Multivariate Behavioral Research

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

Joint Modeling of Ability and Differential Speed Using Responses and Response Times

Jean-Paul Fox & Sukaesi Marianti

To cite this article: Jean-Paul Fox & Sukaesi Marianti (2016): Joint Modeling of Ability and Differential Speed Using Responses and Response Times, Multivariate Behavioral Research

To link to this article: http://dx.doi.org/10.1080/00273171.2016.1171128

Published online: 07 Jun 2016.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=hmbr20

Download by: [Universiteit Twente.] **Date:** 07 June 2016, At: 23:56

MULTIVARIATE BEHAVIORAL RESEARCH
2016, VOL. 0, NO. 0, 1–14
http://dx.doi.org/10.1080/00273171.2016.1171128

Routledge
Taylor & Francis Group

Joint Modeling of Ability and Differential Speed Using Responses and Response Times

Jean-Paul Fox[a] and Sukaesi Marianti[b]

[a]Department of Research Methodology, Measurement and Data Analysis, University of Twente; [b]Department of Psychology, University of Brawijaya

ABSTRACT
With computerized testing, it is possible to record both the responses of test takers to test questions (i.e., items) and the amount of time spent by a test taker in responding to each question. Various models have been proposed that take into account both test-taker ability and working speed, with the many models assuming a constant working speed throughout the test. The constant working speed assumption may be inappropriate for various reasons. For example, a test taker may need to adjust the pace due to time mismanagement, or a test taker who started out working too fast may reduce the working speed to improve accuracy. A model is proposed here that allows for variable working speed. An illustration of the model using the Amsterdam Chess Test data is provided.

KEYWORDS
Joint model; latent growth model; response times; variable speed

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

Introduction

Responses to items in a test of ability reveal information about the accuracy of the responses (i.e., the degree of correctness), which is related to ability. With the introduction of computer-based testing, both responses and response times (RTs), or the amount of time taken to respond to produce an answer, can be collected. RTs will reveal information about the working speed of the respondent. Traditionally, in psychological research, a speed–accuracy trade-off applies, with fast-working test takers often producing a greater number of incorrect responses compared to test takers who work slower. This is referred to as a *within-person relationship* between speed and ability, and in educational research this relationship is assumed (e.g., van der Linden, 2007). When speed and ability are assumed to be constant, however, no such relationship can be studied. The *between-person relationship* between ability and speed has also been studied, building on the information that test takers differ from one another in ability and working speed. Various studies have reported a negative correlation, estimated at the population level, between speed and ability of test takers. Empirical examples of Klein Entink, Fox, and van der Linden (2009) showed that higher-ability test takers tended to work at a slower speed than lower-ability test takers. Klein Entink, Kuhn, Hornke, and Fox (2009), Roberts and Stankov (1999), and van der Linden and Fox (2015) also have reported a negative correlation between ability and speed in their empirical examples.

In the common lognormal RT model of van der Linden (2006), it is assumed that the working speed of a test taker is constant throughout the test. The general item response theory (IRT) models are based on the principle that a test taker will use his or her cognitive knowledge to respond to the test items. Therefore, the relationship between ability and speed is assumed to be constant for each test taker working with a constant speed level.

The assumption of a constant (latent) speed parameter corresponds to the assumption of a constant (latent) ability. However, it is reasonable to assume that test takers will vary their working speed during a test. Changes in time management could be required to finish the test in time, or test takers could decide to work slower to improve their level of accuracy. Working speed can also vary when test takers show aberrant response behavior, such as cheating or guessing (Marianti, Fox, Avetisyan, Veldkamp, & Tijmstra, 2014).

Evidence of variable working speed can also be found in psychological testing, where test takers are asked to do different performance tasks. Manipulating the experimental conditions results in changes in test takers' response behavior. For example, test takers will work faster when the time pressure is increased, but the level of accuracy may or may not change; the speed–accuracy

CONTACT Jean-Paul Fox J.P.Fox@utwente.nl Department of Research Methodology, Measurement and Data Analysis, University of Twente, PO Box 217, 7500AE, Enschede, The Netherlands.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hmbr.
© 2016 Taylor & Francis Group, LLC

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

trade-off is different across different levels of the time-pressure condition. For example, Vandekerckhove, Tuerlinckx, and Lee (2011) defined a hierarchical diffusion model for two-choice RTs, where the parameters of the response process can vary across persons, items, and experimental conditions to model the underlying response process. Assink, van der Lubbe, and Fox (2015) used the hierarchical drift diffusion model to identify tunnel vision (i.e., tendency to focus exclusively on a limited view) due to time pressure. In an experiment, they found an interaction effect between the time-pressure condition and the RT but not the response accuracy.

In educational testing, different joint models for ability and speed assume a constant speed parameter for persons. The hierarchical latent variable modeling of responses and RTs (Fox, Klein Entink, & van der Linden, 2007; Klein Entink, Fox, et al., 2009; van der Linden, 2007; van der Linden & Glas, 2010) and the generalized linear IRT approach (Molenaar, Tuerlinckx, & van der Maas, 2015a) both assume a constant latent working speed parameter for each individual. The constant speed parameter is also assumed in the IRT approach of categorical RTs (e.g., De Boeck & Partchev, 2012; Partchev & De Boeck, 2012; Ranger & Kuhn, 2012) and the nonlinear regressions of IRT parameters on RTs (Ferrando & Lorenzo-Seva, 2007a, 2007b).

To model nonconstant working speed, a latent growth modeling approach is defined for the speed parameter. For each test taker, the within-person systematic differences in observed RTs conditional on the time intensities (i.e., the population average time needed to complete each item) are modeled using latent variable modeling. An individual speed process is assumed, describing the changes in speed across items. Thus, test takers can work with different levels of speed during the test. Each individual speed process will be defined using random effects to model correlations between the RTs for each test taker. A linear (within-person) relationship is defined between individual RTs and random effects.

Furthermore, random effects are also used to define differences in speed process between test takers. This will generalize the common lognormal speed model, where a random intercept is used to define differences in speed across test takers. The latent growth speed process will be a second level of the lognormal speed model.

In latent growth curve analysis, a time scale is needed to model the speed process and to define the individual variation in initial status and growth rate. In the present approach, the order in which the items are solved will define the underlying time scale describing the sequence of observed item RTs. Each item functions as a measurement occasion for speed, and each pattern of RTs is treated as longitudinal RT data with respect to the speed process. The measurement occasions (defining the time scale of the speed process) are defined on a scale from 0 to 1. The chosen time scale values are arbitrary and only represent the order in which the items are solved and reflect that the observations are made at equidistant time points. The time variable will be defined on this scale, where the first (last) measurement corresponds to the response to the first (last) item.

It will be shown that the latent growth model for working speed can be integrated with an IRT model for ability. Under the variable speed model, the ability parameter is influenced by the speed process parameters. This generalizes the univariate relationship between ability and a single speed variable within a test since multiple speed components are involved in this relationship. In this approach, ability will be influenced by a weighted average of the person-specific speed process parameters.

The Markov chain Monte Carlo (MCMC) is used for parameter estimation, which enables joint estimation of all model parameters. The developed MCMC method is built on the estimation methods of Klein Entink, Fox, et al. (2009) and Fox (2010), who developed MCMC schemes for joint models for responses and RTs assuming a constant working speed model.

Simulated and real data examples will be given to illustrate the modeling framework. Data from the Amsterdam Chess Test (ACT; van der Maas & Wagenmakers, 2005) are used to model variable working speed using a linear and a quadratic speed component. A direct comparison is made with the hierarchical model of van der Linden (2007) and Klein Entink, Fox, et al. (2009).

The variable working speed model

Van der Linden (2006, 2007) proposed a lognormal model for RTs using two parameters to describe item and individual variations in RTs. An item factor is defined that represents the time intensity of an item, and each time-intensity parameter represents the population-average time needed to complete the item. A person parameter is defined that represents the constant working speed as the systematic differences in RTs given the time intensities. For example, a test taker works slower (faster) than the average level in the population when the differences between RTs and time intensities are all positive (negative) since, over items, more (less) time is needed than the population-average time.

Let T_{ik} denote the response time of person i $(i = 1, \ldots, N)$ on item k $(k = 1, \ldots, K)$. A lognormal response time distribution is assumed, to account for the positively skewed characteristic of RT distributions,

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

which leads to

$$\ln T_{ik} = \lambda_k - \zeta_i + \varepsilon_{ik},\ \varepsilon_{ik} \sim N\left(0, \sigma^2_{\varepsilon_k}\right). \quad (1)$$

The time-intensity parameter is represented by λ_k and the common speed parameter by ζ_i. The time-intensity parameter represents the population-average time needed to complete the item. The test takers are assumed to be randomly selected from a population. Therefore, the speed parameter is assumed to follow a normal population distribution

$$\zeta_i \sim N\left(\mu_\zeta, \sigma^2_\zeta\right). \quad (2)$$

In Fox et al. (2007) and Klein Entink, Fox, et al. (2009), a time-discrimination parameter, ϕ_k, has been included as a slope parameter for speed. The time-discrimination parameter characterizes the sensitivity of the item for different speed levels of the test takers. This leads to the following specification of the lognormal speed model:

$$\ln T_{ik} = \lambda_k - \phi_k\zeta_i + \varepsilon_{ik},\ \varepsilon_{ik} \sim N\left(0, \sigma^2_{\varepsilon_k}\right). \quad (3)$$

From Equation (3), it follows that the time-discrimination parameter is also used to model the unexplained heterogeneity between time-pattern responses. This follows from the fact that the covariance between the RT to item k and l of person i includes the time discriminations, which are given by

$$\begin{aligned} \operatorname{cov}(T_{ik}, T_{il}) &= \operatorname{cov}(\lambda_k - \phi_k\zeta_i + \varepsilon_{ik}, \lambda_l - \phi_l\zeta_i + \varepsilon_{il}) \\ &= \operatorname{cov}(\phi_k\zeta_i + \varepsilon_{ik}, \phi_l\zeta_i + \varepsilon_{il}) \\ &= \operatorname{cov}(\phi_k\zeta_i, \phi_l\zeta_i) \\ &= \phi_k \operatorname{var}(\zeta_i)\,\phi_l = \phi_k\sigma^2_\zeta\phi_l. \end{aligned} \quad (4)$$

Van der Linden (2015) defined the time-discrimination parameter to be a measurement error variance parameter such that $\sigma^2_\varepsilon = 1/\phi_k$. In that case, the time discrimination (or error variance parameter) will not influence the covariance between RTs.

In latent growth modeling, a time scale is required for the observed responses. The times that the observations were made are represented on this scale. From this perspective, the items in a test can be viewed as measurement occasions for speed and ability. Therefore, a natural time scale can be defined from the collected RTs since each RT also defines the time between two subsequent response observations. Subsequently, the RTs of each test taker, in the order in which the items were solved, also define the time between his or her measurement occasions. This time scale would be inappropriate if the test taker were permitted to take a break after finishing an item and before moving on to the next question.

When estimating all model parameters simultaneously, generating a unique timeline for each test taker increases the computational burden. Furthermore, the latent growth model is applied to a test situation that might take a few hours, such that the nonequidistant property of the time scale can hardly influence the results. Therefore, an equidistant timeline is defined, which leads to a common time scale for all test takers.

Let a time variable, $\mathbf{X}_i = X_{i1}, X_{i2}, \ldots, X_{iK}$, where $X_{i1} = 0$, represent the measurement occasions of test taker i. This time variable represents only the order of the items, which is used to model the speed process over time on an equidistant scale. The measurement of speed from the first item observation is defined as the intercept, and subsequent item observations can be used to model change in speed. Let $\mathbf{X}_{(i)} = X_{(i1)}, X_{(i2)}, \ldots, X_{(iK)}$ denote the order in which the K items are made by person i. Then, a convenient time scale is defined by $X_{ik} = (X_{(ik)} - 1)/K$. The times are defined on a scale from 0 to 1, where 1 is the upper bound representing an infinite number of items.

Note that the scale on which the latent variable working speed is measured is arbitrary. Therefore, the numerical values of the time scale for the speed process only need to address the order in which the items were solved and the assumed equidistant property of the measurements.

The lognormal random linear variable speed model

To introduce the latent growth model for speed, the lognormal RT model is extended with a linear growth term. This model will not be of particular interest in itself since it is not realistic to assume that test takers will accelerate their speed of working in a linear way. However, the linear trend component can be used in combination with higher-order time components to model more complex processes of working speed.

The lognormal RT model with a linear trend for speed can be defined using the time variable X; it follows that

$$\begin{aligned} &\ln T_{ik} = \lambda_k - \phi_k(\zeta_{i0} + \zeta_{i1}X_{ik}) + \varepsilon_{ik} \\ &\begin{pmatrix}\zeta_{i0}\\ \zeta_{i1}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma^2_{\zeta_0} & \rho_{\zeta_{01}}\\ \rho_{\zeta_{01}} & \sigma^2_{\zeta_1}\end{pmatrix}\right). \end{aligned} \quad (5)$$

The parameter ζ_{i0} represents the value of speed measured with the first item solved, also referred to as the *initial value of speed*. The parameter ζ_{i1} represents the random slope in speed, which means that test takers can differ in their growth rate of speed. Note that both random effects have a population mean of zero. This means that the average of time intensities defines the average time to complete the test. Furthermore, the population-average speed trajectory is constant, and shows no changes in speed, since the means of the random effects are zero. Thus, the population-average trajectory with zero values for the random speed variables represents a constant population-average level of speed throughout the test. Test takers can

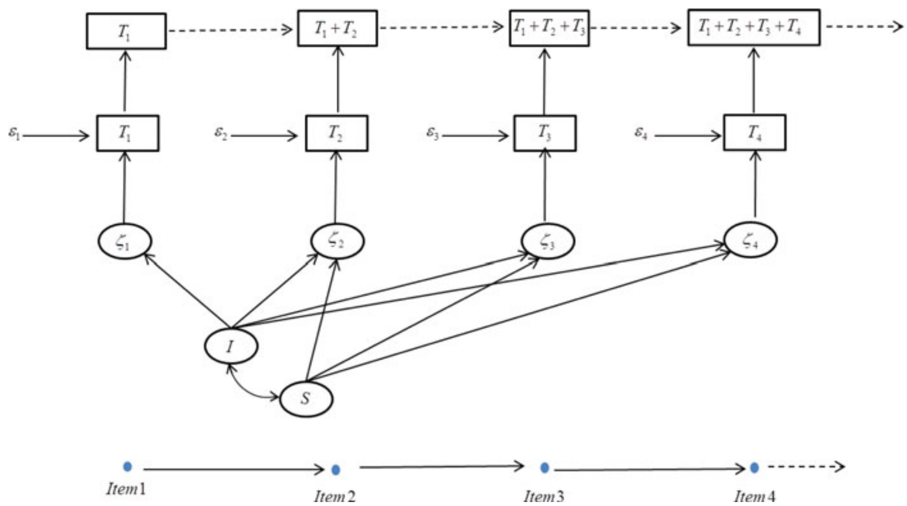

Figure 1. The lognormal random linear variable speed model. I = the random intercept; S = the growth rate; T = response time.

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

work faster than this average level, which corresponds to a positive initial speed value. Furthermore, test takers can show an increasing or decreasing trend in their speed rate, which is represented by a positive or negative growth rate, respectively.

Figure 1 represents this (random linear) variable speed model. The bottom scale represents the order in which the items are solved. The upper axis represents the real-time scale. For each item, an RT is observed to measure speed, and for each RT observation, an error term represents the measurement error that is involved in measuring speed. The latent speed measurements are modeled using a random intercept, referred to as *I*, and a random growth rate, referred to as *S*. The average level of speed *I* is measured by all item observations, where the growth rate is measured by all item observations excluding the first item. The variances of the growth model variables, *I* and *S*, define the between-person variability in initial speed value and growth rates. A covariance term is specified between the growth model variables. Test takers who worked too slowly at the start of the test might improve their speed to finish the test in time. Test takers who started working very fast might later decrease their speed (possibly improving their accuracy level) since by working fast initially they would have sufficient time to finish the test. This corresponds to a negative correlation between the growth model parameters.

The lognormal random quadratic variable speed model

As stated, to define a more complex speed process, the linear trend component is extended with a quadratic term. The linear trend can be used to model the speed processes of a test taker who starts to work faster and continues to work fast until the end of the test. However, a quadratic term can be used to decelerate or accelerate this linear trend. For example, a positive linear trend for speed can be decelerated by a negative quadratic term.

A random quadratic time component is included to define person-specific growth parameters. Then, each trajectory of working speed is modeled by an intercept, a linear trend, and a quadratic time component using individual parameters. The lognormal model with a random quadratic time variable is represented by

$$\ln T_{ik} = \lambda_k - \phi_k\left(\zeta_{i0} + \zeta_{i1}X_{ik} + \zeta_{i2}X_{ik}^2\right) + \varepsilon_{ik}$$
$$\begin{pmatrix}\zeta_{i0}\\ \zeta_{i1}\\ \zeta_{i2}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}\sigma_{\zeta_0}^2 & \rho_{\zeta_{01}} & \rho_{\zeta_{02}}\\ \rho_{\zeta_{01}} & \sigma_{\zeta_1}^2 & \rho_{\zeta_{12}}\\ \rho_{\zeta_{02}} & \rho_{\zeta_{12}} & \sigma_{\zeta_2}^2\end{pmatrix}\right). \quad (6)$$

In Figure 2, a graphical representation of the model is given. The bottom scale represents the order of responses to the items. The upper scale is the true time scale. The random intercept refers to the initial or average speed level; the linear trend is given by *S*; and a quadratic time component is given by *Q*. The random growth components are assumed to be correlated with common covariances across persons, according to the covariance matrix in Equation (6). In this model, the individual speed trajectories are modeled using three random effects, each with a mean of zero, such that the average time intensities define the average time to complete the test.

Joint model for responses and response times

Besides observing RTs, let Y_{ik} denote the response of person $i (i = 1, \ldots, N)$ to item k $(k = 1, \ldots, K)$. An IRT model is considered to model the item responses and to measure ability of each test taker. When considering binary response data, a two-parameter normal ogive model with item discrimination parameter a_k and

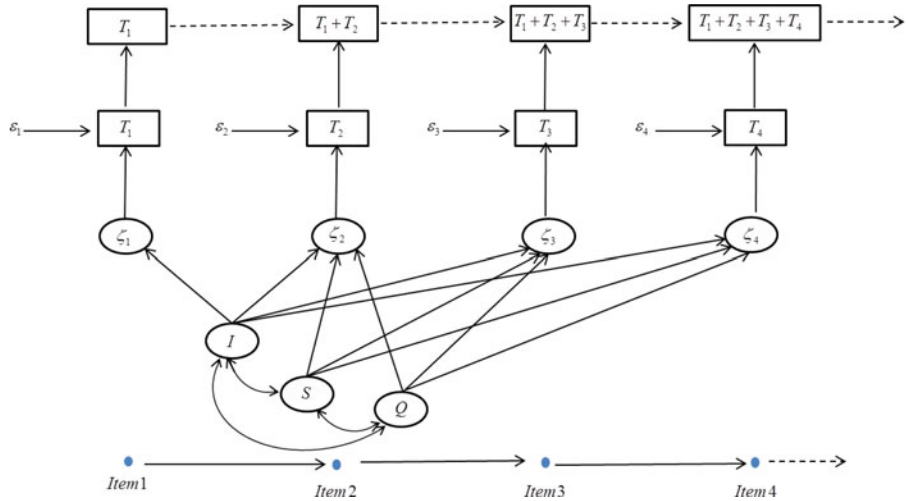

Figure 2. The lognormal random quadratic variable speed model. I = the random intercept; S = the growth rate; Q = the quadratic time component; T = the response time.

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

difficulty parameter b_k. Using the underlying latent response formulation, a latent response Z_{ik} is used, which is normally distributed with mean $a_k\theta_i - b_k$ and variance 1, and truncated from below (above) by zero when the response is correct (incorrect). Thus, a correct response, $Y_{ik} = 1$, is the indicator that Z_{ik} is positive. The joint model for responses and RTs, allowing for variable speed, is given by

$$Z_{ik} = a_k\theta_i - b_k + \omega_{ik}, \omega_{ik} \sim N(0, 1)$$
$$\ln T_{ik} = \lambda_k - \phi_k\left(\zeta_{i0} + \zeta_{i1}X_{ik} + \zeta_{i2}X_{ik}^2\right) + \varepsilon_{ik},$$
$$\varepsilon_{ik} \sim N\left(0, \sigma^2_{\varepsilon_k}\right)$$
$$\begin{pmatrix}\theta_i\\ \zeta_{i0}\\ \zeta_{i1}\\ \zeta_{i2}\end{pmatrix} \sim N\left(\begin{pmatrix}\mu_\theta\\ 0\\ 0\\ 0\end{pmatrix}, \begin{pmatrix}\sigma^2_\theta & \rho_{\theta\zeta_0} & \rho_{\theta\zeta_1} & \rho_{\theta\zeta_2}\\ \rho_{\theta\zeta_0} & \sigma^2_{\zeta_0} & \rho_{\zeta_{01}} & \rho_{\zeta_{02}}\\ \rho_{\theta\zeta_1} & \rho_{\zeta_{01}} & \sigma^2_{\zeta_1} & \rho_{\zeta_{12}}\\ \rho_{\theta\zeta_2} & \rho_{\zeta_{02}} & \rho_{\zeta_{12}} & \sigma^2_{\zeta_2}\end{pmatrix}\right). \tag{7}$$

The prior distribution of the person parameters $(\theta_i, \boldsymbol{\zeta}_i) = (\theta_i, \zeta_{0i}, \zeta_{1i}, \zeta_{2i})$ can be given as

$$\begin{pmatrix}\theta_i\\ \boldsymbol{\zeta}_i\end{pmatrix} \sim N\left(\begin{pmatrix}\mu_\theta\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\sigma^2_\theta & \boldsymbol{\Sigma}_{\theta\zeta}\\ \boldsymbol{\Sigma}_{\zeta\theta} & \boldsymbol{\Sigma}_\zeta\end{pmatrix}\right). \tag{8}$$

The relationship between speed and ability is defined by the covariance between ability and the speed components and is given by $\boldsymbol{\Sigma}_{\theta\zeta}$. The ability parameter is influenced by the different speed components. This follows directly from the conditional distribution of ability given the speed variables. This distribution is given by

$$\theta_i \left| \boldsymbol{\zeta}_i \sim N\left(\mu_\theta + \boldsymbol{\Sigma}_{\theta\zeta}\boldsymbol{\Sigma}_\zeta^{-1}\left(\boldsymbol{\zeta}_i - \mu_\zeta\right), \sigma^2_\theta - \boldsymbol{\Sigma}_{\theta\zeta}\boldsymbol{\Sigma}_\zeta^{-1}\boldsymbol{\Sigma}_{\zeta\theta}\right)\right. \tag{9}$$

Ability is influenced by the weighted average of the speed components, where the weights are defined by the covariance matrix $\boldsymbol{\Sigma}_{\theta\zeta}$ times the inverse of the variance of speed components. When test takers do not vary their speed, only the first diagonal component of $\boldsymbol{\Sigma}_\zeta$ will be larger than zero, showing the variability in constant speed values across test takers. The linear trend and quadratic change in speed will be around zero, which leads to a negligible influence of the remaining variable speed components on ability. When test takers vary their speed according to the quadratic variable speed model, the diagonal components of $\boldsymbol{\Sigma}_\zeta$ will be larger than zero and, together with the covariance matrix $\boldsymbol{\Sigma}_{\theta\zeta}$, will define the relation with ability. It follows that the constant speed model is generalized by allowing variable speed components to influence ability.

It will depend on the application whether changes in working speed will improve the accuracy of the responses. By measuring changes in working speed and modeling the relationship between speed and ability, it is possible to estimate the speed trajectories of test takers with different levels of ability. High-ability test takers may have different speed trajectories than low-ability test takers. The speed trajectories of test takers may also differ across tests. It will be possible to investigate the effects of time limits on test takers' speed changes, including those of proficient test takers. However, the benefits of estimating speed trajectories in relation to ability will depend on the application.

Identification

The latent scale of ability and speed needs to be identified. The mean and variance of the ability scale can be restricted to identify the scale. This can be accomplished by restricting the sum of item difficulties and product of discriminations or by directly restricting the mean and variance of the ability parameter.

Next, for the variable speed model, the scale of the latent speed variable needs to be identified. This can also be accomplished by two restrictions. In the present description of the model, the mean of the speed parameter is set to zero to identify the mean of the speed scale. The average of the time-intensity parameters represents the population-average time needed to finish the test given an average working speed of zero. The variance of the speed scale is identified by fixing this variance directly or by restricting the product of time discriminations to one. For the joint model, the mean of each person parameter is restricted to zero, and the product of discriminations and time discriminations is restricted to one. These identification restrictions are also used by Klein Entink, Fox, et al. (2009) and Fox (2010).

In the variable working speed model, an additional restriction is required since the covariance between speed components is modeled by the time-intensity parameters and by the covariance matrix of the speed components, $\boldsymbol{\Sigma}_\zeta$. As mentioned previously, the time-intensity parameters will influence the correlation between the RTs, which leads to an indeterminacy between the covariance parameters of speed and the time-intensity parameters. Therefore, as an additional constraint, the covariance matrix of the speed components is restricted to have zero nondiagonal terms, and the covariance between speed components is modeled by the time-intensity parameters. When the time-intensity parameters are all fixed to one, the covariance matrix of the speed components is a free matrix, and no additional restriction is required. The residual errors are assumed to be independently distributed and do not influence the covariance modeling structure. When the ability and speed scale are identified, all higher-level model parameters will also be identified.

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

Parameter estimation

The model parameters can be estimated using MCMC. The MCMC algorithms for the joint model with variable speed will follow the algorithms for the constant speed-ability joint models. In Fox (2010), the MCMC steps are fully explained for the so-called (constant speed) RTIRT model. The MCMC method was implemented in a modified version of the *cirt* R-program of Fox et al. (2007). The following sampling steps are required. At iteration $m = 1, \ldots, M$,

1. For $k = 1, \ldots, K$, sample item parameters from $p(\phi_k, \lambda_k, a_k, b_k | \mathbf{z}_k, \mathbf{t}_k, \boldsymbol{\theta}, \boldsymbol{\zeta}, \boldsymbol{\mu}_I, \boldsymbol{\Sigma}_I)$, using a multivariate normal prior with mean $\boldsymbol{\mu}_I$ and covariance matrix $\boldsymbol{\Sigma}_I$.
2. For $k = 1, \ldots, K$, sample the residual variance in the lognormal model from $p(\sigma^2_{\varepsilon_k} | \mathbf{t}_k, \boldsymbol{\zeta}, \lambda_k, \phi_k)$.
3. For $I = 1, \ldots, N$, sample the ability parameter from $p(\theta_i | \boldsymbol{\zeta}_i, \mu_\theta, \boldsymbol{\Sigma}_{\theta\zeta}, \mathbf{z}_i, \mathbf{t}_i)$.
4. Sample the hyperparameters $\boldsymbol{\mu}_I$ and $\boldsymbol{\Sigma}_I$ from $p(\boldsymbol{\mu}_I | \phi, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma}_I)$ and $p(\boldsymbol{\Sigma}_I | \phi, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b})$.
5. For the constant speed model, sample the hyperparameter $\boldsymbol{\Sigma}_{\theta\zeta}$ from $p(\boldsymbol{\Sigma}_{\theta\zeta} | \boldsymbol{\theta}, \boldsymbol{\zeta}, \phi, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b})$.

For the variable speed model, several additional sampling steps are required. With an identification restriction on the covariance matrix of the person parameters, the sampling of the speed components $\boldsymbol{\zeta}$ and the free parameters of the covariance matrix requires a stepwise approach. The speed components are a priori independently and normally distributed. Each diagonal component of the covariance matrix $\boldsymbol{\Sigma}_\zeta$ is inverse-gamma distributed with an inverse gamma prior with parameters g_1 and g_2. The conditional distribution of the variance parameter of speed component $\boldsymbol{\zeta}_j$ $(j = 0,1,2)$ is given by

$$\sigma^2_{j\zeta} \left| \boldsymbol{\mu}_\zeta, \boldsymbol{\zeta}_j, \mathbf{T} \sim IG\left(\frac{N}{2} + g_2, \sum_i \left(\zeta_{ji} - \mu_{\zeta j}\right)^2 / 2 + g_1\right)\right., \tag{10}$$

and the three variance parameters define the diagonal of the covariance matrix $\boldsymbol{\Sigma}_\zeta$. The speed components are conditionally normally distributed, and it follows that

$$\boldsymbol{\zeta}_i \left| \theta_i \sim N\left(\boldsymbol{\mu}_\zeta + \boldsymbol{\Sigma}_{\zeta\theta}\sigma_\theta^{-2}\left(\theta_i - \mu_\theta\right), \boldsymbol{\Sigma}_\zeta - \boldsymbol{\Sigma}_{\zeta\theta}\sigma_\theta^{-2}\boldsymbol{\Sigma}_{\theta\zeta}\right)\right.. \tag{11}$$

In this conditional distribution, the covariance $\boldsymbol{\Sigma}_{\zeta\theta}$ in the mean term is considered to be a regression parameter. The conditional distribution of this parameter is normal with variance

$$\begin{aligned}\boldsymbol{\Omega} &= E\left(\boldsymbol{\Sigma}_{\zeta\theta}\right) = \left(\sigma_\theta^2 - \boldsymbol{\Sigma}_{\theta\zeta}\boldsymbol{\Sigma}_\zeta^{-1}\boldsymbol{\Sigma}_{\zeta\theta}\right)^{-1} \\ &\quad \times \left(\left(\boldsymbol{\Sigma}_\zeta^{-1}\left(\boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta\right)\right)^t \boldsymbol{\Sigma}_\zeta^{-1}\left(\boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta\right)\right) + \boldsymbol{\Sigma}_0^{-1}\end{aligned} \tag{12}$$

and mean

$$\begin{aligned}\boldsymbol{\Sigma}_{\zeta\theta} &= \boldsymbol{\Omega}^{-1}\left(\boldsymbol{\Sigma}_\zeta^{-1}\left(\boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta\right)^t\left(\theta_i - \mu_\theta\right)\right. \\ &\quad \left.\times \left(\sigma_\theta^2 - \boldsymbol{\Sigma}_{\theta\zeta}\boldsymbol{\Sigma}_\zeta^{-1}\boldsymbol{\Sigma}_{\zeta\theta}\right)^{-1}\right).\end{aligned} \tag{13}$$

From the conditional distribution of $\boldsymbol{\theta}_i$ given $\boldsymbol{\zeta}_i$, the distribution of the variance parameter can be derived. This variance parameter, $\sigma = \sigma_\theta^2 - \boldsymbol{\Sigma}_{\theta\zeta}\boldsymbol{\Sigma}_\zeta^{-1}\boldsymbol{\Sigma}_{\zeta\theta}$, is inverse-gamma distributed with scale parameter

$$SS = \sum_i \left(\left(\theta_i - \mu_\theta\right) - \boldsymbol{\Sigma}_{\theta\zeta}\boldsymbol{\Sigma}_\zeta^{-1}\left(\boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta\right)\right)^2 / 2 + g_1 \tag{14}$$

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

Table 1. Simulated and estimated parameter values (over 50 data replications) of the joint model with a lognormal random quadratic variable speed component for 1,000 test takers and with $K = 20$ and $K = 40$ items.

			LNIRTQ			
			$K = 20$		$K = 40$	
Variance Components		True *Mean*	*Mean*	*SD*	*Mean*	*SD*
Person covariance matrix						
Ability	σ^2_θ	1	1.031	0.059	1.009	0.052
	$\rho_{\theta\zeta_0}$	0.7	0.704	0.046	0.699	0.044
Speed	$\sigma^2_{\zeta_0}$	1	0.998	0.054	1.001	0.052
	$\rho_{\theta\zeta_1}$	0.2	0.194	0.052	0.187	0.039
	$\sigma^2_{\zeta_1}$	0.5	0.508	0.083	0.486	0.060
	$\rho_{\theta\zeta_2}$	0.1	0.111	0.048	0.107	0.037
	$\sigma^2_{\zeta_2}$	0.5	0.486	0.092	0.526	0.071
Item covariance matrix						
Discrimination	$\sum_{11}$	0.05	0.065	0.033	0.057	0.016
	$\sum_{12}$	0	− 0.005	0.067	− 0.002	0.043
	$\sum_{13}$	0	0	0.019	0.000	0.010
	$\sum_{14}$	0	0.001	0.073	0.002	0.042
Difficulty	$\sum_{22}$	1	1.054	0.365	1.040	0.244
	$\sum_{23}$	0	0	0.063	− 0.005	0.040
	$\sum_{24}$	0	0.034	0.258	− 0.017	0.163
Time discrimination	$\sum_{33}$	0.05	0.068	0.027	0.059	0.014
	$\sum_{34}$	0	0.018	0.069	0.007	0.040
Time intensity	$\sum_{44}$	1	1.122	0.388	0.981	0.230

Note. SD = standard deviation; *Cor.* = correlation; LNIRTQ = lognormal-IRT model with quadratic time component.

and shape parameter $N/2 + g_2$. Subsequently, from the sampled variance parameter, σ, a sampled value of the variance parameter σ^2_θ can be obtained using the sampled value for $\boldsymbol{\Sigma}_\zeta$.

Without the identification restriction on the covariance matrix, $\boldsymbol{\Sigma}_\zeta$, the values of the complete covariance matrix of the person parameters,

$$\boldsymbol{\Sigma}_P = \begin{pmatrix} \sigma^2_\theta & \boldsymbol{\Sigma}_{\theta\zeta} \\ \boldsymbol{\Sigma}_{\zeta\theta} & \boldsymbol{\Sigma}_\zeta \end{pmatrix}, \tag{15}$$

are sampled from an inverse-Wishart distribution with scale matrix

$$\sum_i \begin{pmatrix} \theta_i - \mu_\theta \\ \boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta \end{pmatrix} \begin{pmatrix} \theta_i - \mu_\theta \\ \boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta \end{pmatrix}^t + g_1 \tag{16}$$

and degrees of freedom $N+Q$, where Q is the number of random effects. The mean of the speed components, $\boldsymbol{\mu}_\zeta$, is fixed to zero.

Simulation study

In this simulation study, attention was focused on a variable speed process in the joint modeling of responses and RTs. The RTs were modeled according to a lognormal random quadratic variable speed model. This RT model included a random trend and a random quadratic time variable, which is represented by

$$\ln T_{ik} = \lambda_k - \left(\zeta_{i0} + \zeta_{i1} X_{ik} + \zeta_{i2} X^2_{ik}\right) + \varepsilon_{ik},$$
$$\varepsilon_{ik} \sim N\left(0, \sigma^2_{\varepsilon_k}\right). \tag{17}$$

The random speed components had a mean of zero to identify the time intensities. The responses were modeled according to a two-parameter normal ogive model. The ability and speed parameters were assumed to be multivariate normally distributed, according to Equation (8).

The joint model for responses and RTs was used to generate the data, and a modified version of the *cirt* program of Fox et al. (2007) was used to estimate all model parameters. The item parameters were simulated from a multivariate normal distribution with the covariance matrix given in Table 1. The mean of the time discriminations and discrimination parameters was set to one, and the mean of the difficulty and time-intensity parameters was set to zero. The measurement error variance was set to .50 for each item. Furthermore, the model was identified by restricting the covariance between random speed components to zero and restricting the product of time discrimination and discrimination parameters to one.

To evaluate the performance of the developed MCMC algorithm, a total of 50 data sets were simulated according to the joint model for 1,000 test takers responding to 20 and 40 items. A burn-in period of 1,000 iterations was

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

used, and a total of 5,000 iterations were made to estimate all model parameters.

In Table 1, the true and final parameter estimates are given across the 50 simulated data sets. Under the header "LNIRTQ," the parameter estimates are represented for replicated data sets of 20 and 40 items. The true value of the covariance matrix of the person parameters shows that the speed trajectories differ a lot across test takers. The random variation over test takers in the trend and quadratic components was around .50, given a variance of one across the average levels of speed. Furthermore, there was a positive covariance simulated between the random person components.

For the 20-item test-length condition, the estimated values of the person covariance matrix are close to the true values, despite the high level of variation in simulated speed trajectories. The variability in speed trajectories, which differed in their trend and quadratic components, was accurately estimated. Estimates of the item covariance parameters also showed a good recovery of the true parameter values.

Subsequently, the test length was increased from 20 to 40 items. Thus, twice as many RT observations were generated for each speed trajectory. In Table 1 under the label "$K = 40$," the final estimates of the person and item covariance parameters for the 40-item test are given. It can be concluded that the estimated parameter values are close to the true values. Furthermore, it follows directly that the estimated standard deviations are smaller than those based on the 20-item test data. Although not shown, the item parameters of the 20- and 40-item test and the measurement error variance were also accurately recovered across 50 simulated data sets.

Modeling variable speed in the Amsterdam Chess Test data

The Amsterdam Chess Test (ACT) data of Van der Maas and Wagenmakers (2005) were used to identify the variable speed trajectories of 259 test takers who responded to 40 chess tasks. The chess items were divided over three sections: tactical skill (20 items), positional skill (10 items), and end-game skill (10 items). Each item concerned a chessboard situation, and the problem-solving task was to select the best possible move. The dichotomous response observations, 1 (correct) and 0 (incorrect), as well as the RTs were stored. Fox (2010, p. 253), using the joint model of Klein Entink, Fox, et al. (2009), analyzed the data using the RTIRT model to identify items not fitting the data.

The purpose of the present study was to investigate whether test takers worked with variable speed and what type of speed trajectories could be identified. Furthermore, the complex between-person relationship between ability and random speed components was investigated. In a different approach, Molenaar et al. (2015b) considered a function of ability on speed in their generalized linear model for responses and RTs. They found a common curvilinear effect of ability on speed for the end-game items, where higher-ability test takers tended to use more time in contrast to lower-ability test takers, who started to answer faster at the end of the test. In their approach, higher-order interaction terms between ability and speed were used to obtain more insight into the relation between speed and ability, but the higher-order ability components were assumed to be fixed deterministic components (with a common effect across test takers), and it was assumed that speed did not have an influence on ability.

The RTIRT lognormal (constant) speed model and the lognormal (random quadratic) variable speed model (Equation [7]) were used to analyze the data. The MCMC algorithm was run for 10,000 iterations to estimate all model parameters, where a burn-in period of 1,000 iterations was used. In Figure 3, for the variable speed model, trace plots of four (variance and covariance) person parameters are given to show the fast convergence and the stable behavior of the MCMC chains. The MCMC algorithm converged rapidly without specifying informative starting values. The other trace plots showed similar behavior. The R-coda package (Plummer, Best, Cowles & Vines, 2006) was used to investigate the chains, and the commonly used convergence diagnostics (e.g., Geweke, Heidelberger, and Welch) did not show any issues.

The RTIRT with a constant speed factor was fitted to the data. In Figure 4, the item parameter estimates of the 40-item test are given. It can be seen that there is sufficient variation in difficulty and item intensity to measure the ability and speed factors accurately at the different levels of the scale. The time discriminations are higher for the first 10 items (they define the tactical skill cluster), which means that responses to those items show more variation between slow- and fast-working test takers. The time discriminations for the end-game items were not as high, indicating less power to discriminate between the working speeds of the test takers. Most items discriminated sufficiently between test takers' ability levels; only around five items had a low discriminating value of < 0.5.

The covariance estimates of the test taker's random factors and the item parameters of the joint model with a constant speed factor are given in Table 2. There is substantial between-item variation in difficulty and intensity but less in discrimination and time discrimination.

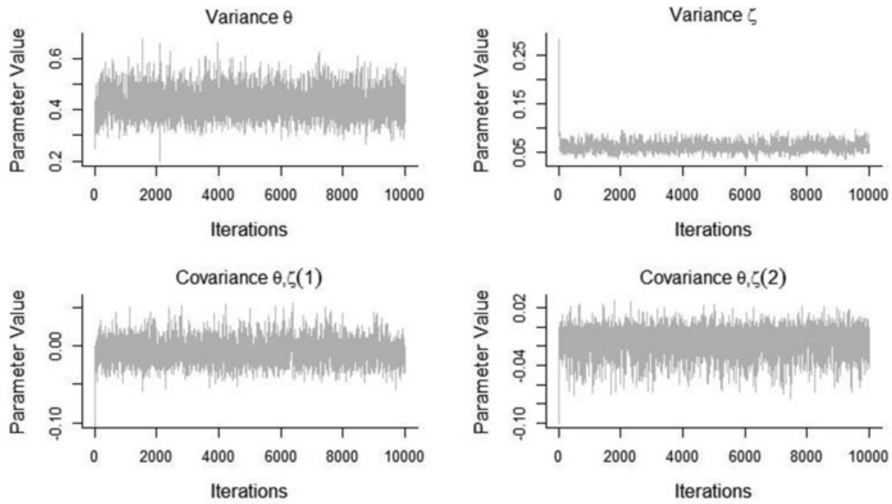

Figure 3. Trace plots of the ability and average speed population variance parameters and the covariance between ability and the slope and quadratic speed components.

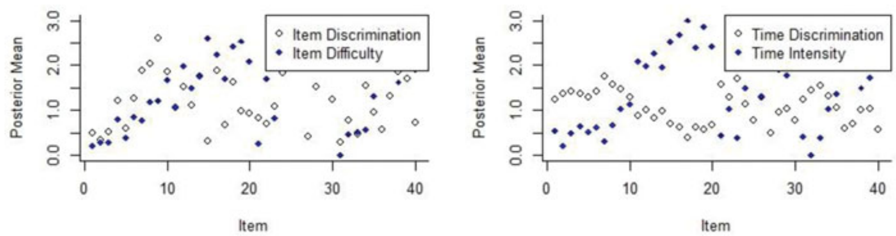

Figure 4. Item parameter estimates of the 40 chess items. The left-hand plot shows the item discrimination (open diamond symbol) and item difficulty (closed diamond symbol) estimates; the right-hand plot shows the time discrimination (open diamond symbol) and time intensity (closed diamond symbol) estimates.

Table 2. Amsterdam chess test: covariance components and correlation estimates.

Variance components		Constant speed			Variable speed		
		Mean	*SD*	*Cor.*	*Mean*	*SD*	*Cor.*
Person covariance matrix							
Ability	σ^2_θ	0.317	0.040		0.425	0.139	
	$\rho_{\theta\zeta_0}$	0.105	0.013	0.640	0.115	0.016	0.720
Speed	$\sigma^2_{\zeta_0}$	0.085	0.008		0.06	0.009	
	$\rho_{\theta\zeta_1}$				− 0.004	0.014	− 0.018
	$\sigma^2_{\zeta_1}$				0.113	0.024	
	$\rho_{\theta\zeta_2}$				− 0.014	0.015	− 0.086
	$\sigma^2_{\zeta_2}$				0.062	0.015	
Item covariance matrix							
Discrimination	$\sum_{11}$	0.479	0.175		0.533	0.188	
	$\sum_{12}$	0.252	0.138	0.352	0.289	0.150	0.371
	$\sum_{13}$	0.030	0.052	0.108	0.019	0.064	0.053
	$\sum_{14}$	0.067	0.151	0.079	0.147	0.07	0.408
Difficulty	$\sum_{22}$	1.073	0.259		1.141	0.271	
	$\sum_{23}$	− 0.268	0.085	− 0.645	− 0.317	0.105	− 0.603
	$\sum_{24}$	1.008	0.272	0.794	0.414	0.111	0.786
Time discrimination	$\sum_{33}$	0.161	0.047		0.242	0.076	
	$\sum_{34}$	− 0.450	0.123	− 0.915	− 0.130	0.047	− 0.536
Time intensity	$\sum_{44}$	1.503	0.375		0.243	0.058	

Note. SD = standard deviation; *Cor.* = correlation.

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

The mean RT residual variance was around .25 and ranged from .15 to .50. The correlation between discrimination and difficulty was around .35, and between time discrimination and intensity was around -.92. This strong negative correlation of -.92 showed that, for high time-intensive items, the speed factor did not explain much variation in RTs, whereas the speed factor did explain it for low time-intensive items. According to the model, for a time-intensive item an increase in working speed does not have much effect on RT due to the low time-discrimination parameter. The influence on the RT due to a change in speed is much higher for low time-intensive items, which have high time discriminations.

The strong correlation of around .80 between item difficulty and item intensity was also apparent. The difficult items were clearly taking much more time to solve than the easy items.

For the person parameters, there was not much variation in speed levels (around .085) or in ability levels (around .32). Under the constant working speed assumption, the correlation between ability and speed was around .65, which showed that high-ability test takers were also completing the items faster. They were able to identify the solution to the chess problem faster than the low-ability test takers.

This covariance structure holds under the assumption that test takers were working with a constant speed. To investigate variable speed trajectories of test takers, the joint model with the random quadratic variable speed model was also fitted. In Table 2, the covariance estimates are given under the header "Variable speed." For the covariance between item parameters, it can be seen that the strong correlation between time discrimination and time intensity diminished to −.54. Apparently, the additional speed components explained the greater variation in RTs, reducing the strong correlation between time discrimination and intensity. The correlation between the item discrimination and time intensity increased to .41. The chess items that were highly discriminating in ability were also the time-intensive items. This relates to the positive correlation between ability and speed. It is likely that the test takers showed different speed behavior in responding to well-discriminating items depending on their ability.

The correlation between the average speed level and ability was around .72, which was slightly higher than it was under the constant speed model. The corresponding 95% highest posterior density (HPD) interval was [0.637, 0.785]. A slightly negative correlation of −.02 was estimated between ability and the random slope speed component (95% HPD interval equaled [−0.157, 0.099]). This means that high-ability test takers were more likely to decrease their speed in a linear way. The correlation between ability and the quadratic speed component was around −.09 (95% HPD interval equaled [−0.266,0.035]), which means that high-ability test takers were more likely to show an acceleration in the negative trend in speed. However, both estimated correlation parameters were not significantly different from zero, since zero was included in the 95% HPD intervals.

From the trace plots of the covariance parameters, see Figure 3, it can also be seen that the drawn covariance values are not significantly different from zero. Since the correlations with ability were not significantly different from zero, the characteristics of the speed trajectories could not be explained by differences in ability.

In Figure 5, the estimated ability estimates are plotted against the random components of speed. It can be seen that there is a strong positive relation between ability and average speed, where the relation between ability and the slope and quadratic speed components is not significantly different from zero. The estimated average speed component was conditionally estimated on the two other random speed components, which accounted for nonconstant-speed behavior. The positive correlation between the linear and quadratic speed component showed that a more negative (positive) trend in speed was accelerated, leading to an even slower (higher) working speed.

In Figure 6, from the total sample of N = 259 test takers, the fitted item-specific working-speed measurements of 20 high- and low-ability test takers were plotted. The test takers started working at different speeds: The high-ability group started to work faster than the low-ability group. Some high-ability test takers increased their working speed toward the end of the test, but others decreased their level of working speed around halfway through the test. The low-ability test takers showed the opposite behavior. Most of the low-ability test takers increased their working speed halfway through the test; only a few showed a constant decrease in working speed. It is possible that high-ability test takers (the 10% highest scoring test takers) were more focused and eager to ensure that all items were correct, whereas low-ability test takers (the 10% lowest-scoring test takers) might have been less motivated halfway through the test, proceeding more quickly through the latter half of the test. Molenaar et al. (2015b) reported on this pattern, based on a higher-order interaction effect between ability and speed. With the quadratic variable speed model, each test taker's trajectory of working speed was estimated, showing the patterns while controlling for the correlation with ability. This made it possible to estimate the variable working speed behavior of each test taker.

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

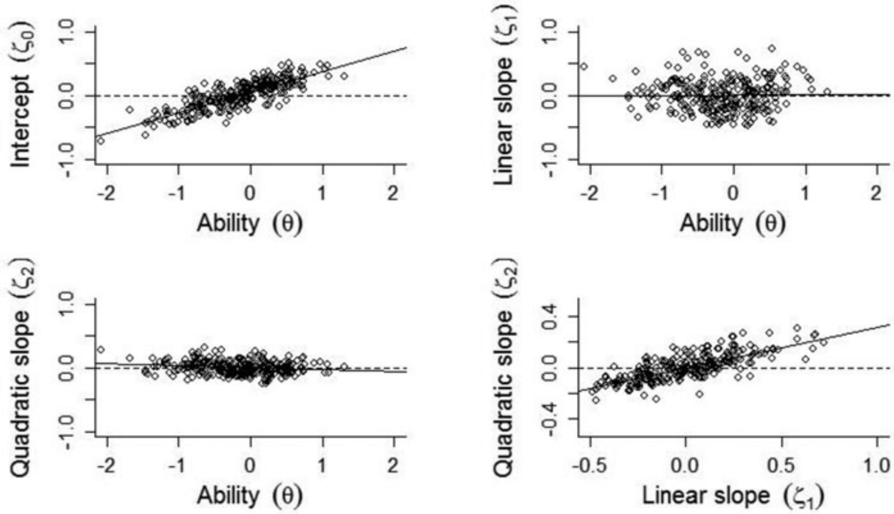

Figure 5. Random person parameter estimates; the average speed (ζ_0), slope speed (ζ_1), and quadratic speed (ζ_2) component plotted against ability (θ). The slope of speed is plotted against the quadratic speed component.

Discussion

Computer-based testing makes it possible to collect RT information as well as response information by simply recording the total time spent on each item and the response to each item, respectively. RTs can be used to make more accurate inferences about test takers' ability (e.g., van der Linden, Klein Entink, & Fox, 2010) and item characteristics. RTs can also reveal new information about test characteristics, test takers' response behavior, and test takers' ability that would not be identified when using response information alone.

The latent growth model for working speed can be used to measure variable working speed according to a time scale defined by the order in which the items were answered. In the present model, the random slope of speed and random quadratic speed components were added to model deviations from a constant speed model. The extension to higher-order random effect components can be made. However, this will require a sufficient number of item observations to estimate all model parameters. The higher-order terms can also be included for groups of test takers using discrete latent random effects.

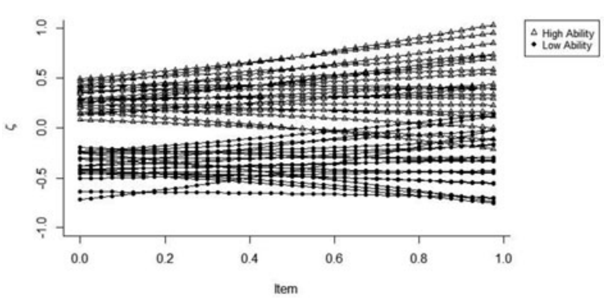

Figure 6. Fitted latent speed trajectories over items of high- and low-ability test takers.

When the trajectory of speed includes higher-order components, the relation between working speed and ability can be defined as the weighted correlation between ability and all the speed components. In that case, the additional random speed components are used to control for non-constant speed to improve the estimation of the relation between speed and ability.

Note that the order in which the items are answered does not influence the ability estimate. The estimation of the latent speed trajectories depends on both the RT information and the order in which the items are answered. In the estimation of each speed trajectory, the information both for each item-specific RT observation and for the relationships between RTs is used. Thus, if the order of observed RTs were changed, a different trajectory would be estimated since the relationships between RTs would be different.

This model, which is a generalization of the constant speed model proposed by van der Linden (2007), can be used to measure a more complex relationship between ability and speed. From a model-building perspective, it is recommended that one first evaluate the fit of the hierarchical model for responses and RTs before fitting a differential speed component. More research is needed to develop information criteria (e.g., Bayesian information criterion [BIC], deviance information criterion [DIC]) that are able to identify the best joint model for responses and RTs among a set of competing models. For example, a straightforward implementation of the DIC is not going to produce reliable results since the estimation of the number of effective parameters, which is required to compute the DIC, is very complex when the model contains many random effects, multiple outcomes of different types (categorical and continuous), and multiple

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

link functions (linear and nonlinear). Future research will focus on the procedure of Klein Entink, Fox, et al. (2009) and Fox (2010, p. 241–242), who considered a DIC based on the integrated likelihood (e.g., see Berger, Brunero, & Wolpert, 1999), where ability and speed were integrated out. This simplifies the computation of the penalty term, since it is no longer based on the random person parameters, and leads to a more accurate estimate of the number of model parameters.

The so-called cross relation between speed and accuracy was also modeled by Molenaar et al. (2015b), who considered different functions of higher-order ability components on speed. They introduced two person factors: ability and speed. In the proposed model, several random person variables were introduced to better describe this relationship by assuming a variable speed model. The MCMC algorithm developed in this article (see section titled "Parameter estimation") can handle numerous random effects since it is a simulation-based estimation procedure.

The situation where test takers are limited in their responding due to time constraints is referred to as *test speededness*. However, when speed is not of interest, it should also not interfere with the measurement of ability. Speededness is considered to be a threat to the validity of the test scores; it inadvertently interferes with the performance level of the test takers. Research has focused on detecting test speededness as a threat to test validity. Chang, Tsai, and Hsu (2014) and Goegebeur, De Boeck, Wollack, and Cohen (2008) considered a mixture modeling approach and defined a speeded class to identify test takers whose performance is affected by the time limit. The general idea is that the time limit influences the probability of an item being answered, without considering RT information. Shao and Cheng (2015) considered a change-point model to identify speeded test takers and considered removing the speeded responses to improve ability estimation. Given the RT and response information, the joint model for ability and speed, using a latent growth model for speed, can provide insight into test speededness. Test takers' fitted speed trajectories can be used to identify (strategic) speed behavior while accounting for differences in ability. An increase in speed at the end of the test would indicate the influence of a time limit on the test performance, whereas a decrease in speed would show the opposite.

Test speededness has a negative influence on test-taker performance. Test fraud, in contrast, is intended to have an opposite effect on performance. Test takers may try to positively distort their responses to improve their achievement scores, which overestimates the test takers' true achievement level. Nowadays, there is an increased interest in test fraud detection (e.g., Kingston & Clark, 2014), where attention is focused on test-taking effort. Test takers may show solution behavior or rapid guessing behavior, where the guessed responses provide no information about the true achievement level of the test takers. Schnipke and Scrams (1997) considered the use of RTs to identify rapid guessing behavior for a speeded test. Wise and Kong (2005) also considered RTs to measure RT effort, which addresses the proportion of true solution behavior in contrast to rapid guessing behavior. Wang and Xu (2015) developed a hierarchical mixture model to determine whether test takers' response strategies could be identified as rapid guessing behavior or solution behavior. Although these approaches consider two different strategies that a test taker might use, they do not consider the actual speed trajectory that underlies the observed RTs. Under the joint model, the speed trajectory in relation to ability will give a more accurate description of the test engagement of the test taker. Therefore, extreme speed trajectories, indicating that responses are given without evaluating the meaning of the question, can indicate rapid guessing behavior, where the speed components are related to ability. Furthermore, a test taker's quadratic time-component effect will show whether a test taker accelerates or decelerates rapidly during the test. This information can be used to address inconsistencies in response patterns. More generally, potential threats to the validity of the test scores (e.g., guessing, cheating) can be evaluated by exploring the speed trajectories of test takers in relation to ability. Statistical tests similar to the tests for aberrant speed behavior reported by Marianti et al. (2014) could be developed to identify extreme changes in speed behavior.

Several model extensions could be considered to make this model suitable for multiple group or multiple latent group settings or for polytomous or nominal response data. The multiple group modeling approach of Azevedo, Andrade, and Fox (2012) might be used to extend the joint modeling of responses and RTs to a multiple group setting. Another interesting extension would be to consider a latent growth model for ability. This would lead to a multivariate latent growth modeling framework for ability and speed, to model changes in the factor variables (speed and ability) over time. Then, changes in speed and ability could be jointly modeled to investigate, for example, changes in the accuracy–speed trade-off over time.

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

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016

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 not supported.

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 ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions is not intended and should not be inferred.

References

Assink, N., van der Lubbe, R. H., & Fox, J. P. (2015). Does time pressure induce tunnel vision? An examination with the Eriksen Flanker Task by applying the hierarchical drift diffusion model. In Y. Wang, P. Borne, & I. Rudas (Eds.), *New developments in computational intelligence and computer science* (pp. 30–40). Vienna, Austria: Proceedings of the International Conference on Applied Physics, Simulation and Computers.

Azevedo, C. L., Andrade, D. F., & Fox, J.-P. (2012). A Bayesian generalized multiple group IRT model with model-fit assessment tools. *Computational Statistics & Data Analysis*, *56*(12), 4399–4412. doi:10.1016/j.csda.2012.03.017

Berger, J. O., Brunero, L., & Wolpert, R. L. (1999). Integrated likelihood methods for eliminating nuisance parameters. *Statistical Science, 14*, 1–28. doi:10.1214/ss/10092118049

Chang, Y. W., Tsai, R. C., & Hsu, N. J. (2014). A speeded item response model: Leave the harder till later. *Psychometrika*, *79*(2), 255–274. doi:10.1007/S11336-007-9031-2

De Boeck, P., & Partchev, I. (2012). IRTrees: Tree-based item response models of the GLMM family. *Journal of Statistical Software*, *48*(1), 1–28. doi:10.18637/jss.v048.c01

Ferrando, P. J., & Lorenzo-Seva, U. (2007a). An item response theory model for incorporating response time data in binary personality items. *Applied Psychological Measurement*, *31*(6), 525–543. doi:10.1177/0146621606295197

Ferrando, P. J., & Lorenzo-Seva, U. (2007b). A measurement model for Likert responses that incorporates response time. *Multivariate Behavioral Research*, *42*(4), 675–706. doi: 10.1080/00273170701710247

Fox, J.-P. (2010). *Bayesian Item Response Modeling: Theory and Applications*. New York, NY: Springer. doi: 10.1007/978-1-4419-0742-4

Fox, J.-P., Klein Entink, R. H., & van der Linden, W. J. (2007). Modeling of responses and response times with the package CIRT. *Journal of Statistical Software*, *20*(7), 1–14.

Goegebeur, Y., De Boeck, P., Wollack, J. A., & Cohen, A. S. (2008). A speeded item response model with gradual process change. *Psychometrika*, *73*(1), 65–87. doi: 10.1007/S11336-007-9031-2

Kingston, N., & Clark, A. (2014). *Test Fraud: Statistical Detection and Methodology*. London: Taylor & Francis.

Klein Entink, R. H., Fox, J.-P., & van der Linden, W. J. (2009). A multivariate multilevel approach to the modeling of accuracy and speed of test takers. *Psychometrika*, *74*(1), 21–48. doi: 10.1007/S11336-008-9075-Y

Klein Entink, R. H., Kuhn, J.-T., Hornke, L. F., & Fox, J.-P. (2009). Evaluating cognitive theory: A joint modeling approach using responses and response times. *Psychological Methods*, *14*, 54—75. doi: 10.1037/a0014877

Marianti, S., Fox, J.-P., Avetisyan, M., Veldkamp, B. P., & Tijmstra, J. (2014). Testing for aberrant behavior in response time modeling. *Journal of Educational and Behavioral Statistics*, *39*(6), 426–451. doi: 10.3102/1076998614559412

Molenaar, D., Tuerlinckx, F., & van der Maas, H. L. (2015a). A bivariate generalized linear item response theory modeling framework to the analysis of responses and response times. *Multivariate Behavioral Research*, *50*(1), 56–74. doi: 10.1080/00273171.2014.962684

Molenaar, D., Tuerlinckx, F., & van der Maas, H. L. (2015b). A generalized linear factor model approach to the hierarchical framework for responses and response times. *British Journal of Mathematical and Statistical Psychology*, *68*(2), 197–219. doi:10.1111/bmsp.12042

Partchev, I., & De Boeck, P. (2012). Can fast and slow intelligence be differentiated? *Intelligence*, *40*(1), 23–32. doi:10.1016/j.intell.2011.11.002

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

Ranger, J., & Kuhn, J. T. (2012). A flexible latent trait model for response times in tests. *Psychometrika*, *77*(1), 31–47. doi: 10.1007/S11336-011-9231-7

Roberts, R. D., & Stankov, L. (1999). Individual differences in speed of mental processing and human cognitive abilities: Toward a taxonomic model. *Learning and Individual Differences*, *11*, 1—120. doi:10.1016/S1041-6080(00)80007-2

Schnipke, D. L., & Scrams, D. J. (1997). Modeling item response times with a two-state mixture model: A new method of measuring speededness. *Journal of Educational Measurement*, *34*(3), 213–232. doi:10.1111/j.1745-3984.1997.tb00516.x

Shao, C., Li, J., & Cheng, Y. (2015). Detection of test speededness using change-point analysis. *Psychometrika*, 1–24. Advance online publication. doi: 10.1007/s11336-015-9476-7

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, *16*(1), 44. doi: 10.1037/a0021765

van der Linden, W. J. (2006). A lognormal model for response times on test items. *Journal of Educational and Behavioral Statistics, 31*(2), 181–204.

van der Linden, W. J. (2007). A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika, 72*(3), 287–308. doi: 10.1007/s11336-006-1478-z

van der Linden, W. J. (2015). Lognormal response-time model. In W. J. van der Linden (Ed.), *Handbook of Item Response Theory: Vol 1. Models*. Boca Raton, FL: Chapman & Hall/CRC. doi: 10.1007/978-1-4757-2691-6

van der Linden, W. J., & Fox, J.-P. (2015). Joint hierarchical modeling of responses and response times. In W. J. van der Linden (Ed.), *Handbook of Item Response Theory: Vol 1. Models*. Boca Raton: FL: Chapman & Hall/CRC.

van der Linden, W. J., & Glas, C. A. W. (2010). Statistical tests of conditional independence between responses and/or response times on test items. *Psychometrika*, *75*(1), 120–139. doi: 10.1007/S11336-009-9129-9

van der Linden, W. J., Klein Entink, R. H., & Fox, J.-P. (2010). IRT parameter estimation with response times as collateral information. *Applied Psychological Measurement, 34*, 327–347. doi: 10.1177/0146621609349800

van der Maas, H. L., & Wagenmakers, E. J. (2005). A psychometric analysis of chess expertise. *The American Journal of Psychology*, *118*, 29–60. doi: 10.1111/bmsp.12054

Wang, C., & Xu, G. (2015). A mixture hierarchical model for response times and response accuracy. *British Journal of Mathematical and Statistical Psychology*, *68*, 456–477. doi: 10.1111/bmsp.12054

Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, *18*, 163–183. doi:10.1207/s15324818ame1802_2

Downloaded by [Universiteit Twente.] at 23:56 07 June 2016