study examining both complex and simple RT also included age as a third-degree polynomial. Other similarities included: a relatively large *n* = 7,000, both sexes, an 18-94 years-old age-range, and several RT findings[46].

For the MindCrowd, UKBb MindCrowd, and UK Biobank cohorts, we used the R package "glmulti" (v1.0.8)[48] to define our GLM models. *"glmulti"* uses full information criterion model selection versus shrinkage regression methods (e.g., LASSO or LAR)[48]. We used *"glmulti"* to avoid the pitfalls of stepwise selection methods or unintentional biased introduced via manual or *p*-value-based model selection. We had *"glmulti"* define the "best" (i.e., lowest *BIC*) MindCrowd and UK Biobank models separately using its *"genetic"* algorithm method with *"marginality"* set to True. We chose *BIC* as opposed to other information criterion methods because *BIC* punishes for model complexity. Two rounds of model selection were run to find pairwise interactions due to package limitations (i.e., millions of potential models). For round 1, the optimal model contained only the main effects when all 22 factors were included. In round 2, the only factors selected in the optimal main effects model were then included to select an optimal model, including two-way interactions.

Code availability

The code supporting each analysis, figure, table, or other material is publicly available. The directory containing these files is included in this study's data repository (see Data availability for more information).

Results

MindCrowd

As of March 13th, 2020, after filtering (see Data Quality Control in Materials and Methods), MindCrowd, had recruited 75,666 qualified participants (see Table 1 for Sociodemographic Characteristics and Supplementary Figure 1A for a histogram of age). We modeled svRT as a function of Age^3 and PAL $Performance^3$ (i.e., curvilinear associations) as well as 20 other factors (see Supplementary Figure 2 for diagnostic regression plots, and Table 2 for all coefficient *n*s). The omnibus model was significant ($F_{omnibus}$[58, 73406] = 858.20, $p_{omnibus}$ < 2.2e-16, Adjusted R^2 = 0.40).

Curvilinear associations: Age and paired-associate learning (PAL) total correct

Our model revealed that all three Age polynomials were significantly associated with svRT. Age^1 (i.e., linear association, first-degree polynomial, aka slope), Age^2 (i.e., quadratic association, second-degree polynomial), and Age^3 (i.e., cubic association, third-degree polynomial). On average from younger to older Age, a one-year difference (X=1). (1) Age^1 (shift in Y; p_{Age2} = 3.23E-17) was associated with 7 msec longer svRT. (2) Age^2 (shift in Age^1; p_{Age1} = 3.06E-18) was associated with 0.15 msec of additional svRT length (i.e., 7 + .15 msec/year, Figure 1A). (3) Age^3 (shift in Age^2; p_{Age3} = 1.46E-34) was associated with a negligible 1.20E-04 msec shift in additional svRT length (i.e., 7 + (0.15 + 1.20E-4) msec/year, Figure 1A). In contrast to Age's association with longer svRT, each word pair correct for PAL Performance. (1) PAL^1 (p_{PAL1} = 3.77E-34) was associated 9 msec shorter svRT. (2) PAL^2 (p_{PAL2} = 2.35E-14) was associated with 0.32 msec of additional svRT shortening (i.e., 9 + 0.32 msec/year, Figure 1B). (3) PAL^3 (p_{PAL3} = 7.28E-09) was associated with a slight 7.14E-04 msec shift in additional svRT shortening (i.e., 9 + (0.32 + 7.14E-04) msec/year Figure 1B).

Biological sex, educational attainment, and handedness

Biological Sex was a significant predictor of svRT (p_{Sex} = 1.26E-163). Being a man was associated with an average of 34 msec (9.63%) faster svRT response than being a woman (Figure 1D). Educational Attainment was also a significant factor associated with svRT. Compared to "No High

School Diploma," participants who had "Some College" ($p_{College}$ = 8.74E-05), or a "College Degree" ($p_{CDegree}$ = 2.95E-17, Figure 2A) were faster. Attending college and attaining a college degree was associated with a respective near 15 (4.14%) and 32 (8.92%) msec shorter svRT compared to not graduating from high school. Handedness was also associated with svRT. Left-handed participants had a near 4 msec (1.09%) faster svRT (p_{Left} = 0.03, Figure 2B). This association was present in individuals 20 to 40 years old (p_{40Left} < 0.01, Figure 2C) but not in individuals 40 to 60 years old (p_{60Left} = 0.07, Figure 2D).

Health, medical, and lifestyle factors

For health and medical factors associated with svRT, we found that Smoking Status ($p_{Smoking}$ = 1.26E-03, Figure 3A) and Reported Dizziness (p_{Dizzy} = 0.04, Figure 3B) were both significant predictors of svRT. Smoking Status was associated with 7 msec (1.99%) slower svRT, and Reported Dizziness was associated with nearly 5 msec (1.37%) slower svRT. When compared to participants reporting "no daily medications," taking "Two" (p_{Meds2} = 2.00E-03), "Three" (p_{Meds3} < 0.01), and "Four" (p_{Meds4} = 3.51E-16) Daily Medications were associated with an approximately 5 (1.64%), 6 (1.76%), and 18 (5.01%) msec longer svRTs, respectively (Figure 3C). Further, Diabetes Mellitus ($p_{Diabetes}$ < 3.36E-05, Figure 3D), and Stroke (p_{Stroke} = 3.59E-04, Figure 3E) were related to 11 (3.16%) and 20 (5.73%) msec longer svRTs respectively when compared to participants not reporting either condition. Of note, in this model, both a first-degree family history of Alzheimer's disease (FHAD; p_{FHAD} = 0.78) and Hypertension (p_{Hyper} = 0.52) were not significant predictors of svRT performance (Supplementary Figure 3A-B).

Two-way interactions

For interactions, we found Age significantly interacted with PAL Performance. Age x PAL Performance ($p_{Age*PAL}$ = 9.93E-62). Analysis of simple slopes suggests that each word-pair correct was associated with shorter svRT from younger to older ages. That is, at 20 ($p_{Age20*PAL}$<0.00), 40

($p_{Age40*PAL}$<0.00), 60 ($p_{Age60*PAL}$<0.00), and 80 ($p_{Age80*PAL}$<0.00, Figure 1C) years old. There was a significant Biological Sex x Age interaction ($p_{Age*Sex}$ = 4.61E-08), indicating that the associated slowing of svRT at younger and older ages in men, compared to women, was 0.36 msec (0.08%) less per one year difference in age (Figure 1C). These data suggest that men's svRT lengthened at a slower rate when compared to women. Of interest, in both women and men, we found significant Age x Educational Attainment interactions. Compared to Age x "No High School Diploma", participants reporting having "Some College" ($p_{Age*College}$ = 4.20E-04) or a College Degree" ($p_{Age*CDegre}$ = 2.07E-12) was associated with slower RTs from young to an older age. These results suggest that attending college or getting a college degree was associated with a 0.65 (0.15%) and 1.31 (0.30%) msec faster svRT performance per year of life, respectively (Figure 2A). The MindCrowd model revealed a significant Age x Reported Stroke interaction ($p_{Age*Stroke}$ = 2.72E-06). Participants who Reported Stroke were about 2 msec (0.37%) slower on svRT per year of life than those without Reported Stroke (Figure 3E). Lastly, we found a significant Age x Smoking Status interaction ($p_{Age*Smoke}$ = 5.67E-07). This interaction suggests that Smoking Status lengthens svRT by adding 0.57 msec (0.11%) per year difference in age (Figure 3A). See Table 3 for a complete list of the significant predictors, covariates, and interactions from the MindCrowd analysis.

Mobile Device

Participants who were identified as using a mobile device to take MindCrowd (i.e., using a touchscreen, n = 7,603, age M = 54.06 SD = 14.54 years) had slower svRTs and were older ($\beta_{Age\sim Mobile}$ = 14.13, $p_{Age\sim Mobile}$ < 2e-16) than those who did not use a mobile device (n = 76,775, age M = 45.54 SD = 18.43 years, see Supplementary Figure 8).

UKBb MindCrowd and UK Biobank

Of the total 75,666 MindCrowd participants, 39,759 between the ages of 40 and 70 were selected to mirror the UK Biobank. This subset is called UKBb MindCrowd from here on to differentiate it from MindCrowd. After filtering (see Data Quality Control), the UK Biobank cohort had 158,249 participants, derived from a data-request we received on 9-19-2019 (See Table 1 for Sociodemographic Characteristics and Supplementary Figure 1B-C for age histograms). We model both the UKBb MindCrowd's svRT (see Supplementary Figure 4 for regression diagnostic plots) as well UK Biobank's cvrRT (see Supplementary Figure 5 for regression diagnostic plots) as a function of 11 shared survey questions (see Table 4 for the *n*s used to estimate each coefficient). The omnibus UKBb MindCrowd (F_{mcomni}[20, 38871] = 1039, p_{mcomni} < 2.2e-16, Adjusted R^2 = 0.08) and UK Biobank ($F_{ukbbbomni}$[20, 157903] = 1038, $p_{ukbbomni}$ < 2.2e-16, Adjusted R^2 = 0.13) LMs were both significant. Table 5 lists the effects for each cohort side by side.

Age and biological sex

The UKBb MindCrowd cohort revealed Age as a significant predictor of svRT (p_{Age} = 2.00E-16). The parallel analysis (see Statistical Methods) of Age in the UK Biobank cohort was also significant (p_{Age} = 2.00E-16) for complex, visual, recognition reaction time (cvrRT). For the association of Age and RT, UKBb MindCrowd and the UK Biobank showed longer RTs or worse RT performance from younger to older ages, with nearly 6 and 3 msec slower RT per year difference of age, respectively (Figure 4A-B). For UKBb MindCrowd, Biological Sex was a significant predictor of RT (β_{Sex}=-40.00, p_{Sex}=8.03E-71), which was also the case in the UK Biobank (β_{Sex}=-18.28, p_{Sex}=2.00E-16). Being a man in both cohorts was associated with shorter RTs compared to being a woman (Figure 4C-D). Here the effect of Biological Sex on RT between UKBb MindCrowd was 40 msec (20.46%), and the UK Biobank was 18 msec (5.16%).

Education and handedness

Akin to the association found in the MindCrowd analysis, the UKBb MindCrowd and UK Biobank comparison found that Educational Attainment was a significant RT predictor. Indeed, for UKBb MindCrowd a "High School Diploma" ($\beta_{HSDiploma}$ = -9.68, $p_{HSDiploma}$ = 2.70E-01, 1.87%, "Some College" ($\beta_{College}$ = -26.08, $p_{College}$ = 1.61E-03, 5.06%) and a "College Degree" ($\beta_{CDegree}$ = -49.07, $p_{CDegree}$ = 1.91E-09, 9.51%), and in the UK Biobank a "High School Diploma" ($\beta_{HSDiploma}$ = -9.68, $p_{HSDiploma}$ = 1.23E-33, 1.74%), "Some College" ($\beta_{College}$ = -10.26, $p_{College}$ = 1.30E-40, 1.85%), or a "College Degree" ($\beta_{CDegree}$ = -11.71, $p_{CDegree}$ = 1.46E-41, 2.11%) were all significantly different from "No High School Diploma" (Figure 5A-B). Here, both UKBb MindCrowd and UK Biobank large cohorts reported shorter RT was associated with more education. Lastly, unlike the MindCrowd analyses, both the UKBb MindCrowd ($p_{Handedness}$ = .40) and the UK Biobank ($p_{Handedness}$ = 0.36) cohorts between the ages of 40-70 did not find Handedness to be a significant predictor of RT performance (Supplementary Figure 6).

Health, medical, and lifestyle factors

In terms of health factors associated with RT, in the UKBb MindCrowd cohort, Diabetes ($\beta_{\mathrm{Diabetes}}$=11.48, $p_{Diabetes}$=3.31E-03, 5.87%), Stroke (β_{Stroke}=18.47, p_{Stroke}=4.00E-02, 9.45%), ($\beta_{Hypertension}$=7.99, $p_{Hypertension}$=3.16E-03, 3.58%), and Dizziness (β_{Dizzy}=12.13, p_{Dizzy}=2.52E-03, 6.19%) were all significantly associated with longer svRTs. These associations were recapitulated by the UK Biobank. To that end, Diabetes Mellitus ($\beta_{Diabetes}$ = 5.48, $p_{Diabetes}$ = 4.80E-07, 1.55%, Figure 5C-D), Reported Stroke (β_{Stroke} = 10.61, p_{Stroke} = 6.15E-07, 2.99%, Figure 6A-B), Reported Hypertension ($\beta_{Hypertension}$ = 1.14, $p_{Hypertension}$ = 0.02, 0.31%, Figure 6C-D), and Reported Dizziness (β_{Dizzy} = 3.21, p_{Dizzy} = 3.71E-14, 0.91%, Figure 7A-B) were all significantly related to longer cvrRTs; however, the association between Reported Hypertension and cvrRT was small (Figure 6D). Further, in agreement with the MindCrowd analysis, Smoking Status was significantly related to svRT (β_{Smoke}=10.18, p_{Smoke}=0.01, 5.21%) in the UKBb MindCrowd cohort; however, FHAD (β_{FHAD}=-0.07, p_{FHAD}=0.97) was not. This

pattern of associations were reversed in the UK Biobank; that is, Smoking Status was not a significant predictor of cvrRT (β_{Smoke}=0.47, p_{Smoke}=0.22, Figure 7C-D), but FHAD was (β_{FHAD}=2.36, p_{FHAD}=3.35E-05, 0.69%, Figure 8A-B).

Two-way interactions

The "*glmulti*" [48] R package defined two interactions in the UKBb MindCrowd and UK Biobank analysis. In UKBb MindCrowd we found a significant Age x Biological Sex interaction ($p_{Age*Sex}$ = 2.00E-02, 0.33%). We found a comparable significant Age x Biological Sex interaction for cvrRT ($p_{Age*Sex}$ = 4.18E-28, 0.16%) in the UK Biobank. Across both MindCrowd and the UK Biobank, these interactions indicated that RT was lengthened in men from younger to older ages compared to women, was over 0.5 msec shorter per one year difference in age (Figure 4C-D). In addition, the UK Biobank analysis revealed a significant Biological Sex x Educational Attainment interaction not found in UKBb MindCrowd. Here, men with a "High School Diploma" ($p_{Sex*HSDiploma}$ = 3.30E-12, 3.35%), "Some College" ($p_{sex*College}$ = 8.71E-12, 3.14%) or a "College Degree" ($p_{sex*CDegre}$ = 1.68E-13, 3.18%) were significantly associated with faster cvrRT performance when compared to women having "No High School Diploma" (Figure 8C). Follow-up analyses of the simple effects via estimated marginal means (EMM, see Statistical Methods) revealed that men who did not graduate high school (*EMM* = 558.91 msec), compared to men with more education (*EMMs* = 542.99, 542.98, and 540.41 msec), had markedly faster cvrRTs, more in line with the women's cvrRT performance (*EMMs* = 566.80, 562.11, 561.40, and 561.09 ms). The associated difference in cvrRT for men with "No High School Diploma" compared to men with a "High School Diploma" (β_{Men}=15.92, p_{Men} = 2.00E-16, 1.75%) was more substantial than between women with "No High School Diploma" compared to women with a "High School Diploma" (β_{Women}=4.68, p_{Women} = 3.12E-04, 1.69%, see Supplementary Figure 7 for a graph of the EMM).

Discussion

Slowing of Reaction Time (RT) from Younger to Older Ages: Factors and Modifiers

Our study's results illuminate a portion of the intricate relationship between age and RT performance by identifying demographic, health, medical, and lifestyle factors associated with either attenuation or exacerbation of RT lengthening from younger to older ages (see figure 9 for an illustrative summary of the results). A large body of work on RT, leading back to Sir Francis Gaulton in 1890[49], has consistently demonstrated an age-associated shift in RT[50]. It is not surprising that MindCrowd, UKBb MindCrowd, and UK Biobank models revealed slowing of simple visual (svRT, MindCrowd) and complex visual recognition RT (cvrRT, UK Biobank) from younger to older ages. Likewise, the MindCrowd model found that the relationship between svRT and age was modestly curvilinear (Figure 1A). While this curvilinear relation between RT and age has been noted previously[51], both cohorts' large sample size combined with our application of an algorithm-based model definition[48] revealed a notable addition to this picture. Specifically, in the MindCrowd analysis, we observed an interaction between age and education (Figure 2A) and smoking (Figure 3A). Here, less education and smoking were related to the additional slowing of simple visual RT (svRT) on top of the svRT slowing associated with transitioning from younger to older ages. Lastly, the age by reported stroke interaction modeled in MindCrowd was associated with longer RTs (Figure 3E). This study's large sample size and broad age and surveyed data range places it as one of the most substantial cross-sectional RT evaluations across the aging spectrum. Our findings suggest that smoking and stroke (i.e., cardiovascular health) and amount of education (i.e., cognitive demand or reserve) are factors, modifiable across aging, that influence age-associated RT slowing.

A First-Degree Family History of Alzheimer's Disease (FHAD), Cognition, and RT: Intertwined systems

In the UK Biobank cohort[27], we found an association between having an FHAD and slower (2.43 msec) cvrRT (Figure 8B). This effect of FHAD on RT or more so underlying process speed is in line with our episodic verbal memory task (i.e., paired-associate learning [PAL]) finding [34], where we found FHAD was linked to lower PAL performance. Furthermore, a prior functional magnetic resonance imaging (fMRI) study examining medial-temporal lobe activation using a cvrRT task found a ~100 msec RT lengthening in 68 (mean age of 54) FHAD participants[52]. These data suggest that genetic and environmental factors relating to AD risk are present in individuals with an FHAD. Indeed, the first-degree relatives of identified FHAD participants consisted of familial early-onset AD or late-onset AD, which also has a high heritability of 79%, suggesting that they are a higher risk category for developing AD. Thus, such shared AD and FHAD factors may relate to sensorimotor function and processing speed (i.e., RT) analogous to alterations in cognition and memory (i.e., PAL).

We found a correlation between svRT and PAL performance (Figure 1B). This finding was in line with many prior studies, the dependence of episodic memory on processing speed, a dependence that grows with age and incident of age-related disease (e.g., AD)[5-8,53]. The association of svRT with PAL may highlight distributed systems and networks that underlie RT performance. For example, svRT performance could depend on the functioning of many cognitive areas that PAL requires and vice versa. Another possibility is that properly functioning memory, and cognitive networks correspond to better RT performance. Evidence for this is suggested by the fact that higher intelligence is related to faster RT[54-56]. However, diverse factors (e.g., exercise, time of day, meal proximity, and individual assessing RT) affect RT performance[53,56,57]. Thus, making an accurate estimation of RT's effect on intelligence and vice versa changeling. Together with our prior results, these findings suggest that RT performance could be used as a metric to assess potential AD risk. However, further research, including longitudinal

studies, replication, and corroboration of RT's link to age-related cognitive decline and disease, are necessary to support this notion.

The difference in the effects of FHAD between UK Biobank and MindCrowd (for both MindCrowd and UKBb MindCrowd analyses, Supplementary Figure 3A and Figure 8A) could be due to the vast difference in the fraction of FHAD participants in the UK Biobank (FHAD = 13% of total) compared to UKBb MindCrowd (FHAD = 35% of total). While noting that we target those with FHAD for recruitment into MindCrowd, this substantial disparity could also be due to the accuracy of the UK Biobank's FHAD. The UK Biobank was calculated from three separate questions (i.e., mother, father, and siblings AD diagnosis). Adding to this, the United Kingdom uses a massive, detailed, nation-wide electronic health record system facilitating respondent health and medical survey accuracy. Compare this to our single question in MindCrowd, asking participants of all ages to remember if a relative was diagnosed with AD. Another possibility is the RT paradigm used; that is, MindCrowd's test of svRT compared to the UK Biobank's use of cvrRT. The fact that complex reaction time, requiring recognition and the choice to "respond or not respond," rather than just stimulus-response, may underly this difference. UKBb MindCrowd svRT performance was consistently faster across age than UK Biobank cvrRT performance; an observation noted in a prior study also measured both simple and complex RT[58]. These findings are consistent with the idea that complex RT requires more processing time[12,59], and prior work found the age-associated slowing of RT was higher for choice RT than simple RT [60]. Lastly, the differences in FHAD associations across MindCrowd and the UK Biobank are perhaps due to the UK Biobank had over 2X the number of participants compared to MindCrowd (i.e., 158K vs. 76K). The larger sample is expected to produce better model definitions and increased statistical power. Increased statistical power may also have enhanced accuracy, validity, and reproducibility.

Biological Sex and Educational Attainment

Numerous RT studies have found sex differences in RT performance, which does not appear to be reduced by practice[16,20,58]. Consistent with others[58,60], men exhibited shorter RTs in each model across cohorts (Figures 1D and 4C-D). In addition, the analysis of the UKBb MindCrowd and UK Biobank implicated biological sex affecting RT slowing from younger to older ages. The age interaction with biological sex suggests that being a woman from younger to older ages is associated with longer RT compared to being a man. These results largely replicate a previous sizable study (i.e., 7,000 participants) evaluating RT[58]. Similar to our own, this study found that 1) men consistently outperformed women on all RT measures across age, 2) differences in RT performance from younger to older ages were nonlinear, 3) including a third-degree polynomial for age provided the best model fit, and 3) compared to men, women displayed longer RTs consistently from younger to older ages[58]. Collected with our prior study of PAL performance [34], these associations replicate prior work and suggest that biological sex affects RT and age-associated shift in RT.

Educational attainment was associated with svRT in both MindCrowd cohorts and cvrRT in the UK Biobank. Overall, having more education (i.e., reporting higher milestones) was related to shorter svRT and cvrRT (Figures 2A and 5A-B). However, it is unclear if individuals with higher processing speed naturally seek more education and what other factors confound this relationship. Further work utilizing both cohorts is necessary to shed light on the effects and modifiers of FHAD and cross-cohort discrepancies. The model of the UK Biobank revealed an interaction between biological sex and education on RT performance. The breakdown revealed that men had similar RT performance if they attained a high school diploma and above. However, men who did not graduate high school showed markedly slower RTs, which brought them in line with women's RT performance. However, the associated lengthening of RT for less-educated men was vastly more than that found in less-educated

women (reported "No High School Diploma" made up 3.83% of women and 3.11% of men; see Supplementary Figure 7).

Handedness

In MindCrowd, handedness, specifically being left-handed, was associated with shorter svRTs. Prior studies have reported similar associations, where left-handedness was correlated with shorter svRT[19,59,61]. Hemispheric asymmetries in spatial processing are thought to underly faster svRT for the left hand[60,62]. Handedness was not associated with svRT in UKBb MindCrowd or cvrRT in the UK Biobank. One explanation for the divergent findings is that MindCrowd includes younger participants (i.e., 18–40-year-olds). Indeed, in MindCrowd, the association appears to diminish from younger to older ages. Specifically, in Figure 2B, the separation of the regression lines between left- and right-handed participants shrinks and eventually crosses around the 4th decade of life. Figure 2C shows that the left- and right-handed regression lines separate in 20 to 40-years-olds, while Figure 2D shows that these regression lines are not separate in 40 to 60-year-olds. While purely speculative, differences in social conventions may have played a role. For example, some older participants were forced to be right-handed, whereas younger participants were not. In doing so, upping the amount of unexplained variance in older, but not younger, participants across MindCrowd and the UK Biobank.

Health, Medical, and Lifestyle Factors

The MindCrowd analysis incorporated all 13 available health, medical, and lifestyle-related factors, of which six were present and incorporated into the shared UKBb MindCrowd/UK Biobank model (Tables 2 and 4). Before the launch of MindCrowd, these factors were carefully selected based upon their known relation to 1) age-associated alterations, 2) RT performance, and 3) PAL Performance. Of the 13 health, medical, and lifestyle factors evaluated in the MindCrowd analysis, we found associations between svRT and the number of daily medications, reported dizziness, smoking status,

reported stroke, and diabetes mellitus. Each health and medical factor were associated with longer svRTs (Figure 3). We should note that the number of daily medications is a serving as a proxy for overall health. That is, the worse one's health, the worse one's performance, the increased number of medications treating the underlying health conditions. The UKBb MindCrowd (svRT) and the UK Biobank (cvrRT) analyses found *similar* associations between reported dizziness, reported stroke, diabetes mellitus, indicating hypertension. Although each association differed in magnitude between the two older cohorts, each was related to slower RT. The UKBb MindCrowd to the UK Biobank found a different association for FHAD (UKBb MindCrowd = no association; UK Biobank = 2.43 msec slower), smoking status (UKBb MindCrowd = 10 msec slower; UK Biobank = no association). Interestingly, despite some differences, only a few coefficient signs differed between the UKBb MindCrowd and UK Biobank; indeed, most estimations were well within an order of magnitude between the two cohorts (e.g., age, educational attainment, and age by biological sex interaction, Table 5).

Many factors are likely to account for the different associations between smoking and FHAD between UKBb MindCrowd and the UK Biobank (Figures 7C-D and 8A-B). Some of these include differences in demographics, genetic heterogeneity, and age[26]. However, candidates include the fractions of participants reporting each factor (e.g., for diabetes mellitus: MindCrowd = 1%, UKBb MindCrowd = 7%, and UK Biobank = 3%). Another factor is that the UK Biobank's participant number is twice the size of MindCrowd and four times the size of UKBb MindCrowd. Despite our study's size, the observational and cross-sectional method means that we cannot rule out effects due to confounding variables.

Consequently, while numbers may be close, we do not assume that the UKBb MindCrowd is similar and can be compared to the UK Biobank. Furthermore, we observed that UKBb MindCrowd consistently reported larger estimates and standard errors than the UK Biobank. For example, the

MindCrowd cohort's estimation of the sex-difference was consistently slower (~40 msec), even in the UKBb MindCrowd cohort, when looking at the UK Biobank (~19 msec). This difference demonstrates why the study of neuropsychological traits and disease requires large sizes to provide accurate estimations driving better predictive validity.

Opportunities and Limitations of Our Research

Opportunities: Unique impact

We strongly advocate for large-scale efforts like ours, the UK Biobank[63], and others[22]. Indeed, studies of this kind have characteristics that provide the unique impact necessary to move the fields of aging and age-related diseases forward. These include: 1) statistical control, as our MindCrowd analysis incorporated all 24 available factors, 11 of which were used in the UKBb MindCrowd and UK Biobank model. 2) The inclusion of each predictor controlled for its association on RT, which potentially removed variability (noise), thus enhancing statistical power. 3) The two models used for each of the analyses were selected with little human input by automated application of specific statistical criteria (see Inclusion of polynomials and automatic model selection in Statistical Methods). This likely decreases bias, the probability of overfitting, and multicollinearity. 4) For the present study, MindCrowd had over 76K and the UK Biobank over 158K participants. Large sample sizes in each cohort were expected to help reduce variance, enhance estimation, select better models, and in turn, enhance statistical power. Expanded statistical power may then enhance accuracy, validity, and reproducibility.

Lastly, a recent genome-wide association study examining associations between RT and single nucleotide polymorphisms (SNP) in the UK Biobank and CHARGE and COGENT consortia noted weak correlations between the reported cognitive-associated SNPs among US and UK cohorts[64]. Here, MindCrowd presents a future opportunity to resolve these weak associations and get a better picture of potential cohort effects. Taken together, these characteristics increase the likelihood of making accurate

inferences regarding associations while boosting predictive validity. These are both necessary and vital attributes when searching for genetic associations and the structure underlying healthy brain aging.

Limitations: Internet self-report and others

There are potential concerns that arise from web-based studies[65]. Indeed, limitations of this study include the cross-sectional design and the partial discrepancy in MindCrowd's svRT test compared to the UK Biobank's cvrRT test and info collected between the UK Biobank and MindCrowd (e.g., the omission of "prefer not to answer" choices for race and education questions). Acknowledging these drawbacks, we believe that the advantage of meaningfully higher participant numbers and enriched cohort diversity facilitated via online research remediates some disadvantages. For example, the range of error reported in recent internet-based studies of self-reported quantitative traits like height and weight was between 0.3% and 20% [66-69]. Previously, we ran simulations on the association between FHAD and PAL by randomly shuffling the FHAD responses (e.g., Yes to No, and No to Yes), introducing increasing sequential amounts of "error." We found that even with a subtle effect such as FHAD on PAL performance, due to our cohort size, 24% error would still have only made us commit a Type1 error 50% of the time[34]. In line with this notion, another publication demonstrated that online RT studies produce reproducible results[70].

Further, we developed an extensive and automated data filtering pipeline (see Data Quality Control and Supplementary Figures 9-10) to address these concerns and enhance validity and accuracy. For example, we controlled for svRT delays generated using the touchscreen or mobile device by adding this information as a variable to the regression model (see Supplementary Figure 8). These data (i.e., real or filtered) were excluded before analysis (i.e., listwise deletion). Exclusion resulted in dropping 0.3 % and 6.1% of MindCrowd and UK Biobank participants, respectively. One of the 25 critical factors had over 5% missing data (see Supplementary Figure 11 and Supplementary Tables 4 and 5). Reported

Dizziness in the UK Biobank had 64.36% missing data. Hence, interpretation of this factor's association with cvrRT should only be considered for "hypothesis-generation[71]." Evaluation of selection bias between retained and excluded participants revealed an overall lower probability of exclusion in MindCrowd and higher likelihood in the UK Biobank (see Supplementary Tables 6 and 7). Notable groups with a higher probability of exclusion included those in the highest age ranges and those reporting hypertension and dizziness. These higher probability groups were found in both study's cohorts.

Lastly, it is essential to note that our internet-based svRT task was not designed to mirror conventional face-to-face RT testing paradigms directly. Indeed, we find higher RTs and steeper slopes across aging than studies assessing svRT via the gold standard, laboratory-based assessments (e.g., [60,72]). However, these paradigm differences are not likely to alter our svRT test's validity or reliability. One reason being our test is only interpreted within MindCrowd to identify associated factors and reveal individual differences. Despite test paradigm differences, we believe that large cross-sectional studies like MindCrowd, utilizing internet-based testing and remote biosample collection, are vital to moving the field of aging and age-related disease forward (see Opportunities: Unique impact above and [63]).

Conclusion

Understanding the modifiable and non-modifiable variables associated with RT and related cognitive function will begin to deconstruct the underlying architecture of elements accounting for the vast heterogeneity seen in individual trajectories of age-associated cognitive decline. Only then will it be possible to develop a healthy brain aging model that is both valid and reliable[73]. Such a model holds immense potential to attenuate age- and disease-related cognitive deficits, thus enhancing cognitive healthspan. Any extension of cognitive healthspan, better aligning it to the human lifespan, would be invaluable and increasingly vital when aggregated across the aging population. Mitigating age- or

disease-related cognitive decline, allowing maintenance of independence by even only a few years, would have many benefits. For example, the U.S. could save billions of dollars in health care costs and lost productivity of caregivers while improving the quality of life for the aging population[63]. In this study, we revealed several potential factors related to aging and processing speed. Of those, smoking and education, as potentially modifiable factors throughout life, were associated with slower and faster RTs, respectively, from younger to older ages. With MindCrowd recruitment ever-increasing, our goal is to continue supplying and refining the knowledge necessary to optimize cognitive performance throughout life.

References

1. Boisgontier, M.P., Wittenberg, G.F., Fujiyama, H., Levin, O. & Swinnen, S.P. Complexity of central processing in simple and choice multilimb reaction-time tasks. *PLoS One* **9**, e90457 (2014).
2. Deary, I.J. & Der, G. Reaction time, age, and cognitive ability: Longitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neuropsychology, and cognition* **12**, 187-215 (2005).
3. Deary, I.J., Allerhand, M. & Der, G. Smarter in middle age, faster in old age: a cross-lagged panel analysis of reaction time and cognitive ability over 13 years in the West of Scotland Twenty-07 Study. *Psychol Aging* **24**, 40-7 (2009).
4. Welford, A.T. Reaction time, speed of performance, and age. *Ann N Y Acad Sci* **515**, 1-17 (1988).
5. Verhaeghen, P. & Salthouse, T.A. Meta-analyses of age-cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. *Psychol Bull* **122**, 231-49 (1997).
6. Salthouse, T.A. & Babcock, R.L. Decomposing adult age differences in working memory. *Developmental psychology* **27**, 763 (1991).
7. Salthouse, T.A. Aging and measures of processing speed. *Biol Psychol* **54**, 35-54 (2000).
8. Sliwinski, M. & Buschke, H. Processing speed and memory in aging and dementia. *J Gerontol B Psychol Sci Soc Sci* **52**, P308-18 (1997).
9. Bryan, J. & Luszcz, M.A. Speed of information processing as a mediator between age and free-recall performance. *Psychology and aging* **11**, 3 (1996).
10. Baudouin, A., Clarys, D., Vanneste, S. & Isingrini, M. Executive functioning and processing speed in age-related differences in memory: contribution of a coding task. *Brain Cogn* **71**, 240-5 (2009).
11. Lindenberger, U., Mayr, U. & Kliegl, R. Speed and intelligence in old age. *Psychol Aging* **8**, 207-20 (1993).
12. Balakrishnan, G. *et al.* A comparative study on visual choice reaction time for different colors in females. *Neurol Res Int* **2014**, 301473 (2014).

13. Fan, J., Wu, Y., Fossella, J.A. & Posner, M.I. Assessing the heritability of attentional networks. *BMC Neurosci* **2**, 14 (2001).
14. Evarts, E.V., Teravainen, H. & Calne, D.B. Reaction time in Parkinson's disease. *Brain* **104**, 167-86 (1981).
15. Pearce, S.C., Stolwyk, R.J., New, P.W. & Anderson, C. Sleep disturbance and deficits of sustained attention following stroke. *J Clin Exp Neuropsychol* **38**, 1-11 (2016).
16. Spierer, D.K., Petersen, R.A., Duffy, K., Corcoran, B.M. & Rawls-Martin, T. Gender influence on response time to sensory stimuli. *J Strength Cond Res* **24**, 957-63 (2010).
17. Blough, P.M. & Slavin, L.K. Reaction time assessments of gender differences in visual-spatial performance. *Percept Psychophys* **41**, 276-81 (1987).
18. Lupinacci, N.S., Rikli, R.E., Jones, C.J. & Ross, D. Age and physical activity effects on reaction time and digit symbol substitution performance in cognitively active adults. *Res Q Exerc Sport* **64**, 144-50 (1993).
19. Dane, S. & Erzurumluoglu, A. Sex and handedness differences in eye-hand visual reaction times in handball players. *Int J Neurosci* **113**, 923-9 (2003).
20. Tonnessen, E., Haugen, T. & Shalfawi, S.A. Reaction time aspects of elite sprinters in athletic world championships. *J Strength Cond Res* **27**, 885-92 (2013).
21. Kaur, P., Paul, M. & Sandhu, J.S. Auditory and visual reaction time in athletes, healthy controls, and patients of type 1 diabetes mellitus: A comparative study. *International Journal of Diabetes in Developing Countries* **26**(2006).
22. Vincent, A.S., Roebuck-Spencer, T., Gilliland, K. & Schlegel, R. Automated Neuropsychological Assessment Metrics (v4) Traumatic Brain Injury Battery: military normative data. *Mil Med* **177**, 256-69 (2012).
23. Vincent, A.S. *et al.* Reference data from the automated Neuropsychological Assessment Metrics for use in traumatic brain injury in an active duty military sample. *Mil Med* **173**, 836-52 (2008).
24. Krieg, E.F., Jr. *et al.* Neurobehavioral test performance in the third National Health and Nutrition Examination Survey. *Neurotoxicol Teratol* **23**, 569-89 (2001).
25. Button, K.S. *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* **14**, 365-76 (2013).
26. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).
27. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine* **12**(2015).
28. Rawle, M.J., Cooper, R., Kuh, D. & Richards, M. Associations between polypharmacy and cognitive and physical capability: a British birth cohort study. *Journal of the American Geriatrics Society* **66**, 916-923 (2018).
29. Grimby, A. & Rosenhall, U. Health-related quality of life and dizziness in old age. *Gerontology* **41**, 286-98 (1995).
30. Sabia, S., Marmot, M., Dufouil, C. & Singh-Manoux, A. Smoking history and cognitive function in middle age from the Whitehall II study. *Arch Intern Med* **168**, 1165-73 (2008).
31. Biessels, G.J., Deary, I.J. & Ryan, C.M. Cognition and diabetes: a lifespan perspective. *Lancet Neurol* **7**, 184-90 (2008).
32. Kilander, L., Nyman, H., Boberg, M., Hansson, L. & Lithell, H. Hypertension is related to cognitive impairment: a 20-year follow-up of 999 men. *Hypertension* **31**, 780-786 (1998).

33. Patel, M.D., Coshall, C., Rudd, A.G. & Wolfe, C.D. Cognitive impairment after stroke: clinical determinants and its associations with long-term stroke outcomes. *J Am Geriatr Soc* **50**, 700-6 (2002).
34. Talboom, J.S. *et al.* Family history of Alzheimer's disease alters cognition and is modified by medical and genetic factors. *Elife* **8**(2019).
35. Bashore, T.R. & Ridderinkhof, K.R. Older age, traumatic brain injury, and cognitive slowing: some convergent and divergent findings. *Psychol Bull* **128**, 151-98 (2002).
36. Team, R.C. R: A language and environment for statistical computing. (Vienna, Austria, 2013).
37. Racine, J.S. RStudio: a platform-independent IDE for R and Sweave. (JSTOR, 2012).
38. Wickham, H. *et al.* Welcome to the Tidyverse. *Journal of Open Source Software* **4**, 1686 (2019).
39. JA, L. Interactions: Comprehensive, User-Friendly Toolkit for Probing Interactions. R package version 1.1.0. (2019).
40. Zeileis, A., Lumley, T., Berger, S., Graham, N. & Zeileis, M.A. Package 'sandwich'. (2020).
41. Long, J.A. & Long, M.J.A. Package 'jtools'. (2017).
42. Lenth, R., Singmann, H. & Love, J. Emmeans: Estimated marginal means, aka least-squares means. *R package version* **1**(2018).
43. Harrison, E., Drake, T. & Ots, R. finalfit: Quickly Create Elegant Regression Results Tables and Plots when Modelling. *R package version 0.9. 0, URL https://CRAN. R-project. org/package= finalfit* (2019).
44. Tierney, N. visdat: Visualising whole data frames. *Journal of Open Source Software* **2**, 355 (2017).
45. Tierney, N., Cook, D., McBain, M. & Fay, C. naniar: data structures, summaries, and visualisations for missing data. R package version 0.5. 0. (2020).
46. Wagenmakers, E.-J. & Farrell, S. AIC model selection using Akaike weights. *Psychonomic bulletin & review* **11**, 192-196 (2004).
47. Ritz, C. & Spiess, A.-N. qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* **24**, 1549-1551 (2008).
48. Calcagno, V. & de Mazancourt, C. glmulti: an R package for easy automated model selection with (generalized) linear models. *Journal of statistical software* **34**, 1-29 (2010).
49. Galton, F. Exhibition of instruments (1) for testing perception of differences of tint, and (2) for determining reaction-time. *The Journal of the Anthropological Institute of Great Britain and Ireland* **19**, 27-29 (1890).
50. Bleecker, M.L., Bolla□Wilson, K., Agnew, J. & Meyers, D.A. Simple visual reaction time: sex and age differences. *Developmental Neuropsychology* **3**, 165-172 (1987).
51. Kosinski, R.J. A literature review on reaction time. *Clemson University* **10**, 337-344 (2008).
52. Johnson, S.C. *et al.* The influence of Alzheimer disease family history and apolipoprotein E epsilon4 on mesial temporal lobe activation. *J Neurosci* **26**, 6069-76 (2006).
53. Moradi, A. & Esmaeilzadeh, S. Simple reaction time and obesity in children: whether there is a relationship? *Environ Health Prev Med* **22**, 2 (2017).
54. Nettelbeck, T. Factors affecting reaction time: Mental retardation, brain damage, and other psychopathologies. *Reaction times*, 355-401 (1980).
55. Lee, J.J. & Chabris, C.F. General cognitive ability and the psychological refractory period: individual differences in the mind's bottleneck. *Psychol Sci* **24**, 1226-33 (2013).
56. Taimela, S. Factors Affecting Reaction-Time Testing and the Interpretation of Results. *Perceptual and Motor Skills* **73**, 1195-1202 (1991).

57. Jain, A., Bansal, R., Kumar, A. & Singh, K.D. A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *Int J Appl Basic Med Res* **5**, 124-7 (2015).
58. Der, G. & Deary, I.J. Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. *Psychol Aging* **21**, 62-73 (2006).
59. Barthelemy, S. & Boulinguez, P. Manual reaction time asymmetries in human subjects: the role of movement planning and attention. *Neurosci Lett* **315**, 41-4 (2001).
60. Woods, D.L., Wyma, J.M., Yund, E.W., Herron, T.J. & Reed, B. Factors influencing the latency of simple reaction time. *Front Hum Neurosci* **9**, 131 (2015).
61. Barthelemy, S. & Boulinguez, P. Orienting visuospatial attention generates manual reaction time asymmetries in target detection and pointing. *Behav Brain Res* **133**, 109-16 (2002).
62. Gonzalez, C.L., Ganel, T. & Goodale, M.A. Hemispheric specialization for the visual control of action is independent of handedness. *J Neurophysiol* **95**, 3496-501 (2006).
63. Huentelman, M.J., Talboom, J.S., Lewis, C.R., Chen, Z. & Barnes, C.A. Reinventing Neuroaging Research in the Digital Age. *Trends Neurosci* **43**, 17-23 (2020).
64. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun* **9**, 2098 (2018).
65. Kraut, R. *et al.* Psychological research online: report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *Am Psychol* **59**, 105-17 (2004).
66. Elgar, F.J. & Stewart, J.M. Validity of self-report screening for overweight and obesity. Evidence from the Canadian Community Health Survey. *Can J Public Health* **99**, 423-7 (2008).
67. Shields, M., Connor Gorber, S. & Tremblay, M.S. Estimates of obesity based on self-report versus direct measures. *Health Rep* **19**, 61-76 (2008).
68. Ikeda, N. Validity of Self-Reports of Height and Weight among the General Adult Population in Japan: Findings from National Household Surveys, 1986. *PLoS One* **11**, e0148297 (2016).
69. Yoon, K., Jang, S.N., Chun, H. & Cho, S.I. Self-reported anthropometric information cannot vouch for the accurate assessment of obesity prevalence in populations of middle-aged and older Korean individuals. *Arch Gerontol Geriatr* **59**, 584-92 (2014).
70. Reimers, S. & Stewart, N. Adobe Flash as a medium for online experimentation: A test of reaction time measurement capabilities. *Behavior research methods* **39**, 365-370 (2007).
71. Madley-Dowd, P., Hughes, R., Tilling, K. & Heron, J. The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology* **110**, 63-73 (2019).
72. Woods, D.L., Wyma, J.M., Yund, E.W., Herron, T.J. & Reed, B. Age-related slowing of response selection and production in a visual choice reaction time task. *Front Hum Neurosci* **9**, 193 (2015).
73. Ryan, L. *et al.* Precision Aging: Applying Precision Medicine to the Field of Cognitive Aging. *Front Aging Neurosci* **11**, 128 (2019).

Data Availability

All programs, software, and other materials described herein are publicly or commercially available. R code used for all analyses and figure generation was included as "Source Data" R

Markdown notebook. The R code includes each model selected and the full list of parameters used. The data that support each analysis, figure, and table is freely available at Dryad (https://datadryad.org).

Acknowledgments

A substantial porting of this research was conducted using the generous data and resources provided by the UK Biobank (https://www.ukbiobank.ac.uk/published-papers). We acknowledge the Mueller Family Charitable Trust and Mike Mueller their significant donations. The State of Arizona DHS also supported this work in support of the Arizona Alzheimer's Consortium (PI; Eric Reiman), the Flinn Foundation (PI; Matt Huentelman), The McKnight Brain Research Foundation, NIH-NIA grant R01-AG049465 (PI; Carol Barnes), and NIH-NIA grants UH2 AG064706, U19 AG023122, U24 AG051129, and U24 AG051129-04S1 (PI; Nicholas J. Schork). We also recognize the support of many anonymous donors and institutional support of TGen, the City of Hope, and the University of Arizona.

Competing Interest

The authors have no known competing or conflicting interests to declare.

Materials and Correspondence

Matt Huentelman, Ph.D., TGen, 445 N Fifth Street, Phoenix, AZ 85004, 602-343-8730 (office), mhuentelman@tgen.org.

Author Contributions

J.S.T wrote and edited much of the manuscript, analyzed much of these data, and supervised several project management areas. J.S.T also contributed to data analysis, interpretation, and synthesis. M.D.DB oversaw much of the information technology needed for this study, analyzed a portion of these data, and supplied critical project support. M.A.N analyzed a part of these data, generated some figures, and provided critical statistical support. A.M.S. helped write and edit the manuscript as well as conduct critical portions of the statistical analyses. C.R.L. aided in writing and editing the manuscript, data

analysis, and general project oversight. W.M.J. aided in writing and editing the manuscript, data analysis, and general project oversight. A.K.H wrote and helped write portions of the manuscript and serve as a primary editor. Further, A.K.H. gave critical early data analysis and project support as well as data interpretation and synthesis. T.R. edited the manuscript and provided critical clinical insight and project management support. B.E.L. edited the manuscript and gave vital clinical insight. S.H. edited the manuscript and supported data analysis. Y.B. edited the manuscript and provided critical clinical insight and project management support. R.D.B. was instrumental in planning, designing, and executing this research. N.J.S. collaborated to provide access to the UK Biobank and helped with data analysis as well as project insight generation. M.H. supplied critical early data analysis and project support, as well as data interpretation and synthesis. C.A.B. helped write and edit the manuscript. In addition, C.A.B. aided several aspects of study design and helped with overall project planning and data interpretation and synthesis. E.G., aided in writing and editing the manuscript and study design and data interpretation and synthesis. L.R. aided in writing and editing the manuscript and study planning, design, and data interpretation and synthesis. M.J.H. was the principal investigator of this project. M.J.H oversaw most study planning, design, data collection and analysis, and data interpretation and synthesis. MJH played a substantial role in writing and editing the manuscript.

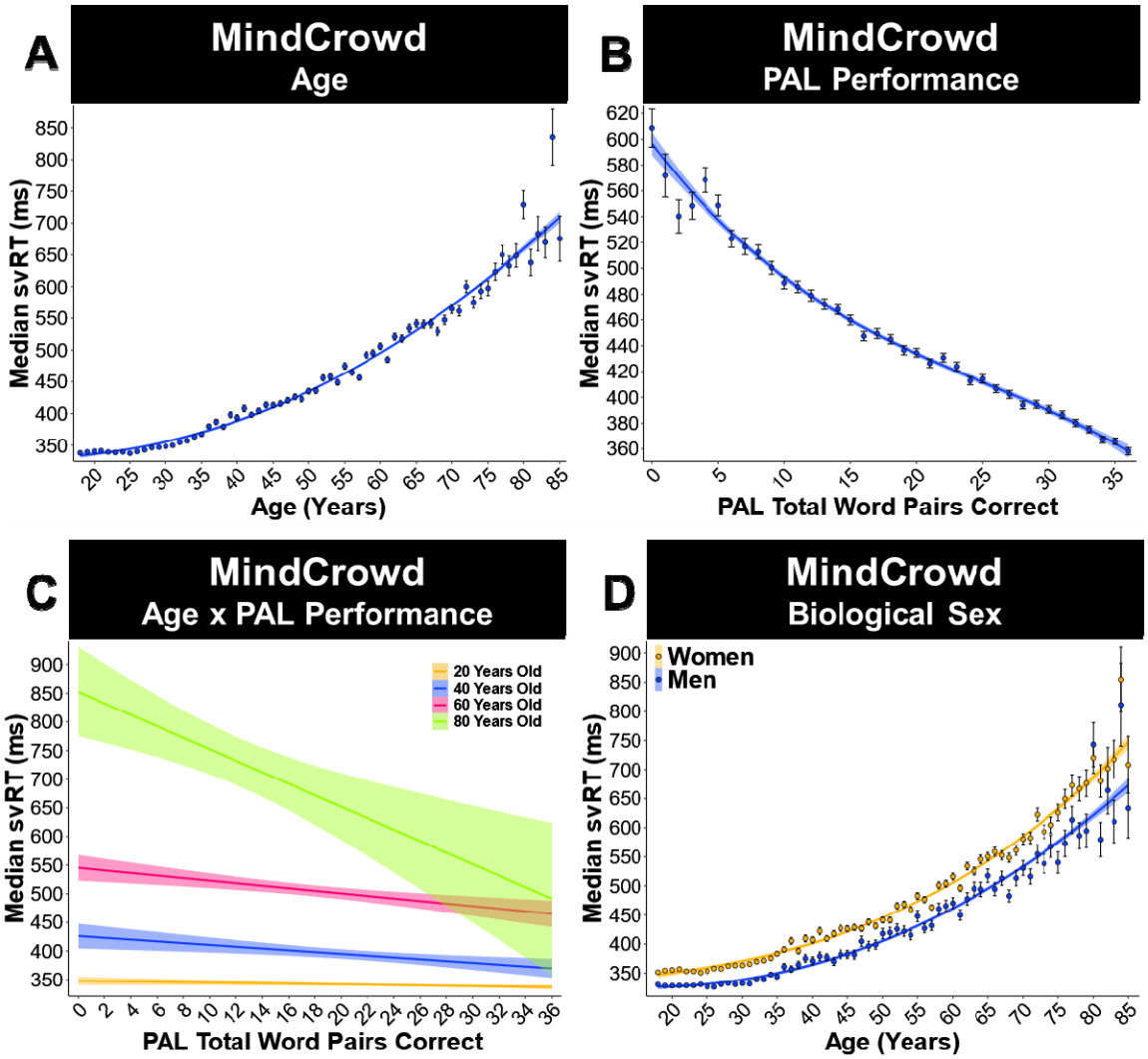

Figure 1. Age, paired associates learning (PAL) performance, and Biological Sex were associated with simple visual reaction time (svRT) in the MindCrowd analysis (ages 18-85). (A) Linear model fits (line fill ± 95% CI, error bars ± SEM) of the median svRT by Age3 (curvilinear model). There was a curvilinear relationship between svRT and Age^1 ($\beta_{Age}1$ = 7.07, $p_{Age}1$ = 3.06E-18), Age^2 (β_{Age2} = -0.15, p_{Age2} = 3.23E-17), and Age^3 (β_{Age3} = 1.47E-03, p_{Age3}< 1.46E-34, n = 75,666). (B) Linear model fits (line fill ± 95% CI, error bars ± SEM) of the median svRT by Age3 (curvilinear model). There was a curvilinear relationship between svRT and paired-associate learning (PAL) performance PAL $Performance^1$ (β_{PAL1} = -8.89, p_{PAL1} = 3.77E-34), PAL $Performance^2$ (β_{PAL2} = 0.32, p_{PALl2} = 2.35E-14), and PAL $Performance^3$ (β_{PAL3} = -4.13E-03, p_{PAL3} = 7.28E-09, n = 75,666). (C) Simple slope analysis of the Linear model fits (line fill ± 95% CI, error bars ± SEM) of the median svRT by PAL Performance ($\beta_{Age*PAL}$ = -0.07, $p_{Age*PAL}$ = 1.26E-59). At 20 ($\beta_{Age20*PAL}$ = -5.48, $p_{Age20*PAL}$ < 0.00, n = 1985), 40 ($\beta_{Age40*PAL}$ = -6.41, $p_{Age40*PAL}$ < 0.00, n = 739), 60 ($\beta_{Age60*PAL}$ = -7.75, $p_{Age60*PAL}$ < 0.00, n = 1789), and 80 ($\beta_{Age80*PAL}$ = -9.08, $p_{Age80*PAL}$ < 0.00, n = 344) years of age. (D) Linear model fits (line fill ± 95% CI, error bars ± SEM) of the median svRT by Age^3 (curvilinear model) with lines split by Biological Sex. Women demonstrated longer svRT compared to men from younger to older ages (β_{Sex} = -34.26, p_{Sex} = 1.26E-163, n_{Women} = 47,700, n_{Men} = 27,966).

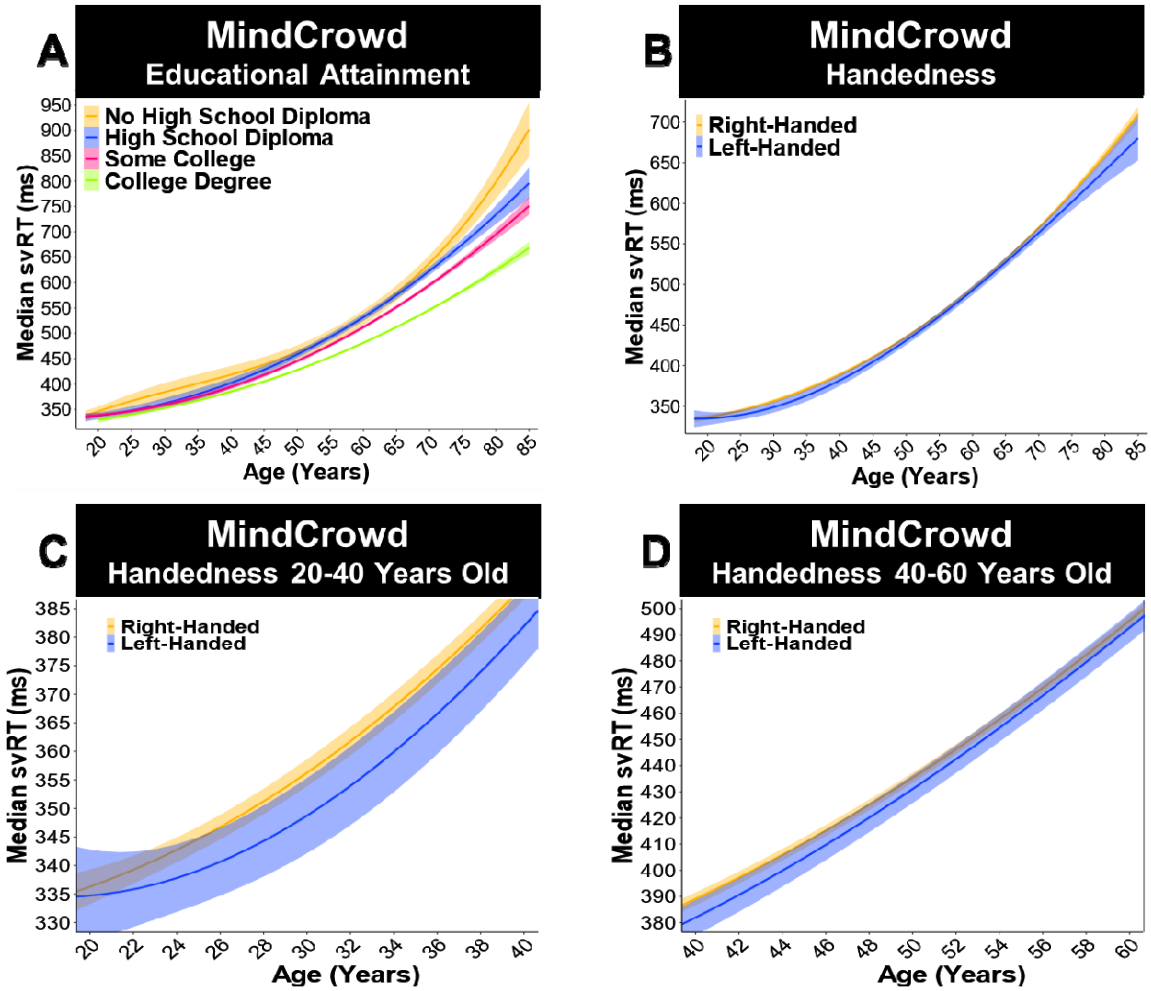

Figure 2. Educational attainment and handedness altered simple visual reaction time (svRT) performance. MindCrowd analysis (ages 18-85). (A) Linear model fits (line fill ± 95% CI) of the median svRT by Age^3 (curvilinear model) with lines split by Educational Attainment. Participants who had "Some College" ($\beta_{College}$ = -14.73, $p_{College}$ = 8.74E-05, n = 22,950), or a "College Degree" ($\beta_{CDegree}$ = -31.78, $p_{CDegree}$ = 2.95E-17, n = 44,140) were faster than those with "No High School Diploma" (n = 1,881). (B-D) Linear model fits (line fill ± 95% CI) of the median svRT by Age^3 (curvilinear model) with lines split by Handedness. (B) From 18-85 years old, left-handed participants showed slightly faster svRTs (β_{Left} = -3.87, p_{Left} < 0.03), (C) an association found in 20-40 years olds (β_{40Left} = -3.16, p_{40Left} < 0.01, Left-Handed n = 8,449), (D) but not 40-60 years olds (β_{60Left} = -2.69, p_{60Left} = 0.07, Right-Handed n = 66,903).

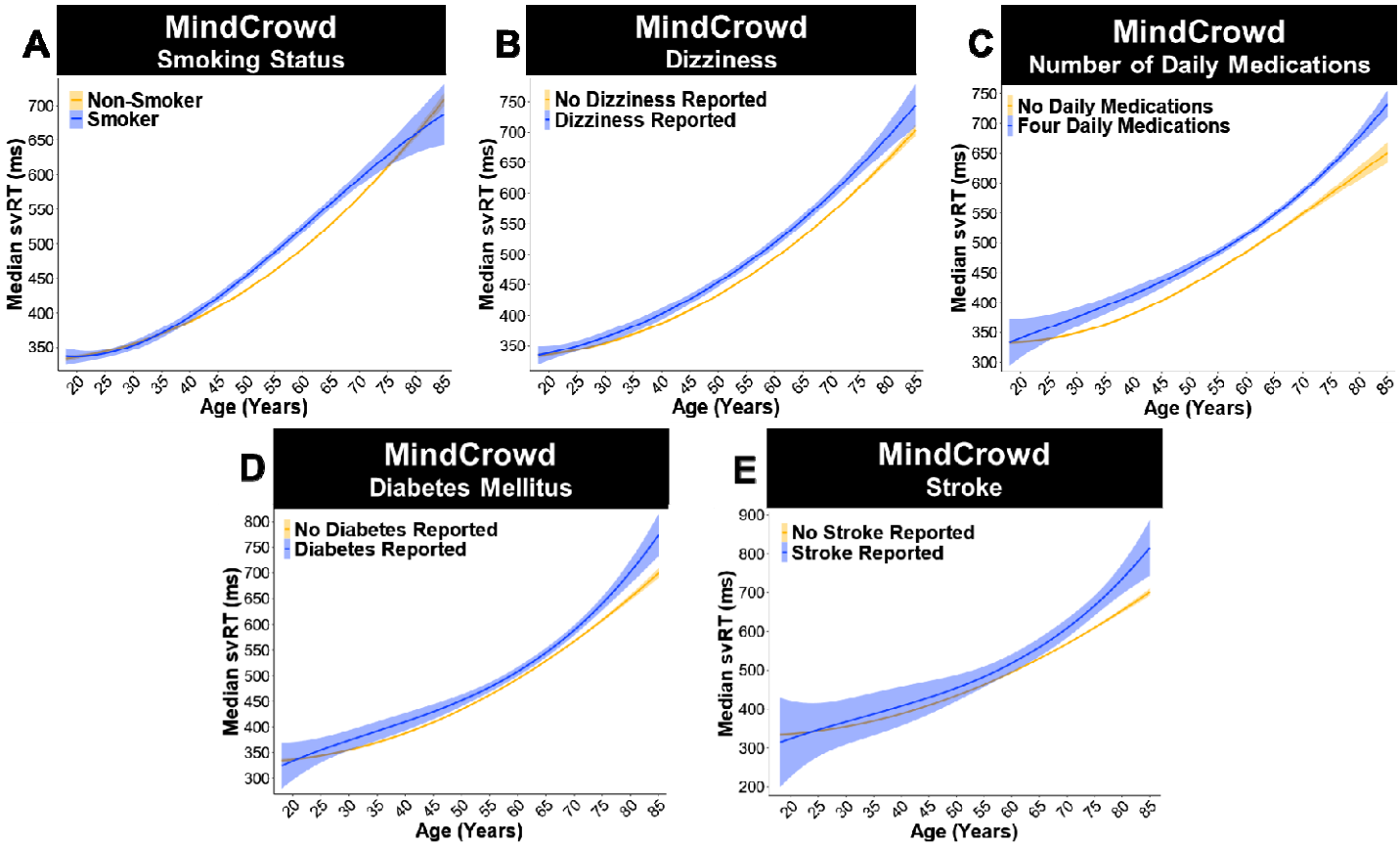

Figure 3. ***A-B*** Simple visual reaction times (svRT) were slower for those who smoked or Reported Dizziness. MindCrowd analysis (ages 18-85). (A) Linear model fits (line fill ± 95% *CI*) of the median svRT by Age^3 (curvilinear model) with lines split by Smoking Status. Participants identifying as a smoker showed slower svRT ($\beta_{Smoking}$ = 7.07, $p_{Smoking}$ = 1.26E-03, Smoker n = 5,793, Non-Smoker n = 69,873). (B) Linear model fits (line fill ± 95% *CI*) of the median svRT by Age^3 (curvilinear model) with lines split by Reported Dizziness. Participants reporting Reported Dizziness showed slower svRT performance (β_{Dizzy} = 4.87, p_{Dizzy} = 0.04, Reported Dizziness Reported n = 4,749, No Reported Dizziness Reported n = 70,917). ***C-E*** Health factors such as daily medications taken, diabetes, and Reported Stroke all slowed simple visual reaction time (svRT). MindCrowd analysis (ages 18-85). (C) Linear model fits (line fill ± 95% *CI*) of the median svRT by Age^3 (curvilinear model) with lines split by Daily Medications Taken. Compared to participants reporting "no daily medications" (n = 33,672), taking "Two" (β_{Meds2} = 5.84, p_{Meds2} = 2.00E-03, n = 9,651), "Three" (β_{Meds3} = 6.24, p_{Meds3}<0.01, n = 6,656), and "Four" (β_{Meds4} = 17.82, p_{Meds4} = 3.51E-16, n = 10,769) slowed svRT. (D) Linear model fits (line fill ± 95% *CI*) of the median svRT by Age^3 (curvilinear model) with lines split by diabetes mellitus. Participants reporting having diabetes had slower svRT performance ($\beta_{Diabetes}$ = 11.23, $p_{Diabetes}$<3.36E-05, Diabetes Reported n = 3,887, No Diabetes Reported n = 71,779). (E) Linear model fits (line fill ± 95% *CI*) of the median svRT by Age^3 (curvilinear model) with lines split by Reported Stroke. Reporting having had a Reported Stroke impaired (slowed) svRT performance (β_{Stroke} = 20.38, p_{Stroke} = 3.59E-04, Reported Stroke Reported n = 765, No Reported Stroke Reported n = 74,901).

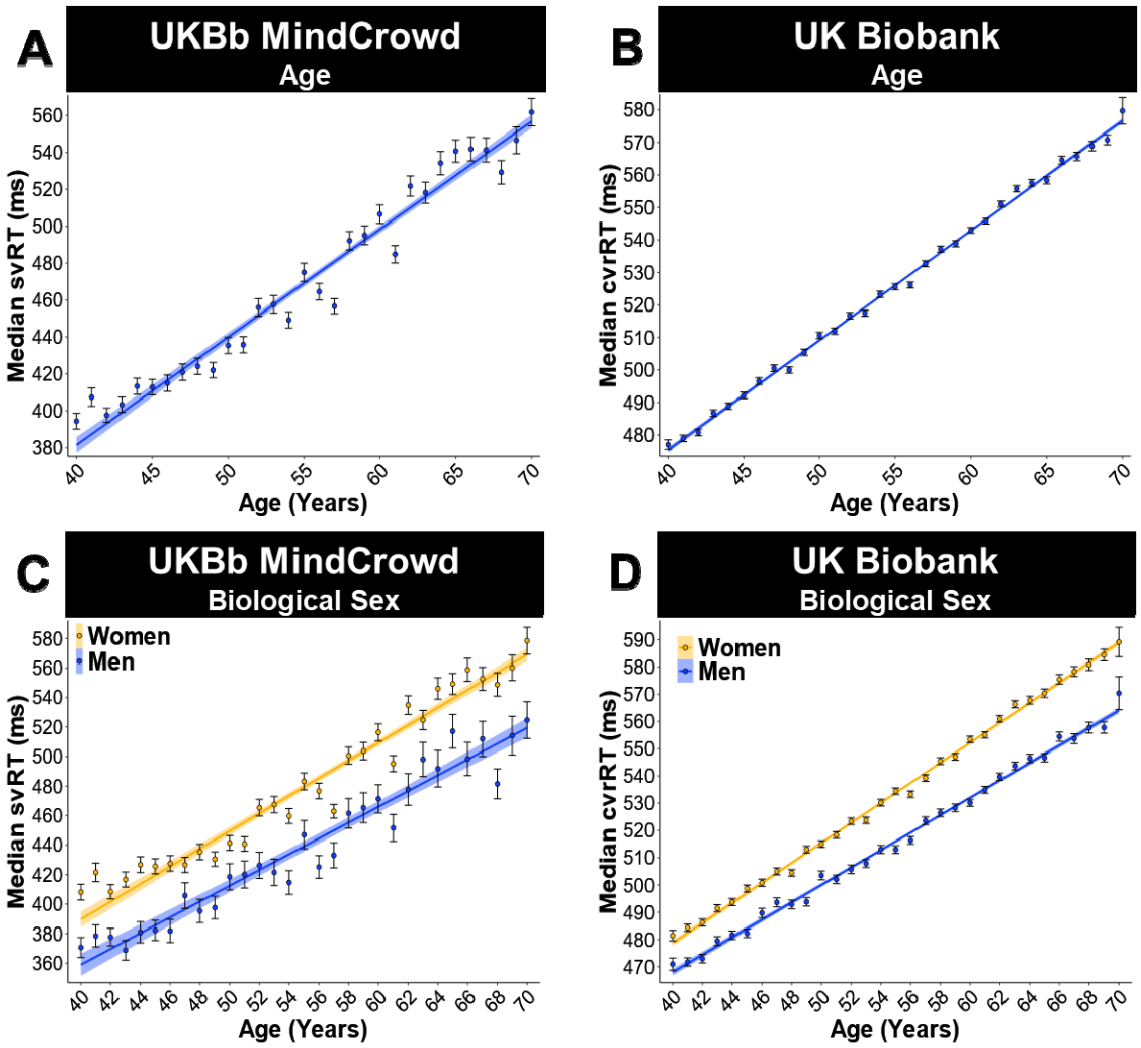

Figure 4. ***A-B*** Age was linearly associated with visual reaction time (RT).UK Biobank analysis (ages 40-70). (A) UKBb MindCrowd linear model fits (line fill ± 95% *CI*, error bars ± *SEM*) of median simple visual RT (svRT) across Age. (B) UK Biobank linear model fits (line fill ± 95% *CI*, error bars ± *SEM*) of median complex, visual, recognition RT (cvrRT) across Age. UKBb MindCrowd svRT (β_{Age} = 5.75, p_{Age} = 2.00E-16, n = 39,795) and UK Biobank cvrRT (β_{Age} = 3.40, p_{Age} = 2.00E-16, n = 158,245) showed similar slowing from younger to older ages, but overall higher RT. The average 50 msec difference between UKBb MindCrowd svRT (M = 478.66 msec) and UK Biobank (M = 528.74 msec) is likely due to the choice component (i.e., do cards match or not > press button) of the UK Biobank's cvrRT task compared to UKBb MindCrowd's stimulus-response (i.e., the pink sphere appears > press button) svRT. ***C-D*** Visual reaction time (RT) was faster in men. UK Biobank analysis (ages 40-70). (C) UKBb MindCrowd linear model fits (line fill ± 95% *CI*, error bars ± *SEM*) of median simple visual RT (svRT) across Age with lines split by Biological Sex. (D) UK Biobank linear model fits (line fill ± 95% *CI*, error bars ± *SEM*) of median complex, visual, recognition RT (cvrRT) across Age with lines split by Biological Sex. Both UKBb MindCrowd svRT (β_{Sex} = -40.00, p_{Sex} = 8.03E-71, 20.46%, Women [M = 489.75 msec, n = 29,640], Men [M = 446.28 msec, n = 10,155]) and UK Biobank cvrRT (β_{Sex} = -18.28, p_{Sex} = 2.00E-16, 5.16%, Women [M = 534.98 msec, n = 89,331], Men [520.66 msec, n = 68,914) found men consistently outperforming women from 40-70 years of age.

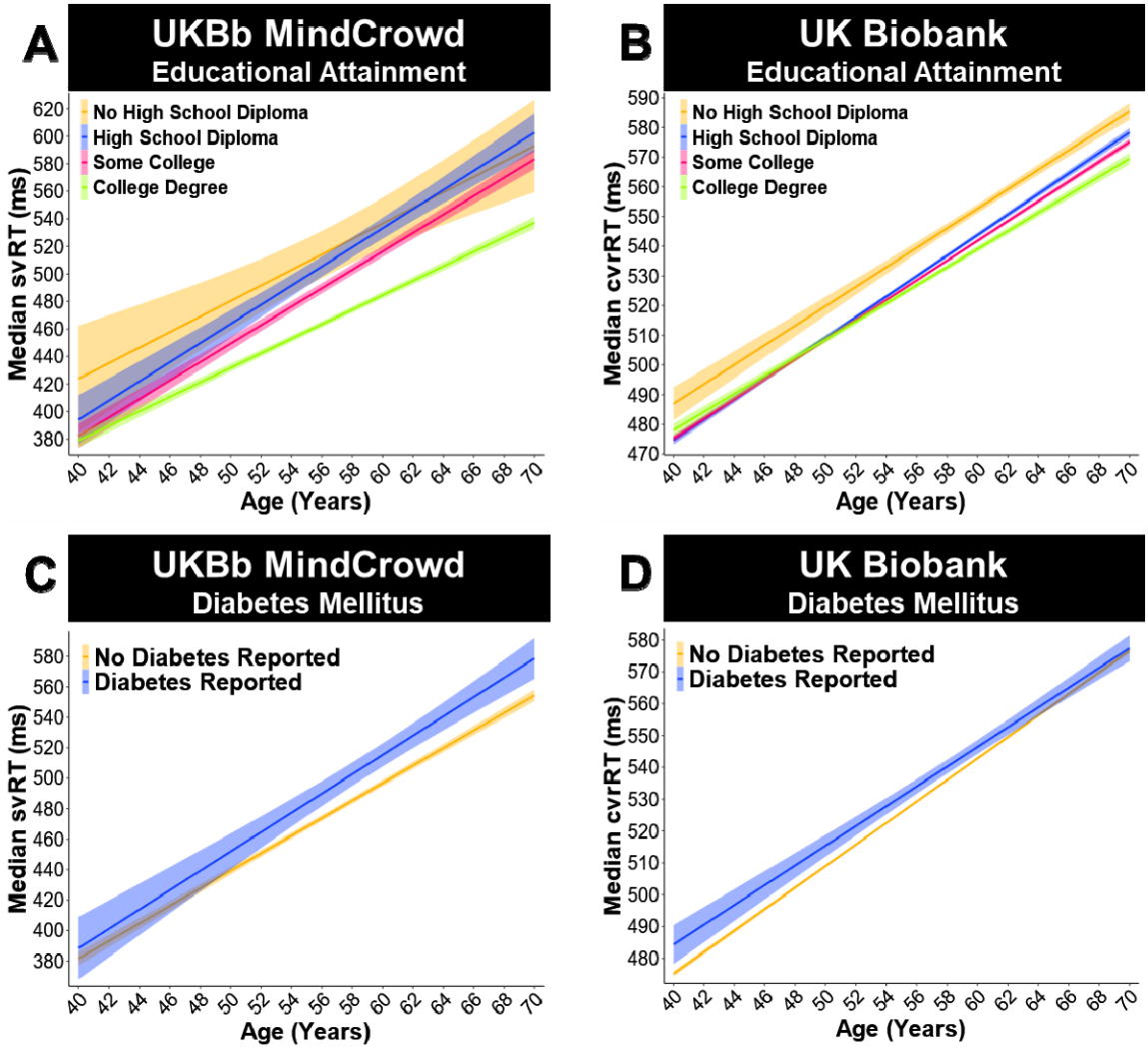

Figure 5. ***A-B*** More education was related to faster visual reaction time (RT). UK Biobank analysis (ages 40-70). (A) UKBb MindCrowd linear model fits (line fill ± 95% *CI*) of median simple visual RT (svRT) across Age with lines split by Educational Attainment. Participants who had a "High School Diploma" ($\beta_{HSDiploma}$ = -9.68, $p_{HSDiploma}$ = 2.70E-01, 1.87%, *n* = 3,176), "Some College" ($\beta_{College}$ = -26.08, $p_{College}$ = 1.61E-03, 5.06%, *n* = 11,139), or a "College Degree" ($\beta_{CDegree}$ = -49.07, $p_{CDegree}$ = 1.91E-09, 9.51%, *n* = 24,875) were faster than those with "No High School Diploma" (*n* = 605). (B) UK Biobank linear model fits (line fill ± 95% *CI*) of median complex, visual, recognition RT (cvrRT) across Age with lines split by Educational Attainment. Like the UKBb MindCrowd cohort, participants who had a "High School Diploma" ($\beta_{HSDiploma}$ = -9.68, $p_{HSDiploma}$ = 1.23E-33, 1.74%, *n* = 46,247), "Some College" ($\beta_{College}$ = -10.26, $p_{College}$ = 1.30E-40, 1.85%, *n* = 77,270), or a "College Degree" ($\beta_{CDegree}$ = -11.71, $p_{CDegree}$ = 1.46E-41, 2.11%, *n* = 23,750) were faster than those with "No High School Diploma" (*n* = 10,978). ***C-D*** Diabetes mellitus slowed visual reaction time (RT). UK Biobank analysis (ages 40-70). (C) UKBb MindCrowd linear model fits (line fill ± 95% *CI*) of median simple visual RT (svRT) across Age with lines split by diabetes mellitus. (D) UK Biobank linear model fits (line fill ± 95% *CI*) of median complex, visual, recognition RT (cvrRT) across Age with lines split by diabetes mellitus. Participant visual RT performance for UKBb MindCrowd svRT ($\beta_{Diabetes}$ = 11.48, $p_{Diabetes}$ = 3.31E-03, Diabetes Reported *n* = 2,807, No Diabetes Reported *n* = 36,988) and UK Biobank cvrRT ($\beta_{Diabetes}$ = 5.48, $p_{Diabetes}$ = 4.80E-07, Diabetes Reported *n* = 4,969, No Diabetes Reported *n* = 153,276), was slowed with Diabetes Mellitus.

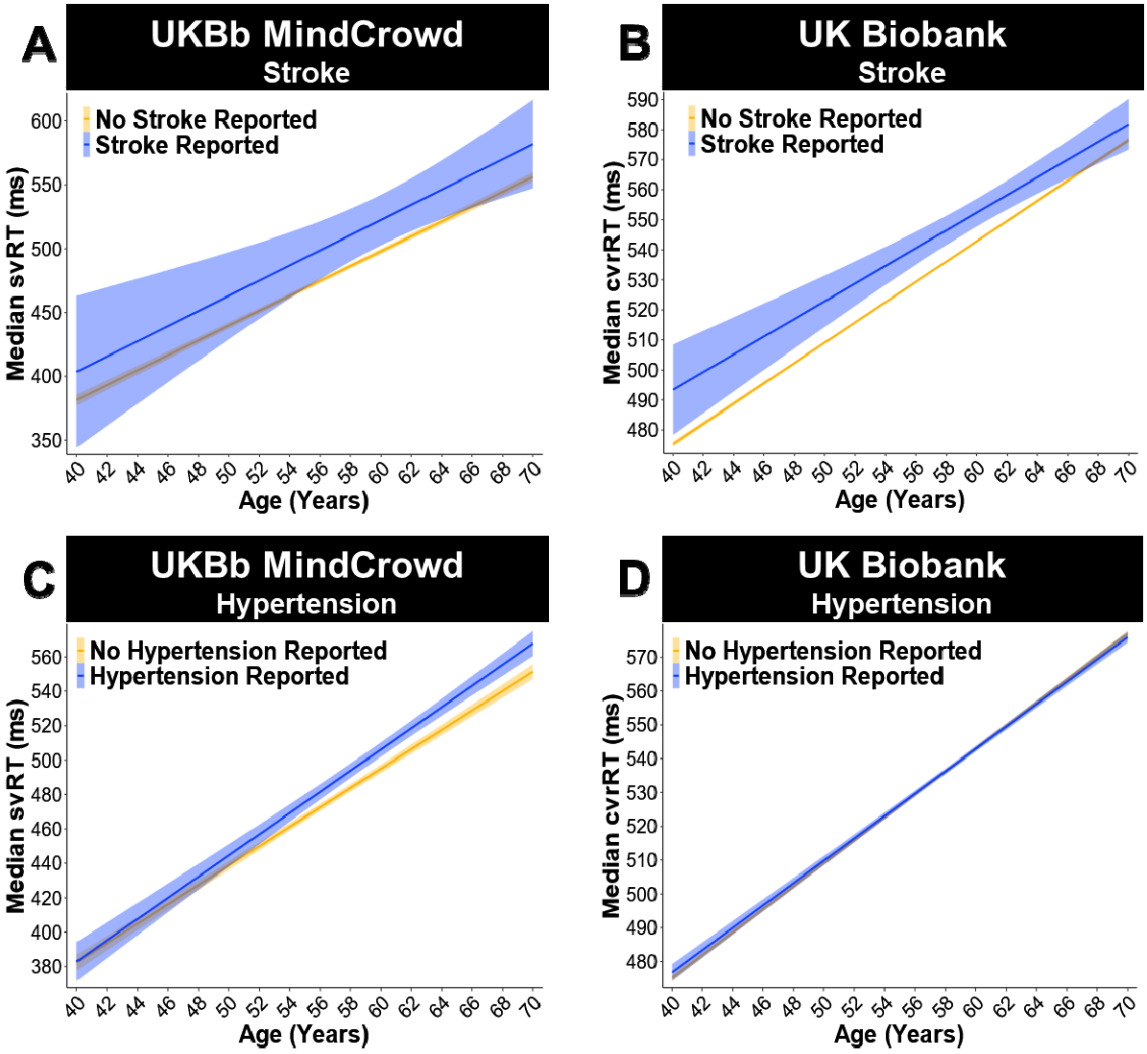

Figure 6. ***A-B*** Visual reaction time (RT) was impaired if a Reported Stroke was reported. UK Biobank analysis (ages 40-70). (A) UKBb MindCrowd linear model fits (line fill ± 95% *CI*) of median simple visual RT (svRT) across Age with lines split by Reported Stroke. (B) UK Biobank linear model fits (line fill ± 95% *CI*) of median complex, visual, recognition RT (cvrRT) across Age with lines split by Reported Stroke. In both the UKBb MindCrowd svRT (β_{Stroke} = 18.47, p_{Stroke} = 4.00E-02, Reported Stroke Reported n = 2,807, No Reported Stroke Reported n = 39,318) and UK Biobank cvrRT (β_{Stroke} = 10.61, p_{Stroke} = 6.15E-07, Reported Stroke Reported n = 1,237, No Reported Stroke Reported n = 157,008) analysis, experiencing a Reported Stroke was associated with slower visual RT. ***C-D*** Reported Hypertension slowed both types of visual reaction time (RT). UK Biobank analysis (ages 40-70). (C) UKBb MindCrowd linear model fits (line fill ± 95% *CI*) of median simple visual RT (svRT) across Age with lines split by Reported Hypertension. (D) UK Biobank linear model fits (line fill ± 95% *CI*) of median complex, visual, recognition RT (cvrRT) across Age with lines split by Reported Hypertension. Unlike the MindCrowd analysis, hypertension was related to slower svRT in UKBb MindCrowd ($\beta_{Hypertension}$ = 7.99, $p_{Hypertension}$ = 3.16E-03, Hypertension Reported n = 9,676, No Hypertension Reported n = 30,119) and cvrRT in the UK Biobank ($\beta_{Hypertension}$ = 1.14, $p_{Hypertension}$ = 0.02, Hypertension Reported n = 32,593, No Hypertension Reported n = 125,652).

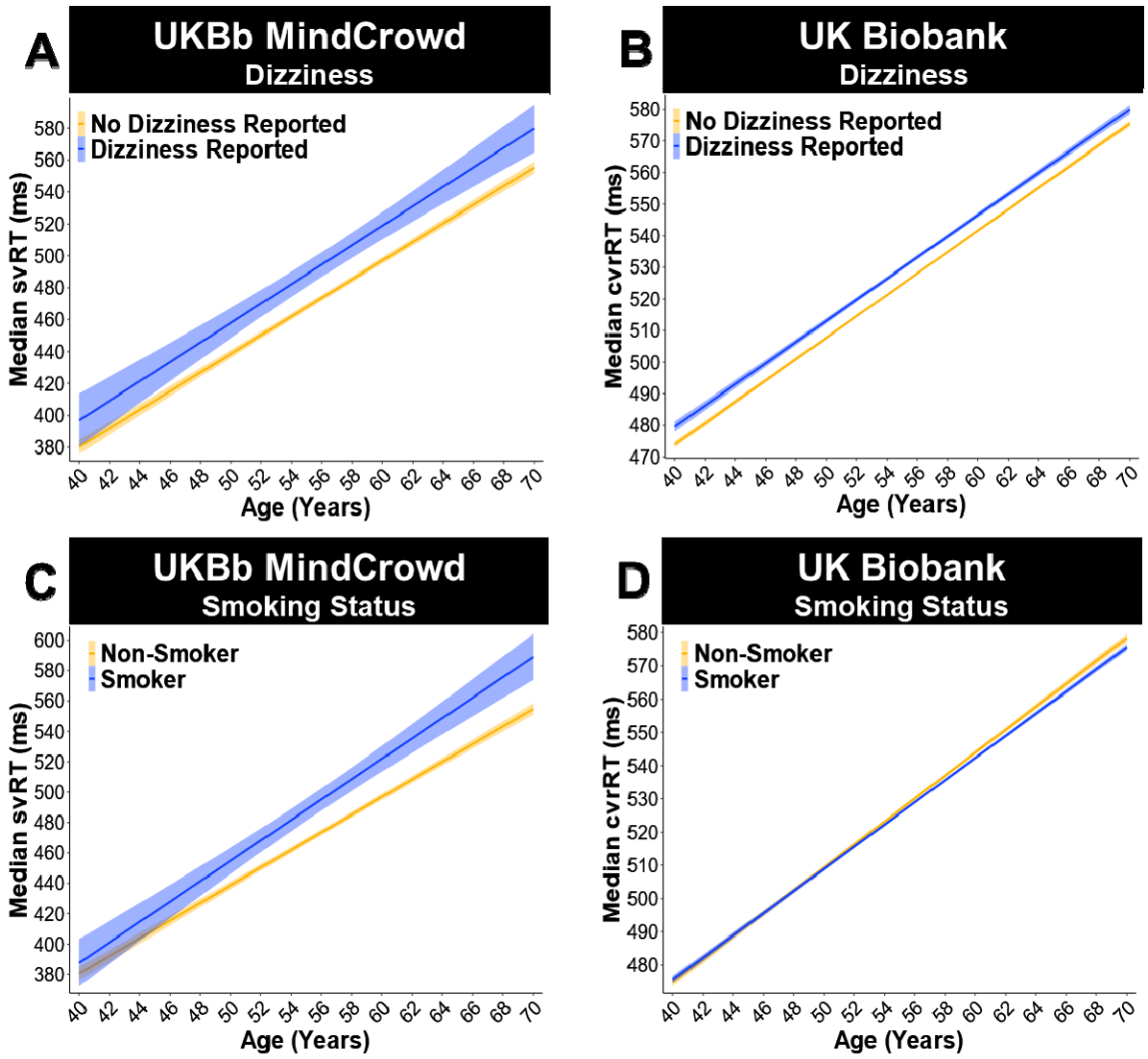

Figure 7. ***A-B*** Reported Dizziness impaired visual reaction time (RT). UK Biobank analysis (ages 40-70). (A) UKBb MindCrowd linear model fits (line fill ± 95% *CI*) of median simple visual RT (svRT) across Age with lines split by Reported Dizziness. (B) UK Biobank linear model fits (line fill ± 95% *CI*) of median complex, visual, recognition RT (cvrRT) across Age with lines split by Reported Dizziness. UKBb MindCrowd svRT (β_{Dizzy} = 12.13, p_{Dizzy} = 2.52E-03, Reported Dizziness Reported *n* = 2,543, No Reported Dizziness Reported *n* = 37,252) and UK Biobank cvrRT (β_{Dizzy} = 3.21, p_{Dizzy} = 3.71E-14, Reported Dizziness Reported *n* = 42,210, No Reported Dizziness Reported *n* = 116,035) were slowed if participants reported Reported Dizziness. ***C-D*** Smoking Status is related to impaired MindCrowd simple visual reaction time (svRT), not UK Biobank complex, visual recognition reaction time (cvrRT). UK Biobank analysis (ages 40-70). (C) UKBb MindCrowd linear model fits (line fill ± 95% *CI*) of median svRT across Age with lines split by Smoking Status. Compared to non-smokers, smokers demonstrated slower svRT (β_{Smoke} = 10.18, p_{Smoke} = 0.01, Smoker *n* = 2,783, Non-Smoker *n* = 37,012). (D) UK Biobank linear model fits (line fill ± 95% *CI*) of median cvrRT across Age with lines split by Smoking Status. An association between Smoking Status and cvrRT was not found in the UK Biobank (β_{Smoke} = 0.47, p_{Smoke} = 0.22, Smoker *n* = 91,312, Non-Smoker *n* = 66,923).

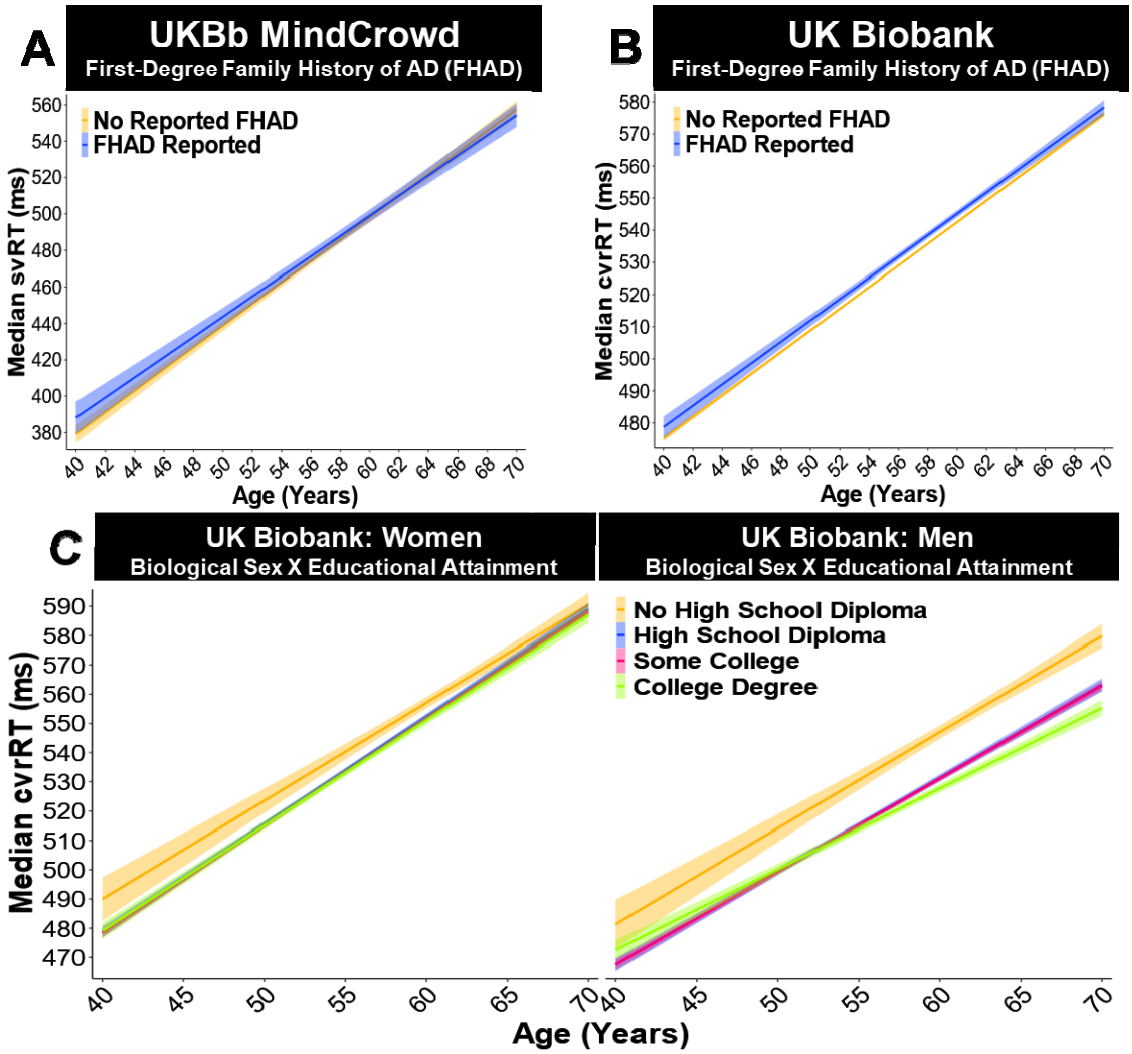

Figure 8. ***A-B*** A first-degree family history of Alzheimer's disease (FHAD) was related to slower UK Biobank complex, visual recognition reaction time (cvrRT), UKBb MindCrowd simple visual reaction time (svRT). UK Biobank analysis (ages 40-70). (A) UKBb MindCrowd linear model fits (line fill ± 95% *CI*) of median svRT across Age with lines split by reported FHAD. An association between FHAD and svRT was not found in the UKBb MindCrowd cohort (β_{FHAD} = -0.07, p_{FHAD} = 0.97, FHAD Reported *n* = 13,748, No FHAD Reported *n* = 26,047). (B) UK Biobank linear model fits (line fill ± 95% *CI*) of median cvrRT across Age with lines split by reported FHAD. Compared to those reporting No FHAD, FHAD was related to worse cvrRT performance in the UK Biobank (β_{FHAD} = 2.36, p_{FHAD} = 3.35E-05, FHAD Reported *n* = 19,741, No FHAD Reported *n* = 138,504). ***C*** In the UK Biobank, biological sex modified the association of educational attainment and complex, visual, recognition reaction time (cvrRT) (ages 40-70). Linear model fits (line fill ± 95% *CI*) of the median cvrRT by Age with lines split by Educational Attainment and Biological Sex. Compared to women having "No High School Diploma" (*n* = 6,056), men with a "High School Diploma" ($\beta_{Sex*HSDiploma}$ = -11.24, $p_{Sex*HSDiploma}$ = 3.30E-12, *n* = 18,505), "Some College" ($\beta_{sex*College}$ = -10.54, $p_{sex*College}$ = 8.71E-12, *n* = 34,230) or a "College Degree" ($\beta_{sex*CDegree}$ = -12.8, $p_{sex*CDegre}$ = 1.68E-13, *n* = 11,257) were all significantly associated with faster cvrRT. See Supplementary Figure 7, displaying simple effects parsed using estimated marginal means (*EMM*).

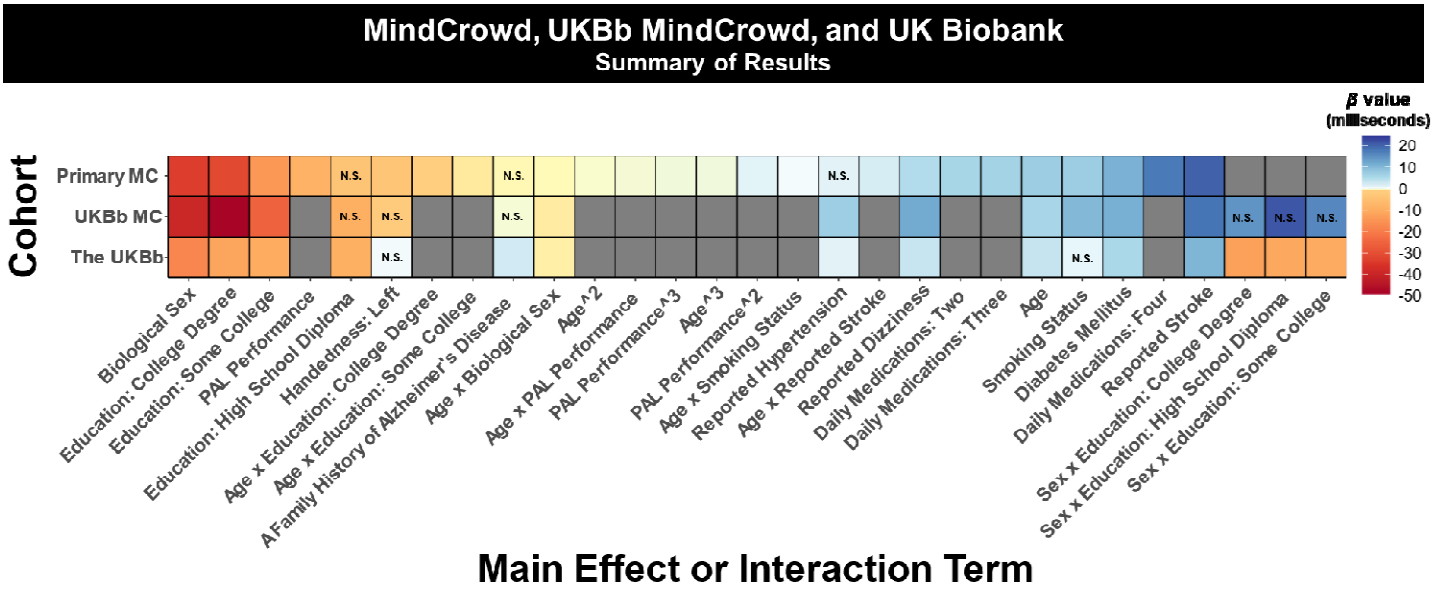

Figure 9. An illustrative summary of the results. Data are shown across the MindCrowd (MC), UKBb MindCrowd, and the UK Biobank (UKBb). The color (i.e., red = low negative and blue = high positive) indicates the size of the β (beta coefficient) estimate." N.S.," indicates if the estimated β value was not statistically significant (α = 0.05).

Table 2

MindCrowd Cohort Sample Sizes (n) for Each Factor

Factor	***n***
Biological Sex	Women = 47,700 Men = 27,966
Educational Attainment	No High School Diploma = 1,881 High School Diploma = 6,695 Some College = 22,950 College Degree = 44,140
Handedness	Left-Handed = 8,449 Right-Handed = 66,903
Daily Medications Taken	None = 33,672 One = 14,409 Two = 9,651 Three = 6,656 Four =10,769
Reported Dizziness	Dizziness Reported = 4,749 No Dizziness Reported = 70,917
Smoking Status	Smoker = 5,793 Non-Smoker = 69,873
Diabetes Mellitus	Diabetes Reported = 3,887 No Diabetes Reported = 71,779
Reported Stroke	Stroke Reported = 765 No Stroke Reported = 74,901

Table 2. Summary of sample size numbers (*n*) of each multiple regression coefficients from MindCrowd.

Table 3
MindCrowd Summary of Results

Factor	Value	β	*SE*	*t*	*p*
Intercept		355.88	11.99	29.68	1.89E-192
Age[1]	Per Year Change in Slope	7.07	0.81	8.71	3.06E-18
Age[2]	Per Year Change in How Much the Slope Is Altered	-0.15	0.02	-8.44	3.23E-17
Age[3]	Per Year Change in How Much the Slope Is Altered	1.20E-04	0	12.27	1.46E-34
PAL Performance[1]	PAL Word Pairs Correctly Remembered	-8.89	0.73	-12.19	3.77E-34
PAL Performance[2]	Per Word Pair Change in Slope	0.32	0.04	7.63	2.35E-14
PAL Performance[3]	Per Year Change in How Much the Slope Is Altered	7.14E-04	0	-5.78	7.28E-09
Biological Sex	Men	-34.26	1.25	-27.33	1.26E-163
Educational Attainment	Some College	-14.73	3.75	-3.92	8.74E-05
Educational Attainment	College Degree	-31.78	3.76	-8.45	2.95E-17
Handedness	Left	-3.87	1.76	-2.2	0.03
Daily Medications Taken	Two	5.84	1.89	3.09	2.00E-03
Daily Medications Taken	Three	6.24	2.26	2.76	0.01
Daily Medications Taken	Four	17.82	2.18	8.16	3.51E-16
Reported Dizziness	TRUE	4.87	2.35	2.07	0.04
Smoking Status	TRUE	7.07	2.19	3.23	1.26E-03
Diabetes Mellitus	TRUE	11.23	2.71	4.15	3.36E-05
Reported Stroke	TRUE	20.38	5.71	3.57	3.59E-04
Age x PAL Performance	Trajectory of age-related change in svRT performance is altered buy PAL Performance	-7.56E-04	4.41E-05	-1.72E+01	8.34E-66
Age x Smoking Status	Smoking Status alters the trajectory of age-related changes to svRT	0.58	0.12	4.67	2.95E-06
Age x Reported Stroke	Reported Stroke adds to age-related decrements in svRT	1.87	0.42	4.46	8.17E-06
Age x Biological Sex	Men have a shifted trajectory of age-related svRT changes	-0.37	0.07	-5.59	2.23E-08
Age x Educational Attainment	Having Some College vs. No High School Diploma alters the trajectory of age-related alterations to svRT	-0.76	0.18	-4.09	4.28E-05
Age x Educational Attainment	A College Degree vs. No High School Diploma alters the trajectory of age-related modifications to svRT performance	-1.53	0.19	-8.14	4.08E-16

Table 3. Summary of the main findings from MindCrowd, listing main associations and interactions. β = unstandardized regression coefficient, *SE* = standard error, *t* = value of *t*-statistic, *p* = *p*-value.

Table 4

UKBb MindCrowd and UK Biobank Cohort Sample Sizes (n)

Characteristic	**UKBb MindCrowd *n*s**	**UK Biobank *n*s**
Biological Sex	Women = 29,640 Men = 10,155	Women = 89,331 Men n = 68,914
Diabetes Mellitus	Diabetes Reported = 2,807 No Diabetes Reported = 36,988	Diabetes Reported = 4,969 No Diabetes Reported = 153,276
Handedness	Left-Handed = 4,520 Right-Handed = 35,034	Left-Handed = 15,287 Right-Handed = 142,958
Reported Stroke	Reported Stroke Reported = 2,807 No Reported Stroke Reported = 39,318	Reported Stroke Reported = 1,237 No Reported Stroke Reported = 157,008
Reported Hypertension	Hypertension Rep. = 9,676 No Hypertension Rep. = 30,119	Hypertension Rep. = 32,593 No Hypertension Rep. = 125,652
Smoking Status	Smoker = 2,783 Non-Smoker = 37,012	Smoker = 91,312 Non-Smoker = 66,923
Reported Dizziness	Reported Dizziness Reported = 2,543 No Reported Dizziness Reported = 37,252	Reported Dizziness Reported = 42,210 No Reported Dizziness Reported = 116,035
Educational Attainment	No High School Diploma = 605 High School Diploma = 3,176 Some College = 11,139 College Degree = 24,875	No High School Diploma = 10,978 High School Diploma = 46,247 Some College = 77,270 College Degree = 23,750
FHAD	FHAD Reported = 13,748 No FHAD Reported = 26,047	FHAD Reported = 19,741 No FHAD Reported = 138,504
Educational Attainment x Biological Sex	Women No High School Diploma = 426 High School Diploma = 2,496 Some College = 8,740 College Degree = 17,978 Men No High School Diploma = 179 High School Diploma = 680 Some College = 2,399 College Degree = 12,493	Women No High School Diploma = 6,056 High School Diploma = 27,742 Some College = 43,040 College Degree = 23,750 Men No High School Diploma = 4,922 High School Diploma = 18,505 Some College = 34,230 College Degree = 11,257

Table 4. Summary of sample size numbers (*n*s) of each multiple regression coefficient from the UKBb MindCrowd and UK Biobank.

Table 5
UKBb MindCrowd and the UK Biobank Summary of Results

		UKBb MindCrowd				The UK Biobank				Outcome
Variable	**Value**	**β**	***SE***	***t***	***p***	**β**	***SE***	***t***	***p***	
Intercept		189.77	10.65	18.35	6.46E-75	335.93	1.69	209.93	2.00E-16	Agree
Age	Per Year	5.75	0.12	46.57	2.00E-16	3.40	0.03	132.87	2.00E-16	Agree
Biological Sex	Men	-40.00	2.24	-17.83	8.03E-71	-18.28	0.38	-47.70	2.00E-16	Agree
Diabetes Mellitus	TRUE	11.48	3.91	2.94	3.31E-03	5.48	1.09	5.03	4.80E-07	Agree
Handedness	Left	-2.58	3.05	-0.85	4.00E-01	0.58	0.63	0.92	0.36	N.S. in both - but MindCrowd subtle effect
Reported Stroke	TRUE	18.47	9.02	2.05	4.00E-02	10.61	2.13	4.99	6.15E-07	Agree
Reported Hypertension	TRUE	6.99	2.37	2.95	3.16E-03	1.09	0.48	2.27	2.00E-02	Agree
Smoking Status	TRUE	10.18	3.87	2.63	0.01	0.40	0.38	1.05	0.29	MindCrowd Finding: Both Primary and UKBb
Reported Dizziness	TRUE	12.10	4.00	3.02	2.52E-03	3.21	0.42	7.57	3.71E-14	Agree
Educational Attainment	High School Diploma	-9.68	8.77	-1.10	2.70E-01	-9.68	0.80	-12.09	1.23E-33	Agree
Educational Attainment	Some College	-26.08	8.27	-3.15	1.61E-03	-10.26	0.77	-13.35	1.30E-40	Agree
Educational Attainment	College Degree	-49.07	8.17	-6.01	1.91E-09	-11.71	0.87	-13.51	1.46E-41	Agree
First-Degree Family History of Alzheimer's Disease	TRUE	-0.07	2.06	-0.04	0.97	2.43	0.57	4.26	2.04E-05	UK Biobank Finding
Age x Biological Sex	Men's vs. Women's per year change in Age slope	-0.62	0.26	-2.38	2.00E-02	-0.55	0.05	-10.99	4.18E-28	Agree
Biological Sex by Educational Attainment	Men with High School Diploma vs. Women with No High School Diploma	21.63	19.62	1.10	0.27	-11.24	1.61	-6.96	3.30E-12	UK Biobank Finding
Biological Sex by Educational Attainment	Men with Some College vs. Women with No High School Diploma	16.25	18.28	0.89	0.37	-10.54	1.54	-6.83	8.71E-12	UK Biobank Finding
Biological Sex by Educational Attainment	Men with a College Degree vs. Women with No High School Diploma	14.79	17.94	0.82	0.41	-12.8	1.74	-7.37	1.68E-13	UK Biobank Finding

Table 5. Summary of the main findings from the UKBb MindCrowd and the UK Biobank cohorts. Here factor names, factor definitions, β = unstandardized regression coefficients as well as estimate *SE* = standard error, *t* = value of *t*-statistic, and *p* = *p*-value across both cohorts is shown. The last column displays model outcomes in terms of the agreement between the cohorts.

Table 6

UK Biobank Factor Generation

Index #	Field Description	Field ID	Count (*n*)	Type	Description	Column Label	Used to Generate
6	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=31	31	502,524	Categorical (single)	Sex	Sex	Biological Sex
7	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=34	34	502,524	Integer	Year of birth	Birth Year	Age
264	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=404	404	202,464	Integer	Duration to the first press of snap-button in each round	Snap-Button Time to First Press (Instance 0)	Median cvrRT (8 out 12 Rounds)
395	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=1707	1707	501,626	Categorical (single)	Handedness (chirality/laterality)	Handedness	Handedness
422	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=2443	2443	501,593	Categorical (single)	Diabetes diagnosed by doctor	Diabetes mellitus	Diabetes mellitus
426	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=2453	2453	501,593	Categorical (single)	Cancer diagnosed by doctor	Cancer	Cancer
482	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=2966	2966	134,645	Integer	Age high blood pressure diagnosed	Age Hypertension Diagnosed	Hypertension Diag. via MD
616	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=4056	4056	7,579	Integer	Age Reported Stroke diagnosed	Age Reported Stroke Diagnosed	Reported Stroke Diag. via MD
2104	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=20107	20107	487,790	Categorical (single)	Illnesses of father	(Multiple Columns)	FHAD
2144	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=20110	20110	492,928	Categorical (single)	Illnesses of mother	(Multiple Columns)	FHAD
2188	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=20111	20111	433,922	Categorical (single)	Illnesses of siblings	(Multiple Columns)	FHAD
2325	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=20116	20116	501633	Categorical (single)	Smoking status	Smoking status	Smoking Status
2356	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=21022	21022	502,524	Integer	Age at recruitment	Age at Recruitment	Age (verify)
120	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=401	401	495,702	Categorical (single)	Index for card A in round	Needed to Know if Snap-Button Should Have Been Pressed	Median cvrRT
168	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=402	402	495,702	Categorical (single)	Index for card B in round	Needed to Know if Snap-Button Should Have Been Pressed	Median cvrRT
1635	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=6138	6138	497,883	Categorical (multiple)	Qualifications	Education	Educational Attainment
2357	http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=21053	21053	174774	Categorical (single)	Degree bothered by dizziness in the last three months	Dizziness	Dizziness

Table 6. List of derived factors from the UK Biobank. UK Biobank's "Index Number," "Field Description," "Field ID" and sample size (*n*s) for data field used to generate factors evaluated via multiple linear regression. "Data fields" were chosen to match to the questions posed in MindCrowd. If multiple instances (e.g., 0-4, timepoints) were available for data, only the first (instance 0) was used. For cvrRT and FHAD, multiple columns were compiled to generate the factor. For the UKBb MindCrowd and UK Biobank analyses,

participants were removed if any of their responses to a demographic, medical, health, and lifestyle question did not match the other cohort.

Table 1

Sociodemographic Characteristics of Each Cohort

Cohort	**Descriptive or Factor Level**	***n***	**%**
1. MindCrowd 18-85yr: Age	*M* = 46.54 *SD* = 18.40	75,666	100
UKBb MindCrowd 40-70yr: Age	*M* = 56.59 *SD* = 8.16	39,795	100
UK Biobank 40-70yr: Age	*M* = 55.81 *SD* = 7.72	158,249	100
2. MindCrowd 18-85yr: Biological Sex	Women	47,700	63.08
	Men	27,966	36.91
UKBb MindCrowd 40-70yr: Biological Sex	Women	29,640	74.51
	Men	10,155	25.49
UK Biobank 40-70yr: Biological Sex	Women	89,333	56.45
	Men	68,916	43.55
3. MindCrowd 18-85yr: Race	Native American	445	0.6
	Asian	3,491	4.74
	Black/African American	1,555	2.11
	Native Hawaiian/Pacific Islander	270	0.37
	Mixed	490	0.66
	White	67,433	91.52
UKBb MindCrowd 40-70yr: Race	Asian	743	1.90
	Black/African American	732	1.87
	Mixed	182	0.47
	White	37,235	95.03
UK Biobank 40-70yr: Race	Asian	1,612	1.02
	Black/African American	979	0.62
	Mixed	847	0.54
	White	154,809	97.83
4. MindCrowd 18-85yr: FHAD	TRUE	17,382	23.59
	FALSE	56,302	76.41
UKBB MindCrowd 40-70yr: FHAD	TRUE	13,554	34.59
	FALSE	25,627	65.41
UK Biobank 40-70yr: FHAD	TRUE	19,741	12.48
	FALSE	138,506	87.52
5. MindCrowd 18-85yr: Handedness	Left	8,276	11.23
	Right	65,408	88.77
UKBb MindCrowd 40-70yr: Handedness	Left	4,474	11.42
	Right	34,707	88.58
UK Biobank 40-70yr: Handedness	Left	15,287	9.66
	Right	142,960	90.34
6. MindCrowd 18-85yr: Educational Attainment	No High School Diploma	1,792	2.43
	High School Diploma	6,470	8.78
	Some College	22,335	30.31
	College Degree	43,087	58.48
UKBb MindCrowd 40-70yr: Educational Attainment	No High School Diploma	577	1.47
	High School Diploma	3,108	7.93
	Some College	10,955	27.96
	College Degree	24,541	62.63
UK Biobank 40-70yr: Educational Attainment	No High School Diploma	10,979	6.94
	High School Diploma	46,247	29.23
	Some College	77,271	48.83
	College Degree	23,750	15.01

Table 1. Sociodemographic Characteristics Across Each of the Three Cohorts. Study *n*s and related percentages of commonly reported sociodemographic factors from each of the three cohorts (i.e., MindCrowd and UKBb MindCrowd and the UK Biobank)**.**