Intelligence 39 (2011) 108–119

Contents lists available at ScienceDirect

Intelligence

# Speed of reasoning and its relation to reasoning ability☆

Frank Goldhammer[a,*], Rinke H. Klein Entink[b]

[a] German Institute for International Educational Research (DIPF), Germany
[b] University of Twente, Netherlands

A R T I C L E I N F O

*Article history:*
Received 21 October 2010
Received in revised form 4 February 2011
Accepted 4 February 2011
Available online 21 March 2011

*Keywords:*
Reasoning ability
Reasoning speed
IRT model
Response time model
Cognitive covariates

A B S T R A C T

The study investigates empirical properties of reasoning speed which is conceived as the fluency of solving reasoning problems. Responses and response times in reasoning tasks are modeled jointly to clarify the covariance structure of reasoning speed and reasoning ability. To determine underlying abilities, the predictive validities of two cognitive covariates, namely perceptual and executive attention, are investigated. A sample of N = 230 test takers completed a reasoning test, Advanced Progressive Matrices (APM), and attention tests indicating perceptual and executive attention. For modeling responses the two-parameter normal ogive model, and for modeling response times the two-parameter lognormal model was applied. Results suggest that reasoning speed is a unidimensional construct representing significant individual differences, and that reasoning speed and ability are negatively correlated but clearly distinguishable constructs. Perceptual and executive attention showed differential effects on reasoning speed and reasoning ability, i.e., reasoning speed is explained by executive attention only, while reasoning ability is explained by both covariates. Implications for the assessment of reasoning are discussed.



## 1. Introduction

Models of the structure of human cognitive abilities paint a complex picture of mental capabilities (cf. Carroll, 1993; Horn & Blankson, 2005; Roberts & Stankov, 1999; Stankov, 2000). Such models include the well-known higher-order ability (level) factors, like fluid intelligence (Gf) and crystallized intelligence (Gc), as well as cognitive speed factors whose hierarchical structure has become more and more differentiated. In Carroll's (1993) seminal work about cognitive abilities in factor-analytic research, the cognitive ability domain of reasoning includes several reasoning ability factors, i.e., deductive, inductive, and quantitative reasoning, and, furthermore, the domain of cognitive speed is assumed to comprise a specific *speed of reasoning* factor. Although the existence of reasoning speed as cognitive speed factor is acknowledged in Carroll's framework and following extensions (for an overview see McGrew, 2005, 2009), only limited or inconsistent empirical evidence is available about the existence of reasoning speed and how it is related to reasoning ability (cf. Carroll, 1993).

The major goal of the study is to clarify the relationship between reasoning ability and reasoning speed. Therefore, a recently developed method for the joint modeling of responses and response times is applied (Klein Entink, Fox, & van der Linden, 2009) to obtain information about both the test taker's level of ability and speed when completing reasoning tasks. First, the covariance structure of reasoning ability and reasoning speed is investigated. Moreover, the study aims to clarify whether two cognitive abilities that have been shown to predict reasoning ability (for an overview see e.g., Schweizer, 2005) also predict reasoning speed to the same extent or differently. That is, the relationship between the two constructs is further clarified by investigating and comparing the predictive validities of two cognitive covariates, perceptual and executive attention, with reasoning ability and speed.

☆ We are grateful to Roberto Colom and two anonymous reviewers for helpful comments on an earlier version of this manuscript.

* Corresponding author at: German Institute for International Educational Research (DIPF), Schloßstr. 29, 60486 Frankfurt/Main, Germany. Tel.: +49 69 24708 323; fax: +49 69 24708 444.
*E-mail address:* Goldhammer@dipf.de (F. Goldhammer).

0160-2896/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2011.02.001

### 1.1. Reasoning ability

Since Spearman's (1923) definition of general intelligence as the ability to extract correlates and relations among a set of entities, reasoning has played a highly important role in the domain of intelligence. In Carroll's (1993) three-stratum model of cognitive abilities, the general intelligence factor, g, at stratum III is identified by a range of broad ability factors at stratum II. One of them is fluid intelligence, Gf, which itself is defined by level factors of reasoning ability, like inductive and deductive reasoning, at stratum I. Following Carroll (1993), inductive reasoning can be conceptualized as the cognitive ability to induce a rule or common characteristics for observed entities and the relations among them (i.e., the conclusion includes more information than the premises from which the conclusion has been derived), while deductive reasoning refers to the ability to draw inferences from given premises to provide a conclusion or to evaluate the correctness of conclusions (i.e., the conclusion does not include more information than that already provided by the premises).

The mental model approach (Johnson-Laird, 1994a,b) conceptualizes reasoning as the creation and manipulation of models representing entities, their properties, and the relations among these entities. The process of inductive reasoning is assumed to comprise three phases (Johnson-Laird, 1994a). The first stage includes the determination of the premises (e.g., by perceptual observations) which at the second stage enable the formulation of a tentative conclusion. Finally, at the third stage the conclusion is evaluated which may result in keeping, updating, or abandoning it. The third stage includes a detection of inconsistencies between conclusion and evidence, retracting the conclusion or doubting the premises, and finally finding explanations for detected inconsistencies (Johnson-Laird, Girotto, & Legrenzi, 2004).

Measures of inductive reasoning, e.g., Raven's (1962) Advanced Progressive Matrices (APM) that are being used in this study, usually require the test taker to generate the logical rules governing the entities and their relations included in the task's stimulus. Once the rules have been found, the task requires at least one deductive step when applying the induced rules, i.e., drawing inferences to give a response (Carroll, 1993). Individual results are usually obtained by number correct scores or person parameter estimates as defined by an item response model. For the APM, Carpenter, Just, and Shell (1990) could identify the abilities to induce abstract relations and to generate, maintain, and monitor the attainment of (sub)goals in working memory as major sources of individual differences in reasoning ability. Both abilities can be conceived as crucial building blocks for a mental model as described by Johnson-Laird (1994a) in that inducing relations clarifies how observed entities are related and the management of multiple (sub)goals is necessary to keep track of the observed figural attributes and derived rules in the APM matrix.

### 1.2. Reasoning speed

The construct of reasoning speed is perceived as an indication of the fluency in performing reasoning tasks. From an individual differences perspective, individuals are assumed to differ not only in their ability but also in the speed level at which they complete the reasoning tasks.

Carroll (1993) reports a few studies providing evidence for a speed of reasoning factor with rate-of-work measures or response times in items as indicators. The three-stratum model locates the speed of reasoning at stratum I as a cognitive speed factor of fluid intelligence (Gf) indicating the efficiency in achieving a cognitive goal. In their factor analytical research work, Roberts and Stankov (1999) relate the reasoning speed factor (*Induction speed*) at stratum I to a general *Psychometric speed* factor (encompassing processes of Carroll's *Broad cognitive speediness*) between strata I and II; the Psychometric speed factor itself serves as indicator of a general speed factor being located at stratum II. This multifaceted mental speed framework extending Carroll's cognitive speed domain suggests that the structure of cognitive speed may be as complex as the structure of cognitive ability (cf. McGrew, 2005; Stankov, 2000).

Individual differences in reasoning speed can be expected in several respects. First, individuals taking exactly the same processing steps may differ in their general processing speed which affects the time needed across the various stages of the reasoning process; for instance, in APM tasks one important aspect of speed is rule generation speed (Verguts, De Boeck, & Maris, 1999). This general source of individual differences in response times is reflected by Roberts and Stankov's (1999) Psychometric speed factor which also explains Induction speed. Moreover, when controlling for the general processing speed there may be further differences in the time spent for the third stage of reasoning, i.e., validating tentative conclusions by looking for inconsistencies and if needed modifying the logical argument. As indicated by Johnson-Laird (1994a), a prudent person will continuously evaluate the (tentative) conclusions, and if needed revise them. This of course will take more time than reasoning without a cautious validation of the conclusion.

### 1.3. Relation between reasoning ability and reasoning speed

Previous findings on the relation between reasoning ability and reasoning speed are limited and inconsistent. Carroll's (1993) overview refers to some datasets (e.g., Kyllonen, 1985) showing speed factors along with the ability factors. He summarizes the role of speed in intelligence stating that there are individual differences in the time needed to perform cognitive tasks, and that these times show low or zero correlations with levels of intelligence. Roberts and Stankov (1999) report a weak positive correlation of the Induction speed factor (indicated by average response times in number and letter series tests) with the corresponding ability factor of Inductive reasoning. In a factor analysis of speed and level factor scores Induction speed showed a significant positive loading not only on a common factor interpreted as overarching broad speed (Gt), but also on another common factor GF which is marked by fluid intelligence (Gf) and Inductive reasoning (IR).

Further empirical evidence is provided by Acton and Schroeder (2001). They assessed the trait *Quickness in seeing relations* (Inductive speed) and found a moderate correlation with analytical reasoning (as the ability to arrange concepts into a logical sequence). In a neuroimaging study by Haier, Schroeder, Tang, Head, and Colom (2010) using the same test battery, a confirmatory factor model was tested with Inductive speed and Analytical reasoning as indicators of a common Reasoning factor; the latter loaded substantially on a second-order factor (which was interpreted as general intelligence, g factor). Most interestingly, speed of reasoning showed a relatively strong and specific pattern of gray matter correlations.

The empirical example provided by Klein Entink, Fox et al. (2009) for their hierarchical modeling approach shows that quantitative and scientific reasoning and associated reasoning speed are negatively correlated indicating that test takers showing higher levels of ability tended to spend more time on completing the tasks. The study by Klein Entink, Kuhn, Hornke, and Fox (2009) also reports a substantial negative correlation between figural reasoning ability and speed. These findings are in line with the above mentioned results of a positive correlation between ability factors and speed factors in that the latter are actually slowness factors given the applied parameterization in factor analysis. It is important to note that in these two empirical examples the relation between the person parameters was estimated at the population or between-person level, i.e., the obtained negative correlation cannot be interpreted as speed-accuracy trade-off which is considered to be a phenomenon at the within-person level.

### 1.4. Effects of person level covariates on reasoning ability and speed

The explanation of individual differences in intelligence and reasoning, respectively, has a long tradition in cognitive psychology. A lot of research has been devoted to the relation of basic cognitive abilities to reasoning ability. One major goal of this vast amount of research was to regress intelligence on various cognitive bases like mental or perceptual speed (cf. the cognitive correlates approach using elementary cognitive tasks, e.g., Jensen, 1982, 1987; Neubauer, 1991), attention (e.g., Schweizer, Moosbrugger, & Goldhammer, 2005; Stankov, 1983), executive attention (e.g., Kane et al., 2004), working memory (e.g., Kyllonen & Christal, 1990; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002), and others.

As regards reasoning speed, to our knowledge hardly any studies are available that address the effect of person level covariates on reasoning speed. In the empirical example provided by Klein Entink, Fox et al. (2009) a negative effect was observed for self-reported test effort, i.e., test takers who care more about their results take more time to complete the reasoning tasks.

### 1.5. Goals and hypotheses

The major goal of the present study is to clarify the properties of reasoning speed from an individual differences perspective. More specifically, we investigate the relation of reasoning speed to reasoning ability, and the predictive validity of attention abilities with reasoning speed and ability. For testing the following hypotheses, the well-validated Advanced Progressive Matrices (APM) test has been selected to assess (figural) inductive reasoning.

Hypothesis 1: Previous empirical research and related theoretical frameworks assume that a speed of reasoning factor exists within the domain of cognitive speed (e.g., Carroll, 1993; Roberts & Stankov, 1999). Based on this research work, we assume that the test takers' response times in APM items reflect one common reasoning speed dimension, i.e., we assume a unidimensional measurement model with one latent speed variable that is sufficient to capture all response time covariance across items; moreover, we assume that test takers differ in their individual level of speed, i.e., the variance of this latent reasoning speed variable is expected to be significant.

Hypothesis 2: We assume that reasoning ability and reasoning speed can be distinguished empirically. Previous findings showed varying degrees of commonalities ranging from weak correlations close to zero (cf. Carroll, 1993) to substantial correlations (cf. Doerfler & Hornke, 2010; Klein Entink, Fox et al., 2009; Klein Entink, Kuhn, et al., 2009). This variability may be accounted for to some extent by specificities of the studies (i.e., samples, speed indicators and modeling approaches). Taken together, we assume reasoning speed and ability to be moderately and negatively related as suggested by previous research work (including also corresponding positive correlations between ability and slowness factors).

Hypothesis 3: To further clarify the expected uniqueness of reasoning speed and ability, underlying cognitive abilities are investigated and compared.

Based on the previous research on the cognitive basis of intelligence (e.g., Jensen, 1987; Kane et al., 2004) we expect perceptual and executive attention to show significant predictive validity with reasoning ability.

Perceptual and executive attention have been selected as covariates because they proved to be major factors underlying individual differences in various attention-related cognitive tasks as suggested by the confirmatory factor model proposed by Moosbrugger, Goldhammer, and Schweizer (2006). Perceptual attention indicates processing speed when performing elementary cognitive tasks including perceptual stimuli, and, therefore, it is assumed to reflect mental speed. Executive attention refers to superordinate control processes that are needed if the task set needs to be (re)configured within or between tasks according to the task goal, e.g., to switch from a primary to a secondary task, to deal with inconsistent stimulus–response mapping and interference between (sub)task goals etc. (cf. Logan & Gordon, 2001).

Most important, the present study aims to investigate whether perceptual attention and executive attention show predictive validity with reasoning speed as expected for reasoning ability. To our knowledge, no empirical evidence is yet available that clarifies the cognitive basis of reasoning speed.

## 2. Method

### 2.1. Participants

A sample of 230 high school and university students completed a computer-based test battery including Raven's Advanced Progressive Matrices (APM) as well as scales assessing executive attention and perceptual attention. There were 65.70% females and 34.30% males aged 19 to 40 years ($M = 23.99$, $SD = 4.00$). Four participants were excluded because for them no measures for executive attention and perceptual attention were available. Participants were assessed individually or in pairs.

### 2.2. Measures

#### 2.2.1. Reasoning scale

Reasoning was assessed by computer-based versions of Raven's (1962) Advanced Progressive Matrices (APM). The figural APM items consist of 3×3 matrices composed of geometrical elements. For each item, one element is missing, and the task is to select the missing element from a set of eight figures so that the rule indicated by the first eight elements in each item is fulfilled. In the present study, form 2, Set II of the APM, consisting of 36 items was used. The APM test was administered without time limit.

### 2.2.2. Measures of perceptual and executive attention

The following attention measures were administered to determine factor scores for perceptual and executive attention in confirmatory factor analysis (CFA).

Four subtests of the Test for Attentional Performance (TAP) (Zimmermann & Fimm, 2000) were used. The *alertness* task is a simple reaction time task. The test taker responds to the appearance of the target ("x") by pressing the response key as fast as possible. The *focused attention* task requires test takers to respond selectively to the appearance of a target, and in the case of a non-target no reaction is required. Stimuli are five regular textures included in a square (two targets and three non-targets). In the *attentional switching* task a letter and a number are presented to the left and to the right of a fixation point. In the first trial the participant detects whether the letter has appeared to the left or to the right and presses the corresponding response key. In the next trial the participant needs to look for the number. In the *sustained attention task*, combinations of a beep (high or low) and one capital letter are presented on each trial. If a low beep is followed by an "E" or a high beep followed by a "N", the participant has to press the response key. In all four scales the result was the median time between the presentation of the critical stimulus and the response.

The *Frankfurt Adaptive Concentration Test* (*FACT*) (Moosbrugger & Goldhammer, 2007) requires test takers to respond selectively to figural targets and non-targets by pressing one of two response buttons. The administered test form *FACT-SR* is characterized by the simultaneous presentation of ten stimuli on the screen. An arrow moving from left to right indicates the next stimulus. The individual FACT-SR result is the inverted average reaction time.

Finally, from the *Multi-dimensional Attention Test (MAT)* (Heyden, 1999) the scale *skill-based interference* was used. In each task two letters appear above and below the center of a square, and two digits to the left and to the right of the center. The participant has to perform simultaneously on two demands: press the first key if the letter channel includes either "D"or "F", otherwise press the second key; press the third key if the digit channel includes either "3"or "5", otherwise press the fourth key. The result was the mean time elapsing between the presentation of the stimulus and the response.

All attention scales were assumed to assess perceptual attention because they require participants to process figural stimulus material. A subset of attention tests additionally requires executive attention, i.e., switching the mental set during task completion because of changing stimulus–response (SR) mapping (TAP attentional switching task), categorizations based on two stimulus dimensions (FACT task), and interfering stimulus dimensions in a dual task (MAT skill-based interference task).

## 2.3. Joint modeling of responses and response times

To address the research questions, a modeling framework is needed that allows for the joint analysis of reasoning speed and ability and their relationship with person-level covariates. The joint modeling approach as proposed by Klein Entink, Fox et al. (2009; see also Klein Entink, Kuhn, et al., 2009) includes measurement models for ability and speed of test takers. At a higher level, the relationship between these constructs is modeled and covariates can be introduced to explain individual differences in ability and speed.

### 2.3.1. Measurement models at level 1

The response model used in this study is the two-parameter normal ogive (2PNO) model which defines the probability that test taker $i$ answers item $k$ correctly as function of the test taker's ability $\theta_i$ as well as the item's difficulty $b_k$ and discrimination $a_k$, is given by

$$P(Y_{ik} = 1 \mid \theta_i, a_k, b_k) = \Phi(a_k \cdot \theta_i + b_k), \tag{1}$$

where $\Phi()$ denotes the normal cumulative distribution function.

Similarly, the two-parameter log-normal (2PLNO) model for response times models the log response time $T_{ik}$ as function of the test taker's speed $\zeta_i$ and the item's time intensity $\lambda_k$ and time discrimination $\phi_k$, i.e.,

$$T_{ik} = -\phi_k \cdot \zeta_i + \lambda_k + \varepsilon_{ik}, \tag{2}$$

where $\varepsilon_{ik}$ denotes the residual which is distributed $\varepsilon_{ik} \sim N(0, \tau_k^2)$. This parameterization enables to disentangle person effects (i.e., speed), and item effects (i.e., intensity and discrimination) on response time.

### 2.3.2. Modeling item and person parameters at level 2

At the second level, the joint distribution of person parameters is specified. From a Bayesian viewpoint, this bivariate normal distribution of speed and ability can be considered to be a common prior for the person parameters:

$$(\theta, \zeta) = \mu_P + e_P, \mu_P = \left(\mu_\theta, \mu_\zeta\right), e_P \sim N(0, \Sigma_P)$$

where $\Sigma_P$ is the covariance matrix given by:

$$\Sigma_P = \begin{bmatrix} \sigma_\theta^2 & \sigma_{\theta\zeta} \\ \sigma_{\theta\zeta} & \sigma_\zeta^2 \end{bmatrix}.$$

The covariance parameter, $\sigma_{\theta\zeta}$, is an important parameter since it reflects the possible dependencies between ability and speed within the population of test takers. Its value reflects to what extend ability and speed are different constructs. Similarly, the variance parameters provide information about individual differences in ability and speed in the population.

Regarding the items, a multivariate normal distribution is in the same way specified for the item parameters of the response and response time models. The covariance structure of this joint distribution provides information about dependencies between item parameters, e.g., the assumption that more difficult items may also show higher time intensity values can be checked by the estimate of $\sigma_{b\lambda}$.

### 2.3.3. Model assumptions

The model is supposed to hold for scales that can be completed within generous time limits. Accordingly, test takers are expected to be able to complete items at a fixed level of speed and accuracy, respectively, i.e., they are not assumed to change the levels of speed and accuracy during testing due to strict time limits (stationarity assumption). Moreover, conditional independence of observations is assumed, i.e., responses and response times, respectively, are expected to be independent across items conditional on the respective person parameter. Based on this, also responses and response times are supposed to be conditionally independent within an item, i.e., the levels of speed and ability presumably capture the covariance between responses and response times to an item completely.

#### 2.3.4. Estimation and software

Statistical inferences for the joint model were performed within the Bayesian statistical framework. In the Bayesian approach, a model parameter is assumed to be a random variable. That is, there is uncertainty about its value, which is reflected by specifying a probability distribution for the parameter. This distribution is called the prior distribution and it reflects the subjective belief of the researcher about admissible values for the parameter before seeing the data. Subsequently, data is collected and the prior is updated according to Bayes' rule, resulting in the posterior distribution of the model parameter, on which inferences can be based. For an introduction into Bayesian statistics, see Gelman, Carlin, Stern, and Rubin (2004).

For estimation of the modeling framework for responses and response times, the CIRT package version 2.5 (Klein Entink, 2010; Fox, Klein Entink, & van der Linden, 2007) for use in the R environment version 2.10.1 (R Development Core Team, 2009) was used. The CIRT package employs a Bayesian Markov Chain Monte Carlo (MCMC) algorithm to obtain parameter estimates by posterior simulation from the joint distribution of the model parameters given the observed data (cf. Gelfand & Smith, 1990). For model identification, the variance of ability, $\sigma_\theta^2$, and the product of time discriminations are fixed to be 1, and the means of ability and speed are fixed to be 0. For all other model parameters, the default non-informative priors as implemented in the CIRT package were used.

For CFA modeling and deriving factor scores of perceptual attention and executive attention Mplus software version 4.21 (Muthén & Muthén, 2006) was used. Model parameters were estimated by means of Maximum likelihood.

### 2.4. Model fit and model selection

For model comparison within the Bayesian approach, the Deviance Information Criterion (DIC) can be used which is the sum of a deviance measure and a penalty term for the effective number of parameters in the model (Spiegelhalter, Best, Carlin, & van der Linde, 2002).

Alternatively, Bayes factors can be computed for selecting the most explanatory model (Kass & Raftery, 1995). The Bayes factor is defined as the ratio of the marginal likelihood of the data under a model $M_0$ and the marginal likelihood of the data under a model $M_1$. The marginal likelihood is the average of the density of the data taken over all possible parameter values admissible by the prior. A Bayes factor of about 1 indicates that both models are equally likely; a value of $\geq 3$ is considered to be strong evidence in favor of model $M_0$, while a value near 0 favors model $M_1$ as the better explanation for the data.

To evaluate the fit of the response model Bayesian posterior predictive checks were done with respect to appropriate test statistics (cf. Sinharay, Johnson, & Stern, 2006). By simulating replicated data sets $\mathbf{x}^{rep}$ from the posterior predictive distribution of the model, a posterior predictive distribution of a test statistic can be constructed. In such cases, the check consists of comparing the replicated data to the observed data and the probability of the model-predicted test statistic being greater than the observed test statistic is assessed (posterior predictive $p$ value).

Regarding the response model, the odds ratio statistic was used to evaluate the conditional independencies among items. If items are pair-wise independent, higher order dependencies are highly implausible (McDonald & Mok, 1995). To assess the overall fit of the response model, the observed sum score distribution was evaluated by comparing it with the sum score distribution as predicted by the response model.

Bayesian residual analysis was used to assess the fit of the response time model (cf. Klein Entink, Fox, et al., 2009). For this, the observed response time of a test taker in a particular item is evaluated under the respective posterior density of response times. More specifically, the probability of a (model-predicted) response time being smaller than the observed one is determined. Following the probability integral transform theorem, under a good fitting RT model for each item these probabilities should be distributed U(0, 1) across test takers.

#### 2.4.1. CFA modeling

The CFA model was considered to show a good model fit if the following criteria were met (cf. Schermelleh-Engel, Moosbrugger, & Müller, 2003): $\chi^2$/df values <2; root mean square error of approximation (RMSEA) values ≤.05; comparative fit index (CFI) and non-normed fit index (NNFI) values ≥.97, and SRMR values <.05.

## 3. Results

### 3.1. Model estimation and fit

First, four models were iteratively fitted to the data to explore the required number of item parameters in the response and in the response time model, respectively. For model estimation 5000 iterations of the Gibbs sampler were used; the final estimates were based on the last 4000 iterations, i.e., the first 1000 iterations were considered as burn-in phase and discarded. Four models were tested with one or two item parameters in the measurement models. Table 1 shows that the most restrictive model $M_1$ including only the item parameters difficulty and time intensity shows the highest DIC. When introducing one of the two discrimination parameters in model $M_2$ and $M_3$, respectively, the DIC is substantially reduced. Adding time discrimination to response time model gave rise to greater decrease in the DIC value than adding discrimination to the response model. However, the best performing model as indicated by the DIC was model $M_4$, which was obtained by including two item parameters in both the response and the response time model.

**Table 1**
DIC values for models fitted to the reasoning scale.

| Model | DIC |
|---|---|
| $M_1$: 1PNO, 1PLN | 19,791.14 |
| $M_2$: 2PNO, 1PLN | 19,731.91 |
| $M_3$: 1PNO, 2PLN | 19,638.26 |
| $M_4$: 2PNO, 2PLN | 19,565.65 |

Note. DIC = Deviance Information Criterion, PNO = parameter normal ogive model, and PLN = parameter log normal model.

To check the stationary speed assumption, i.e., that a test taker does not change his or her speed level during the test, the standardized response time residuals, $e_{ik} = (t_{ik} - (\lambda_k - \phi_k \cdot \zeta_i))/\tau_k$, were computed for each test taker. A systematic trend in the residual pattern of a test taker across test items can be an indication of the violation of this assumption. Fig. 1 shows the residual pattern of 12 selected test takers. An aberrant residual pattern is shown for the 12th test taker who responded much faster to the last seven items than predicted by the RT model. However, the graphical check of residual patterns of all test takers did not show general aberrancies from the stationary assumption. Therefore, we concluded that there was no indication for speededness of the test.

The fit of the response model was investigated by Bayesian posterior predictive checks. For the overall model fit, we checked whether the endorsed model $M_4$ is able to reproduce the observed sum score distribution. In Fig. 2 the observed frequency of sum scores from 0 to 36 items is presented by the line. For 1000 data sets replicated under the posterior density of $M_4$ the sum scores and their frequencies were computed. The solid points represent the average model-predicted frequencies for each sum score. All observed frequencies fall within the .95 highest posterior density (HPD) interval of the model-predicted frequencies suggesting that a 2-parameter response model fits the data well.

Furthermore, the local independence assumption was assessed by means of the odds ratio statistic. For some item pairs the posterior p-values of the odds ratio statistic were <.025 or >.975 which may indicate a violation of the local independence assumption. However, these significant p-values were mainly associated with items showing a high portion of correct responses, and in sum, only a small portion of item combinations were affected (less than 5%).

Finally, the fit of the RT model was assessed by a Bayesian residual analysis for each item. Fig. 3 shows these probability values plotted against their expected values under the U(0, 1) distribution for the first twelve items of the reasoning scale. As shown by these QQ plots the RT model fits the RT data quite well for the first twelve items; for item 12 the RT model slightly underpredicted faster responses. For the other items, there was also almost no deviation from the identity line suggesting that the overall model fit for the RT model was highly acceptable.

### 3.2. Estimated covariance components

The estimated variance and covariance components of the response and the response time model show the relations of parameters within and between measurement models. Table 2 presents the posterior mean (EAP) of the respective component of the covariance structure, the posterior standard deviation (SD), as well as the correlation for the covariance components. The estimated correlation between reasoning speed and ability of $\rho(\theta, \zeta) = -.36$ indicates that more able test takers tend to complete reasoning items more slowly than lower-ability test takers.

Results also gave insight into the relationships between item parameters included in response and response time measurement models. There was a strong positive relationship between difficulty and time intensity, $\rho(b, \lambda) = .63$. This is in line with the plausible assumption that more difficult items require more processing steps and, therefore, more processing time than less complex items. Moderate correlations were obtained for time discrimination and difficulty, $\rho(b, \phi) = .38$, as well as for time discrimination and time intensity, $\rho(\phi, \lambda) = .35$, suggesting that more difficult and time intensive items discriminate better between test takers with different levels of speed. The correlations between discrimination and the other item parameters were negligible.

### 3.3. Testing hypotheses

Hypothesis 1. In Hypothesis 1 we assumed unidimensionality of the reasoning speed and the existence of individual differences in reasoning speed. The fit statistics presented above provide strong support to the assumption that reasoning speed and reasoning ability, respectively, represent unidimensional constructs. As shown in Table 2 the variance of the reasoning speed was estimated to be $\sigma_\zeta^2 = .11$ with a standard deviation of the posterior density of SD = .01. The related .95 HPD interval of [.09, .14] clearly indicates that $\sigma_\zeta^2$ is significantly bigger than zero, i.e., test takers differ in their speed level.

Hypothesis 2. In Hypothesis 2 we assumed that the reasoning ability and speed are negatively correlated but still clearly distinguishable. The estimated correlation of $\rho(\theta, \zeta) = -.36$ fits this assumption, and suggests that test takers showing a higher ability tended to take more time to complete the reasoning tasks than those showing lower ability levels. The amount of shared variance between ability and speed is small, i.e. a large portion of the test takers' speed variance cannot be traced back to their ability levels. Thus, individual reasoning speed is considered to be distinct from the reasoning ability.

To assess the certainty of this estimate the .95 HPD interval for $\rho(\theta, \zeta)$ was computed based on the last 4000 draws of the Gibbs sampling. The .95 HPD interval was [−.49, −.22] suggesting that $\rho(\theta, \zeta)$ deviates significantly from −1 (perfect dependency) and 0 (perfect independency).

Hypothesis 3. In Hypothesis 3 we assumed that individual reasoning ability is significantly predicted by the person level covariates perceptual attention (PA) and executive attention (EA):

$$\theta_i = \gamma_{00} + PA_i \cdot \gamma_{01} + EA_i \cdot \gamma_{02} + e_{0i},$$

$$\zeta_i = \gamma_{10} + PA_i \cdot \gamma_{11} + EA_i \cdot \gamma_{12} + e_{1i},$$

where $(e_{0i}, e_{1i}) \sim N(0, \Sigma_P)$.

The individual levels of perceptual attention and executive attention were determined as factor scores of a confirmatory factor model with test scores of attention ability scales as indicator variables. Following the modeling approach by Moosbrugger et al. (2006), a measurement model was specified with a general factor (Perceptual attention) accounting for

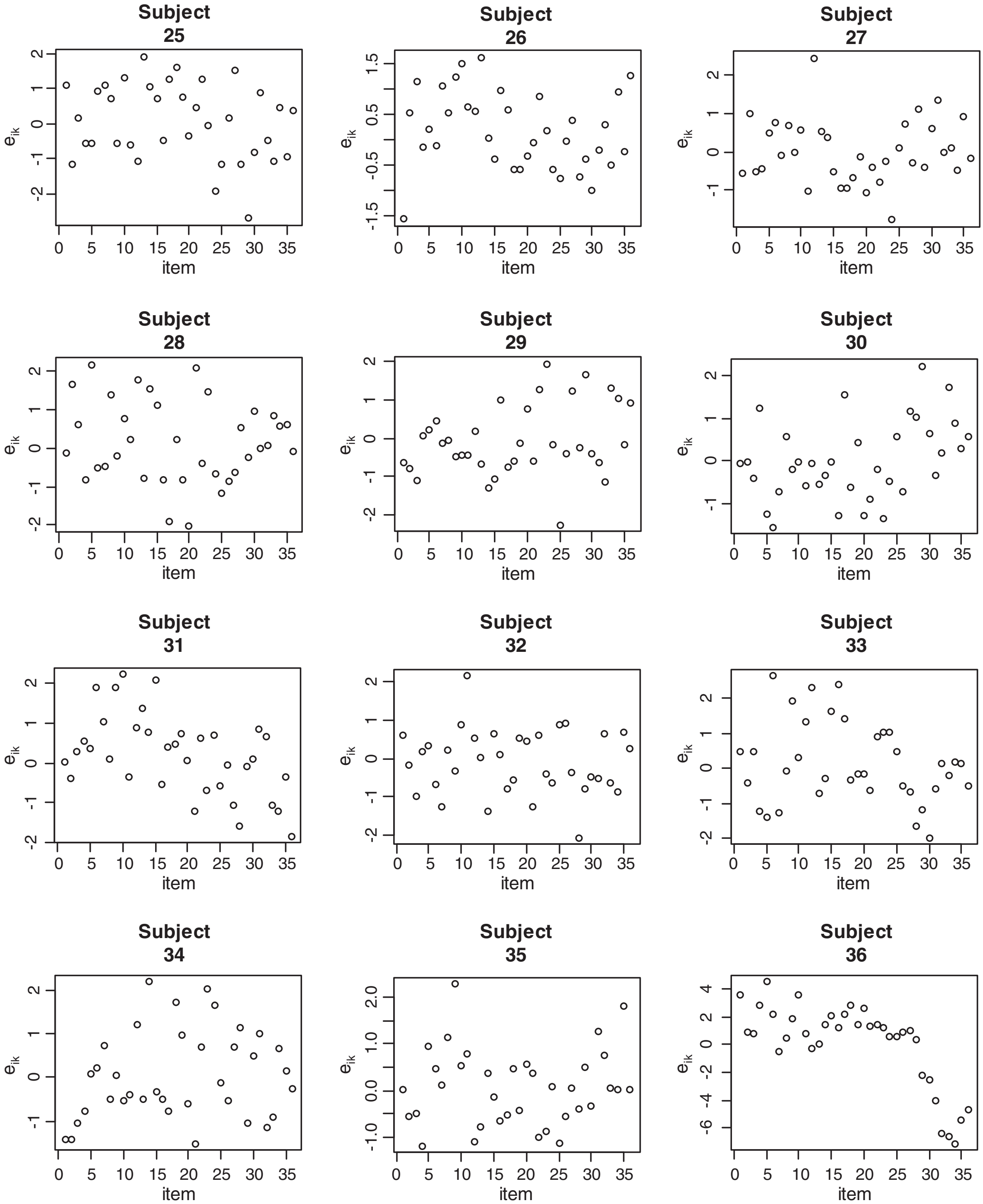


**Fig. 1.** Standardized RT residuals $e_{ik}$ of 12 selected test takers.

variance in all attention measures, and, in addition, with an independent group factor (Executive attention) accounting for the residual variance in attention measures requiring executive control (i.e., the measures MAT Skill-based interference, FACT-SR, TAP Attentional switching). The two-dimensional CFA model fit the data very well, $\chi^2(6)=1.39$, $p=.97$, RMSEA $<.001$, SRMR $=.01$, CFI $=1.00$, NNFI $=1.04$, AIC $=3643.33$ (as opposed to the more parsimonious unidimensional model: $\chi^2(9)=55.65$, $p<.01$, RMSEA $<.149$, SRMR $=.08$, CFI $=.86$, NNFI $=.76$, AIC $=3691.59$ with $\Delta\chi^2(3)=54.26$, $p<.01$). The estimated factor scores for Perceptual attention and Executive attention were used in the following analyses.

Model $M_{4a}$ was specified by introducing perceptual attention and executive attention as person level covariates into Model $M_4$. Table 3 shows the estimated relationships of reasoning ability and speed with the covariates. To compare

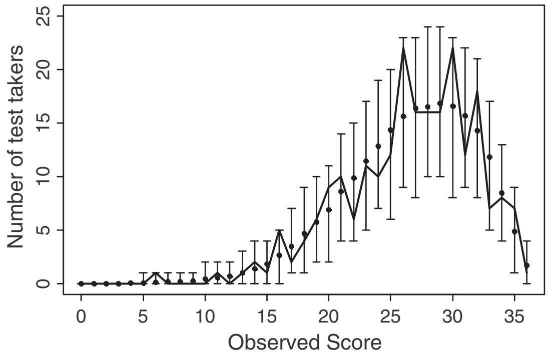


**Fig. 2.** Observed sum scores (line) and model predicted sums scores (points with .95 HPD interval).

the effects of covariates on the person parameters, the regression coefficients were standardized by standardizing each draw of the regression coefficients in the last 4000 iterations of the Gibbs sampler. As assumed in Hypothesis 3, both covariates showed significant effects on reasoning ability. The effect of perceptual attention was estimated to be $\gamma_{01} = .37$, and the effect of executive attention to be $\gamma_{02} = .21$. Their .95 HPD intervals clearly indicate that the effects deviate from 0.

Table 3 also includes the effects of the covariates on reasoning speed. Unlike reasoning ability, the effect of perceptual attention on reasoning speed was around zero, $\gamma_{11} = -.07$ and the .95 HPD also suggests that this effect is not significant. However, there was a significant effect of executive attention on reasoning speed of $\gamma_{12} = .19$.

Model $M_{4b}$ was specified with $\gamma_{11}$ restricted to be zero. As shown in Table 3, the obtained effects for $\gamma_{01}$,$\gamma_{02}$, and $\gamma_{12}$ remain quite the same. To compare models $M_{4a}$ and $M_{4b}$ a

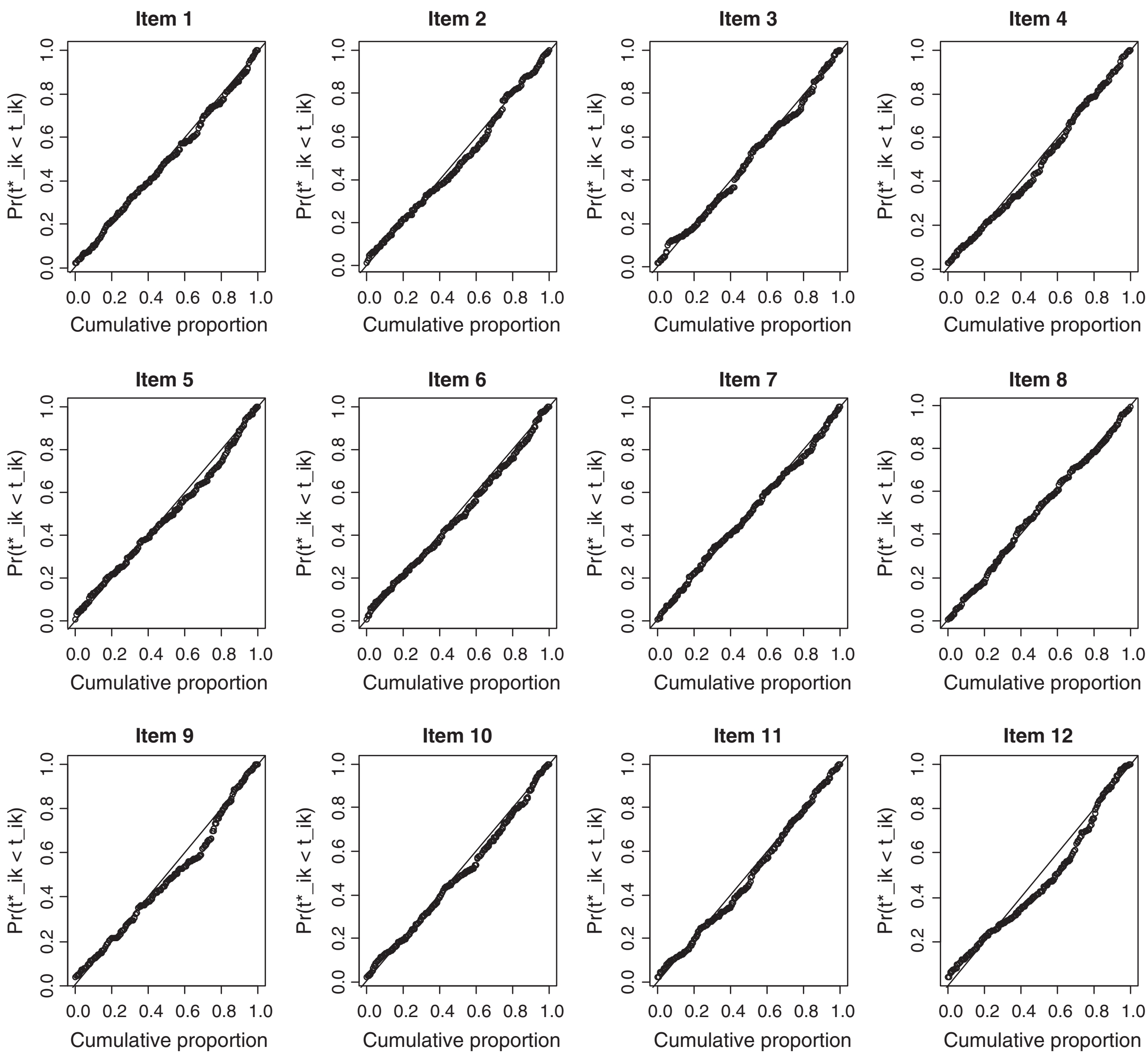


**Fig. 3.** Distribution of $P(t_{ik}{}^{*} < t_{ik}|y, t)$ against its expected distribution U(0, 1) for 12 selected items of the reasoning scale.

**Table 2**
Estimated covariance components and correlations obtained for Model $M_4$.

| Variance component | | EAP | SD | Correlation |
|---|---|---|---|---|
| Person parameters | $\sigma^2_\theta$ | 1.00 | – | 1.00 |
| | $\sigma_{\theta\zeta}$ | −.12 | .02 | −.36 |
| | $\sigma^2_\zeta$ | .11 | .01 | 1.00 |
| Item parameters | $\sigma^2_a$ | .36 | .09 | 1.00 |
| | $\sigma_{ab}$ | .00 | .13 | .00 |
| | $\sigma_{a\phi}$ | .03 | .07 | .08 |
| | $\sigma_{a\lambda}$ | .04 | .09 | .08 |
| | $\sigma^2_b$ | 1.23 | .31 | 1.00 |
| | $\sigma_{b\phi}$ | .26 | .13 | .38 |
| | $\sigma_{b\lambda}$ | .58 | .19 | .63 |
| | $\sigma^2_\phi$ | .38 | .10 | 1.00 |
| | $\sigma_{\phi\lambda}$ | .18 | .09 | .35 |
| | $\sigma^2_\lambda$ | .68 | .17 | 1.00 |

Note. θ = reasoning ability, ζ = reasoning speed; a = discrimination, b = difficulty, ϕ = time discrimination, and λ = time intensity.

Bayes factor (BF) was computed. As the two models are nested, the BF is just the ratio of $p(\gamma_{11}=0|y, M_{4a})$ and $p(\gamma_{11}=0 \mid M_{4a})$, i.e., the evaluation of $\gamma_{11}=0$ under marginal posterior density is divided by the evaluation of $\gamma_{11}=0$ under the prior density with $\gamma_{11} \sim N(0, 1)$. The obtained Bayes factor was BF = 21.78, i.e., the update of the prior density of $\gamma_{11}$ by observed data **y** clearly shows that $\gamma_{11}=0$ in the restricted model $M_{4b}$ is appropriate. As Bayes factors depend on the chosen prior density, e.g. the BF tend to favor an alternative model when the prior density of the restricted parameter is vague, the Bayes factor was also computed with a more informative prior, $\gamma_{11} \sim N(0, .20)$, and a less informative prior, $\gamma_{11} \sim N(0, 2)$. The obtained BFs of 4.32 and 42.92 were both well above the threshold of 3 indicating strong evidence in favor of the restricted model $M_{4b}$.

Taken together, the model comparison shows that reasoning ability and reasoning speed need to be distinguished with respect to the predictive validity of basic cognitive abilities. Reasoning ability is explained by both individual perceptual attention and executive attention, while reasoning speed is only explained by executive attention. This result suggests that test takers with higher levels of executive attention tend to show more accurate and faster responses in reasoning tasks (the positive effects of executive attention on both speed and ability while speed and ability are negatively correlated is possible since the correlation between ability and speed is only moderate and far from 1). Moreover, the fact that perceptual attention was related to reasoning ability but not speed, provides further support to Hypothesis 2, assuming that reasoning speed and reasoning ability are indeed different constructs.

## 4. Discussion

The present study investigated the relation between reasoning ability and reasoning speed to come to a more complete understanding of the domain of reasoning as a major part in the complex structure of human cognitive abilities and (corresponding) speed factors.

In Hypothesis 1 we assumed that a speed of reasoning factor exists within the domain of cognitive speed. The obtained results clearly show that APM response time data is accounted for by a unidimensional measurement model of reasoning speed, and reasoning speed proved to be a person parameter that varies significantly across test takers.

Carroll (1993) assigned the speed of reasoning (at stratum I) to the broader ability factor of fluid intelligence (at stratum II). The association of (figural) reasoning speed with reasoning ability found in the present study may suggest that the reasoning speed can be considered as a factor in the domain of reasoning. However, the extent to which the reasoning speed is a more general processing speed factor, e.g., Psychometric speed as suggested by Roberts and Stankov (1999), or a speed factor that is specific and corresponding to reasoning ability (cf. Bates & Shieles, 2003, discussing parcelated speed effects on group factors of ability; McGrew, 2005; Stankov, 2000), needs to be further addressed empirically. Such studies would include indicators for manifold cognitive abilities and would compare the covariance structure of ability and related speed factors with the covariance among the speed factors.

Hypothesis 2 assumed that the reasoning ability and reasoning speed can be distinguished empirically. Overall, a moderate negative correlation was estimated which proved to be well above zero and well below a perfect correlation. The negative relationship suggests that more able test takers complete tasks at a lower speed level and vice versa. It is important to note that test takers may show differences in the chosen speed level at which they complete test items, which in turn determines the accuracy level as predicted by the speed-accuracy tradeoff (cf. Wickelgren, 1977). For instance, one test taker may complete the tasks at high speed and therefore obtains a lower ability score than a test taker with the same cognitive capability focusing on accuracy and therefore selecting a lower level of speed. This, however, does not mean that test takers can increase their reasoning performance arbitrarily by just taking more time. The

**Table 3**
Estimated standardized effects of perceptual attention (PA) and executive attention (EA) in model $M_4$.

| Model | $\gamma_{01}$ | | $\gamma_{02}$ | | $\gamma_{11}$ | | $\gamma_{12}$ | |
|---|---|---|---|---|---|---|---|---|
| | PA | | EA | | PA | | EA | |
| | EAP | .95 HPD | EAP | .95 HPD | EAP | .95 HPD | EAP | .95 HPD |
| $M_{4a}$ | .37 | [.23, .58] | .21 | [.07, .35] | −.07 | [−.20, .06] | .19 | [.05, .32] |
| $M_{4b}$ | .36 | [.22, .50] | .21 | [.08, .37] | – | – | .18 | [.05, .31] |

Note. EAP = expected A posteriori estimate, and HPD = highest posterior density interval, — indicates that the effect was restricted to be zero.

potential increase of the assessed ability level by taking more time is limited by the level and trajectory of the monotonically decreasing intra-individual speed-ability function (cf. van der Linden, 2009).

The negative correlation raises the question of how test takers benefit from taking more time. Johnson-Laird's (1994a) mental model approach suggests that especially the processes assumed in the third stage could explain why test takers obtaining higher ability scores show lower speed levels. To ensure a correct response, it is necessary that (tentative) conclusions are repeatedly monitored and validated (see also Carpenter et al., 1990). Monitoring, evaluating, and, if necessary, withdrawing hypothesized rules implies a lower speed level; however, it also decreases the probability of incorrect reasoning. This interpretation is in line with the previous finding that test takers who care more about their results take more time to complete reasoning tasks (see Klein Entink, Fox, et al., 2009).

To further investigate the relation of trait variables to reasoning speed and ability, it would be interesting to extend the current analysis approach by including confidence judgements on the accuracy of answers. Previous research shows that confidence is moderately related to latencies in reasoning tasks (cf. Pallier et al., 2002; Stankov & Crawford, 1997), suggesting that test takers with lower levels of speed tend to show higher confidence.

Finally, in Hypothesis 3 we expected perceptual and executive attention to predict reasoning ability and explored the effects of these covariates on the reasoning speed. The ability to reason correctly was assumed to depend on perceptual attention as suggested by the cognitive correlates approach assuming that the speed of information processing is basic to general intelligence (cf. Jensen, 1982, 1987). We used a set of six attentional tests requiring perceptual processing to define a perceptual attention factor that was marked by three tests being similar to simple response time measures (and which were not supposed to require executive control). Perceptual attention moderately predicted the reasoning ability as expected. Furthermore, reasoning ability was hypothesized to rely on executive attention (cf. Kane et al., 2004; Logan & Gordon, 2001). Accordingly, the executive attention factor represented by the subset of three attentional tests requiring also executive control processes significantly explained variance in reasoning ability.

Results show that reasoning speed was not related to perceptual attention, i.e., higher levels of perceptual attention do not advance the reasoning speed. However, executive attention proved to be a source of individual differences in reasoning speed. This finding suggests that the distinctness of the reasoning speed and reasoning ability is also shown by their different cognitive bases.

One means of explaining this result pattern would be to further elaborate the already proposed interpretation that reasoning speed mainly reflects the time spent in the last phase of the mental modeling process as described by Johnson-Laird (1994a). In this phase of evaluating and modifying the current conclusion, the test taker induces alternative rules which require to reconfiguring perceptual processes to determine alternative correspondence among figures and their attributes, if necessary the test taker withdraws hypothesized rules and he or she needs to monitor and coordinate associated (sub)task goals (cf. Carpenter et al., 1990; Johnson-Laird, 1994a). This phase may also involve executive functioning that is responsible for monitoring and coding incoming information and revising appropriately the content of working memory (cf. Miyake et al., 2000). Here, updating the mental model would depend on whether incoming information is confirmatory and/or contradictory to the current mental model. Thus, assuming that reasoning speed mainly reflects time spent for validation and evaluation, it seems plausible that the efficiency of executive attention shows a stronger effect on the reasoning speed than perceptual attention, and, thereby, determines to some extent how long the reasoning process for a particular reasoning task takes. To further validate this interpretation, the result pattern needs to be replicated in independent samples in future studies.

Finally, the nonsignificant effect of perceptual attention supports the notion that reasoning speed is a distinct cognitive speed factor in that the latency-based factor of perceptual attention predicts reasoning ability but not reasoning speed. This is in line with Neubauer's (1990) finding that speed in simple response time measures (Hick paradigm tasks) is unrelated to speed in intelligence suggesting that speed in the Hick paradigm and speed in responding to the intelligence test items reflect different processes.

Carroll (1993) assumed that intelligence is mainly a *level* ability. His conclusion is based on significant but only small correlations between intelligence as a level factor and elementary cognitive tasks, as well as weak or zero correlations between level and speed of intelligence. Nevertheless, the present study shows that the collection of response times provides an important source of information about individual differences and that a measurement model for the response times contributes to understanding how test takers differ in solving cognitive tasks. This means that the test taker's performance level can be described more specifically in that the information about the ability level is enriched by the respective speed level, i.e., for instance, two test takers may show substantial differences in speed even though they show the same level of ability. The speed level provides additional information which can be used to create an individual efficiency profile including both the test taker's ability and speed level. Such profiles could be interpreted in the sense of Thorndike, Bregman, Cobb, and Woodyard's (1926) assumption that "other things being equal, the more quickly a person produces the correct response, the greater is his intelligence" (p. 24). From this perspective, individual reasoning speed and ability provide information about the test taker's efficiency in solving reasoning problems. This additional information may improve the predictive validity of a reasoning test as a selection device (see e.g., the study by Doerfler & Hornke, 2010, using response latencies in APM tasks to explain the lower APM test score of extravert test takers). However, the question of whether reasoning speed shows incremental validity when reasoning ability is, for example, used to predict educational or occupational success, still needs to be addressed empirically. If the reasoning speed explained variance above and beyond reasoning ability, the relevance of the speed parameter would be emphasized also with respect to criterion validity.

## Appendix A

Descriptive statistics and Pearson correlations for measures (test scores) of reasoning and attention.

| | Measure | *M* | *SD* | Skewness | Kurtosis | Pearson correlation | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | APM Set II, number correct | 26.95 | 4.61 | −.38 | −.47 | 1.00 | | | | | | |
| 2 | APM Set II, total time [min.] | 41.35 | 15.79 | .80 | 1.02 | −.27 | 1.00 | | | | | |
| 3 | TAP Alertness [ms] | 243.58 | 39.81 | 1.72 | 5.18 | .31 | −.15 | 1.00 | | | | |
| 4 | TAP Focused attention [ms] | 488.32 | 56.76 | .66 | .68 | .38 | −.16 | .57 | 1.00 | | | |
| 5 | TAP Attentional switching [ms] | 715.73 | 184.40 | 1.57 | 3.89 | .19 | .02 | .24 | .23 | 1.00 | | |
| 6 | TAP Sustained attention [ms] | 429.75 | 82.84 | 1.25 | 2.01 | .29 | −.11 | .56 | .62 | .21 | 1.00 | |
| 7 | FACT-SR [100/s] | 161.84 | 37.58 | .19 | .08 | .24 | .07 | .15 | .18 | .24 | .13 | 1.00 |
| 8 | MAT Skill-based interference [ms] | 3142.08 | 610.47 | .14 | −.09 | .35 | .16 | .27 | .30 | .47 | .26 | .28 |

Note. For the computation of correlations, all measures representing time information (i.e., APM Set II total time, Alertness, Focused attention, Attentional switching, Sustained attention, Skill-based interference) were multiplied by (−1) so that all measures have a positive orientation.

## References


Acton, G. S., & Schroeder, D. H. (2001). Sensory discrimination as related to general intelligence. *Intelligence*, *29*, 263−271.

Bates, T. C., & Shieles, A. (2003). Crystallized intelligence as a product of speed and drive for experience: The relationship of inspection time and openness to g and Gc. *Intelligence*, *31*, 275−287.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*, 404−431.

Carroll, J. B. (1993). *Human cognitive abilities.* New York: Cambridge University Press.

R Development Core Team (2009). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing 3-900051-07-0. http://www.R-project.org.

Doerfler, T., & Hornke, L. H. (2010). Working style and extraversion: New insights into the nature of item response latencies in computer-based intelligence testing. *Journal of Research in Personality.*, *44*, 159−162.

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).

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, *85*, 398−409.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Washington, DC: Chapman & Hall/CRC.

Haier, R. J., Schroeder, D. H., Tang, C., Head, K., & Colom, R. (2010). Gray matter correlates of cognitive ability tests used for vocational guidance. *BMC Research Notes*, *3*, 206.

Heyden, M. (1999). *Entwicklung und Erprobung einer multidimensionalen Aufmerksamkeits-Testbatterie* [Development and validation of a multidimensional attention test battery]. Dissertation Fachbereich Psychologie. Mikrofiche-Ausgabe. Frankfurt a. M.: Senckenbergische Bibliothek.

Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 41−68). (2nd ed). New York: Guilford Press.

Jensen, A. R. (1982). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in research on intelligence*, *1*, Hillsdale, NJ: Erlbaum.

Jensen, A. R. (1987). Individual differences in the Hick paradigm. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 101−175). Norwood, NJ: Ablex.

Johnson-Laird, P. N. (1994a). A model theory of induction. *International Studies in the Philosophy of Science*, *8*, 5−29.

Johnson-Laird, P. N. (1994b). Mental models and probabilistic thinking. *Cognition*, *50*, 189−209.

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review*, *111*, 640−661.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*, 189−217.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773−795.

Klein Entink, R. H. (2010*). R-package for response-response time analysis. CIRT version 2.5 (beta).* Retrieved February 15, 2010, from http://www.kleinentink.eu/download/cirt_2.5.

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*, 21−48.

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.

Kyllonen, P. C. (1985). *Dimensions of information processing speed (Tech. Rep. No. AFHRL-TP-84-56).* Brooks Air Force Base, TX: Air Force Systems Command.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389−433.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*, 393−434.

McDonald, R. P., & Mok, M. M. -C. (1995). Goodness of fit in item response models. *Multivariate Behavioral Research*, *30*, 23−40.

McGrew, K. S. (2005). The Cattell–Horn–Carroll (CHC) theory of cognitive abilities: Past, present and future. In D. Flanagan, & Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136−202). New York: Guilford Press.

McGrew, K. S. (2009). Editorial: CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1−10.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49−100.

Moosbrugger, H., & Goldhammer, F. (2007). *FAKT-II. Frankfurter Adaptiver Konzentrationsleistungs-Test. Grundlegend neu bearbeitete und neu normierte 2. Auflage des FAKT von Moosbrugger und Heyden (1997) [FAKT-II. Frankfurt adaptive concentration-performance test. Second, completely revised and renormed edition of the FAKT by Moosbrugger and Heyden (1997)].* Bern: Huber.

Moosbrugger, H., Goldhammer, F., & Schweizer, K. (2006). Latent factors underlying individual differences in attention measures: Perceptual and executive attention. *European Journal of Psychological Assessment*, *22*, 177−188.

Muthén, L. K., & Muthén, B. O. (1998–2006). *Mplus user's guide* (Fourth Edition). Los Angeles, CA: Muthén & Muthén.

Neubauer, A. C. (1990). Speed of information processing in the hick paradigm and response latencies in a psychometric intelligence test. *Personality and Individual Differences*, *11*, 147−152.

Neubauer, A. C. (1991). Intelligence and RT: A modified Hick paradigm and a new RT paradigm. *Intelligence*, *15*, 175−192.

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *Journal of General Psychology*, *129*, 257−299.

Raven, J. C. (1962). *Advanced progressive matrices.* London: Lewis & Co. Ltd.

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.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*, 23−74.

Schweizer, K. (2005). An overview of research into the cognitive basis of intelligence. *Journal of Differential Psychology*, *26*, 43−51.

Schweizer, K., Moosbrugger, H., & Goldhammer, F. (2005). The structure of the relationship between attention and intelligence. *Intelligence*, *33*, 589−611.

Sinharay, S., Johnson, M. S., & Stern, H. S. (2006). Posterior predictive assessment of item response theory models. *Applied Psychological Measurement*, *30*, 298−321.

Spearman, C. (1923). *The nature of intelligence and the principles of cognition.* London: Macmillan.

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*, *64*, 583−639.

Stankov, L. (1983). Attention and intelligence. *Journal of Educational Psychology*, *75*, 471−490.

Stankov, L. (2000). Structural extensions of a hierarchical view on human cognitive abilities. *Learning and Individual Differences*, *22*, 35−51.

Stankov, L., & Crawford, J. D. (1997). Self-confidence and performance on tests of cognitive abilities. *Intelligence*, *25*, 93−109.

Süß, H. -M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability — and a little more. *Intelligence*, *30*, 261−288.

Thorndike, E. L., Bregman, E. O., Cobb, M. V., & Woodyard, E. (1926). *The measurement of intelligence.* New York: Bureau of Publications, Teachers College, Columbia University.

van der Linden, W. J. (2009). Conceptual issues in response-time modeling. *Journal of Educational Measurement*, *46*, 247−272.

Verguts, T., De Boeck, P., & Maris, E. (1999). Generation speed in Raven's progressive matrices test. *Intelligence*, *27*, 329−345.

Wickelgren, W. A. (1977). Speed-accuracy trade-off and information processing dynamics. *Acta Psychologica*, *41*, 67−85.

Zimmermann, P., & Fimm, B. (2000). *Test for attentional performance (TAP).* Herzogenrath: PSYTEST.