Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: https://www.tandfonline.com/loi/hmbr20

Age Differences in Day-To-Day Speed-Accuracy Tradeoffs: Results from the COGITO Study

Paolo Ghisletta, Emilie Joly-Burra, Stephen Aichele, Ulman Lindenberger & Florian Schmiedek

To cite this article: Paolo Ghisletta, Emilie Joly-Burra, Stephen Aichele, Ulman Lindenberger & Florian Schmiedek (2018) Age Differences in Day-To-Day Speed-Accuracy Tradeoffs: Results from the COGITO Study, Multivariate Behavioral Research, 53:6, 842-852, DOI: 10.1080/00273171.2018.1463194

To link to this article: https://doi.org/10.1080/00273171.2018.1463194

Published online: 23 Apr 2018.

MULTIVARIATE BEHAVIORAL RESEARCH
2018, VOL. 53, NO. 6, 842–852
https://doi.org/10.1080/00273171.2018.1463194

Age Differences in Day-To-Day Speed-Accuracy Tradeoffs: Results from the COGITO Study

Paolo Ghisletta[a,b], Emilie Joly-Burra[a], Stephen Aichele[b], Ulman Lindenberger[c], and Florian Schmiedek[d]

[a]Faculty of Psychology and Educational Sciences, University of Geneva; [b]Universities of Lausanne and of Geneva; [c]Center for Lifespan Psychology, Max Planck Institute for Human Development; [d]German Institute for International Educational Research (DIPF)

ABSTRACT

We examined adult age differences in day-to-day adjustments in speed-accuracy tradeoffs (SAT) on a figural comparison task. Data came from the COGITO study, with over 100 younger and 100 older adults, assessed for over 100 days. Participants were given explicit feedback about their completion time and accuracy each day after task completion. We applied a multivariate vector auto-regressive model of order 1 to the daily mean reaction time (RT) and daily accuracy scores together, within each age group. We expected that participants adjusted their SAT if the two cross-regressive parameters from RT (or accuracy) on day *t-1* of accuracy (or RT) on day *t* were sizable and negative. We found that: (a) the temporal dependencies of both accuracy and RT were quite strong in both age groups; (b) younger adults showed an effect of their accuracy on day *t-1* on their RT on day *t*, a pattern that was in accordance with adjustments of their SAT; (c) older adults did not appear to adjust their SAT; (d) these effects were partly associated with reliable individual differences within each age group. We discuss possible explanations for older adults' reluctance to recalibrate speed and accuracy on a day-to-day basis.

KEYWORDS

Reaction time; multilevel vector autoregressive model; speed-accuracy tradeoff; COGITO

When making decisions in daily life, we know that if we are under time pressure to solve a problem, we risk reducing accuracy, whereas if we want to produce an accurate solution we need time. We thus face the dilemma of sacrificing time for accuracy, or accuracy for response time. This is the so-called speed-accuracy trade-off (SAT; Fitts, 1954; Garrett, 1922; Norman & Bobrow, 1975; Pew, 1969; Yellott, 1971). The SAT can be observed in various decision-making processes not only in humans, but also in insects, rodents, and primates, among others (for a review see Heitz, 2014). There are, however, situations in which taking a longer time may increase doubt, leading to an inaccurate answer. This is true for some psychophysical and recognition memory tasks, for which fast responses tend to be more accurate than slow responses (Norman & Brobow, 1975). Here, however, we focus on SAT in the context of choice reaction time tasks only, where slower responses generally tend to be more accurate than faster ones.

Methodologically, the SAT can be studied via experimental conditions aimed at influencing subjects' decision criteria. Common experimental manipulations include verbal instructions (emphasizing either speed or accuracy), payoffs (rewarding correct answers and penalizing errors), temporal deadlines for task completion, varying response-to-stimulus intervals, and stimulus strength (Heitz, 2014; Palmer, Huk, & Shadlen, 2005; Wickelgren, 1977). The purpose of such experimental manipulations is to influence subjective decision criteria in the hope of gaining insight about the decision process. Existing research on the SAT spans various methodological and statistical approaches. In his review, Heitz (2014) discusses several mathematical models that have been proposed to study how subjects (human or not) adjust to different experimental conditions, while trading off accuracy for speed during choice reaction time tasks. Probably most popular are models assuming that individuals sequentially sample information during a

CONTACT Paolo Ghisletta paolo.ghisletta@unige.ch Faculty of Psychology and Educational Sciences, 4136 UniMail, Boulevard du Pont d'Arve 40, 1211 Genève, Switzerland.

Full affiliation information for all authors is listed as follows: Paolo Ghisletta, Faculty of Psychology and Educational Sciences, University of Geneva, Switzerland; Swiss Distance Learning University, Switzerland; Swiss National Centre of Competence in Research LIVES – Overcoming vulnerability: Life course perspectives, Universities of Lausanne and of Geneva, Switzerland. Emilie Joly-Burra, Faculty of Psychology and Educational Sciences, University of Geneva, Switzerland. Stephen Aichele, Swiss National Centre of Competence in Research LIVES – Overcoming vulnerability: Life course perspectives, Universities of Lausanne and of Geneva, Switzerland. Ulman Lindenberger, Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany; Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany; European University Institute, San Domenico di Fiesole, Italy. Florian Schmiedek, German Institute for International Educational Research(DIPF), Germany; Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany.

Parts of these analyses have been presented at the workshop "The COGITO study: Looking at 100 day ten years after" (Oct. 2–6, 2016), Max Planck Institute for Human Development, Berlin, Germany.

decision process and continuously update prior knowledge or expectations. The amount of evidence that needs to be accumulated to elicit a decision between alternative choices (i.e., the response criterion) can vary across and within individuals. The accumulated evidence affects the duration of the decision process and also the likelihood of making a correct decision, and therefore reflects the SAT. One popular model is the drift-diffusion model, which allows translating common behavioral indicators (like accuracy, mean response times, and response time distributions) into components of cognitive processing (like efficiency of evidence accumulation and response criterion setting) to refine theoretical understanding of decision processes in two-choice reaction time tasks (Ratcliff, 1978; Ratcliff & McKoon, 2008).

The drift-diffusion model has been applied to aging studies, typically by comparing younger and older adults' performance with respect to the SAT. Starns and Ratcliff (2010) concluded that older adults consistently prioritize accuracy over speed, even when instructed to prioritize faster response time, whereas younger adults focus on balancing speed and accuracy, thereby aiming for an optimum SAT. In the drift-diffusion model, the separation of the decision boundaries in a two-choice decision task represents the amount of information needed to choose one of the response alternatives. The greater this parameter, the more accurate, but also the slower, the answer. For a given amount of time, there is an optimal decision boundary that maximizes accuracy. Across multiple decision tasks, and under different instructions (emphasizing either speed or accuracy), the authors found that younger adults used boundaries closer to the optimum compared to older adults, whereas older adults consistently used boundaries that were much wider than the x optimum.

Moreover, younger adults benefited from accuracy feedback and practice to more closely approximate the optimal decision boundary, whereas older adults did not. In a follow-up study, Starns and Ratcliff (2012) further studied age differences in SAT strategies by adding so called "fixed-time blocks," that is, blocks of trials that lasted a fixed amount of time, regardless of the number of trials completed (in contrast to fixed-trial blocks, composed of a fixed number of trials, independent of the time needed for completion). Participants were instructed to aim at getting as many correct trials as possible within the (same) time limit, and that to do so they should simply speed up, given that they were not penalized for errors. The authors expected that, if older and younger adults set different task goals, the fixed-time condition should at least alleviate the age difference in boundary width. Results showed that older adults consistently used boundaries farther away from the optimal decision boundary than younger adults, but that both age groups reduced boundary width in the fixed-time condition compared to the fixed-trial condition. Furthermore, the age difference in boundary width was the same across both conditions, meaning that the observed effect is not entirely due to differences in group-specific task goals. The authors concluded that younger adults were better able to adjust their SAT to the new performance objective than older adults.

The relation between speed and accuracy in cognitive performance has been studied intensively at the micro-level, that is, across trials within a unique test session (Heitz, 2014). Indeed, the vast majority of studies of the SAT follow a traditional experimental design composed of multiple trials within a limited number of sessions, all administered on the same day. In real life, however, we often need to make multiple decisions of the same type within the same day; correspondingly, SAT has been shown to vary across repeated assessments of the same task within the same day (e.g., Gueugneau, Pozzo, Darlot, & Papaxanthis, 2017). SAT on an abstract two-choice reaction time task may also vary across multiple days. To examine this possibility would require a study design in which the same task is administered at least once per day across multiple days. Additionally, feedback concerning speed and accuracy performance on a given day may invoke strategies that influence the tradeoff between speed and accuracy on the following day (Siegler, 1994, 2007).

The COGITO study (Schmiedek, Bauer, Lövdén, Brose, & Lindenberger, 2010a; Schmiedek, Lövdén, & Lindenberger, 2010b) used a very intensive multiple-measure cognitive intervention within a sample of over 100 younger and 100 older adults, who were assessed each day during a period of at least 100 days. We focus here on the most difficult of three two-choice comparison reaction time tasks. This task required participants to compare two complex figural stimuli and decide whether these were identical or different. As such, the task differed from a two-alternative forced choice paradigm, based on inspection of a single stimulus, such as an odd-even decision for a number. Hence, the present task is not well suited for application of the standard drift-diffusion model. Instead, we analyze daily mean RT and daily accuracy scores to examine how these trade off with each other at the macro-level, that is, over the entire testing period.

Method

Participants

The COGITO study included 101 younger (51.5% women, age: 20–31 years) and 103 older (49.5% women, age: 65–80 years) adults, recruited in Berlin, Germany. All participants were living independently and in good health, and their general cognitive functioning was representative of their age group and comparable to those of larger population-based studies (Schmiedek et al., 2010b). Participants were financially remunerated.

Procedure

The study consisted of a pretest phase, an intensive training phase lasting at least 100 days, and a posttest phase. During the pre- and posttest phases, a broad cognitive battery with tasks on reasoning, episodic memory, perceptual speed, and working memory was administered. During the intensive training phase, some of the same perceptual speed, working memory, and episodic memory tasks were administered. In this article we consider participants' performance on a two-choice perceptual speed task with figural stimuli during the intensive training phase. All tasks were computerized and participants were tested on individual workstations in rooms with at most five other people. Participants could not communicate with each other. At the end of each day, the computer program provided feedback on how accurate and fast the participants' responses were, by showing them their overall accuracy (sum of correct items) and mean reaction time, and giving participants the opportunity to print out and take home a sheet with the results of the day (Schmiedek et al., 2010a, 2010b).

Two-choice figural comparison task

This figural comparison task consisted of two "fribbles," three-dimensional colored objects comprising several connected parts, presented on either side of the screen. The fribbles came from a battery developed by the laboratory of professor Michael J. Tarr (Carnegie Mellon University, Pittsburgh, PA, USA, http://www.tarrlab.org). Participants had to decide as quickly as possible if the two fribbles were identical or different. When the two fribbles differed, they did so by only one element.

At pretest, five blocks of 40 trials each were shown (total of 200 trials). This high number of trials assured that participants became familiar with the fribbles. The intensive training phase consisted of at least 100 days, with two blocks of 40 trials each per day. Thus, during the training phase each participant received a total of at least 8000 training trials (40 trials × 2 blocks × 100 days). Reaction times (RTs; in ms) and accuracy (0 = wrong vs 1 = correct) were measured.

Analyses

Preliminary data preparation and transformations

Because our interest was in SAT adjustment on a day-to-day basis (possibly due to the daily speed and accuracy feedback provided to participants), for each individual we combined the 80 daily RT and accuracy measurements, so as to obtain (a) a single mean RT and (b) a single accuracy proportion score (ranging from 0 to 1) for each of the 100 days. Thus, for each individual we obtained a time series of 100 RT means and a time series of 100 accuracy proportion scores. Prior to calculating the RT means, we eliminated all RTs below 100 ms (<1%), because we considered these to be anticipatory responses, rather than responses to the experimental stimulus. We obtained daily accuracy proportion scores by calculating the proportion of correct items on each day. To normalize the data, we computed (a) the natural log of all daily mean RTs (i.e., ln(mean RT)) and (b) the logit of the accuracy proportion scores $[\ln(\hat{p} / (1 - \hat{p})]$, where $\hat{p}$ is the proportion of correct responses; in case of perfect accuracy ($\hat{p} = 1$), to calculating $\ln[1 / (1 - 1)]$, we replaced ($\hat{p} = 1.00$) with ($\hat{p} = \{[(max - 1) + max] / 2\} / max$), where *max* is the maximum number of items presented at a given day (almost always 80); that is, we replaced the perfect accuracy proportion 1.00 with the proportion of the average of the perfect total score and one-less-than-perfect total score ($\hat{p} = \{[(80 - 1) + 80] / 2\} / 80 = 0.99375$, when $max = 80$). There were no observations with $\hat{p} = 0.00$, which would also have been computationally problematic (i.e., $\ln[0 / (1 - 0)]$).

Multilevel first-order vector autoregressive model

We conceived the process of SAT adjustment as the influence that a RT or accuracy score of a given day could exert on the accuracy or RT score of the following day, respectively, over and above the autoregressive effect of either score. That is, we examined the extent to which one's RT on day *t-1* influences the accuracy on day *t*, possibly because feedback about being slower than usual would motivate the participant to be faster, and consequently also less accurate, on day *t*. Likewise, knowing that one's accuracy on a given day was below one's previous scores may stimulate the participant to be more accurate, but consequently also slower, on the following day. We thus expected that negative cross-regressive effects between RT and accuracy would indicate an adjustment in SAT, meaning that higher accuracy on day *t-1* would lead one to try to be faster (but also potentially somewhat less accurate) on day *t*, and lower accuracy on day *t-1* would lead one to slower but more accurate responses on day *t*. Likewise, a slower RT on day *t-1* (and therefore higher accuracy on that day) may stimulate one to be less accurate, but faster, on day t, whereas faster (but less accurate) responses on day t-1 would lead to more accurate (but slower) responses on day t.

To examine the dynamic associations between the RT and the accuracy time series, which we defined as the auto- and cross-regressive effects of lag 1 day (between day t-1 and day t), we used a first-order vector autoregressive (VAR(1)) model. We kept the lag at 1 day, because we assumed that participants may have had an immediate reaction to the feedback about their daily performance. Typical VAR models are defined and estimated on single individual's intensive data, that is, on $N = 1$. In the present application, this would imply estimating a separate VAR(1) model for each of the over 200 COGITO participants. Recently, multilevel extensions of the VAR(1) model have been proposed to take advantage of the nested structure of the data: The intensive repeated assessments at level-1 are nested within the individuals at level-2. This model decomposes the daily RT and accuracy scores into between-person and within-person parts, and at the same time estimates the average dynamic association among the within-person parts, allowing them to vary between persons. This model can be estimated using the latest version of the latent variable modeling program Mplus (version 8; Muthén & Muthén, 1998–2017) as illustrated in Hamaker, Asparouhov, Brose, Schmiedek, and Muthén, in press. The parameters are estimated within the Bayesian framework, and thus point estimates are defined by posterior means (or medians) and inferential conclusions can be reached by examining the credible intervals of the posterior distribution. We estimated a multilevel VAR(1) (MLVAR(1)) model separately in each age group. For further information on this model, see Asparouhov, Hamaker, and Muthén (2018) and the Mplus manual (Muthén & Muthén, 1998–2017). Appendix A presents the commented Mplus (version 8) syntax of this model.

The model can be represented as follows:

$$\begin{aligned} RT_{i,t} &= \mu_{RT,i} + \phi_{RT\rightarrow RT,i} RT_{i,t-1} + \phi_{A\rightarrow RT,i} A_{i,t-1} + \varepsilon_{RT,i,t} \\ A_{i,t} &= \mu_{A,i} + \phi_{A\rightarrow A,i} A_{i,t-1} + \phi_{RT\rightarrow A,i} RT_{i,t-1} + \varepsilon_{A,i,t}, \end{aligned} \tag{1}$$

where RT and A are reaction time and accuracy (transformed) scores, respectively, t-1 and t indicate any two successive days, and μ is the overall level or mean score across the 100 days. Note that the hypotheses cited above about SAT adjustment reflect a pure within-person process, rather than a between-person phenomenon. As a person learns about her/his slow response time or low accuracy level on day t-1, she/he is hypothesized to adjust the answer style to reduce the response time or increase the accuracy level on day t. This within-person process is clearly fundamentally different from a between-person phenomenon that posits, for instance, that participants that are (generally) slow are (generally) also accurate, compared to fast participants. Thus, the bulk of the model is about within-person deviations, in both RTs and accuracy, from the general level of performance. In Equation (1), $\mu_{RT,i}$ and $\mu_{A,i}$ represent the between-person differences in performance, and characterize solely interindividual differences, whereas the remaining elements represent the within-person, or intraindividual, variability, which represents person-specific daily fluctuations around the overall mean. This decomposition of RT and accuracy scores into between-person and within-person components holds if the time series is stationary (either around zero or a constant mean), but breaks down if the time series is subject to a trend (cf. Curran & Bauer, 2011).

The within-person component of the time series is then modeled according to a VAR model: ϕ represents either an autoregressive ($\phi_{RT\rightarrow RT,i}$ and $\phi_{A\rightarrow A,i}$) or a cross-regressive ($\phi_{A\rightarrow RT,i}$ and $\phi_{RT\rightarrow A,i}$) parameter, and ε is the remaining dynamic error on day t. The simultaneous application of the model to all participants rests on the multilevel structure of the data, where the time measurements, denoted by t, are nested within the individuals' data, denoted i. Level-2 (individual specific) random effects are defined for all components of the model with a subscript i, and these are allowed to correlate. More precisely, the person-specific random effects around the overall level scores ($\mu_{RT,i}$ and $\mu_{A,i}$), the autoregressive ($\phi_{RT\rightarrow RT,i}$ and $\phi_{A\rightarrow A,i}$) and the cross-regressive ($\phi_{A\rightarrow RT,i}$ and $\phi_{RT\rightarrow A,i}$) parameters correlate with each other at the between-person level. Equation (1) represents the model at the within-person level (level-1), where the dynamic errors ($\varepsilon_{RT,i,t}$ and $\varepsilon_{A,i,t}$) are allowed to correlate with each other at the within-person level.

Stationarity

Time series data may display trends, such as an initial decreasing pattern in RT or increasing pattern in accuracy, signifying that participants learn during the first trials and as a result become faster and more accurate. To detect such trends, we visually inspected the time series and also computed the Augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1979). The ADF is a popular test for assessing stationarity around zero, around a constant value (drift), and around a constant value and a regular change pattern (drift and trend). The ADF test assesses weak or second-order stationarity, that is, the constancy of the mean and the variance of the outcome across repeated assessments (Wei, 2012). We excluded from the subsequent analyses all individuals with non-stationary time series, mainly because the methods for detrending time series are varied and can obfuscate the time-dependency information therein (Wang, Hamaker, & Bergeman, 2012). This assures that the decomposition of the time series into between-person and within-person components as specified in Equation (1) holds.

Irregular testing schedules and missing data

Although the full sample of 204 participants was tested on 100 days, these were not consecutive days. Testing never took place on weekends, holidays often meant interruptions of over one week, and the testing was not intended to interfere heavily with participants' life. Thus, in reality, the 100 daily administrations took place on average across 161.62 days in younger adults (ranging from 117 to 373 days) and 150.00 in older adults (from 114 to 276 days). Standard statistical software presupposes that time series are composed of regularly spaced measurements, which implies no missing observations. This assumption was clearly violated in the COGITO data. Nevertheless, Mplus version 8 allows estimating autoregressive parameters even in the presence of irregular testing schedules and missing data via Bayesian estimation. The software applies an algorithm that, if needed, discretizes the time variable, thereby allowing approximating a continuous time series by a discrete time series. In our case, this step was not necessary, because participants were tested on a given day. We just needed to define the individual time series such that time did not mean the number of the occasion of measurement. Rather, we needed to specify time as the number of days elapsed since the first occasions, starting at 1, for every participant. Hence, a participant first tested on a Monday, then on the following Wednesday would have the values $t = 1$ and $t = 3$. This participant missed the Tuesday assessment, so that $t = 2$ corresponds to a missing RT and a missing accuracy value. The natural interval width for this discrete time variable is 1 (day). In the end, this time metric implies that t does not span from 1 to 100, but from 1 to (a) at least 114 and 177 and (b) at most 373 and 276 in younger and older adults, respectively.

As is common in Bayesian estimation, missing data are then considered just like any other parameter in the model, as unknown random variables for which a posterior distribution can be estimated by relying on their prior distribution and the likelihood function (Gelman et al., 2013). Thus, missing data add little complexity to the estimation process. The Markov chain Monte Carlo algorithm is the usual simulation-based approach to estimating the parameters' and missing values' posterior distributions.

Results

Stationarity

Given that the dependent variables of the time series considered here are daily mean RTs and daily accuracy scores of a speed task, it is not surprising that stationarity around zero was clearly refuted across all participants. Stationarity across a constant, non-zero mean was deemed acceptable in all but 16 (eight younger and eight older adults) and 3 participants (all older adults) for RT and accuracy, respectively. Thus, after removing a constant mean from the individual time series of the remaining 188 RT (93 younger and 95 older adults) and 201 (101 younger and 100 older adults) accuracy series, the resulting time series can be considered weakly stationary. Because we applied a VAR model and used Bayesian estimation, the final sample sizes were $n = 101$ younger and $n = 102$ older adults (in fact, for some rare participants data were collected on only either RT or accuracy scores, but not on both).

Rather than removing the mean, we estimated a MLVAR(1) model that included intercept terms (cf. $\mu_{RT,i}$ and $\mu_{A,i}$ in Equation 1). The number of non-stationary series during the training phase might be expected to be low, because typically participants become faster and more accurate with repeated exposure to the same cognitive tasks. However, participants underwent a rather extensive pretest phase before the training phase, so that most of their learning occurred during this pretest. Thus, for the vast majority of participants no further systematic learning (i.e., trend) occurred after pretest. Graphical inspection of the time series confirmed this conclusion: The stationary series displayed no systematic trend, whereas the non-stationary series showed an increasing trend for the accuracy scores and a decreasing trend for the RTs.

MLVAR(1)

Bayesian estimation procedures

We estimated the model with 50000 iterations and a thinning of 10, thereby basing our results on 5000 iterations (cf. Hamaker et al., in press). We specified the Gibbs random walk sampler algorithm for the Markov chain Monte Carlo estimation and ran two parallel chains. We used the default specification for priors (i.e., normal distribution with mean = 0 and variance = 10^{10}). This corresponds to a highly diffused (non-informative) prior, which results in a greater weight on the likelihood than on the prior distribution during the estimation procedure. We concluded that the models converged in both age groups because (a) the potential scale reduction statistic was relatively close to 1 (1.08 in both younger and older adults, meaning that in both age groups the two chains converged in their results, despite having different starting values), (b) the posterior distributions of all parameters were symmetrical, (c) trace plots of all posterior parameters indicated stability in the final solution, and (d) the autocorrelation plot for each parameter indicated values quickly dropping to zero (Gelman et al., 2013; Plummer, Best, Cowles, & Vines, 2006). Estimation was computationally quite intensive (i.e., estimation of each model took from about two to over seven hours on a PC with two 2.39-GHz processors of with 8 GB of RAM).

Fixed-effects and variance estimates

The model's estimates of fixed effects and variances are shown in Table 1, with their 95% credible intervals (CIs) for both age groups. As expected, younger adults appear overall both faster (μ_{RT} of 7.16 vs. 7.71, corresponding to $e^{7.16} = 1287$ ms vs. $e^{7.71} = 2231$ ms) and more accurate (μ_A *of* 3.11 vs. 2.71, corresponding to proportions of correct responses of $e^{3.11}$ / $(1 + e^{3.31}) = .96$ vs. $e^{2.11}$ / $(1 + e^{2.71}) = .94$) than older adults. There are reliable individual differences in both overall RT and accuracy levels for both age groups, as indicated by the CIs of the variances of μ_{RT} and μ_A.

For the auto- and cross-regressive parameters, we present the estimates in two metrics: raw (in the upper line of each cell) and within-level (i.e., within-person) standardized (lower line). The latter metric corresponds to the *average* within-person standardized values and is particularly useful when comparing the two cross-regressive parameters to each other. Their magnitudes can be considered equivalent when standardized, and thus conclusions about asymmetrical spill-over effects can be reached (for more details, see Schuurman, Ferrer, de Boer-Sonnenschein, & Hamaker, 2016, and Hamaker et al., in press). In particular here, the RTs and accuracy scores (albeit their transformations prior to analyses) are in different metrics. This implies that their cross-regressive effects (RT influencing upcoming accuracy scores and vice versa) should be examined in the within-person standardized metric, whereas for the autoregressive effects (each variable influencing itself at a later time) the two metrics yield very similar estimates. The autoregressive parameters of both RT and accuracy are quite strong in both age groups, but younger adults seem to displayed a stronger RT carryover effect than older adults (0.61 vs. 0.53, respectively), whereas older adults appear to show more regularity in accuracy than younger adults (0.44 vs. 0.26, respectively). All autoregressive parameters show between person variance indicating that individual differences exist.

The younger adults show negative spill-over effects, which is coherent with our expectations about SAT adjustment. In particular, the effect that accuracy on day *t-1* exerts on RT on day *t*, over and above the autoregressive effect of RT, appears to differ from zero with a negative sign at the group level (−0.09 in standardized metric). The opposite effect, also in the expected negative direction, did not reliably differ from zero in this age group. Of interest, the random effects around $\phi_{A\to RT}$ do not differ from zero, whereas the random effects around $\phi_{RT\to A}$ are quite large. Overall, in younger adults, RT on day *t-1* does not appear to influence accuracy the next day, but there is evidence that for some younger individuals this effect is likely to be at play and negative in sign.

The older adults, however, show positive spill-over effects in both directions. Thus, in this age group accuracy on day *t-1* influences RT on day *t* ($\phi_{A\to RT} = 0.07$), whereas RT on a given day influences accuracy on the next day ($\phi_{RT\to A} = 0.05$). This result indicates that higher RTs and greater accuracy on day *t-1* implies even higher

Table 1. Posterior means and 95% credible intervals of fixed effects and variances for younger and older adults estimated from the VAR model. Auto- and cross-regressive parameters are shown in both raw (upper line) and within-level standardized (lower line) metrics.

	Fixed Effects		Random Variances	
	Younger Adults	Older Adults	Younger Adults	Older Adults
μ_{RT}	7.16 [7.10 – 7.22]	7.71 [7.67 – 7.75]	0.09 [0.07 – 0.13]	0.04 [0.03 – 0.06]
μ_A	3.11 [2.89 – 3.33]	2.71 [2.53 – 2.89]	1.25 [0.89 – 1.55]	0.77 [0.56 – 1.05]
$\phi_{RT\to RT}$	0.61 [0.56 – 0.66]	0.53 [0.48 – 0.57]	0.05 [0.04 – 0.07]	0.05 [0.03 – 0.07]
	0.61 [0.58 – 0.63]	0.52 [0.50 – 0.54]		
$\phi_{A\to A}$	0.26 [0.21 – 0.32]	0.44 [0.40 – 0.48]	0.06 [0.04 – 0.80]	0.03 [0.02 – 0.05]
	0.26 [0.24 – 0.29]	0.44 [0.42 – 0.46]		
$\phi_{A\to RT}$	−0.03 [−0.04 – −0.01]	0.01 [0.00 – 0.01]	0.00 [0.00 – 0.01]	0.00 [0.00 – 0.00]
	−0.09 [−0.11 – −0.07]	0.07 [0.05 – 0.09]		
$\phi_{RT\to A}$	*−0.15 [−0.35 – 0.04]*	0.57 [0.25 – 0.89]	0.59 [0.41 – 0.88]	1.84 [1.51 – 2.43]
	−0.03 [−0.06 – 0.01]	0.05 [0.03 – 0.07]		

Notes. The parameters refer to the model of Equation 1. μ_{RT} and μ_A are the overall person-specific means of reaction time and accuracy scores, respectively; $\phi_{RT\to RT}$ and $\phi_{A\to A}$ are the auto-regressive parameters of reaction time and accuracy, respectively, whereas $\phi_{A\to RT}$ and $\phi_{RT\to A}$ are the cross-regressive effects; estimates that are italicized correspond to parameters that are not different from zero.

Table 2. Correlations among the between-person (level-2) random effects of the multivariate VAR(1) model.

	μ_{RT}	μ_A	$\phi_{RT\to RT}$	$\phi_{A\to A}$	$\phi_{A\to RT}$	$\phi_{RT\to A}$
μ_{RT}		.53	−.49	.30	*.03*	*.10*
μ_A	.76		−.37	.57	*−.29*	*−.21*
$\phi_{RT\to RT}$	*−.26*	*−.12*		−.35	*−.18*	*−.16*
$\phi_{A\to A}$	*.16*	.28	−.45		*.13*	.29
$\phi_{A\to RT}$	*.22*	*.13*	−.48	.53		.68
$\phi_{RT\to A}$	*−.19*	−.50	*.25*	*−.16*	*.06*	

Notes. The parameters refer to the model of Equation 1. μ_{RT} and μ_A are the overall person-specific means of reaction time and accuracy scores, respectively; $\phi_{RT\to RT}$ and $\phi_{A\to A}$ are the auto-regressive parameters of reaction time and accuracy, respectively, whereas $\phi_{A\to RT}$ and $\phi_{RT\to A}$ are the cross-regressive effects; lower entries are correlations for the younger adults; upper entries are correlations for the older adults; estimates that are italicized correspond to parameters that are not different from zero.

accuracy and RTs on day *t*. The magnitude of these cross-regressive effects in older adults are much weaker than those of the auto-regressive effect, leading the dynamic system to remain within realistic boundaries.

Finally, with respect to the within-person part of the model, the dynamic errors of RT correlate positively with those of accuracy in both age groups: $r(\varepsilon_{RT}\text{-}\varepsilon_A) = .10$ and .38 in the younger and older adults, respectively (not shown in Tables). Thus, the residuals of daily within-person fluctuations (around the overall level) in RT and in accuracy were nearly independent in younger adults, and only weakly associated in older adults.

Covariance estimates

Table 2 shows the between-person correlations of the level and auto- and cross-regressive effects. When relevant, we compare correlations across the two age groups by computing the Fisher *r*-to -*z* transformations for both younger and older adults and then testing their difference. If applicable, we report the *p*-value corresponding to the null hypothesis of no difference between the correlation of the younger and that of the older adults. Not surprisingly, person mean levels of RT and of accuracy correlate positively. Across the 100 days, individuals with greater accuracy tended to have greater RTs, thus to be slower, whereas faster individuals tended to be less accurate. This effect is stronger in younger $(r(\mu_{RT}\text{-}\mu_A) = .76)$ than in older adults $(r(\mu_{RT}\text{-}\mu_A) = .53;\ p = .0048)$[1]. In both age groups, overall faster individuals are also more consistent in their RT behavior $(r(\mu_{RT}\text{-}\phi_{RT\to RT}) = -.26$ vs $-.49$ in younger and in older adults, respectively; these correlations were not different, $p = .0596$). Similarly, individuals with greater overall accuracy are also more consistent in the accuracy behavior, more so in older $(r(\mu_A\text{-}\phi_{A\to A}) = .57)$ than in younger adults $(r(\mu_A\text{-}\phi_{A\to A}) = .28,\ p = .0123)$. In both younger and older adults, displaying more consistency in one performance aspect means displaying less consistency in the other $(r(\phi_{RT\to RT}\text{-}\phi_{A\to A}) = -0.45$ in younger and -0.35 in older adults, respectively; these correlations were not different, $p = .4036$). For younger adults, we also observe that stronger SAT adjustment motivated by accuracy implies both higher consistency in RT performance $(r(\phi_{A\to RT}\text{-}\phi_{RT\to RT}) = -.48)$ and lower consistency in accuracy performance $(r(\phi_{A\to RT}\text{-}\phi_{A\to A}) = .53)$, whereas in older adults this adjustment implies greater overall accuracy levels $(r(\phi_{A\to RT}\text{-}\mu_A) = -.29)$. Finally, in younger adults, higher SAT adjustment motivated by RT is associated to higher overall accuracy levels $(r(\phi_{RT\to A}\text{-}\mu_A) = -.50)$, whereas for older adults the RT-driven SAT adjustment implies both less consistency in accuracy $((r(\phi_{RT\to A}\text{-}\phi_{A\to A}) = .29)$ and more accuracy-driven SAT adjustment $(r(\phi_{RT\to A}\text{-}\phi_{A\to RT}) = .68)$.

Effect size estimates

The average within-person amount of explained variance in fluctuations around the overall mean values were 44% and 41% for daily mean RT and 16% and 32% for daily accuracy in younger and older adults, respectively. That is, the model explained from 1/6 to nearly 1/2 of the variance in daily fluctuations across 100 days in RT and accuracy scores.

To evaluate the importance of the cross-regressive effects it terms of effect size, we estimated the same model as above, but without the cross-regressive effects $\phi_{RT\to A}$ and $\phi_{A\to RT}$. This second model is statistically nested within the previous model; its parameters form a subset of the parameters of the first model. In frequentist estimation, this sort of model comparison can be evaluated in terms of the difference in statistical fit taking into account the difference in parsimony., With Bayesian estimation, it is customary to compare models' deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002). However, the DIC value provided in the context of multivariate VAR models in Mplus version 8 is not suitable for model comparisons. It is highly dependent on the number of latent variables that are treated as parameters, and the two cross-regressive parameters fall in this category (see Asparouhov et al., 2018, for more details). Therefore, we limit ourselves to comparing the average amount of variance in within-person fluctuations of the two models, with and without the cross-regressive parameters. For RT, the R^2 drops from 0.44 to 0.40 in the younger, and from 0.41 to 0.35 in the older adults. For accuracy, the R^2 remains constant at 0.16 for the younger, whereas it drops from 0.32 to 0.28 in the older adults. However, within both age groups, the 95% credible intervals about these effect size measures in the two models

[1] The mean over the 100 days of the daily mean RT of each participant and the mean accuracy score across the 100 days of the daily accuracy scores correlated $r = .74$ in the younger and $r = .51$ in the older adults. We interpret these between subject correlations as being dominated by individual differences in SAT, rather than by individual differences in ability (according to which we would expect that higher levels of ability would correlate positively with higher accuracy and negatively with RTs).

overlap widely, so that possible changes in R^2 must be interpreted with caution.

Discussion

Both younger and older adults appeared to display fairly strong consistency in their day-to-day performance on the two-choice comparison task. The autoregressive parameter for RT was positive, significant, and of substantial magnitude in both age groups. How quickly COGITO participants responded during the figural task on a given day was partially determined by their speed of response on the previous day. For accuracy, the autoregressive parameter, although positive and significant, was weaker in magnitude. It thus appears that temporal (day-to-day) dependency for RT is stronger than for accuracy.

Additionally, younger adults appeared to adjust their SAT from one day to the next, such that their accuracy level on a given day negatively influenced their RT on the next day. It is possible that the daily feedback that participants received about how accurately and quickly they performed the task contributed to this outcome. Younger participants who were told on a given day that they were more accurate, but also slower than usual, would then be motivated to perform more quickly on the following day. This effect is represented by the significant negative cross-regressive effect from accuracy on day *t-1* to RT on day *t*. The opposite effect, from RT on day *t-1* to accuracy on day *t*, was also negative in sign in younger adults, but its credible interval included zero. Nevertheless, the large random-effects of this parameter could indicate that at least some younger adults engaged also in this type of SAT adjustment.

In contrast, older adults displayed positive and significant cross-regressive effects, meaning that high accuracy levels on a given day were followed by high RTs on the next day, and vice versa. This pattern might imply that older adults ignored the daily feedback or simply focused on accuracy, sacrificing faster RT. It thus appears that older adults did not adjust their SAT: A high accuracy score on day *t-1* predicted a high RT on day *t*, meaning that this age group maintained a consistent response pattern.

Typically, in two-choice comparison tasks, older adults prioritize accuracy over speed, whereas younger adults are more willing to sacrifice accuracy for improvement in speed (e.g., Hertzog, Vernon, & Rypma, 1993; Rabbitt, 1979; Salthouse, 1979). This effect has also been observed when participants are explicitly instructed to emphasize speed over accuracy, and when age differences in speed of evidence accumulation are taken into account (e.g., Starns & Ratcliff, 2010, 2012). When speed is emphasized during instructions, both younger and older adults typically adapt their performance to obtain faster responses. What differentiates the two age groups is that younger adults speed up at the cost of accuracy, by adopting a more liberal response style, whereas older adults are less willing to sacrifice accuracy for speed gains. Such results have been interpreted in terms of greater conservative response style in older compared to younger adults: older adults are not trading accuracy for speed, even when pushed to do so by the task instructions (Hertzog et al., 1993). This tendency of older adults to favor accuracy over speed may magnify the difference in perceptual speed generally observed between them and younger adults.

Age differences in goal setting during a two-choice comparison task may not only reflect conscious preference for a conservative versus liberal response style but may also be dictated by physiological characteristics. Forstmann et al. (2010) found that in a sample of young adults, those with stronger structural connections between the presupplementary motor area and striatum were better able to adjust to cues emphasizing either speed or accuracy in a moving-dots task, thereby showing superior SAT capacities. Given that aging affects brain connectivity in general (Madden et al., 2009), this finding may partially account for age differences in abilities to adjust SAT. This is consistent with Forstmann et al. (2011)), who showed that older adults who adopt slower SAT settings than younger adults also have reduced white matter integrity in corticostriatal tracts connecting the presupplementary motor area to the striatum.

Future directions

An alternative approach to investigate age differences in SAT in the COGITO study calls for the application of the drift-diffusion model. Within this model, SAT is primarily captured by the boundary separation parameter (Starns & Ratcliff, 2010): The wider the separation, the more accurate, but also slower, the response. Application of this model typically requires a two-alternative forced choice in which decisions are based on evidence accumulation for a single stimulus. In contrast, our task contained two stimuli that needed to be checked for identity. Such a task is likely to require shifting attention back and forth between each of the two stimuli, a behavior that cannot be captured by standard versions of the drift-diffusion model.

Limitations

A limitation of this study is that we did not directly assess strategies or conscious adaptations of test taking features that participants might have adopted. Behavioral measures, such as eye tracking, or self-report measures could have enhanced our understanding of such strategies. Instead, our interpretation of the data in terms of SAT adjustment and the associated age-related differences was inferred indirectly from the cross-regressive parameters within the MLVAR(1) model. Self-reported strategies were assessed for some of the other COGITO tasks that were administered daily (e.g., Hertzog, Lövdén, Lindenberger, & Schmiedek, 2017), but not for measures of perceptual speed. The duration, location, and sequence of fixations using eye-tracking may help to discriminate between different strategies to reconcile speed and accuracy, and may serve to validate our interpretation of cross-regressive parameters.

Conclusions

Most of the extant research comparing SAT in younger and older adults is based on experiments lasting only a few minutes and in which participants are observed in their SAT adjustment across different conditions that emphasize either speed or accuracy. Evidence showing how this micro-level phenomenon may generalize to a larger temporal scale remains scarce. The uniquely rich COGITO study and recent development of statistical software for multivariate vector autoregressive modeling allowed us to examine how daily SAT adjustments unfolded over several months—and how this process differed in younger and older adults. In general, younger adults appeared to adjust their SAT on a daily basis during a two-choice comparison task, whereas older adults did not appear to do so. Nevertheless, individual differences within both age groups, as evidenced by the magnitudes of the random effects that represent individual differences in the cross-regressive parameters, leave the door open to investigations of age-untypical behavior, such as conservative decisions in younger adults and liberal choices in older adults.

Article Information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was supported by the Grant M.FE.A.BILD0005 from the Max Planck Society, by the Sofja Kovalevskaja Award (to Martin Lövdén) of the Alexander von Humboldt Foundation donated by the German Federal Ministry for Education and Research, by the Grant KFG163 from the German Research Foundation, and by the Gottfried Wilhelm Leibniz Award of the German Research Foundation to Ulman Lindenberger.

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The authors would like to thank the Associate Editor for the Special Issues Stephen G. West for his precious editorial assistance on prior versions of this manuscript. The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions or the funding agencies is not intended and should not be inferred.

References

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. doi:10.1080/10705511.2017.1406803

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*(1), 583–619. doi:10.1146/annurev.psych.093008.100356

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427–431. doi:10.2307/2286348

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381–391. doi:10.1037/h0055392

Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., … Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences*, *107*(36), 15916–15920. doi:10.1073/pnas.1004932107

Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., & Brown, S. (2011). The speed-accuracy tradeoff in the elderly brain: A structural model-based approach. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(47), 17242–17249. doi:10.1523/JNEUROSCI.0309-11.2011

Garrett, H. E. (1922). A study of the relation of accuracy to speed. *Archives of Psychology*, *56*, 1–104.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition* (3rd ed.). Boca Raton: Chapman and Hall/CRC.

Gueugneau, N., Pozzo, T., Darlot, C., & Papaxanthis, C. (2017). Daily modulation of the speed-accuracy trade-off. *Neuroscience*, *356*, 142–150. doi:10.1016/j.neuroscience.2017.04.043

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. O. (in press). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research* [this issue].

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*. doi:10.3389/fnins.2014.00150

Hertzog, C., Vernon, M. C., & Rypma, B. (1993). Age differences in mental rotation task performance: The influence of speed/accuracy tradeoffs. *Journal of Gerontology*, *48*(3), P150–P156. doi:10.1093/geronj/48.3.P150

Hertzog, C., Lövdén, M., Lindenberger, U., & Schmiedek, F. (2017). Age differences in coupling of intraindividual variability in mnemonic strategies and practice-related associative recall improvements. *Psychology and Aging*, *32*(6), 557–571. doi:10.1037/pag0000177

Madden, D. J., Spaniol, J., Costello, M. C., Bucur, B., White, L. E., Cabeza, R., … Huettel, S. A. (2009). Cerebral white matter integrity mediates adult age differences in cognitive performance. *Journal of Cognitive Neuroscience*, *21*(2), 289–302. doi:10.1162/jocn.2009.21047

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus users's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*(1), 44–64. doi:10.1016/0010-0285(75)90004-3

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*(5), 376–404. doi:10.1167/5.5.1

Pew, R. W. (1969). The speed-accuracy operating characteristic. *Acta Psychologica*, *30*(Supplement C), 16–26. doi:10.1016/0001-6918(69)90035-3

Plummer, M., Best, N., Cowles, M. K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*, 7–11.

Rabbitt, P. (1979). How old and young subjects monitor and control responses for accuracy and speed. *British Journal of Psychology*, *70*(2), 305–311. doi:10.1111/j.2044-8295.1979.tb01687.x

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *89*, 59–108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. doi:10.1162/neco.2008.12-06-420

Salthouse, T. A. (1979). Adult age and the speed-accuracy trade-off. *Ergonomics*, *22*(7), 811–821. doi:10.1080/00140137908924659

Schmiedek, F., Bauer, C., Lövdén, M., Brose, A., & Lindenberger, U. (2010a). Cognitive enrichment in old age: Web-based training programs. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, *23*(2), 59–67. doi:10.1024/1662-9647/a000013

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010b). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience*, *2*. doi:10.3389/fnagi.2010.00027

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*(2), 206–221. doi:10.1037/met0000062

Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science*, *3*(1), 1–5. doi:10.1111/1467-8721.ep10769817

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. doi:10.1111/j.1467-7687.2007.00571.x

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639. doi:10.1111/1467-9868.00353

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*(2), 377–390. doi:10.1037/a0018022

Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin & Review*, *19*(1), 139–145. doi:10.3758/s13423-011-0189-3

Wei, W. W. S. (2012). Time series analysis. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods* (Vol. 2, pp. 678–700). New York, NY: Oxford University Press.

Wang, L., Hamaker, E., & Bergeman, C. S. (2012). Investigating inter-individual differences in short-term intra-individual variability. *Psychological Methods*, *17*(4), 567–581. doi:10.1037/a0029317

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67–85. doi:10.1016/0001-6918(77)90012-9.

Yellott, J. I. (1971). Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology*, *8*(2), 159–199. doi:10.1016/0022-2496(71)90011-3

Appendix A

Supplementary Material for the MBR Website: Mplus (version 8) syntax of a MLVAR(1) model

```
  TITLE: DSEM to test speed-accuracy tradeoff in COGITO data.
  To appear in Ghisletta, P., Joly-Burra, E., Aichele, S., Lindenberger, U., &
Schmiedek, F.
  Age differences in day-to-day speed-accuracy tradeoffs: Results from the
COGITO study.
```

```
    DATA: FILE = cogito.dat; ! specify data file in free format
VARIABLE: NAMES = id day mRT acc; ! name variables in data file
          USEVARIABLES = mRT acc; ! name variables to be analyzed in model
          CLUSTER = id; ! name variable denoting level-2 between-person
          identifier
          LAGGED = mRT(1) acc(1); ! specify the time interval
          TINTERVAL = day(1); ! specify the time variable and its interval
          MISSING = .; ! specify how missing data are coded

ANALYSIS: TYPE = TWOLEVEL RANDOM; ! do not modify
          ESTIMATOR = BAYES; ! do not modify
          PROCESSORS = 2; ! do not modify
          BITERATIONS = (50000); ! specify number of iterations
          BSEED = 41; ! specify seed if you want to compare multiple
          estimations
          THIN = 10; ! specify the thinning parameter, which is combined with
          the number of iterations
          ALGORITHM = GIBBS(RW); ! specify the MCMC algorithm

   MODEL: %WITHIN% ! this part concerns the within-person effects
          pRT | mRT ON mRT&1; ! specify the autoregressive (AR) parameter of
          RT and name it pRT
          pacc | acc ON acc&1; ! specify the AR parameter of accuracy (acc)
          and name it pacc
          pRTacc | mRT ON acc&1; ! specify the cross-regressive (CR) parameter
          from acc to RT and name it pRTacc
          paccRT | acc ON mRT&1; ! specify the CR parameter from RT to acc and
          name it paccRT
          %BETWEEN% ! this part concerns the between-person effects
          mRT WITH acc; ! specify the intraindividual mean (iM) of RT to cor-
          relate with the iM of acc
          pRT WITH pacc-paccRT mRT acc; ! specify pRT to correlate with pacc,
          pRTacc, paccRT, mRT, acc
          pacc WITH pRTacc paccRT mRT acc; ! specify pacc to correlate with
          pRTacc, paccRT, mRT, acc
          pRTacc WITH paccRT mRT acc; ! specify pRTacc to correlate with pac-
          cRT, mRT, acc
          paccRT WITH mRT acc; ! specify paccRT to correlate with mRT, acc

    PLOT: TYPE = PLOT2; ! ask for plots for diagnostic purposes

  OUTPUT:  TECH1 TECH8 STANDARDIZED TECH4 RESIDUAL; ! specify output options
```