On the use of growth models to study normal cognitive aging

Ghisletta, Paolo; Mason, Fabio; von Oertzen, Timo; Hertzog, Christopher; Nilsson, Lars-Göran; Lindenberger, Ulman

How to cite

GHISLETTA, Paolo et al. On the use of growth models to study normal cognitive aging. In: International Journal of Behavioral Development, 2020, vol. 44, n° 1, p. 88–96. doi: 10.1177/0165025419851576

This publication URL: https://archive-ouverte.unige.ch/unige:130713

Publication DOI: 10.1177/0165025419851576

On the Use of Growth Models to Study Normal Cognitive Aging

Abstract

Growth models (GM) of the mixed-effects and latent curve varieties have become popular methodological tools in lifespan research. One of the major advantages of GM is their flexibility in studying individual differences in change. We scrutinized the change functions of GM used in five years of publications on cognitive aging. Of the 162 publications that we identified, 88% test linear or quadratic polynomials, and fewer than 5% apply functions that are nonlinear in their parameters, such as exponential decline. This apparent bias in favor of polynomial decomposition calls for exploring what conclusions about individual differences in change are likely to be drawn if one applies linear or quadratic GMs to data simulated under a conceptually and empirically plausible model of exponential cognitive decline from adulthood to old age. Hence, we set up a simulation that manipulated the rate of exponential decline, measurement reliability, number of occasions, interval width, and sample size. True rate of decline and interval width influenced results strongly, number of occasions and measurement reliability exerted a moderate effect, and the effects of sample size appeared relatively minor. Critically, our results show that fit statistics generally do not differentiate misspecified linear or quadratic models from the true exponential model. Moreover, power to detect variance in change for the linear and quadratic GMs is low, and estimates of individual differences in level and change can be highly biased by model misspecification. We encourage researchers to also consider plausible nonlinear change functions when studying behavioral development across the lifespan.

Abstract: 249 words

Running Head: Growth Models in Normal Cognitive Aging

Keywords: Growth Model; long-term change; normal cognitive aging; nonlinear mixed-effects models; longitudinal research designs.

Text, references, figures, tables: ~~5844~~ 6409 words (max=6000)

On the Use of Growth Models to Study Normal Cognitive Aging

The last 25 years have seen major advances in statistical models for characterizing developmental change (e.g., Hertzog & Nesselroade, 2003; Little, 2013; Singer & Willett, 2003). One popular method is the growth model (GM), implemented as a linear mixed-effects model (LMEM; Bryk & Raudenbush, 1987; Laird & Ware, 1982) or a latent curve model (McArdle, 1988; Meredith & Tisak, 1990). The critical feature of the GM is that it models change by specifying an intercept and a slope component describing the overall level and change across time, respectively. For both components, one can estimate the mean parameter and variance parameter reflecting individual deviations around the mean. Given individual differences in level and/or in change, the GM can be readily expanded to test antecedents and/or correlates of such individual differences.

The Growth Model

The *linear* GM specifies a variable Y_{ij} for an individual j at time i $(i=0,1,\ldots,T-1)$ in terms of a level π_{0j} and a linear rate of growth π_{1j} multiplied by age a_{ij} (often centered around its mean) at time i for each individual j, plus a residual E_{ij}:

$$\begin{aligned} Y_{ij} &= \pi_{0j} + \pi_{1j} a_{ij} + E_{ij} \\ \pi_{0j} &= \beta_0 + U_{0j} \\ \pi_{1j} &= \beta_1 + U_{1j} \end{aligned} \qquad (1)$$

Field Cod

The linear GM predicts all individual growth trajectories to be straight lines. The level corresponds to the intercept (prediction when a_{ij}=0) and the rate of growth to the linear slope (predicted change in Y for one unit change in a). Both level and change have a mean value of β_0 and β_1 and a subject-specific deviation from the mean of U_{0j} and U_{1j}, respectively (Bryk & Raudenbush, 1987; Laird & Ware, 1982). β_0 and β_1 are fixed effects and U_{0j} and U_{1j} are random effects, assumed to be normally distributed around zero (with variances σ^2_L and σ^2_C) and to possibly covary (σ_{LC}). The errors E_{ij} are typically assumed independently and normally distributed with time-invariant error variance σ^2_E.

Another popular specification of the GM is the *quadratic GM* of Equation (2) that allows the predicted trajectories to have a single change in direction. This model augments the linear GM by adding a second (quadratic) change component that linearly relates age squared to Y_{ij}:

$$\begin{aligned} Y_{ij} &= \pi_{0j} + \pi_{1j} a_{ij} + \beta_2 a_{ij}^2 + E_{ij} \\ \pi_{0j} &= \beta_0 + U_{0j} \\ \pi_{1j} &= \beta_1 + U_{1j} \end{aligned} \qquad (2)$$

Compared to the linear model, the quadratic additionally estimates one additional parameter, β_2, which allows the predicted trajectories to follow a quadratic polynomial. Although the quadratic change component could also be specified with random effects, their estimation is typically not included in extant applications (in our literature review, of the 40 records that tested a quadratic model, only nine declared having tested random effects of the quadratic slope, and of those only three found positive evidence).

Linear and quadratic GMs are often contrasted with a simpler *no growth* or *level model* that postulates no change whatsoever, with individuals maintaining the same level score over time (Equation (3); Widaman & Thompson, 2003):

$$\begin{aligned} Y_{ij} &= \pi_{0j} + E_{ij} \\ \pi_{0j} &= \beta_0 + U_{0j} \end{aligned}, \qquad (3)$$

to test the hypothesis that Y_{ij} changes over time.

The Growth Model as a Nonlinear Mixed-Effects Model

So far, the specified GMs represent change as a linear combination of level, change(s), and errors, and are thus linear in their parameters (Davidian & Giltinan, 1995). The linearity of the parameters is not to be confused with the linearity of the shape of specified change (Cudeck & Harring, 2007). In particular, the quadratic GM specifies a nonlinear trajectory of change, despite being a LMEM.

LMEMs can be generalized to allow for nonlinear relations between Y_{ij} and the parameters of the growth components (Davidian & Giltinan, 1995), generating nonlinear mixed-effects models (NLMEMs). These models have been successfully applied in various disciplines (e.g., pharmacokinetics, biology, demography, finance), to capture underlying change processes that are inherently nonlinear (e.g., drug effects and population growth often follow a logistic function, accumulation of financial interests is often exponential). One can argue that many psychological constructs change nonlinearly over time. For instance, effects of accumulating age-related insults to an adult's central nervous system could be modeled by exponential decline. Indeed, nonlinear GM have occasionally been applied to study adult cognitive change, both normal (e.g., McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Grimm, McArdle, & Hamagami, 2007) and pathological (e.g., Driscoll, Resnick, Troncoso, An, O'Brien, & Zonderman, 2006).

One variant of the (nonlinear) exponential model, as employed by several authors (e.g., McArdle et al., 2002; Grimm et al., 2007) is:

$$\begin{aligned} Y_{ij} &= \pi_{0j} + \pi_{1j}\, e^{\gamma a_{ij}} + E_{ij} \\ \pi_{0j} &= \beta_0 + U_{0j} \\ \pi_{1j} &= \beta_1 + U_{1j} \end{aligned} \qquad (4)$$

This nonlinear mixed-effects GM specifies an exponential functional form of change, so that as age increases successive gains or losses are not constant but proportional to the current value. Compared to the linear GM, this model estimates one additional parameter, the exponential rate of change γ. Typically, random effects for γ are not estimated in extant applications. However, because of random effects of π_{1j}, which are estimated as the variance of change σ^2_C, individuals are allowed to differ in the amount of exponential decline, ~~but not~~ rather than in its rate.

The Present Study

We start by reviewing five years of scientific literature on applications of the GM to study normal cognitive aging. This review established that the linear and quadratic GMs are very often applied by substantive researchers in this field, whereas nonlinear mixed-effects GMs, on the other hand, are very rarely used.

Justified by this pattern, we evaluated what happens if longitudinal data, created by an underlying exponential model, are analyzed with a linear or quadratic GM. We investigate how well the linear and quadratic GMs preserve or distort individual differences in level and change in such instances. If the individual differences specified by the underlying generating model are largely concordant with those estimated by the misspecified GMs, one would be justified to claim that LMEM GMs are practically useful despite their misspecification. That is, the models would be of the "wrong but useful" kind (cf. Box & Draper, 1987). If, however, the two sets of results lead to discordant inferences about individual differences in change, we would conclude that the misspecified GMs are misleading.

Method

Literature Review

We computed a detailed literature search to study how longitudinal data are commonly analyzed within GMs to study normal cognitive aging. By covering five years of literature we initially obtained 1253 records, which reduced to 162 relevant records for full analysis. The detailed search and screening procedure are described in supplementary material (Appendices 1-2, Table 1, and Figure 1).

Simulation Design

Population model. The exponential population model of Equation (4) was used to generate the data for the Monte Carlo simulation (McArdle, et al., 2002). We specified the existence of individual differences in both level and change, but not in decline rate, γ. This choice was guided by the facts that (a) in our empirical illustration (see below) estimating random effects for the exponential rate failed, (b) extant applications of the exponential model usually exclude random effects for γ, and, more importantly, (c) estimating multiple random effects in both linear and especially NLMEMs requires data denser than those of standard longitudinal panels.

Simulation factors. We based our simulation study on parameter estimates from an exponential model fitted to data from the Betula Project (Nilsson, Bäckman, Erngrund, Nyberg, Adolfsson, Bucht, Karlsson et al., 1997), a well-known study of aging, memory, and dementia. We analyzed the episodic recall scores of 1000 Betula participants assessed four times, at five-year intervals, as a function of age (final range: 35-95 years). We tested the level, linear, quadratic, and exponential GMs and found that the exponential best described the data. We used the estimated population parameters as the basis for the simulation experiment (for full results, see supplementary Table 2 and supplementary Figure 2).

The constant factors were the level mean (β_0 = 20.603), change mean (β_1 = -4.007), level variance (σ^2_L = 30.980), change variance (σ^2_C = 5.041), and level-change covariance (σ_{LC} = -3.163, corresponding to a correlation of -.253). The varying factors were the rate of exponential decline (γ= 0.066, corresponding to the illustration, and γ= 0.033, for a gentler decline) and the Growth Curve Reliability (GCR = .500, .681 - cf. illustration -, or .900, corresponding to σ^2_E = 30.980, 14.541 - cf. illustration -, and 3.442, respectively), defined as the proportion of level variance over total variance at time=0 (age 65; GCR=σ^2_L / (σ^2_L + σ^2_E); McArdle, 1988).

We generated data under six conditions, 2 (γ) by 3 (GCR), to obtain continuous trajectories spanning over a maximum of 71 time points (i.e., from 35 to 105 years). Figure 1 presents 50 random individual trajectories simulated under the six different conditions.

INSERT FIGURE 1 ABOUT HERE

Design factors. In practice, longitudinal studies of long-term change processes often adopt a longitudinal sequence design, which begins with a cross-sectional age-heterogeneous sample of persons and then follows them longitudinally over time. This sampling design generates multiple parallel subsamples differing in initial age, each providing longitudinal data on different segments of the targeted overall age range, with coverage increasing as the number of measurement occasions increases or as the measurement interval widens (Baltes, Reese, & Nesselroade, 1988). We included 10 longitudinal subsamples with 50, 100, or 200 units each, yielding a total sample size of N = 500, 1000, or 2000. Each subsample contained observations representing individuals of initial age 35, 40, 45, etc. to 80 years, observed repeatedly either T = 3 or T = 6 times. Furthermore, we varied Δ, the interval width between adjacent occasions of measurement, going from Δ = 1 to Δ = 5 years in steps of 1 year.

This designs mimics a study where the youngest subsample observed the least (T = 3) at the shortest interval (Δ = 1) would be assessed at ages 35, 36, and 37 years, the next subsample at ages 40, 41, and 42, etc. and the oldest subsample at ages 80, 81, and 82 years. The youngest subsample observed the most (T = 6) at the shortest interval (Δ = 1) would be assessed at ages 35, 36, 37, 38, 39, and 40 years, thus overlapping with the next subsample (ages 40, 41, 42, 43, 44, and 45). The oldest subsample observed the most (T = 6) at the shortest interval (Δ = 1) would have ages 80, 81, 82, 83, 84, and 85, while if Δ increases to 5, the ages would be 80, 85, 90, 95, 100, and 105 years, ultimately covering the entire adult life span. Crossing low numbers of repeated observations (T = 3 or 6) with very short to relatively long intervals within repeated observations (Δ = 1 to 5) more closely mimics existing longitudinal studies than assuming that all units have been observed during the full length of the study and at all occasions (e.g., Schaie & Hofer, 2001). Our illustration from the Betula study counts 10 subsamples, each of 100 individuals (hence N = 1000), of initial age 35, 40, 45, 50, 55, 60, 65, 70, 75, or 80 years, assessed on 4 occasions (T=4), with an interval between adjacent occasions of $\Delta = 5$ years.

Summary. In total, we fully crossed the varying simulation factors and design features to obtain 180 conditions: 2 (γ) by 3 (GCR) by 2 (T) by 5 (Δ) by 3 (N) (see Supplementary Table 3 for a full description of the simulation design). For each condition we generated 100 data sets, for a total of 18000 data sets.

Procedure

Growth model analyses. For each generated data set we estimated a series of GMs. We first applied the exponential decline GM. By doing so we tested whether the estimation procedure recovered the values of the generating function despite sampling limitations on number of occasions and varying interval lengths, and also provided a basis for comparing results obtained from the alternative GMs. We subsequently implemented the level, linear, and quadratic GMs. For the level GM we estimated 3 parameters (β_0, σ^2_L, and σ^2_E; cf. Equation (3)); for the linear GM, 6 parameters (β_0, β_1, σ^2_L, σ^2_C, σ_{LC}, and σ^2_E; cf. Equation (1)); and for the quadratic GM, 7 parameters (β_0, β_1, β_2, σ^2_L, σ^2_C, σ_{LC}, and σ^2_E; cf. Equation (2)); finally, for the exponential decline GM, we estimated 7 parameters (β_0, β_1, γ, σ^2_L, σ^2_C, σ_{LC}, and σ^2_E; cf. Equation (4)).

In all GM analyses we centered the time predictor a_{ij} (age) around its grand (sample) mean, to reduce estimation bias in the level-change covariance (σ_{LC}), and in the level variance (σ^2_L; Mehta & West, 2000; Rovine & Molenaar, 1998; Wainer, 2000).

The level, linear, and quadratic GMs are LMEMs, for which the marginal likelihood can be expressed and estimated without approximation. However, because the exponential model is a NLMEM, we had to apply an approximation method (Davidian & Giltinan, 1995). To rule out that the estimation procedure acted as a generator of divergent results we applied the same estimation procedure to all models. We used PROC NLMIXED in SAS (v. 9.3) to estimate all GMs with the adaptive Gaussian quadrature approximation, which produces the same estimates as maximum likelihood in linear models (Littell, Milliken, Stroup, & Wolfinger, 1996).

Acceptable solutions. We deemed a solution acceptable only if (a) the estimation converged to a solution and (b) the solution contained parameter estimates that were statistically viable. In particular, the estimated variance of change (σ^2_C) had to be non-negative.

Goodness of fit. From each analysis we saved the Schwarz Bayesian Information Criterion (BIC), defined as

$$BIC = 2 \cdot f\left(\hat{\theta}\right) + p \cdot \ln\left(N\right),$$

where f is the negative of the marginal log-likelihood function, $\hat{\theta}$ is the vector of parameter estimates, p is the number of estimated parameters, ln is the natural logarithm, and N is the number of subjects (Schwarz, 1978). This index is not normed, but allows comparing alternative models tested on the same data. A difference in BIC values of more than 10 between competing models can be interpreted as strong evidence in favor of the model with the lower BIC value (Kass & Raftery, 1995).

Power to detect individual differences in change. We estimated the power to detect interindividual differences in change by comparing each GM to a nested GM where the variance in change (σ^2_C) and the level-change covariance (σ_{LC}) were fixed at zero. We then compared each fully specified GM to its nested counterpart, and computed a likelihood ratio test ($\alpha = 5\%$).

Concordance in individual differences. We evaluated the validity of individual difference estimates from each model by inspecting the concordance of individual differences in growth components between the generated data and the results obtained from the GM analyses. To generate the data according to the exponential decline model of Equation (4) we first generated random effects in level and change scores (i.e., U_{0j} and U_{1j} of Equation (4)) in accordance to the constant simulation factors (cf. Supplementary Table 3). We then combined these random growth scores according to Equation (4) and added the randomly generated time-specific error component (E_{ij}), to obtain the individual repeated observations (Y_{ij}).

We analyzed each simulated data set of individual repeated observations with the four GMs and each time saved the empirical Bayes estimates of the random effects of the level and the change component. To inspect the concordance of the individual differences in growth components, we correlated the random level and change scores used to generate the data with the empirical Bayes estimates of the random effects in level and change obtained from the GM analyses.

Results

Literature Review

The literature search produced 1253 records, and after exclusion of duplicate and irrelevant records, we fully analyzed 162 records (see Supplementary Table 1). Of the 162, 102 records (63%) tested only a linear function of change, whereas 40 (25%) tested also a quadratic function, and one a cubic function. Thirteen (8%) records tested a broken-stick function (a.k.a. single-node spline; ten of which compared it to a linear or quadratic function). Of the remaining records, five (3%) tested an exponential function, two (1%) estimated the change function from the data (within a latent curve model), one used a multiple-node spline approach, and one did not specify the change function. See Supplementary Appendix 3 for detailed results.

Overall, then, 142 (88%) of the 162 records either used the linear or the quadratic function, whereas eight (less than 5%) used a function that is nonlinear in its parameters.

Simulation Results

Number of acceptable solutions. Generally, the greater the sample, the number of repeated measurements, the interval of measurements, the variable's reliability, and the rate of exponential decline, the higher the number of acceptable solutions (AS). The percentages of AS across models were: 100, 93.53, 85.57, and 97.74 for the level (which does not estimate σ^2_C), linear, quadratic and exponential model, respectively. Supplementary Figure 3 displays the average number of AS as a function of the design features.

To aid understanding of these effects, we used logistic regression to predict the probability of obtaining an AS as a function of the simulation factors. Model type most strongly influenced the probability of an AS ($\sim R^2$ = 13%-25%), GCR, T, and Δ have low effects ($\sim R^2$ = 3%-8%), and sample size was the least influential factor ($\sim R^2 < 1\%$). See supplementary Appendix 4 for detailed results.

Goodness of ~~F~~fit. We differenced BIC values of the level, linear, and quadratic GMs to against the exponential GM's BIC and compared this index of evidence favoring the exponential as a function of simulation condition and design features. The rejection rates based on BIC differences for the level, linear, and quadratic GMs were, respectively, 100%, 99.955%, and 96.295% for $\gamma = 0.066$ and 99.884%, 69.625%, and 47.191% for $\gamma = 0.033$. Thus, for a shallow declining process, the quadratic model would be judged to fit the data as well as the exponential more than half the time. Supplementary Figure 4 displays this information in detail.

We computed a logistic regression predicting the odds that the linear and quadratic GMs be rejected in favor of the exponential model as a function of GCR, T, Δ, and N, for $\gamma = 0.033$ (and not for $\gamma = 0.066$ or the level GM, because of the extremely high rejection rates). The number of occasions (T), and interval width (Δ) were the most influential factors ($\sim R^2$ = 13%-25%), GCR had a low effect ($\sim R^2$ = 8%-16%), and N was rather unimportant ($\sim R^2$ = 3%-12%). See supplementary Appendix 5 for detailed results.

Power to ~~D~~detect ~~I~~individual ~~D~~differences in ~~C~~change. Supplementary Figure 5 shows the average proportions of times the nested analogous model with $\sigma^2_C = \sigma_{LC} = 0$ was rejected compared to the fully specified linear, quadratic, and exponential model, only for $\gamma = 0.033$. For $\gamma = 0.066$ these rejection curves are virtually at 1 across all other conditions. Generally, the rejection curve of each model (hence estimated power) increased monotonically as Δ increases, with greater values at GCR = .900, $T = 6$, and Δ greater than 3. As GCR decreased, the rejection curves needed wider Δ to achieve high values, especially when $T = 3$. With the lowest GCR, the quadratic model had the least power to detect variance in change, especially when $T = 3$ and with small Δ. The most flagrant disparity between the quadratic model and the others was visible when $T = 3$, $N = 2000$, GCR = .500, and $\Delta = 1$: the power estimated for the quadratic GM in that cell is .105, while for the linear and exponential models it was .625 and .612, respectively.

To understand further the effects influencing detection of change variance we computed a logistic regression for each GM predicting the probability of a significant likelihood ratio test as a function of GCR, T, Δ, and N, for $\gamma = 0.033$. For all models the strongest factor was again Δ (~R^2 = 10%-20%), followed by T for the linear and quadratic models (~R^2 = 7%-14%), and by N for the exponential (~R^2 = 8%-13%). In the linear and exponential models, GCR influenced power the least (~R^2 = 2%-8%), while in the quadratic it was N (~R^2 = 3%-5%). See supplementary Appendix 6 for detailed results.

Concordance in ~~I~~individual ~~D~~differences in ~~L~~level. We computed analyses of variance (ANOVAs) predicting correlations between the level scores used to generate the data and the Bayes estimates of levels from each GM (in all ANOVAs the residuals were normally distributed, centered on zero, and homoscedastic). A first ANOVA testing only main effects obtained that they were significant, although their sizes were moderate to strong for type of model (partial eta squared $\eta_p^2 = .536$), γ ($\eta_p^2 = .464$), Δ ($\eta_p^2 = .112$), GCR ($\eta_p^2 = .159$), but virtually nil for *T* and *N* (both η_p^2 <.001). All two-way interactions about *N* were irrelevant (all η_p^2 = .0001). However, some two-way interactions about *T* were associated to small, yet nonzero, effect sizes. In the end, we display the results graphically with respect to all effects but *N*.

Figure 2a displays the average level correlations for γ= 0.033, by $\varDelta$ (abscissa), *T* (rows) and GCR (columns). In general, all indices are around .8. There appeared to be little effect of the model and of $\varDelta$ when GCR = .500 and *T* = 3, but when both *T* and GCR increased, the level model produces lower concordance rates, especially as $\varDelta$ increased. The correlations of the linear and quadratic models were hardly distinguishable, and they both diminished as $\varDelta$ increased (especially when *T* = 6). For the exponential model the concordance rates remain stable or even rise as $\varDelta$ increased. An ANOVA for the γ= 0.033 condition obtained the following effect size estimates (η_p^2): .308, .097, and .333 for type of model, $\varDelta$, and GCR, and less than .002 for *T* and *N*.

INSERT FIGURE 2a ABOUT HERE

The disparities in concordance rates across the models increased dramatically when γ= 0.066 (cf. Figure 2b). Already when *T* = 3 and GCR = .500, the exponential model produced the highest level correlations. As $\varDelta$ increases the correlations diminished for all models except for the exponential, for which they remained stable or grow even further, approaching 1.0, whereas for the level, linear, and quadratic they dropped at, respectively, .2, .6, and .6-.7. An ANOVA for the γ= 0.066 condition obtained the following effect size estimates (η_p^2): .821, .273, and .244 for type of GM, $\varDelta$, and GCR, and less than .001 for *T* and *N*.

INSERT FIGURE 2b ABOUT HERE

Concordance in ~~I~~individual ~~D~~differences in ~~C~~change. We repeated the same ANOVAs with respect to individual differences in change information. An ANOVA with main effects only obtained moderate to strong effects for all factors but sample size (for type of GM, γ, $\varDelta$, GCR, and T, the η_p^2 estimates were .195, .568, .527, .281, and .472, respectively; for N, η_p^2 = .002). N was also irrelevant with respect to all two-way interactions (all η_p^2 <.001). We again graphically display all effects but sample size, in Figure 3a for γ= 0.033 and 3b for γ= 0.066, respectively.

With γ= 0.033, concordance indices started very low (around .2) also for the exponential model, and only increased considerably as a function of wider $\varDelta$ when GCR increased. The three models obtain similar correlations, except with GCR = .900 and T = 6, where the linear model performed slight worse than the quadratic and exponential models. An ANOVA for the condition γ= 0.033 obtained the following effect size estimates (η_p^2): .014, .599, .389, and .569 for type of GM, $\varDelta$, GCR, and T and less than .005 for N.

INSERT FIGURE 3a ABOUT HERE

When γ= 0.066, in general the lowest correlations started around .4 and increased with wider $\varDelta$. When T = 3, the quadratic and exponential models are again hardly discernable, while the linear model obtains lower correlations. With T = 6, however, starting at $\varDelta$ = 5, 4, and 3, for GCR = .500, .681, and .900, respectively, the exponential model clearly obtained stronger concordance indices about change than the quadratic model (reaching .9 vs. .6). An ANOVA for the condition γ= 0.066 obtained the following effect size estimates (η_p^2): .509, .570, .275, and .487 for type of GM, $\varDelta$, GCR, and T and less than .002 for N.

INSERT FIGURE 3b ABOUT HERE

Prediction Quality in the Betula Project

To evaluate one possible implication of the application of the various GMs considered here, we assessed the quality of the linear, quadratic, and exponential GMs applied to analyze the Betula Project data used in the illustration and that informed our simulation study. For each participant we combined, according to Equations (1), (2), and (4) the estimated level and change scores and their age of assessment to predict their episodic recall scores. We then calculated the mean squared errors (MSE) between the observed and predicted scores under each GM for each participant.

We obtained MSE of11.40, 10.25, and 9.94, respectively, for the linear, quadratic, and exponential GM. Thus, the exponential GM, which adjusted best to the data in terms of BIC (cf. supplementary Table 2), also produced predictions that were closest to the data, thereby showing the highest internal validity.

Survival Prediction in the Betula Project

In terms of predictive validity, we compared the three GMs in terms of the terminal decline hypothesis in the Betula Project data. This hypothesis states that many individuals during the final phase of life manifest accelerated decrements in functional capacities (Gerstorf & Ram, 2013). We used Bayesian factor score estimates of each person's level and change scores on the basis of the linear, quadratic, and exponential GMs to predict survival in the 1000 Betula participants considered here, 529 of whom were deceased by the study's most recent mortality update.

The Cox proportional hazard model predicted survival and the age of either death (for the deceased) or the last assessment (for the right-censored participants). The exponential level and change estimated scores predicted best survival, with a BIC of 5681, whereas the quadratic and linear obtained worse adjustments (BIC values of 5685 and 5691, respectively). The estimated level and change scores from the exponential GM lowered the fit of the survival the most (drop in χ^2 statistic of 32.33 for two degrees of freedom, compared to 28.69 and 22.29 for the quadratic and linear, respectively – all *p*'s < .01).

Discussion

As expected, our literature review found that the vast majority (86%) of GMs applied to study normal cognitive aging implemented a linear or a quadratic function of change, whereas very few (less than 5%) considered a truly nonlinear function of change.

Hypothesized Consequences for Data Analysts

We explored what conclusions are likely to be reached when data generated under an exponential GM are analyzed with a linear or quadratic GM.

First, the exponential model can be implemented without great difficulty, despite its complexity~~, at least for fitting a population where exponential decline characterizes the dependent variable~~. Exponential models using adaptive Gaussian quadrature failed to converge in less than 3% of the 18000 data sets we generated. In contrast, the quadratic model failed to converge 14.43% of the cases. Thus, within the constraints present in our simulation study, the exponential model is a practicable alternative model to the linear and quadratic GMs.

Second, we found that in many conditions of our simulation study, it is not possible to tease apart the quadratic and the exponential model ~~apart~~ based on the commonly used BIC goodness of fit index. Thus, the chances of retaining the quadratic over the exponential model are high. ~~Of course,~~Most analysts typically testing linear and quadratic models will not consider other functions; nevertheless,~~. But~~ those scientists evaluating ~~trying out~~ an exponential model using comparative fit may end up discarding it for the more familiar and equally well-fitting quadratic model.

Third, we evaluated possible consequences of ~~does~~ retaining the misspecified ~~model, as would be the case for accepting the~~ quadratic model in our simulation study. Would doing so~~,~~ necessarily lead to inaccurate substantive conclusions about the change process analyzed? We addressed this question by considering two aspects: the power to detect individual differences in change, and the concordance between the level and change scores used to generate the data and the analogous scores predicted from the growth models. With respect to power, the quadratic GM clearly performed worst (across all conditions at 72.22%), whereas the linear and exponential fared similarly and in a range deemed by many to be satisfactory (82.21% and 81.26%, respectively). This gap increases in many conditions commonly encountered in empirical research. Thus, based on the quadratic model, we ~~may~~would often conclude that the individuals in the sample cannot be discriminated with respect to amount of change, whereas the exponential model may detect that important source of heterogeneity.

In contrast, the characterization of individual differences in growth components is often seriously compromised in quality if we rely on too simple a GM, especially when the decline rate is substantial. Concordance rates ~~a~~were generally higher for the level than for the change component. This ~~is~~ was expected, as level information is more easily estimated than change information. Whereas~~ile for the exponential model~~ the level correlations either remain stable or increase as testing intervals widen for the exponential model, for the level, linear, and quadratic models the concordance rates decrease monotonically. This effect is influenced by measurement reliability (the higher the GCR, the stronger the discordance; this effect would become even more pernicious with multiple indicator latent curve models), but not by number of occasions or by sample size. For change, concordance rates increase with more occasions of measurement and wider intervals, but not considerably with greater sample sizes. The exponential model clearly outperforms both the linear and quadratic model, and is the only model obtaining high concordance rates.

In sum then, under the empirically-based simulation conditions studied here, if a change process is exponential in nature but is~~we~~ analyzed~~ it~~ with a quadratic model, we may not pick up that individuals differ in amount of change. If we do detect variance in change, then we are very likely to obtain biased estimates of individuals' change scores. Hence, not only is the detection of change compromised; even when change is detected, it isn't well characterized by derived individual indices of change.

Finally, in terms of the empirical illustration on the Betula Project data, we found that the exponential model (a) fit the data best, (b), implied the expected individual cognitive scores that were closest to the observed scores, and (c) predicted mortality better than the linear and quadratic models. We deduce that in this sample the exponential GM likely provides the representation that best captures cognitive aging processes relative to its two competitors. We further hypothesize that the advantage of the exponential over both the linear and quadratic models would increase with greater intervals of testing ($\Delta > 5$ years), more reliable variables (GCR >.681), more occasions ($T > 4$), and a greater sample size ($N > 1000$).

Recommendations for Data Analysts

In light of these results, what should one do when analyzing real long-term change data? First, rather than continuing the common practice of solely testing the linear and the quadratic GMs, one might consider alternative, possibly nonlinear functions, which allow for substantive interpretations of their parameters. The exponential model is one such function that has proven useful in many disciplines, but other functions deserve to be considered as well (e.g., Grimm et al., 2007). Omitting nonlinear functions may cloud our understanding of the phenomena under investigation.

Second, when planning a longitudinal study, one should try to use reliable instruments (e.g., Little, Lindenberger, & Nesselroade, 1999) and assess the instruments repeatedly at carefully planned intervals (Hertzog & Nesselroade, 2003; Willett, 1989). What our results show is that even under ideal conditions, such as those of our simulation experiment (e.g., no unwanted retest effects, no longitudinal dropout, and group homogeneity), the discovery of change-related information often requires many widely spaced assessments. One practical implication is the need for patience in longitudinal sampling. A lack of change-related evidence at the second or third wave of a longitudinal study of long-term change does not preclude the sought-after effect manifesting itself after additional data collection (especially when waves are separated by wide time intervals).

Third, we found that overall the effect of sample size was rather minor, compared to the other effects of our simulation experiment. Although~~While in~~ psychologists~~y we~~ have become extremely conscious of the overall importance of sample size, the design features related to density and span of change assessments appear to be more important in enabling accurate characterization of individual differences in change ~~re are other equally, if not more, important design features that need to be considered as well~~ (Brandmaier, von Oertzen, Ghisletta, Lindenberger, & Hertzog, 2018).

In conclusion, much progress has been made in estimating nonlinear growth models and in showing their relevance in lifespan developmental psychology. We argue that the time is right for ending exclusive reliance on linear GMs for describing change. Embracing conceptual, statistical, and computational advances embedded in nonlinear GMs may deepen our understanding of psychological change phenomena.

References

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1988). *Life-span developmental psychology: An introduction to research methods* (Vol. reprint of 1977). Hillsdale, NJ: Lawrence Erlbaum Associates.

Box, G. E. P., & Draper, N. R. (1987). *Empirical model-building and response surfaces*. New York, NY: John Wiley and Sons.

Brandmaier, A. M., von Oertzen, T., Ghisletta, P., Lindenberger, U., & Hertzog, C. (2018). Precision, Reliability, and Effect Size of Slope Variance in Latent Growth Curve Models: Implications for Statistical Power Analysis. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00294

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147-158.

Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, *58*, 615–637. https://doi.org/10.1146/annurev.psych.58.110405.085520

Davidian, W., & Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. London, UK.: Chapman & Hall.

Driscoll, I., Resnick, S. M., Troncoso, J. C., An, Y., O'Brien, R., & Zonderman, A. B. (2006). Impact of Alzheimer's pathology on cognitive trajectories in nondemented elderly. *Annals of Neurology*, *60*(6), 688–695. https://doi.org/10.1002/ana.21031

Gerstorf, D., & Ram, N. (2013). Inquiry into terminal decline: five objectives for future study. *The Gerontologist*, *53*(5), 727–737. https://doi.org/10.1093/geront/gnt046

Grimm, K. J., McArdle, J. J., & Hamagami, F. (2007). Nonlinear growth mixture models in research on cognitive aging. In K. v. Monfort, H. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioural and related sciences* (pp. 267–294). Mahwah, N.J.: Lawrence Erlbaum Associates.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: an overview of methodological issues. *Psychology and Aging*, *18*(4), 639–657. https://doi.org/10.1037/0882-7974.18.4.639

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics, 38*, 963-974.

Littell R, Milliken F, Stroup W, Wolfinger R (1996). *SAS system for mixed models*. SAS Institute, Cary, NC: SAS Institute.

Little, T. D., & Card, N. A. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, *4*(2), 192–211. https://doi.org/10.1037/1082-989X.4.2.192

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade, & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 561-614). New York, NY.: Plenum Press.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38*, 115-142.

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*, 23–43.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107-122.

Nilsson, L.-G., Adolfsson, R., Bäckman, L., Frias, C. M. de, Molander, B., & Nyberg, L. (2004). Betula: A Prospective Cohort Study on Memory, Health and Aging. *Aging, Neuropsychology, and Cognition*, *11*(2–3), 134–148. https://doi.org/10.1080/13825580490511026

Rovine, M. J., & Molenaar, P. C. M. (1998). The covariance between level and shape in the latent growth curve model with estimated basis vector coefficients. *Methods of Psychological Research Online, 3*, 95–107.

Schaie, K. W., & Hofer, S. M. (2001). Longitudinal studies in aging research. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 53–77). San Diego, CA.: Academic Press.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*, 461–464.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis*. Oxford, UK.: Oxford University Press.

Wainer, H. (2000). The centercept: an estimable and meaningful regression parameter. *Psychological Science*, *11*(5), 434–436. https://doi.org/10.1111/1467-9280.00284

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods*, *8*(1), 16–37.

Willett, J. B. (1989). Some results on reliability for the measurement of change: Implications for the design of studies of individual growth. *Educational and Psychological Measurement*, *49*, 587–602.

Acknowledgment

This work benefitted from support of the Lifebrain grant, financed by the European Commission in the Horizon2020 Programme, Grant agreement no: SC1–PM-04-2016. We thank Mikael Stiernstedt and Lars Nyberg for providing us with the Betula data used in the empirical illustration.

Figure 1. Simulated trajectories generated under the exponential decline model.

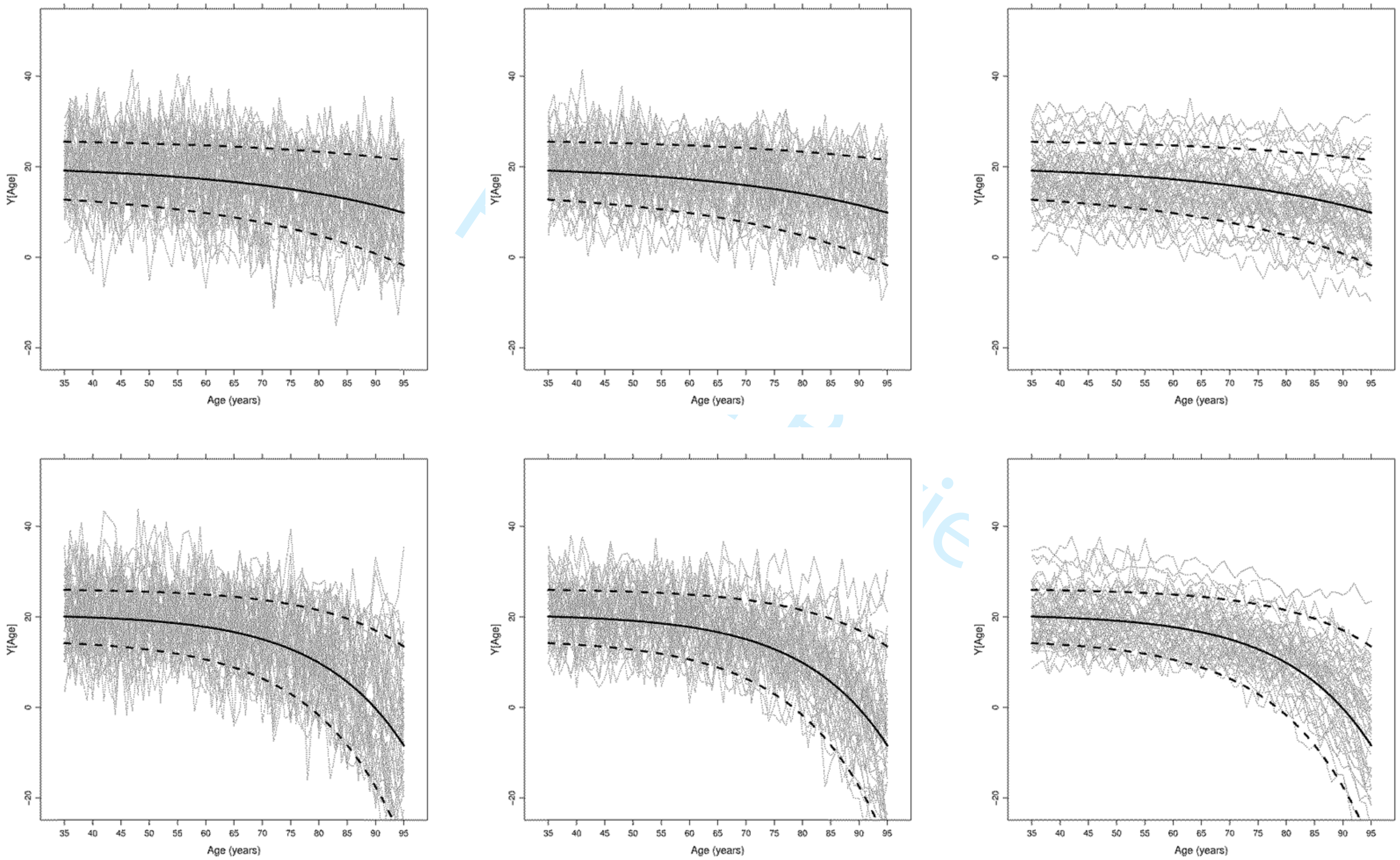
Figure 2a. Concordance in levels by interval width and model, for $\gamma = 0.033$.

Figure 2b. Concordance in levels by interval width and model, for $\gamma = 0.066$.

Figure 3a. Concordance in changes by interval width and model, for $\gamma = 0.033$.

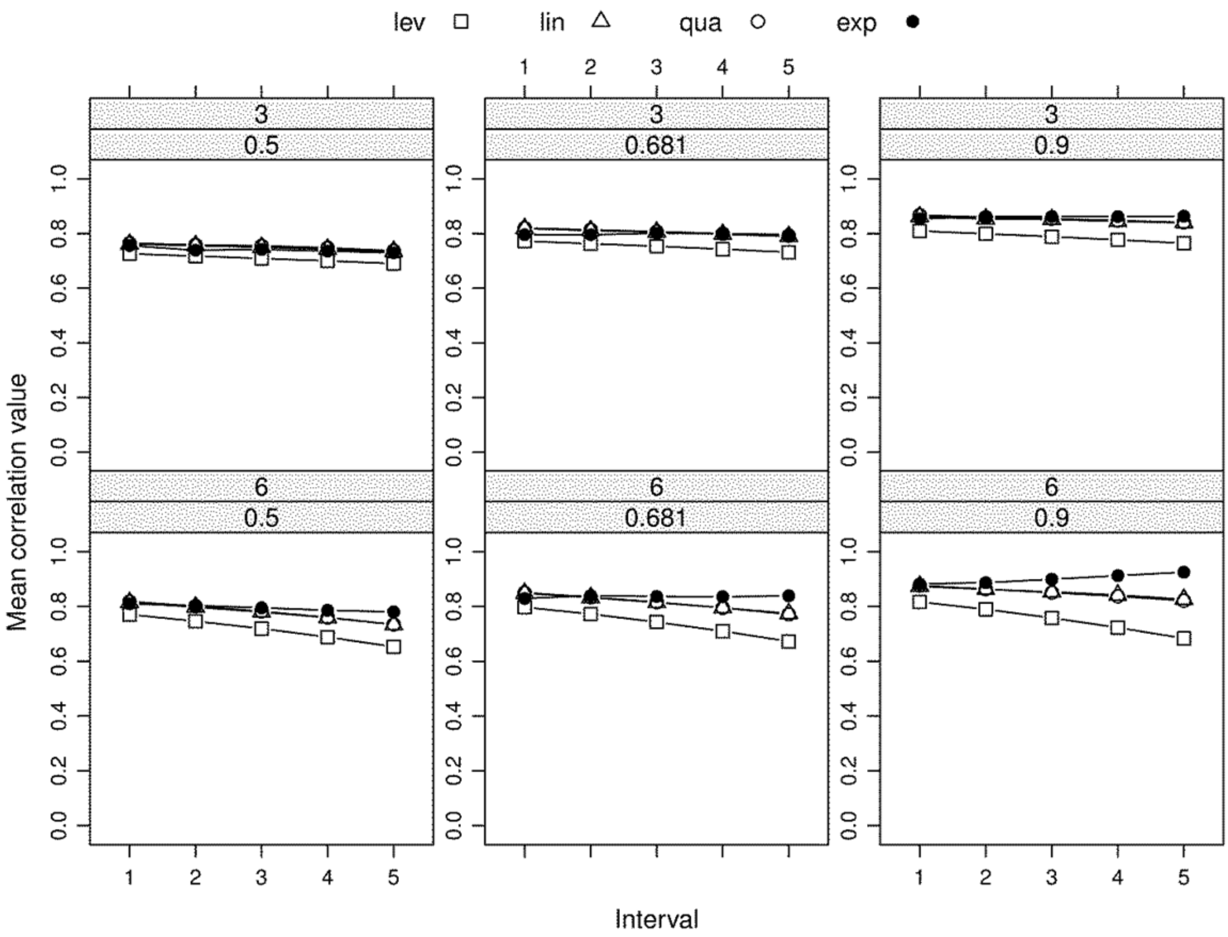
Figure 3b. Concordance in changes by interval width and model, for $\gamma = 0.066$.

Figure 1: Simulated Trajectories Generated under the Exponential Decline Model



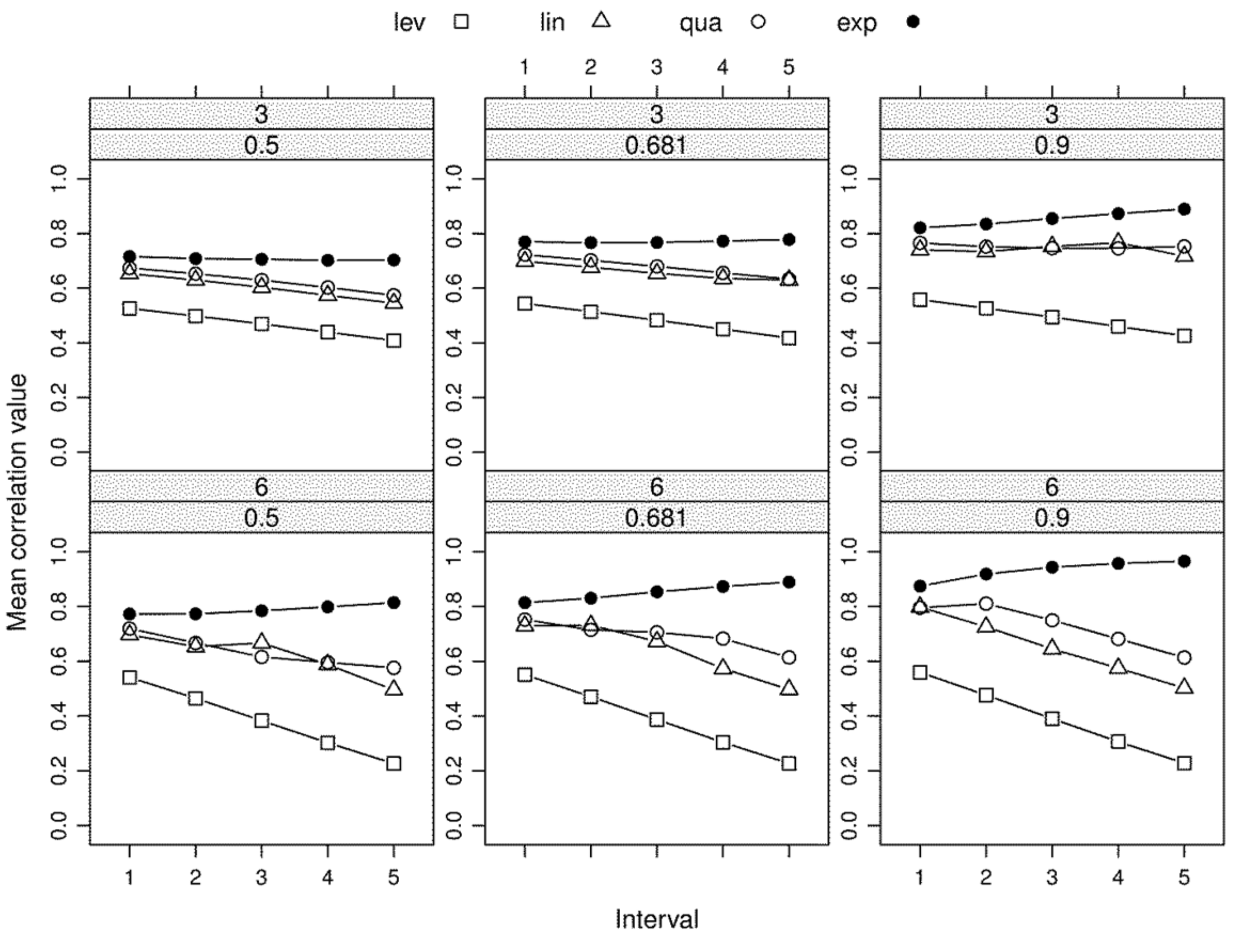
Note. Row 1: γ= 0.033, row 2: γ= 0.066; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900; Thick continuous line is sample average without residuals, thick dashed lines are 1 SD above/below the mean; Thin continuous lines are 50 simulated individual trajectories.

Figure 2a: Concordance in Levels by Interval Width and Model, for γ= 0.033



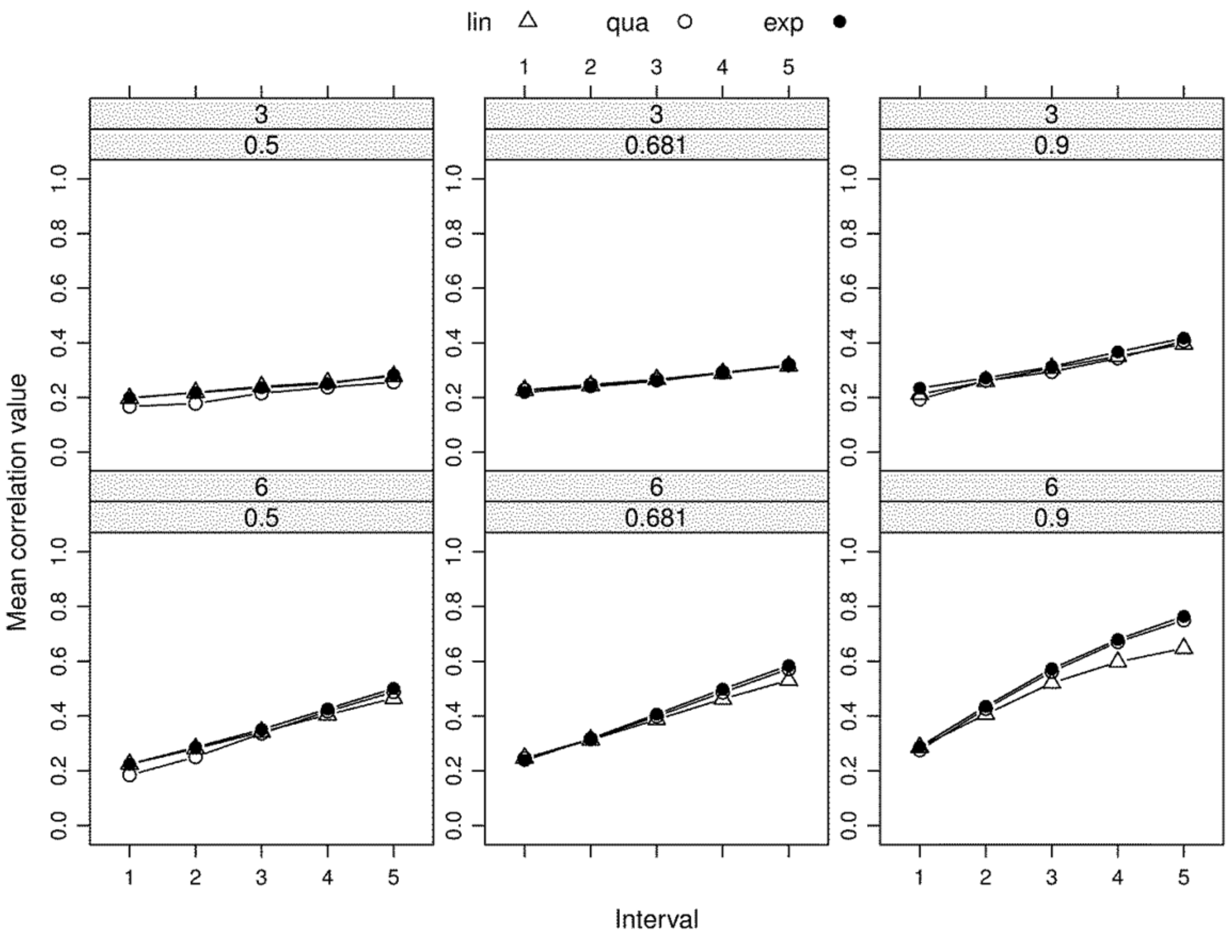
Note. Row 1: T = 3, row 2: T = 6; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900.

Figure 2b: Concordance in Levels by Interval Width and Model, for γ= 0.066



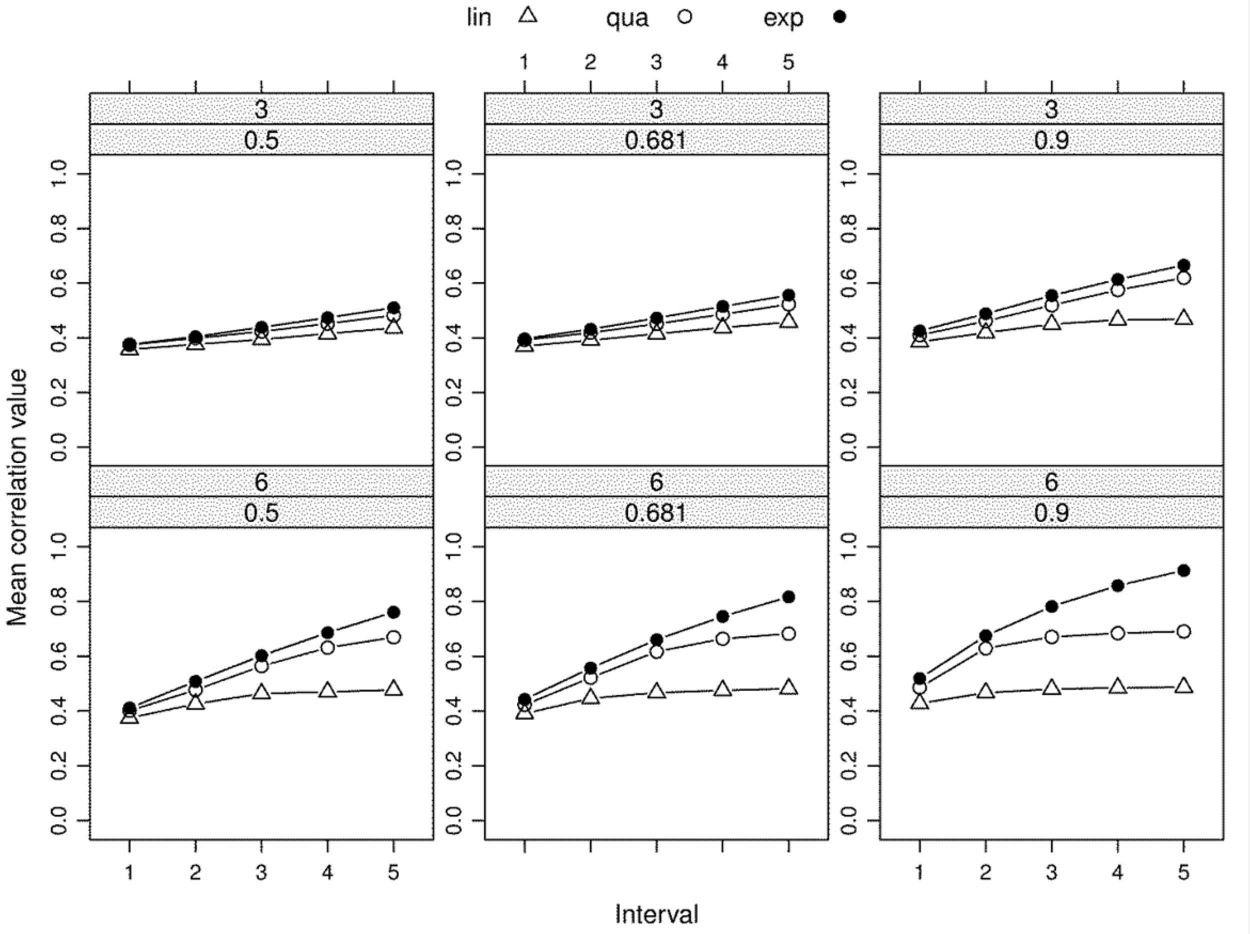
Note. Row 1: *T* = 3, row 2: *T* = 6; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900.

Figure 3a: Concordance in Changes by Interval Width and Model, for γ= 0.033



Note. Row 1: T = 3, row 2: T = 6; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900.

Figure 3b: Concordance in Changes by Interval Width and Model, for γ= 0.066



Note. Row 1: T = 3, row 2: T = 6; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900.

Supplementary Appendix 1

Detailed description of the screening procedure

On December 2015 the first author (FA) and second author (SA) performed a systematic literature search by using three popular databases in psychology and related disciplines: Web of Science (WoS), PsycINFO (PI), and PubMed (PM). We limited our search to the previous five full years, that is, from January 1st, 2010, to December 31st, 2014. We based the search on two groups of terms, one for the statistical model and the other for the substantive field of application. All (statistical and substantive) terms were searched anywhere in the text (e.g., title, abstract, body text). We required that the records returned from the search be journal articles, book chapters, or books, written in English. The aim of the literature search was to evaluate current practice of substantive researchers in non-pathological cognitive aging, who analyzed longitudinal data with some form of growth model. Thus, we excluded hits that were statistical/methodological in nature (that were tutorials, or that focused on studying properties of such models).

Terms concerning the statistical model were: “growth curve model*”; “latent curve model*”; “multilevel model*”; “mixed effects model*”; “linear mixed model*”; “random effects model*”; “hierarchical linear model*”; “nonlinear hierarchical model*”; “generalized linear mixed model*”; “nonlinear mixed model*”. We included the wildcard “*” to include variations (such as modeling, modelling, models, etc.). Growth curve and latent curve models are by definition applied to longitudinal data, whereas those within the multilevel family can also be applied to cross-sectional data (e.g., typically hierarchically organized data, such as children within schools) or to meta-analyses (to account for both across-studies and within-study variations). Thus, for this latter class of models we added the “longitudinal” term. The terms concerning the statistical model were combined with two terms concerning the substantive application: “intelligence” and “cogniti*”. We again included the wild card “*” to assure coverage of related terms such as cognition, cognitive, cognitively, etc. We refrained from searching for specific cognitive abilities (e.g., speed, episodic or working memory, inhibition), because of the very large number of terms that such a search would have required. Moreover, we reasoned that scientific writings focused on specific abilities are very likely to include the terms “intelligence” and/or “cogniti*” in their text.

The supplemental Table 1 shows the number of records from each literature database for each combined use of statistical and substantive search terms. The last column shows the sum across the three databases. The row named Sum shows the total number of records across all search terms from each database: 593 for WoS, 571 for PI, and 89 for PM, for a total of 1253 records. First, the FA and SA independently eliminated first all duplicate records within each database and for this first step obtained full agreement. The resulting numbers of retained records were 536 for WoS, 493 for PI, and 82 for PM, for a total of 1111 records. Then, FA and SA independently eliminated duplicate and non-English records across databases, and obtained a final number of 719 entries, again with full agreement. Duplicates were established based on authors’ names, title and date of record, and name of publication.

Second, FA and SA independently read the abstracts of 182 records (a fourth) to screen them for eligibility, according to the criteria listed below. That is, despite the inclusion of the search terms, a record could be unrelated to the aims of the search (in few cases the abstract was not sufficient and portions of the text had to be read). At this stage, the FA and SA had a 0.73

agreement index. Thus, the FA and SA together analyzed, discussed, and reclassified the discordant records to reach full consensus.

Third, the FA and SA independently read and classified the abstracts of the remaining (719 – 182 =) 537 records. At this stage, the FA and SA had a 0.88 agreement index. They again reclassified the discordant records to reach full consensus. In the end, of the 719 records, the following were excluded:

- n=4, because they were commentary articles (n=2), reviews (n=1), or corrections (n=1);
- n=1, because it was a conference poster and only the abstract was available;
- n=20, because they were meta-analyses;
- n=7, because they did not include longitudinal/repeated measurement data (e.g., mixed-effects models were used on cross-sectional data with heterogeneous variances across conditions, or with siblings or neighborhoods as clusters);
- n=48, because they were of statistical or methodological nature [1];

---

[1] i.e., that appeared in specialized journals (The American Statistician; Annals of Applied Statistics; Applied Psychological Measurement; Behavior research methods; Biometrical Journal; Biometrical Journal; Biometrika; British Journal of Mathematical and Statistical Psychology; Computational Statistics and Data Analysis; Frontiers in Psychology: Quantitative Psychology and Measurement; Journal of the American Statistical Association; Journal of Applied Statistics; Journal of Biometrics and Biostatistics; Journal of Computational and Graphical Statistics; Journal of Educational and Behavioral Statistics; Journal of the Royal Statistical Society, Series A, Statistics in Society; Journal of the Royal Statistical Society, Series C, Applied Statistics; Methodology; Multivariate Behavioral Research; Psychological Methods; Psychometrika; Statistical Methods in Medical Research; Statistical Modelling; Statistics in Medicine; Structural Equation Modeling; Statistical Methodology), that appeared in substantive journals but that focused on methodological/statistical issues (American Journal of Epidemiology; The American Journal of Geriatric Psychiatry; Annals of Epidemiology; Contemporary Clinical Trials; European Journal of Developmental Psychology; Journal of Cognition and Development; Neuroepidemiology; NeuroImage; Psychology and Aging; Sleep Medicine; World Journal of Biological Psychiatry), or that appeared in methodological or statistical books (e.g., Contemporary issues in exploratory data mining in the behavioral sciences; Handbook for advanced multilevel analysis; The handbook of life-span development; Handbook of structural equation modeling; Longitudinal data analysis: A practical guide for researchers in aging, health, and social sciences; Quantitative and qualitative methods in psychotherapy research).

- n=128, because they included and focused on non-normal samples (e.g., patients with Alzheimer, Parkinson, autism, depression, bipolar disorders, breast cancer, MCI, heavy drinkers);
- n=243 did not include samples of elderly individuals (age < 65 years);
- n=349 did not specify cognitive or intelligence performance as dependent variable of analysis (e.g., controlled for cognitive variables but specified non-cognitive dependent variables, or discussed relevance of findings in terms of general cognitive functioning, without explicit assessment or analysis of cognitive variables).

Many records met multiple exclusion criteria (e.g., n=149 did not include samples of elderly individuals and did not specify cognitive or intelligence performance as dependent variable of analysis; n=70 included non-normal samples and did not specify cognitive or intelligence performance as dependent variable of analysis). In the end, we retained and fully analyzed n=162 records. The screening procedure is represented in the Supplemental Figure 1.

Supplementary Appendix 2

Records in final analysis based on the literature review

Abner, E. L., Dennis, B. C., Mathews, M. J., Mendiondo, M. S., Caban-Holt, A., Kryscio, R. J., … Crowley, J. J. (2012). Practice effects in a longitudinal, multi-center Alzheimer's disease prevention clinical trial. *Trials*, *13*, 217. https://doi.org/10.1186/1745-6215-13-217

Ahles, T. A., Saykin, A. J., McDonald, B. C., Li, Y., Furstenberg, C. T., Hanscom, B. S., … Kaufman, P. A. (2010). Longitudinal assessment of cognitive changes associated with adjuvant treatment for breast cancer: impact of age and cognitive reserve. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *28*(29), 4434–4440. https://doi.org/10.1200/JCO.2009.27.0827

Aichberger, M. C., Busch, M. A., Reischies, F. M., Ströhle, A., Heinz, A., & Rapp, M. A. (2010). Effect of physical inactivity on cognitive performance after 2.5 years of follow-up: Longitudinal results from the Survey of Health, Ageing, and Retirement (SHARE). *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, *23*(1), 7–15. https://doi.org/10.1024/1662-9647/a000003

Allerhand, M., Gale, C. R., & Deary, I. J. (2014). The dynamic relationship between cognitive function and positive well-being in older people: A prospective study using the English Longitudinal Study of Aging. *Psychology and Aging*, *29*(2), 306–318. https://doi.org/10.1037/a0036551

Alwerdt, J., Edwards, J. D., Athilingam, P., O'Connor, M. L., & Valdés, E. G. (2013). Longitudinal differences in cognitive functioning among older adults with and without heart failure. *Journal of Aging and Health*, *25*(8), 1358–1377. https://doi.org/10.1177/0898264313505111

Andersson-Roswall, L., Engman, E., Samuelsson, H., & Malmgren, K. (2010). Cognitive outcome 10 years after temporal lobe epilepsy surgery: A prospective controlled study. *Neurology*, *74*(24), 1977–1985. https://doi.org/10.1212/WNL.0b013e3181e39684

Arab, L., Biggs, M. L., O'Meara, E. S., Longstreth, W. T., Crane, P. K., & Fitzpatrick, A. L. (2011). Gender differences in tea, coffee, and cognitive decline in the elderly: the Cardiovascular Health Study. *Journal of Alzheimer's Disease: JAD*, *27*(3), 553–566. https://doi.org/10.3233/JAD-2011-110431

Ball, K. K., Ross, L. A., Roth, D. L., & Edwards, J. D. (2013). Speed of processing training in the ACTIVE study: how much is needed and who benefits? *Journal of Aging and Health*, *25*(8 Suppl), 65S-84S. https://doi.org/10.1177/0898264312470167

Barnes, L. L., Wilson, R. S., Everson-Rose, S. A., Hayward, M. D., Evans, D. A., & de Leon, C. F. M. (2012). Effects of early-life adversity on cognitive decline in older African Americans and whites. *Neurology*, *79*(24), 2321–2327. https://doi.org/10.1212/WNL.0b013e318278b607

Benke, K. S., Carlson, M. C., Doan, B. Q., Walston, J. D., Xue, Q. L., Reiner, A. P., … Fallin, M. D. (2011). The association of genetic variants in interleukin-1 genes with cognition: findings from the cardiovascular health study. *Experimental Gerontology*, *46*(12), 1010–1019. https://doi.org/10.1016/j.exger.2011.09.005

Beydoun, M. A., Ding, E. L., Beydoun, H. A., Tanaka, T., Ferrucci, L., & Zonderman, A. B. (2012). Vitamin D receptor and megalin gene polymorphisms and their associations with longitudinal cognitive change in US adults. *The American Journal of Clinical Nutrition*, *95*(1), 163–178. https://doi.org/10.3945/ajcn.111.017137

Bielak, A. A. M., Anstey, K. J., Christensen, H., & Windsor, T. D. (2012). Activity engagement is related to level, but not change in cognitive ability across adulthood. *Psychology and Aging*, *27*(1), 219–228. https://doi.org/10.1037/a0024667

Bielak, A. A. M., Cherbuin, N., Bunce, D., & Anstey, K. J. (2014). Intraindividual variability is a fundamental phenomenon of aging: Evidence from an 8-year longitudinal study across young, middle, and older adulthood. *Developmental Psychology*, *50*(1), 143–151. https://doi.org/10.1037/a0032650

Bielak, A. A. M., Cherbuin, N., Bunce, D., & Anstey, K. J. (2014). Preserved differentiation between physical activity and cognitive performance across young, middle, and older adulthood over 8 years. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *69*(4), 523–532. https://doi.org/10.1093/geronb/gbu016

Bielak, A. A. M., Gerstorf, D., Anstey, K. J., & Luszcz, M. A. (2014). Longitudinal associations between activity and cognition vary by age, activity type, and cognitive domain. *Psychology and Aging*, *29*(4), 863–872. https://doi.org/10.1037/a0036960

Bindawas, S. M., Graham, J. E., Karmarkar, A. M., Chen, N.-W., Granger, C. V., Niewczyk, P., … Ottenbacher, K. J. (2014). Trajectories in functional recovery for patients receiving inpatient rehabilitation for unilateral hip or knee replacement. *Archives of Gerontology and Geriatrics*, *58*(3), 344–349. https://doi.org/10.1016/j.archger.2013.12.009

Blackwell, T., Yaffe, K., Laffan, A., Ancoli-Israel, S., Redline, S., Ensrud, K. E., … Osteoporotic Fractures in Men (MrOS) Study Group. (2014). Associations of objectively and subjectively measured sleep quality with subsequent cognitive decline in older community-dwelling men: the MrOS sleep study. *Sleep*, *37*(4), 655–663. https://doi.org/10.5665/sleep.3562

Brown, M. T., Wilmoth, J. M., & London, A. S. (2014). Veteran status and men's later-life cognitive trajectories: evidence from the health and retirement study. *Journal of Aging and Health*, *26*(6), 924–951. https://doi.org/10.1177/0898264314534893

Buchman, A. S., Boyle, P. A., Leurgans, S. E., Barnes, L. L., & Bennett, D. A. (2011). Cognitive function is associated with the development of mobility impairments in community-dwelling elders. *The American Journal of Geriatric Psychiatry*, *19*(6), 571–580. https://doi.org/10.1097/JGP.0b013e3181ef7a2e

Buchman, A. S., Yu, L., Wilson, R. S., Shulman, J. M., Boyle, P. A., & Bennett, D. A. (2014). Harm avoidance is associated with progression of parkinsonism in community-dwelling older adults: a prospective cohort study. *BMC Geriatrics*, *14*, 54. https://doi.org/10.1186/1471-2318-14-54

Buracchio, T., Dodge, H. H., Howieson, D., Wasserman, D., & Kaye, J. (2010). The trajectory of gait speed preceding mild cognitive impairment. *Archives of Neurology*, *67*(8), 980–986. https://doi.org/10.1001/archneurol.2010.159

Bürki, C. N., Ludwig, C., Chicherio, C., & de Ribaupierre, A. (2014). Individual differences in cognitive plasticity: An investigation of training curves in younger and older adults. *Psychological Research*, *78*(6), 821–835. https://doi.org/10.1007/s00426-014-0559-3

Chiang, G. C., Insel, P. S., Tosun, D., Schuff, N., Truran-Sacrey, D., Raptentsetsang, S. T., … Weiner, M. W. (2010). Hippocampal atrophy rates and CSF biomarkers in elderly APOE2 normal subjects. *Neurology*, *75*(22), 1976–1981. https://doi.org/10.1212/WNL.0b013e3181ffe4d1

Chiao, C., Botticello, A., & Fuh, J.-L. (2014). Life-course socio-economic disadvantage and late-life cognitive functioning in Taiwan: results from a national cohort study. *International Health*, *6*(4), 322–330. https://doi.org/10.1093/inthealth/ihu046

Choi, M., Lohman, M. C., & Mezuk, B. (2014). Trajectories of cognitive decline by driving mobility: evidence from the Health and Retirement Study. *International Journal of Geriatric Psychiatry*, *29*(5), 447–453. https://doi.org/10.1002/gps.4024

Collins, B., MacKenzie, J., Tasca, G. A., Scherling, C., & Smith, A. (2013). Cognitive effects of chemotherapy in breast cancer patients: A dose–response study. *Psycho-Oncology*, *22*(7), 1517–1527. https://doi.org/10.1002/pon.3163

Comijs, H. C., van den Kommer, T. N., Minnaar, R. W. M., Penninx, B. W. J. H., & Deeg, D. J. H. (2011). Accumulated and differential effects of life events on cognitive decline in older persons: Depending on depression, baseline cognition, or ApoE ε4 status? *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *66B*(Supp 1), 111–120. https://doi.org/10.1093/geronb/gbr019

Crane, P. K., Gruhl, J. C., Erosheva, E. A., Gibbons, L. E., McCurry, S. M., Rhoads, K., … White, L. (2010). Use of spoken and written Japanese did not protect Japanese-American men from cognitive decline in late life. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *65*(6), 654–666. https://doi.org/10.1093/geronb/gbq046

Dahl, A. K., Hassing, L. B., Fransson, E. I., Gatz, M., Reynolds, C. A., & Pedersen, N. L. (2013). Body mass index across midlife and cognitive change in late life. *International Journal of Obesity*, *37*(2), 296–302. https://doi.org/10.1038/ijo.2012.37

Dahl, A., Hassing, L. B., Fransson, E., Berg, S., Gatz, M., Reynolds, C. A., & Pedersen, N. L. (2010). Being overweight in midlife is associated with lower cognitive ability and steeper cognitive decline in late life. *The Journals of Gerontology: Series A: Biological Sciences and Medical Sciences*, *65A*(1), 57–62. https://doi.org/10.1093/gerona/glp035

Daly, M., McMinn, D., & Allan, J. L. (2015). A bidirectional relationship between physical activity and executive function in older adults. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.01044

Danhauer, S. C., Legault, C., Bandos, H., Kidwell, K., Costantino, J., Vaughan, L., … Shumaker, S. (2013). Positive and negative affect, depression, and cognitive processes in the Cognition in the Study of Tamoxifen and Raloxifene (Co-STAR) Trial. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *20*(5), 532–552. https://doi.org/10.1080/13825585.2012.747671

Danthiir, V., Burns, N. R., Nettelbeck, T., Wilson, C., & Wittert, G. (2011). The older people, omega-3, and cognitive health (EPOCH) trial design and methodology: a randomised, double-blind, controlled trial investigating the effect of long-chain omega-3 fatty acids on cognitive ageing and wellbeing in cognitively healthy older adults. *Nutrition Journal*, *10*, 117. https://doi.org/10.1186/1475-2891-10-117

Das, D., Tan, X., Bielak, A. A. M., Cherbuin, N., Easteal, S., & Anstey, K. J. (2014). Cognitive ability, intraindividual variability, and common genetic variants of catechol-O-methyltransferase and brain-derived neurotrophic factor: a longitudinal study in a population-based sample of older adults. *Psychology and Aging*, *29*(2), 393–403. https://doi.org/10.1037/a0035702

Deary, I. J., Johnson, W., Gow, A. J., Pattie, A., Brett, C. E., Bates, T. C., & Starr, J. M. (2011). Losing one's grip: A bivariate growth curve model of grip strength and nonverbal reasoning from age 79 to 87 years in the Lothian birth cohort 1921. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *66B*(6), 699–707. https://doi.org/10.1093/geronb/gbr059

de Frias, C. M., Schaie, K. W., & Willis, S. L. (2014). Hypertension moderates the effect of APOE on 21-year cognitive trajectories. *Psychology and Aging*, *29*(2), 431–439. https://doi.org/10.1037/a0036828

Desikan, R. S., McEvoy, L. K., Thompson, W. K., Holland, D., Roddey, J. C., Blennow, K., … Dale, A. M. (2011). Amyloid-β associated volume loss occurs only in the presence of phospho-tau. *Annals of Neurology*, *70*(4), 657–661. https://doi.org/10.1002/ana.22509

Desikan, R. S., McEvoy, L. K., Holland, D., Thompson, W. K., Brewer, J. B., Aisen, P. S., … Alzheimer's Disease Neuroimaging Initiative. (2013). Apolipoprotein E epsilon4 does not modulate amyloid-β-associated neurodegeneration in preclinical Alzheimer disease. *AJNR. American Journal of Neuroradiology*, *34*(3), 505–510. https://doi.org/10.3174/ajnr.A3267

Desplenter, F., Lavikainen, P., Hartikainen, S., Sulkava, R., & Bell, J. S. (2012). Sedative use and incident cognitive decline among persons aged 75 years and older: a population-based longitudinal study. *International Psychogeriatrics*, *24*(1), 48–54. https://doi.org/10.1017/S1041610211001359

Dodge, H. H., Wang, C.-N., Chang, C.-C. H., & Ganguli, M. (2011). Terminal decline and practice effects in older adults without dementia: The MoVIES project. *Neurology*, *77*(8), 722–730. https://doi.org/10.1212/WNL.0b013e31822b0068

Dodonov, Y. S., & Dodonova, Y. A. (2012). Response time analysis in cognitive tasks with increasing difficulty. *Intelligence*, *40*(5), 379–394. https://doi.org/10.1016/j.intell.2012.07.002

Dowling, N. M., Johnson, S. C., Gleason, C. E., Jagust, W. J., & Alzheimer's Disease Neuroimaging Initiative. (2015). The mediational effects of FDG hypometabolism on the association between cerebrospinal fluid biomarkers and neurocognitive function. *NeuroImage*, *105*, 357–368. https://doi.org/10.1016/j.neuroimage.2014.10.050

Duits, F. H., Kester, M. I., Scheffer, P. G., Blankenstein, M. A., Scheltens, P., Teunissen, C. E., & van der Flier, W. M. (2013). Increase in cerebrospinal fluid F2-isoprostanes is related to cognitive decline in APOE ε4 carriers. *Journal of Alzheimer's Disease: JAD*, *36*(3), 563–570. https://doi.org/10.3233/JAD-122227

Dzierzewski, J. M., Fung, C. H., Jouldjian, S., Alessi, C. A., Irwin, M. R., & Martin, J. L. (2014). Decrease in daytime sleeping is associated with improvement in cognition after hospital discharge in older adults. *Journal of the American Geriatrics Society*, *62*(1), 47–53.

Ellis, K. A., Lim, Y. Y., Harrington, K., Ames, D., Bush, A. I., Darby, D., … AIBL Research Group. (2013). Decline in cognitive function over 18 months in healthy older adults with high amyloid-β. *Journal of Alzheimer's Disease: JAD*, *34*(4), 861–871. https://doi.org/10.3233/JAD-122170

Ellwardt, L., Aartsen, M., Deeg, D., & Steverink, N. (2013). Does loneliness mediate the relation between social support and cognitive functioning in later life? *Social Science & Medicine*, *98*, 116–124. https://doi.org/10.1016/j.socscimed.2013.09.002

Emanuel, J. E., Lopez, O. L., Houck, P. R., Becker, J. T., Weamer, E. A., Demichele-Sweet, M. A. A., … Sweet, R. A. (2011). Trajectory of cognitive decline as a predictor of psychosis in early Alzheimer disease in the cardiovascular health study. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *19*(2), 160–168. https://doi.org/10.1097/JGP.0b013e3181e446c8

Engelman, C. D., Koscik, R. L., Jonaitis, E. M., Okonkwo, O. C., Hermann, B. P., La Rue, A., & Sager, M. A. (2013). Interaction between two cholesterol metabolism genes influences memory: findings from the Wisconsin Registry for Alzheimer's Prevention. *Journal of Alzheimer's Disease: JAD*, *36*(4), 749–757. https://doi.org/10.3233/JAD-130482

Eramudugolla, R., Bielak, A. A. M., Bunce, D., Easteal, S., Cherbuin, N., & Anstey, K. J. (2014). Long-term cognitive correlates of traumatic brain injury across adulthood and interactions with APOE genotype, sex, and age cohorts. *Journal of the International Neuropsychological Society*, *20*(4), 444–454. https://doi.org/10.1017/S1355617714000174

Farias, S. T., Chand, V., Bonnici, L., Baynes, K., Harvey, D., Mungas, D., … Reed, B. (2012). Idea density measured in late life predicts subsequent cognitive trajectories: implications for the measurement of cognitive reserve. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *67*(6), 677–686. https://doi.org/10.1093/geronb/gbr162

Farias, S. T., Mungas, D., Hinton, L., & Haan, M. (2011). Demographic, neuropsychological, and functional predictors of rate of longitudinal cognitive decline in Hispanic older adults. *The American Journal of Geriatric Psychiatry*, *19*(5), 440–450. https://doi.org/10.1097/JGP.0b013e3181e9b9a5

Feng, L., Li, J., Ng, T.-P., Lee, T.-S., Kua, E.-H., & Zeng, Y. (2012). Tea drinking and cognitive function in oldest-old Chinese. *The Journal of Nutrition, Health & Aging*, *16*(9), 754–758. https://doi.org/10.1007/s12603-012-0077-1

Ferguson, A., Spencer, E., Craig, H., & Colyvas, K. (2014). Propositional idea density in women's written language over the lifespan: computerized analysis. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *55*, 107–121. https://doi.org/10.1016/j.cortex.2013.05.012

Finkel, D., Reynolds, C. A., Larsson, M., Gatz, M., & Pedersen, N. L. (2011). Both odor identification and ApoE-ε4 contribute to normative cognitive aging. *Psychology and Aging*, *26*(4), 872–883. https://doi.org/10.1037/a0023371

Fiocco, A. J., Lindquist, K., Ferrell, R., Li, R., Simonsick, E. M., Nalls, M., … Yaffe, K. (2010). COMT genotype and cognitive function: An 8-year longitudinal study in white and black elders. *Neurology*, *74*(16), 1296–1302. https://doi.org/10.1212/WNL.0b013e3181d9edba

Frazier, D. T., Seider, T., Bettcher, B. M., Mack, W. J., Jastrzab, L., Chao, L., … Kramer, J. H. (2014). The role of carotid intima-media thickness in predicting longitudinal cognitive function in an older adult cohort. *Cerebrovascular Diseases (Basel, Switzerland)*, *38*(6), 441–447. https://doi.org/10.1159/000366469

Gale, C. R., Allerhand, M., Sayer, A. A., Cooper, C., & Deary, I. J. (2014). The dynamic relationship between cognitive function and walking speed: the English Longitudinal Study of Ageing. *Age (Dordrecht, Netherlands)*, *36*(4), 9682. https://doi.org/10.1007/s11357-014-9682-8

Ganguli, M., Fu, B., Snitz, B. E., Unverzagt, F. W., Loewenstein, D. A., Hughes, T. F., & Chang, C.-C. H. (2014). Vascular risk factors and cognitive decline in a population sample. *Alzheimer Disease and Associated Disorders*, *28*(1), 9–15. https://doi.org/10.1097/WAD.0000000000000004

Gerritsen, L., Comijs, H. C., Deeg, D. J. H., Penninx, B. W. J. H., & Geerlings, M. I. (2011). Salivary cortisol, APOE-ε4 allele and cognitive decline in a prospective study of older persons. *Neurobiology of Aging*, *32*(9), 1615–1625. https://doi.org/10.1016/j.neurobiolaging.2009.09.007

Ghisletta, P., Kennedy, K. M., Rodrigue, K. M., Lindenberger, U., & Raz, N. (2010). Adult age differences and the role of cognitive resources in perceptual-motor skill acquisition: application of a multilevel negative exponential model. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *65B*(2), 163–173. https://doi.org/10.1093/geronb/gbp126

Ghisletta, P., Rabbitt, P., Lunn, M., & Lindenberger, U. (2012). Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. *Intelligence*, *40*(3), 260–268. https://doi.org/10.1016/j.intell.2012.02.008

Goldhammer, F., Rauch, W. A., Schweizer, K., & Moosbrugger, H. (2010). Differential effects of intelligence, perceptual speed and age on growth in attentional speed and accuracy. *Intelligence*, *38*(1), 83–92. https://doi.org/10.1016/j.intell.2009.07.002

Gooding, L. F., Abner, E. L., Jicha, G. A., Kryscio, R. J., & Schmitt, F. A. (2014). Musical Training and Late-Life Cognition. *American Journal of Alzheimer's Disease and Other Dementias*, *29*(4), 333–343. https://doi.org/10.1177/1533317513517048

Gow, A. J., Avlund, K., & Mortensen, E. L. (2014). Leisure activity associated with cognitive ability level, but not cognitive change. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01176

Gow, A. J., Avlund, K., & Mortensen, E. L. (2014). Occupational characteristics and cognitive aging in the Glostrup 1914 Cohort. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *69*(2), 228–236. https://doi.org/10.1093/geronb/gbs115

Gow, A. J., Johnson, W., Mishra, G., Richards, M., Kuh, D., & Deary, I. J. (2012). Is age kinder to the initially more able?: Yes, and no. *Intelligence*, *40*(1), 49–59. https://doi.org/10.1016/j.intell.2011.10.007

Gow, A. J., Mortensen, E. L., & Avlund, K. (2012). Activity participation and cognitive aging from age 50 to 80 in the Glostrup 1914 Cohort. *Journal of the American Geriatrics Society*, *60*(10), 1831–1838. https://doi.org/10.1111/j.1532-5415.2012.04168.x

Greene, D., Tschanz, J. T., Smith, K. R., Ostbye, T., Corcoran, C., Welsh-Bohmer, K. A., … Cache County Investigators. (2014). Impact of offspring death on cognitive health in late life: the Cache County study. *The American Journal of Geriatric Psychiatry*, *22*(11), 1307–1315. https://doi.org/10.1016/j.jagp.2013.05.002

Gross, A. L., Rebok, G. W., Unverzagt, F. W., Willis, S. L., & Brandt, J. (2011). Word list memory predicts everyday function and problem-solving in the elderly: results from the ACTIVE cognitive intervention trial. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *18*(2), 129–146. https://doi.org/10.1080/13825585.2010.516814

Gurgel, R. K., Ward, P. D., Schwartz, S., Norton, M. C., Foster, N. L., & Tschanz, J. T. (2014). Relationship of hearing loss and dementia: a prospective, population-based study. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *35*(5), 775–781. https://doi.org/10.1097/MAO.0000000000000313

Harvey, P. D., Reichenberg, A., Bowie, C. R., Patterson, T. L., & Heaton, R. K. (2010). The course of neuropsychological performance and functional capacity in older patients with schizophrenia: Influences of previous history of long-term institutional stay. *Biological Psychiatry*, *67*(10), 933–939. https://doi.org/10.1016/j.biopsych.2010.01.008

Hassing, L. B., Dahl, A. K., Pedersen, N. L., & Johansson, B. (2010). Overweight in midlife is related to lower cognitive function 30 years later: A prospective study with longitudinal assessments. *Dementia and Geriatric Cognitive Disorders*, *29*(6), 543–552. https://doi.org/10.1159/000314874

Hayden, K. M., Reed, B. R., Manly, J. J., Tommet, D., Pietrzak, R. H., Chelune, G. J., … Jones, R. N. (2011). Cognitive decline in the elderly: an analysis of population heterogeneity. *Age and Ageing*, *40*(6), 684–689. https://doi.org/10.1093/ageing/afr101

Hendrie, H. C., Murrell, J., Baiyewu, O., Lane, K. A., Purnell, C., Ogunniyi, A., … Gao, S. (2014). APOE ε4 and the risk for Alzheimer disease and cognitive decline in African Americans and Yoruba. *International Psychogeriatrics*, *26*(6), 977–985. https://doi.org/10.1017/S1041610214000167

Hess, T. M., Emery, L., & Neupert, S. D. (2012). Longitudinal relationships between resources, motivation and functioning. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *67B*(3), 299–308. https://doi.org/10.1093/geronb/gbr100

Hill, T. D., Angel, J. L., Balistreri, K. S., & Herrera, A. P. (2012). Immigrant status and cognitive functioning in late-life: An examination of gender variations in the healthy immigrant effect. *Social Science & Medicine*, *75*(12), 2076–2084. https://doi.org/10.1016/j.socscimed.2012.04.005

Hjelm, C., Dahl, A., Broström, A., Mårtensson, J., Johansson, B., & Strömberg, A. (2012). The influence of heart failure on longitudinal changes in cognition among individuals 80 years of age and older. *Journal of Clinical Nursing*, *21*(7–8), 994–1003. https://doi.org/10.1111/j.1365-2702.2011.03817.x

Holtzer, R., Wang, C., Lipton, R., & Verghese, J. (2012). The protective effects of executive functions and episodic memory on gait speed decline in aging defined in the context of cognitive reserve. *Journal of the American Geriatrics Society*, *60*(11), 2093–2098. https://doi.org/10.1111/j.1532-5415.2012.04193.x

Hülür, G., Hertzog, C., Pearman, A., Ram, N., & Gerstorf, D. (2014). Longitudinal associations of subjective memory with memory performance and depressive symptoms: Between-person and within-person perspectives. *Psychology and Aging*, *29*(4), 814–827. https://doi.org/10.1037/a0037619

Jokinen, H., Lipsanen, J., Schmidt, R., Fazekas, F., Gouw, A. A., van der Flier, W. M., … LADIS Study Group. (2012). Brain atrophy accelerates cognitive decline in cerebral small vessel disease: the LADIS study. *Neurology*, *78*(22), 1785–1792. https://doi.org/10.1212/WNL.0b013e3182583070

Jokinen, H., Schmidt, R., Ropele, S., Fazekas, F., Gouw, A. A., Barkhof, F., … LADIS Study Group. (2013). Diffusion changes predict cognitive and functional outcome: the LADIS study. *Annals of Neurology*, *73*(5), 576–583. https://doi.org/10.1002/ana.23802

Jones, R. N., Marsiske, M., Ball, K., Rebok, G., Willis, S. L., Morris, J. N., & Tennstedt, S. L. (2013). The ACTIVE Cognitive Training Interventions and Trajectories of Performance among Older Adults. *Journal of Aging and Health*, *25*(0). https://doi.org/10.1177/0898264312461938

Kaffashian, S., Dugravot, A., Brunner, E. J., Sabia, S., Ankri, J., Kivimäki, M., & Singh-Manoux, A. (2013). Midlife stroke risk and cognitive decline: a 10-year follow-up of the Whitehall II cohort study. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *9*(5), 572–579. https://doi.org/10.1016/j.jalz.2012.07.001

Köhler, S., van Boxtel, M. P. J., van Os, J., Thomas, A. J., O'Brien, J. T., Jolles, J., … Allardyce, J. (2010). Depressive symptoms and cognitive decline in community-dwelling older adults. *Journal of the American Geriatrics Society*, *58*(5), 873–879. https://doi.org/10.1111/j.1532-5415.2010.02807.x

Laughlin, G. A., McEvoy, L. K., von Mühlen, D., Daniels, L. B., Kritz-Silverstein, D., Bergstrom, J., … Barrett-Connor, E. (2011). Sex differences in the association of Framingham Cardiac Risk Score with cognitive decline in community-dwelling elders without clinical heart disease. *Psychosomatic Medicine*, *73*(8), 683–689. https://doi.org/10.1097/PSY.0b013e31822f9089

Laukka, E. J., Macdonald, S. W. S., Fratiglioni, L., & Bäckman, L. (2012). Preclinical cognitive trajectories differ for Alzheimer's disease and vascular dementia. *Journal of the International Neuropsychological Society: JINS*, *18*(2), 191–199. https://doi.org/10.1017/S1355617711001718

Leist, A. K., Glymour, M. M., Mackenbach, J. P., van Lenthe, F. J., & Avendano, M. (2013). Time away from work predicts later cognitive function: differences by activity during leave. *Annals of Epidemiology*, *23*(8), 455–462. https://doi.org/10.1016/j.annepidem.2013.05.014

Lim, Y. Y., Ellis, K. A., Pietrzak, R. H., Ames, D., Darby, D., Harrington, K., … AIBL Research Group. (2012). Stronger effect of amyloid load than APOE genotype on cognitive decline in healthy older adults. *Neurology*, *79*(16), 1645–1652. https://doi.org/10.1212/WNL.0b013e31826e9ae6

Liu, H., Gao, S., Hall, K. S., Unverzagt, F. W., Lane, K. A., Callahan, C. M., & Hendrie, H. C. (2013). Optimal blood pressure for cognitive function: Findings from an elderly African-American cohort study. *Journal of the American Geriatrics Society*, *61*(6), 875–881. https://doi.org/10.1111/jgs.12259

Luchsinger, J. A., Biggs, M. L., Kizer, J. R., Barzilay, J., Fitzpatrick, A., Newman, A., … Kuller, L. (2013). Adiposity and cognitive decline in the cardiovascular health study. *Neuroepidemiology*, *40*(4), 274–281. https://doi.org/10.1159/000345136

MacDonald, S. W. S., DeCarlo, C. A., & Dixon, R. A. (2011). Linking biological and cognitive aging: toward improving characterizations of developmental time. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *66 Suppl 1*, i59-70. https://doi.org/10.1093/geronb/gbr039

MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2011). Aging and the shape of cognitive change before death: Terminal decline or terminal drop? *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *66B*(3), 292–301. https://doi.org/10.1093/geronb/gbr001

Mascherek, A., & Zimprich, D. (2011). Correlated change in memory complaints and memory performance across 12 years. *Psychology and Aging*, *26*(4), 884–889. https://doi.org/10.1037/a0023156

McEvoy, L. K., Laughlin, G. A., Barrett-Connor, E., Bergstrom, J., Kritz-Silverstein, D., Der-Martirosian, C., & von Mühlen, D. (2012). Metabolic syndrome and 16-year cognitive decline in community-dwelling older adults. *Annals of Epidemiology*, *22*(5), 310–317. https://doi.org/10.1016/j.annepidem.2011.12.003

McFall, G. P., Wiebe, S. A., Vergote, D., Jhamandas, J., Westaway, D., & Dixon, R. A. (2014). IDE (rs6583817) polymorphism and pulse pressure are independently and interactively associated with level and change in executive function in older adults. *Psychology and Aging*, *29*(2), 418–430. https://doi.org/10.1037/a0034656

Mengel-From, J., Christensen, K., Thinggaard, M., McGue, M., & Christiansen, L. (2011). Genetic variants in the choline acetyltransferase (chAT) gene are modestly associated with normal cognitive function in the elderly. *Genes, Brain & Behavior*, *10*(8), 876–882. https://doi.org/10.1111/j.1601-183X.2011.00728.x

Mengel-From, J., Thinggaard, M., Lindahl-Jacobsen, R., McGue, M., Christensen, K., & Christiansen, L. (2013). CLU Genetic Variants and Cognitive Decline among Elderly and Oldest Old. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0079105

Mitchell, M. B., Cimino, C. R., Benitez, A., Brown, C. L., Gibbons, L. E., Kennison, R. F., … Piccinin, A. M. (2012). Cognitively Stimulating Activities: Effects on Cognition across Four Studies with up to 21 Years of Longitudinal Data. *Journal of Aging Research*, *2012*, 461592. https://doi.org/10.1155/2012/461592

Muniz-Terrera, G., Matthews, F. E., Stephan, B., & Brayne, C. (2011). Are terminal decline and its potential indicators detectable in population studies of the oldest old? *International Journal of Geriatric Psychiatry*, *26*(6), 584–592. https://doi.org/10.1002/gps.2566

Nishita, Y., Tange, C., Tomida, M., Ando, F., & Shimokata, H. (2013). Does high educational level protect against intellectual decline in older adults?: A 10-year longitudinal study. *Japanese Psychological Research*, *55*(4), 378–389. https://doi.org/10.1111/jpr.12028

Ossenkoppele, R., van der Flier, W. M., Verfaillie, S. C. J., Vrenken, H., Versteeg, A., van Schijndel, R. A., … van Berckel, B. N. M. (2014). Long-term effects of amyloid, hypometabolism, and atrophy on neuropsychological functions. *Neurology*, *82*(20), 1768–1775. https://doi.org/10.1212/WNL.0000000000000432

Ottenbacher, A. J., Snih, S. A., Bindawas, S. M., Markides, K. S., Graham, J. E., Samper-Ternent, R., … Ottenbacher, K. J. (2014). Role of physical activity in reducing cognitive decline in older Mexican-American adults. *Journal of the American Geriatrics Society*, *62*(9), 1786–1791. https://doi.org/10.1111/jgs.12978

Parisi, J. M., Gross, A. L., Rebok, G. W., Saczynski, J. S., Crowe, M., Cook, S. E., … Unverzagt, F. W. (2011). Modeling change in memory performance and memory perceptions: Findings from the ACTIVE study. *Psychology and Aging*, *26*(3), 518–524. https://doi.org/10.1037/a0022458

Paulson, D., Bowen, M. E., & Lichtenberg, P. A. (2014). Does brain reserve protect older women from vascular depression? *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *69*(2), 157–167. https://doi.org/10.1093/geronb/gbt007

Payne, B. R., Gross, A. L., Parisi, J. M., Sisco, S. M., Stine-Morrow, E. A. L., Marsiske, M., & Rebok, G. W. (2014). Modelling longitudinal changes in older adults' memory for spoken discourse: findings from the ACTIVE cohort. *Memory (Hove, England)*, *22*(8), 990–1001. https://doi.org/10.1080/09658211.2013.861916

Pearman, A., Hertzog, C., & Gerstorf, D. (2014). Little evidence for links between memory complaints and memory performance in very old age: Longitudinal analyses from the Berlin Aging Study. *Psychology and Aging*, *29*(4), 828–842. https://doi.org/10.1037/a0037141

Peretz, C., Korczyn, A. D., Shatil, E., Aharonson, V., Birnboim, S., & Giladi, N. (2011). Computer-based, personalized cognitive training versus classical computer games: a randomized double-blind prospective trial of cognitive stimulation. *Neuroepidemiology*, *36*(2), 91–99. https://doi.org/10.1159/000323950

Petkus, A. J., Wetherell, J. L., Stein, M. B., Liu, L., & Barrett-Connor, E. (2012). History of sexual assault is associated with greater declines in executive functioning in older adults with APOE epsilon4. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *67*(6), 653–659. https://doi.org/10.1093/geronb/gbr163

Piccinin, A. M., Muniz, G., Matthews, F. E., & Johansson, B. (2011). Terminal Decline From Within- and Between-Person Perspectives, Accounting for Incident Dementia. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66B*(4), 391–401. https://doi.org/10.1093/geronb/gbr010

Power, M. C., Korrick, S., Tchetgen Tchetgen, E. J., Nie, L. H., Grodstein, F., Hu, H., … Weisskopf, M. G. (2014). Lead exposure and rate of change in cognitive function in older women. *Environmental Research*, *129*, 69–75. https://doi.org/10.1016/j.envres.2013.12.010

Praetorius, M., Thorvaldsson, V., Hassing, L. B., & Johansson, B. (2013). Substantial effects of apolipoprotein E ε4 on memory decline in very old age: longitudinal findings from a population-based sample. *Neurobiology of Aging*, *34*(12), 2734–2739. https://doi.org/10.1016/j.neurobiolaging.2013.06.002

Rabbitt, P., Lunn, M., Pendleton, N., & Yardefagar, G. (2011). Terminal pathologies affect rates of decline to different extents and age accelerates the effects of terminal pathology on cognitive decline. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *66*(3), 325–334. https://doi.org/10.1093/geronb/gbr026

Rajan, K. B., Aggarwal, N. T., Wilson, R. S., Everson-Rose, S. A., & Evans, D. A. (2014). Association of cognitive functioning, incident stroke, and mortality in older adults. *Stroke*, *45*(9), 2563–2567. https://doi.org/10.1161/STROKEAHA.114.005143

Ram, N., Gerstorf, D., Fauth, E., Zarit, S., & Malmberg, B. (2010). Aging, Disablement, and Dying: Using Time-as-Process and Time-as-Resources Metrics to Chart Late-Life Change. *Research in Human Development*, *7*(1), 27–44. https://doi.org/10.1080/15427600903578151

Rast, P. (2011). Verbal knowledge, working memory, and processing speed as predictors of verbal learning in older adults. *Developmental Psychology*, *47*(5), 1490–1498. https://doi.org/10.1037/a0023422

Riediger, M., Voelkle, M. C., Schaefer, S., & Lindenberger, U. (2014). Charting the life course: Age differences and validity of beliefs about lifespan development. *Psychology and Aging*, *29*(3), 503–520. https://doi.org/10.1037/a0036228

Rist, P. M., Dufouil, C., Glymour, M. M., Tzourio, C., & Kurth, T. (2011). Migraine and cognitive decline in the population-based EVA study. *Cephalalgia*, *31*(12), 1291–1300. https://doi.org/10.1177/0333102411417466

Ritchie, S. J., Tucker-Drob, E. M., & Deary, I. J. (2014). A strong link between speed of visual discrimination and cognitive ageing. *Current Biology*, *24*(15), R681–R683. https://doi.org/10.1016/j.cub.2014.06.012

Robitaille, A., Muniz, G., Piccinin, A. M., Johansson, B., & Hofer, S. M. (2012). Multivariate longitudinal modeling of cognitive aging: Associations among change and variation in processing speed and visuospatial ability. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, *25*(1), 15–24. https://doi.org/10.1024/1662-9647/a000051

Rosano, C., Aizenstein, H. J., Newman, A. B., Venkatraman, V., Harris, T., Ding, J., … Health ABC Study. (2012). Neuroimaging differences between older adults with maintained versus declining cognition over a 10-year period. *NeuroImage*, *62*(1), 307–313. https://doi.org/10.1016/j.neuroimage.2012.04.033

Rosenberg, P. B., Mielke, M. M., Xue, Q.-L., & Carlson, M. C. (2010). Depressive symptoms predict incident cognitive impairment in cognitive healthy older women. *The American Journal of Geriatric Psychiatry*, *18*(3), 204–211. https://doi.org/10.1097/JGP.0b013e3181c53487

Royall, D. R., Palmer, R., Chiodo, L. K., & Polk, M. J. (2012). Depressive symptoms predict longitudinal change in executive control but not memory. *International Journal of Geriatric Psychiatry*, *27*(1), 89–96. https://doi.org/10.1002/gps.2697

Runge, S. K., Small, B. J., McFall, G. P., & Dixon, R. A. (2014). APOE moderates the association between lifestyle activities and cognitive performance: Evidence of genetic plasticity in aging. *Journal of the International Neuropsychological Society*, *20*(5), 478–486. https://doi.org/10.1017/S1355617714000356

Sabia, S., Elbaz, A., Dugravot, A., Head, J., Shipley, M., Hagger-Johnson, G., … Singh-Manoux, A. (2012). Impact of smoking on cognitive decline in early old age: the Whitehall II cohort study. *Archives of General Psychiatry*, *69*(6), 627–635. https://doi.org/10.1001/archgenpsychiatry.2011.2016

Salthouse, T. A. (2014). Correlates of cognitive change. *Journal of Experimental Psychology: General*, *143*(3), 1026–1048. https://doi.org/10.1037/a0034847

Sanders, A. E., Wang, C., Katz, M., Derby, C. A., Barzilai, N., Ozelius, L., & Lipton, R. B. (2010). Association of a functional polymorphism in the cholesteryl ester transfer protein (CETP) gene with memory decline and incidence of dementia. *JAMA: Journal of the American Medical Association*, *303*(2), 150–158. https://doi.org/10.1001/jama.2009.1988

Schiepers, O. J. G., Harris, S. E., Gow, A. J., Pattie, A., Brett, C. E., Starr, J. M., & Deary, I. J. (2012). APOE E4 status predicts age-related cognitive decline in the ninth decade: Longitudinal follow-up of the Lothian Birth Cohort 1921. *Molecular Psychiatry*, *17*(3), 315–324. https://doi.org/10.1038/mp.2010.137

Schiepers, O. J. G., van Boxtel, M. P. J., Harris, S. E., Gow, A. J., Pattie, A., Brett, C. E., … Deary, I. J. (2011). MTHFR polymorphisms and cognitive ageing in the ninth decade: the Lothian Birth Cohort 1921. *Genes, Brain, and Behavior*, *10*(3), 354–364. https://doi.org/10.1111/j.1601-183X.2010.00675.x

Schneider, J. A., Arvanitakis, Z., Yu, L., Boyle, P. A., Leurgans, S. E., & Bennett, D. A. (2012). Cognitive impairment, decline and fluctuations in older community-dwelling subjects with Lewy bodies. *Brain: A Journal of Neurology*, *135*(Pt 10), 3005–3014. https://doi.org/10.1093/brain/aws234

Sharp, E. S., Reynolds, C. A., Pedersen, N. L., & Gatz, M. (2010). Cognitive engagement and cognitive aging: is openness protective? *Psychology and Aging*, *25*(1), 60–73. https://doi.org/10.1037/a0018748

Shokouhi, S., Claassen, D., Kang, H., Ding, Z., Rogers, B., Mishra, A., … Alzheimer’s Disease Neuroimaging Initiative. (2013). Longitudinal progression of cognitive decline correlates with changes in the spatial pattern of brain 18F-FDG PET. *Journal of Nuclear Medicine:*

Official Publication, Society of Nuclear Medicine, *54*(9), 1564–1569. https://doi.org/10.2967/jnumed.112.116137

Silbert, L. C., Dodge, H. H., Perkins, L. G., Sherbakov, L., Lahna, D., Erten-Lyons, D., … Kaye, J. A. (2012). Trajectory of white matter hyperintensity burden preceding mild cognitive impairment. *Neurology*, *79*(8), 741–747. https://doi.org/10.1212/WNL.0b013e3182661f2b

Small, B. J., Dixon, R. A., & McArdle, J. J. (2011). Tracking cognition-health changes from 55 to 95 years of age. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *66 Suppl 1*, i153-161. https://doi.org/10.1093/geronb/gbq093

Snitz, B. E., Weissfeld, L. A., Lopez, O. L., Kuller, L. H., Saxton, J., Singhabahu, D. M., … Dekosky, S. T. (2013). Cognitive trajectories associated with β-amyloid deposition in the oldest-old without dementia. *Neurology*, *80*(15), 1378–1384. https://doi.org/10.1212/WNL.0b013e31828c2fc8

Staff, R. T., Hogan, M. J., & Whalley, L. J. (2014). Aging trajectories of fluid intelligence in late life: The influence of age, practice and childhood IQ on Raven's Progressive Matrices. *Intelligence*, *47*, 194–201. https://doi.org/10.1016/j.intell.2014.09.013

Stoykova, R., Matharan, F., Dartigues, J.-F., & Amieva, H. (2011). Impact of social network on cognitive performances and age-related cognitive decline across a 20-year follow-up. *International Psychogeriatrics*, *23*(9), 1405–1412. https://doi.org/10.1017/S1041610211001165

Tangney, C. C., Kwasny, M. J., Li, H., Wilson, R. S., Evans, D. A., & Morris, M. C. (2011). Adherence to a Mediterranean-type dietary pattern and cognitive decline in a community population. *The American Journal of Clinical Nutrition*, *93*(3), 601–607. https://doi.org/10.3945/ajcn.110.007369

Thomas, K. R., & Marsiske, M. (2014). Verbal prompting to improve everyday cognition in MCI and unimpaired older adults. *Neuropsychology*, *28*(1), 123–134. https://doi.org/10.1037/neu0000039

Thorvaldsson, V., Skoog, I., Hofer, S. M., Börjesson-Hanson, A., Östling, S., Sacuiu, S., & Johansson, B. (2012). Nonlinear blood pressure effects on cognition in old age: Separating between-person and within-person associations. *Psychology and Aging*, *27*(2), 375–383. https://doi.org/10.1037/a0025631

Toledo, J. B., Korff, A., Shaw, L. M., Trojanowski, J. Q., & Zhang, J. (2014). Low levels of cerebrospinal fluid complement 3 and factor H predict faster cognitive decline in mild cognitive impairment. *Alzheimer's Research & Therapy*, *6*(3), 36. https://doi.org/10.1186/alzrt266

Trompet, S., de Craen, A. J. M., Jukema, J. W., Pons, D., Slagboom, P. E., Kremer, D., … Westendorp, R. G. J. (2011). Variation in the CBP gene involved in epigenetic control associates with cognitive function. *Neurobiology of Aging*, *32*(3), 549.e1-8. https://doi.org/10.1016/j.neurobiolaging.2009.12.019

Tucker-Drob, E. M. (2011). Global and domain-specific changes in cognition throughout adulthood. *Developmental Psychology*, *47*(2), 331–343. https://doi.org/10.1037/a0021361

Tucker-Drob, E. M., Reynolds, C. A., Finkel, D., & Pedersen, N. L. (2014). Shared and unique genetic and environmental influences on aging-related changes in multiple cognitive abilities. *Developmental Psychology*, *50*(1), 152–166. https://doi.org/10.1037/a0032468

van den Kommer, T. N., Dik, M. G., Comijs, H. C., Lütjohann, D., Lips, P., Jonker, C., & Deeg, D. J. H. (2012). The role of extracerebral cholesterol homeostasis and ApoE e4 in cognitive decline. *Neurobiology of Aging*, *33*(3), 622.e17-28. https://doi.org/10.1016/j.neurobiolaging.2011.02.019

Van Gerven, P. W. M., Van Boxtel, M. P. J., Ausems, E. E. B., Bekers, O., & Jolles, J. (2012). Do apolipoprotein E genotype and educational attainment predict the rate of cognitive decline in normal aging? A 12-year follow-up of the Maastricht Aging Study. *Neuropsychology*, *26*(4), 459–472. https://doi.org/10.1037/a0028685

Vannorsdall, T. D., Kueider, A. M., Carlson, M. C., & Schretlen, D. J. (2014). Higher baseline serum uric acid is associated with poorer cognition but not rates of cognitive decline in women. *Experimental Gerontology*, *60*, 136–139. https://doi.org/10.1016/j.exger.2014.10.013

Velseboer, D. C., Broeders, M., Post, B., van Geloven, N., Speelman, J. D., Schmand, B., … CARPA Study Group. (2013). Prognostic factors of motor impairment, disability, and quality of life in newly diagnosed PD. *Neurology*, *80*(7), 627–633. https://doi.org/10.1212/WNL.0b013e318281cc99

Vemuri, P., Lesnick, T. G., Przybelski, S. A., Machulda, M., Knopman, D. S., Mielke, M. M., … Jack, C. R. J. (2014). Association of lifetime intellectual enrichment with cognitive decline in the older population. *JAMA Neurology*, *71*(8), 1017–1024. https://doi.org/10.1001/jamaneurol.2014.963

von Stumm, S., & Deary, I. J. (2012). Typical intellectual engagement and cognition in the ninth decade of life: The Lothian Birth Cohort 1921. *Psychology and Aging*, *27*(3), 761–767. https://doi.org/10.1037/a0026527

Vos, S. J., Xiong, C., Visser, P. J., Jasielec, M. S., Hassenstab, J., Grant, E. A., … Fagan, A. M. (2013). Preclinical Alzheimer's disease and its outcome: a longitudinal cohort study. *The Lancet. Neurology*, *12*(10), 957–965. https://doi.org/10.1016/S1474-4422(13)70194-7

Wessels, A. M., Lane, K. A., Gao, S., Hall, K. S., Unverzagt, F. W., & Hendrie, H. C. (2011). Diabetes and cognitive decline in elderly African Americans: A 15-year follow-up study. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*(4), 418–424. https://doi.org/10.1016/j.jalz.2010.07.003

West, R. L., & Hastings, E. C. (2011). Self-regulation and recall: growth curve modeling of intervention outcomes for older adults. *Psychology and Aging*, *26*(4), 803–812. https://doi.org/10.1037/a0023784

Wilson, R. S., Aggarwal, N. T., Barnes, L. L., Mendes de Leon, C. F., Hebert, L. E., & Evans, D. A. (2010). Cognitive decline in incident Alzheimer disease in a community population. *Neurology*, *74*(12), 951–955. https://doi.org/10.1212/WNL.0b013e3181d64786

Wilson, R. S., Boyle, P. A., Yu, L., Barnes, L. L., Schneider, J. A., & Bennett, D. A. (2013). Life-span cognitive activity, neuropathologic burden, and cognitive aging. *Neurology*, *81*(4), 314–321. https://doi.org/10.1212/WNL.0b013e31829c5e8a

Wilson, R. S., Hebert, L. E., Scherr, P. A., Dong, X., Leurgens, S. E., & Evans, D. A. (2012). Cognitive decline after hospitalization in a community population of older persons. *Neurology*, *78*(13), 950–956. https://doi.org/10.1212/WNL.0b013e31824d5894

Wilson, R. S., Segawa, E., Hizel, L. P., Boyle, P. A., & Bennett, D. A. (2012). Terminal dedifferentiation of cognitive abilities. *Neurology*, *78*(15), 1116–1122. https://doi.org/10.1212/WNL.0b013e31824f7ff2

Yaffe, K., Falvey, C., Hamilton, N., Schwartz, A. V., Simonsick, E. M., Satterfield, S., … Harris, T. B. (2012). Diabetes, glucose control, and 9-year cognitive decline among older adults without dementia. *Archives of Neurology*, *69*(9), 1170–1175. https://doi.org/10.1001/archneurol.2012.1117

Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., & Manly, J. J. (2014). Bilingualism does not alter cognitive decline or dementia risk among Spanish-speaking immigrants. *Neuropsychology*, *28*(2), 238–246. https://doi.org/10.1037/neu0000014

Zeki Al Hazzouri, A., Haan, M. N., Osypuk, T., Abdou, C., Hinton, L., & Aiello, A. E. (2011). Neighborhood socioeconomic context and cognitive decline among older Mexican Americans: results from the Sacramento Area Latino Study on Aging. *American Journal of Epidemiology*, *174*(4), 423–431. https://doi.org/10.1093/aje/kwr095

Zeki Al Hazzouri, A., Newman, A. B., Simonsick, E., Sink, K. M., Sutton Tyrrell, K., Watson, N., … Health ABC Study. (2013). Pulse wave velocity and cognitive decline in elders: the Health, Aging, and Body Composition study. *Stroke*, *44*(2), 388–393. https://doi.org/10.1161/STROKEAHA.112.673533

Zhang, C., & Pierce, B. L. (2014). Genetic susceptibility to accelerated cognitive decline in the US health and retirement study. *Neurobiology of Aging*, *35*(6), e11–e18. https://doi.org/10.1016/j.neurobiolaging.2013.12.021

Zheng, L., Mack, W. J., Chui, H. C., Heflin, L., Mungas, D., Reed, B., … Kramer, J. H. (2012). Coronary artery disease is associated with cognitive decline independent of changes on magnetic resonance imaging in cognitively normal elderly adults. *Journal of the American Geriatrics Society*, *60*(3), 499–504. https://doi.org/10.1111/j.1532-5415.2011.03839.x

Supplementary Appendix 3

Detailed results of literature review

The literature search produced 1253 records, and after exclusion of duplicate and irrelevant records, we fully analyzed 162 records. See Supplementary Appendix 1 and Supplementary Figure 1 for full details.

Of the 162 records, the data were of panel type for 156 records, and experimental for six.

The number of repeated measurements ranged from 2 to 20, with a mean of 5.071 (median = 5, s.d. = 2.895, interquartile range = 3 to 6; information not available for nine records).

Multilevel/mixed effects type models were used in 135 records (one with a penalized linear spline; one with a thin plate regression spline; one was nonlinear in the parameters; one was a generalized mixed-effects model), whereas a latent curve type model was used in 26 records (three analyzed multiple variables in parallel; one was a 2nd level curve model; two used a structured latent curve model). Finally, the remaining record applied both models.

One hundred and two records (63%) tested only a linear change function, and of these only one had experimental data, whereas the remaining 101 were panel studies. Moreover, the number of occasions ranged from 2 to 17 (median = 4.000, mean = 4.792, s.d. = 2.927). Forty (25%) studies tested also a quadratic function, and of these all were panel studies, with between 2 and 12 occasions (median = 5.000, mean = 5.351, s.d. = 2.163). One study tested a cubic function, and this study was a panel with five occasions.

Thirteen (8%) records tested a broken-stick function (a.k.a. single-node spline; ten of which compared it to a linear or quadratic function). These were all panel studies, with between 3 and 6 occasions (median = mean = 5.000, s.d. = 1.265).

Of the remaining records, five (3%) tested an exponential function (all with experimental data, with from 5 to 20 occasions, median=8.000, mean = 9.800, s.d. = 6.017), two (1%) estimated the change function from the data (within a latent curve model; both were panel studies, with either 3 or 4 occasions), one used a multiple-node spline approach (a panel study with 7 occasions), and one did not specify the change function (a panel study with 6 occasions). Overall, of the eight records that tested a nonlinear function, five were on experimental data, with number of occasions ranging from 5 to 20 (median = 8.000, mean = 9.800, s.d. = 6.017); all used occasions of measurement as the time basis; all applied an exponential function, and one also estimated the change function within a latent curve model.

~~T~~Across all studies, the most frequent time bases were time in study (n = 76), occasions of measurement (n = 41), chronological age (n = 33), and time to event (n = 13; e.g., death, dementia, stroke, hospitalization). One record used a time-varying medical covariate as basis, and three records did not provide this information. Finally, five records used and compared two of the aforementioned bases.

Supplementary Appendix 4

Logistic regression to study the determinants of an acceptable solution

To understand which design feature most strongly influences the probability of obtaining an acceptable solution, we computed a logistic regression analysis. The dichotomous variable was AS vs. non-AS, and the predictors were model type, GCR, T, Δ, and N. We computed this analysis for the shallower exponential decline rate only ($\gamma = 0.033$), given the very high number of AS for $\gamma = 0.066$. For simplicity, we present the analyses including all main effects only (all interactions proved irrelevant). Below we display the type-III likelihood ratio tests concerning each simulation factor, with its degrees of freedom and *p*-value, and the drop in effect size measures when each simulation factor is omitted from the logistic equation. The strongest effect was by far associated to the model type, while all other factors were associated to much weaker effect sizes.

Factor (df)	LRT	$\sim R^2 CS$	$\sim R^2 N$
model (3)	5490.772	.126	.249
GCR (2)	1921.237	.042	.083
T (1)	1845.481	.040	.079
Δ (4)	1239.208	.027	.053
N (2)	225.037	.005	.009
$\sim R^2$ *CS*	.232 (.269)		
$\sim R^2$ *N*	.457 (.529)		

Note. LRT=Type-III likelihood ratio tests associated to each simulation factor and design feature (rows) in a logistic regression predicting the probability of obtaining an acceptable solution (the associated degrees of freedom are in parenthesis; all *p*-values are <.001). $\sim R^2CS$ and $\sim R^2N$ = drop in pseudo R^2 of Cox and Snell and pseudo R^2 of Nagelkerke, respectively, when that effect is removed from the regression equation. The lower panel presents the total effect size indexes of pseudo R^2 of Cox and Snell (tot. R^2 CS) and of Nagelkerke (tot. R^2 N) with only the main effects, and in parentheses the analogous estimates with the addition of all interactions.

Supplementary Appendix 5

Summary of results of analyses on rejections of the linear GM and the quadratic GM when compared to the exponential GM and of the linear GM when compared to the quadratic GM for $\gamma = 0.033$ only (based on difference in BIC greater than 10)

	Linear vs. Exponential			Quadratic vs. Exponential			Linear vs. Quadratic		
Factor (df)	LRT	R^2CS	R^2N	LRT	R^2CS	R^2N	LRT	R^2CS	R^2N
GCR (2)	1575.403	.112	.159	830.301	.084	.113	991.222	.196	.307
T (1)	2443.488	.185	.261	1324.233	.140	.187	2096.576	.251	.382
Δ (4)	2503.807	.190	.269	1236.640	.130	.173	2171.507	.208	.296
N (2)	1200.185	.083	.118	288.879	.028	.037	870.941	.139	.232
tot. R^2 CS	.511 (.539)			.403 (.547)			.514 (.538)		
tot. R^2 N	.723 (.763)			.538 (.730)			.733 (.766)		

Note. LRT=Type-III likelihood ratio tests associated to each simulation factor and design feature (rows) for the linear vs. exponential (columns 2-4), quadratic vs. exponential (columns 5-7), and linear vs. quadratic (columns 8-10) GMs in logistic regressions predicting the probability of rejecting the linear or quadratic GM when compared to the exponential GM or the linear GM when compared to the quadratic GM (the associated degrees of freedom are next to the name of each term; all *p*-values are < .001). R^2CS and R^2N = drop in pseudo R^2 of Cox and Snell and pseudo R^2 of Nagelkerke, respectively, when that effect is removed from the regression equation. The lower panel presents the total effect size indexes of pseudo R^2 of Cox and Snell (tot. R^2 CS) and of Nagelkerke (tot. R^2 N) with only the main effects, and in parentheses the analogous estimates with the addition of all interactions.

Supplementary Appendix 6

Summary of Results of Analyses on Power to Detect Change Variance for γ= 0.033

Because the true variance in change (σ^2_C) is at a boundary (i.e., cannot be negative), we used a 50:50 mixture of a 1- and 2-degree-of-freedom chi square distribution (Self & Liang, 1987; Stoel, Garre, Dolan, & van den Wittenboer, 2006) to compute all likelihood ratio tests. However, given the reservations raised by some about this methodology (e.g., Savalei & Kolenikov, 2008), we also applied a 2-degree-of-freedom test. The results were identical.

	Linear			Quadratic			Exponential		
Factor (df)	LRT	R^2CS	R^2N	LRT	R^2CS	R^2N	LRT	R^2CS	R^2N
GCR (2)	248.320	.024	.043	493.457	.058	.096	546.575	.047	.077
T (1)	702.835	.071	.125	725.443	.087	.143	741.567	.065	.106
Δ (4)	932.453	.096	.168	993.848	.122	.200	1192.894	.107	.176
N (2)	660.821	.067	.117	279.642	.032	.053	901.961	.079	.131
total ~R^2 CS	.241 (.272)			.272 (.318)			.282 (.317)		
total ~R^2 N	.422 (.477)			.448 (.524)			.464 (.522)		

Note. Type-III likelihood ratio tests associated to each simulation factor and design feature (rows) for the linear (columns 2-4), quadratic (column 5-7), and exponential (column 8-10) GMs in logistic regressions predicting the probability of rejecting the alternative models in which $\sigma^2_C = \sigma_{LC} = 0$ (the associated degrees of freedom are next to the name of each term; all *p*-values are <.001). R^2CS and R^2N = drop in pseudo R^2 of Cox and Snell and pseudo R^2 of Nagelkerke, respectively, when that effect is removed from the regression equation. The lower panel presents the total effect size indexes of pseudo R^2 of Cox and Snell (total ~R^2 CS) and of Nagelkerke (total ~R^2 N) with only the main effects, and in parentheses the analogous estimates with the addition of all interactions.

References

Self, G. S., & Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, *82*, 605–610. https://doi.org/10.2307/2289471

Stoel, R. D., Garre, F. G., Dolan, C., & van den Wittenboer, G. (2006). On the likelihood ratio test in structural equation modeling when parameters are subject to boundary constraints. *Psychological Methods*, *11*(4), 439–455. https://doi.org/10.1037/1082-989X.11.4.439

Savalei, V., & Kolenikov, S. (2008). Constrained versus unconstrained estimation in structural equation modeling. *Psychological Methods*, *13*(2), 150–170. https://doi.org/10.1037/1082-989X.13.2.150

Supplementary Table 1

Number of records from each literature database for each combined use of statistical and substantive search terms

Search terms		**Database**			
Statistical	**Substantive**	**WoS**	**PI**	**PM**	**all**
"Growth curve model*"	"cogniti*"	164	128	16	308
"Growth curve model*"	"intelligence"	22	34	4	60
"Latent curve model*"	"cogniti*"	7	5	0	12
"Latent curve model*"	"intelligence"	2	2	0	4
"multilevel model*" and "longitudinal"	"cogniti*"	74	74	1	149
"multilevel model*" and "longitudinal"	"intelligence"	12	30	1	43
"mixed effects model*" and "longitudinal"	"cogniti*"	104	84	26	214
"mixed effects model*" and "longitudinal"	"intelligence"	7	16	2	25
"linear mixed model*" and "longitudinal"	"cogniti*"	98	79	23	200
"linear mixed model*" and "longitudinal"	"intelligence"	6	17	2	25
"random effects model*" and "longitudinal"	"cogniti*"	40	19	9	68
"random effects model*" and "longitudinal"	"intelligence"	0	2	0	2
"hierarchical linear model*" and "longitudinal"	"cogniti*"	48	52	4	104
"hierarchical linear model*" and "longitudinal"	"intelligence"	3	13	0	16
"nonlinear hierarchical model*" and "longitudinal"	"cogniti*"	0	0	0	0
"nonlinear hierarchical model*" and "longitudinal"	"intelligence"	0	0	0	0
"generalized linear mixed model*" and "longitudinal"	"cogniti*"	5	13	1	19

“generalized linear mixed model*” and “longitudinal”	“intelligence”	0	3	0	3
“nonlinear mixed model*” and “longitudinal”	“cogniti*”	1	0	0	1
“nonlinear mixed model*” and “longitudinal”	“intelligence”	0	0	0	0
Sum		593	571	89	1253
Records without intra-base duplicates		536	493	82	1111
Records without inter-bases duplicates			719		
Records retained for full lecture			**162**		

Notes. WoS=Web of Science; PI=PsycINFO; PM=PubMed.

Supplementary Table 2

Parameter Estimates (and Standard Errors) of the Linear, Quadratic, and Exponential GM applied to the Recall Score in the Betula Study.

parameter	level GM	linear GM	quadratic GM	exponential GM
β_0	23.521 (0.251)	23.506 (0.214)	24.615 (0.218)	20.603 (0.439)
β_1	--	-0.254 (0.012)	-0.341 (0.013)	-4.007 (0.451)
β_2	--	--	-0.009 (<0.001)	--
γ	--	--	--	0.066 (0.004)
σ^2_L	55.518 (2.840)	33.758 (2.246)	30.779 (1.983)	30.980 (2.786)
σ^2_C	--	0.032 (0.007)	0.027 (0.006)	5.041 (1.402)
$\sigma_{L,C}$ [$\rho_{L,C}$]	--	0.638 (0.086) [0.619]	0.379 (0.072) [0.413]	-3.163 (1.599) [-0.253]
σ^2_E	18.148 (0.576)	16.444 (0.597)	14.942 (0.541)	14.541 (0.526)
-2LL	19‘453	18’911	18’692	18’686
BIC	19‘473	18’952	18’741	18’734
p	3	6	7	7

Note. GM=growth model. $\rho_{L,C}$ = correlation between level and change. *p* = number of estimated parameters. All models were estimated with age centered at 65 years. Standard errors are in parentheses. For the quadratic model and the exponential model it was not possible to estimate the variance, respectively, of the quadratic slope and of the exponential decline rate.

Supplementary Table 3

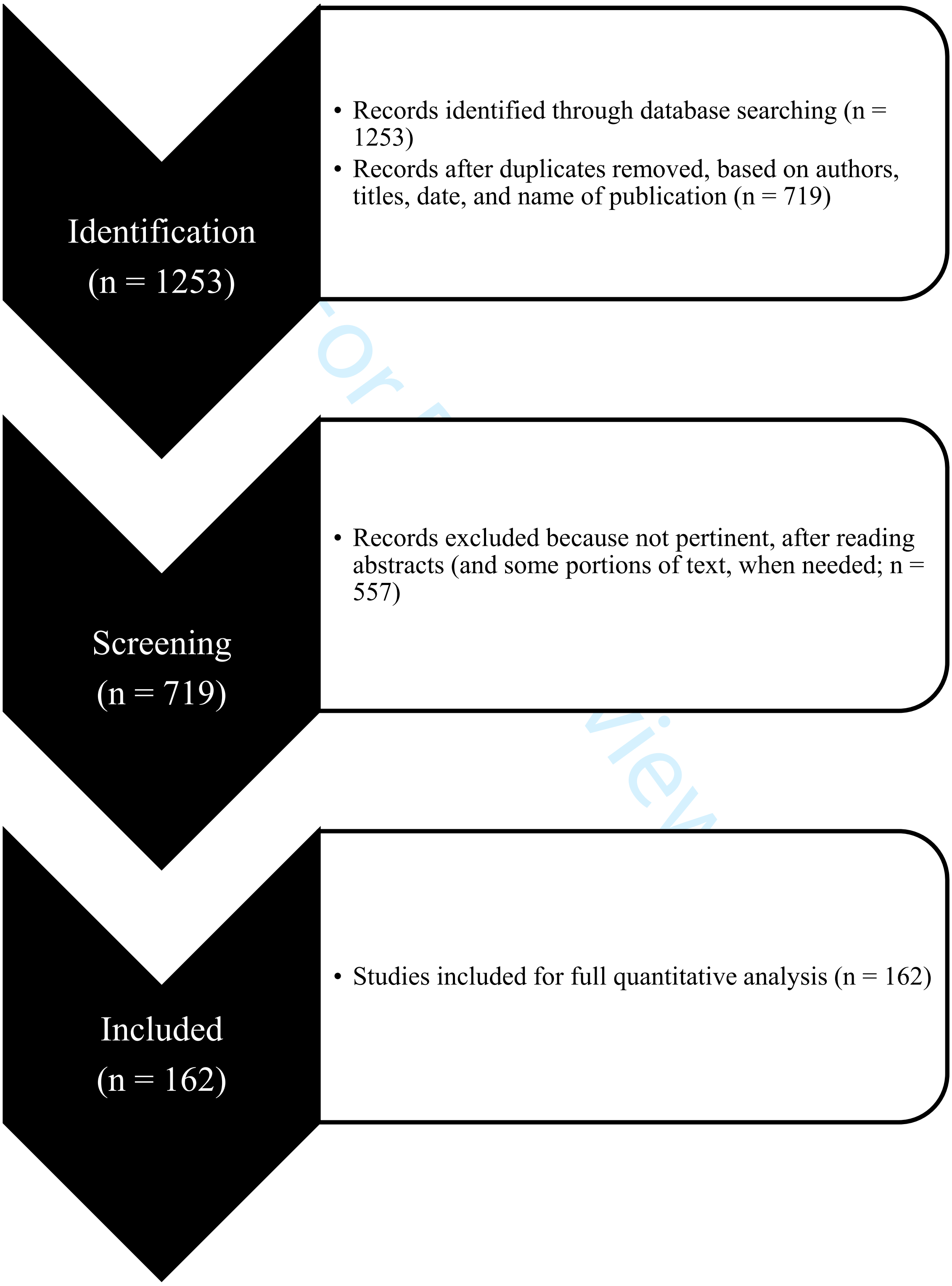
Factors in Simulation Study

Constant Simulation Factors	
Level mean (β_0)	20.603
Change mean (β_1)	-4.007
Level variance (σ^2_L)	30.980
Change variance (σ^2_C)	5.041
Level-change covariance ($\sigma_{L,C}$)	-3.163
Varying Simulation Factors	
Rate of exponential decline (γ)	0.033, 0.066
Growth Curve Reliability $\left(GCR = \frac{\sigma_L^2}{\sigma_L^2 + \sigma_E^2} \right)$	.500, .681, .900
Varying Design Factors	
Occasions (T)	3, 6
Interval (Δ)	1, 2, 3, 4, 5
Total sample size (N)	500, 1000, 2000

Note. The factors refer to the exponential decline function (Equation (4)). The full design includes 5 fully crossed varying factors (γ, GCR, T, Δ, N) for a total of 2 x 3 x 2 x 5 x 3) = 180 conditions. On each condition 4 growth models were tested (level, linear, quadratic and exponential) and each condition contained 100 replications. In the end, we obtained 180 x 4 x 100 = 72000 sets of solutions.

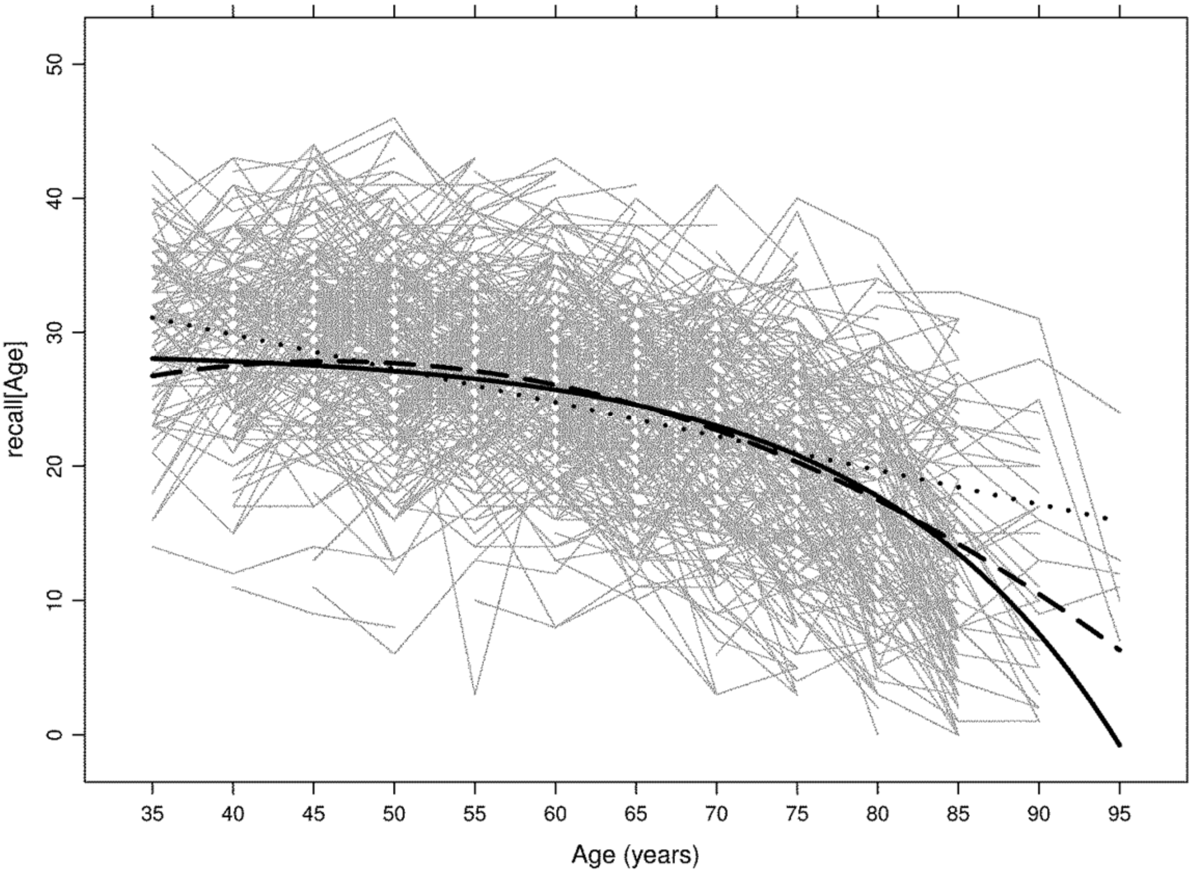
Supplementary Figure 1

Flowchart with screening procedure of articles included in the quantitative analysis



Supplementary Figure 2

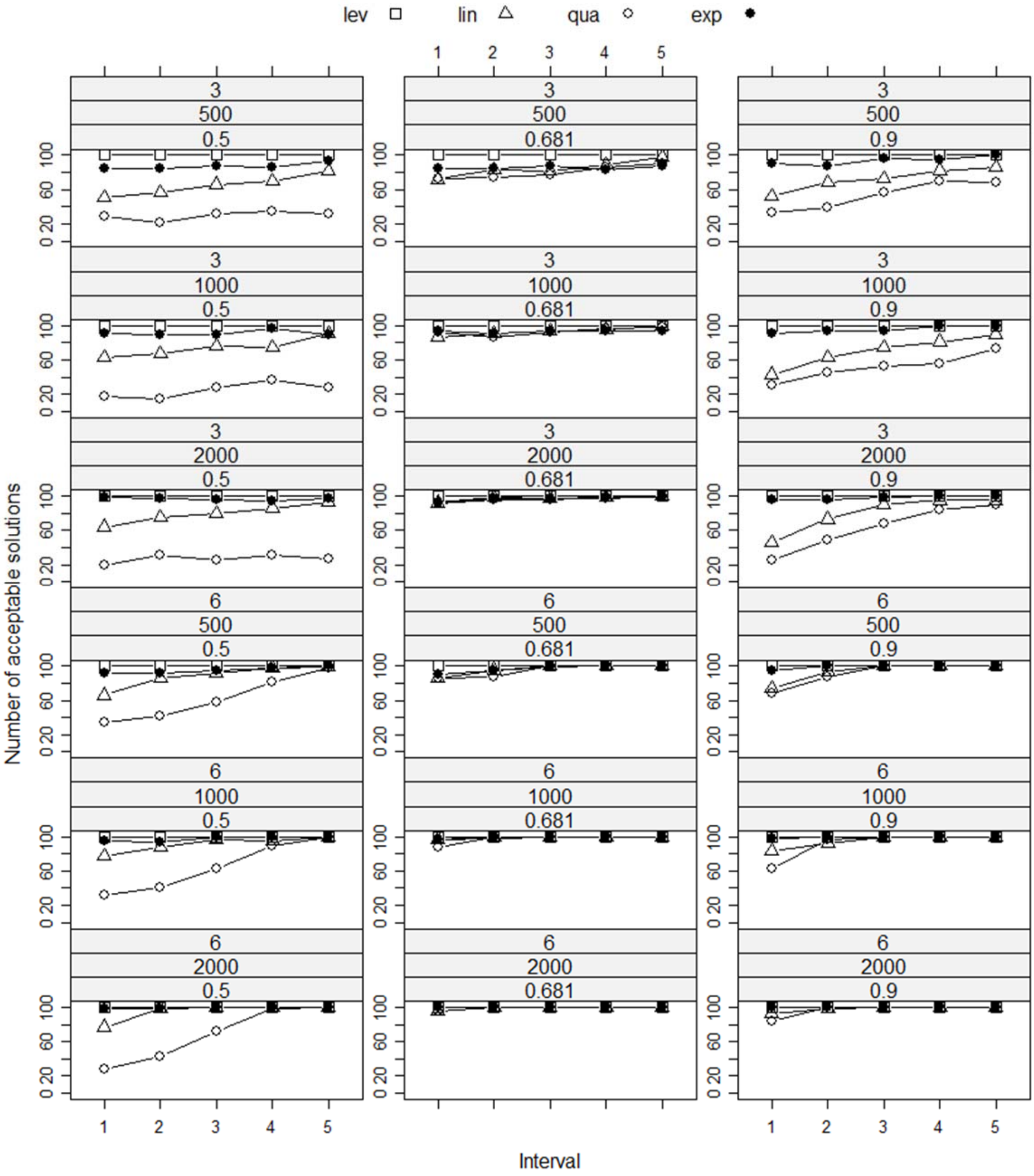
Betula recall scores by age



Note. The sample predicted trajectory is represented by the black dotted line for the linear GM, the black dashed line for the quadratic GM, and the continuous black line for the exponential decline GM. The grey lines represent the observed individual trajectories.

Supplementary Figure 3

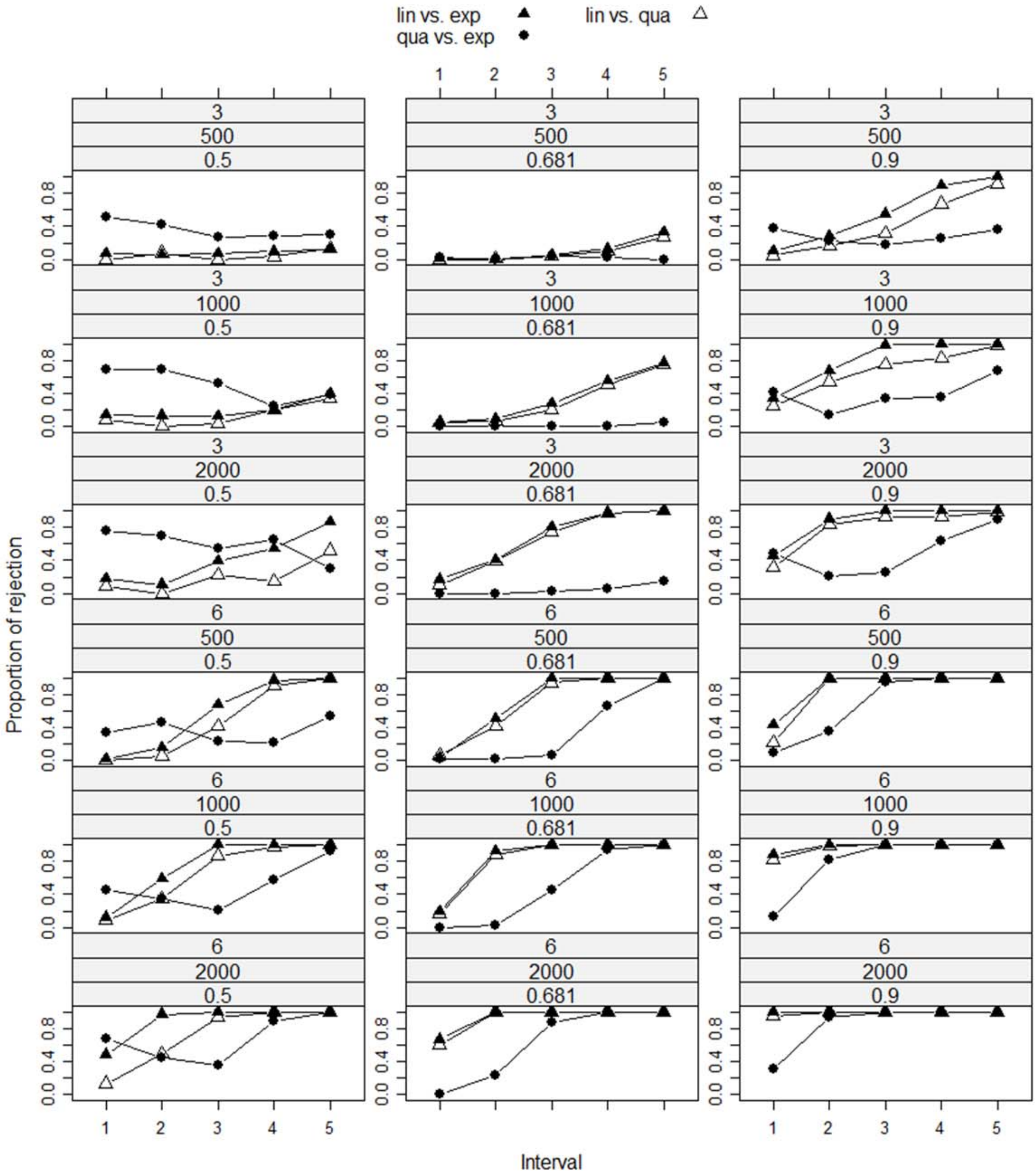
Average number of acceptable solutions by interval width for γ= 0.033, separately for each model



Note. Rows 1, 2, 3: $T = 3$, rows 4, 5, 6: $T = 6$; Rows 1, 4: $N = 500$, rows 2, 5: $N = 1000$, rows 3, 6: $N = 2000$; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900.

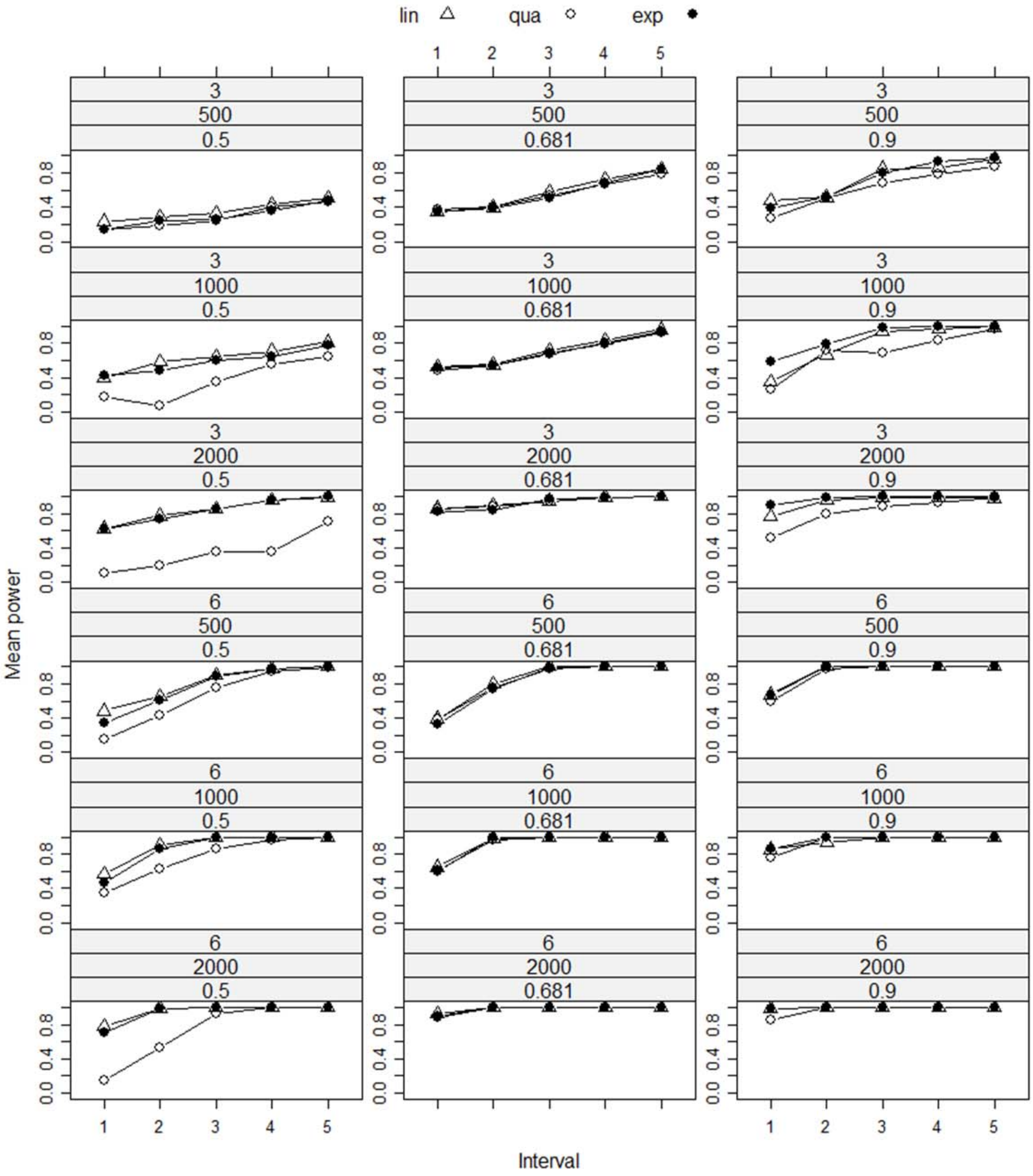
Supplementary Figure 4

Proportions of rejection of the linear GM and the quadratic GM when compared to the exponential GM and of the linear GM when compared to the quadratic GM (based on difference in BIC greater than 10) as a function of interval width for γ= 0.033 only



Note. Rows 1, 2, 3: T = 3, rows 4, 5, 6: T = 6; Rows 1, 4: N = 500, rows 2, 5: N = 1000, rows 3, 6: N = 2000; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900.

Supplementary Figure 5
Power to Detect Variance in Change by Interval Width for γ= 0.033, for each Model



Note. Rows 1, 2, 3: $T = 3$, rows 4, 5, 6: $T = 6$; Rows 1, 4: $N = 500$, rows 2, 5: $N = 1000$, rows 3, 6: $N = 2000$; Column 1: GCR = .500, column 2: GCR = .681, column 3: GCR = .900.