Available online at www.sciencedirect.com
ScienceDirect
Intelligence 35 (2007) 591 – 604
INTELLIGENCE

Investigating the relationship of working memory tasks and fluid intelligence tests by means of the fixed-links model in considering the impurity problem

Karl Schweizer *

Institute of Psychology, Johann Wolfgang Goethe-University, Mertonstr. 17, 60054 Frankfurt a. M., Germany

Received 21 December 2005; received in revised form 29 November 2006; accepted 29 November 2006
Available online 12 January 2007

Abstract

The impurity of measures is considered as cause of erroneous interpretations of observed relationships. This paper concentrates on impurity with respect to the relationship between working memory and fluid intelligence. The means for the identification of impurity was the fixed-links model, which enabled the decomposition of variance into experimental and non-experimental parts. A substantial non-experimental part could be expected to signify impurity. In a sample of 345 participants error scores and reaction times, which were obtained by the Exchange Test, represented working memory, and Advanced Progressive Matrices served as measure of fluid intelligence. The four independent latent variables of the model associated with error scores and reaction times led to a multiple correlation .67 with the latent variable of fluid intelligence. However, there was impurity since the decomposition by means of the fixed-links model showed that only 45% of the common variance was due to working memory.

Keywords: Cognitive process; Fluid intelligence; Structural equation modeling; Latent curve model

The question of whether working memory contributes to intelligence has stimulated a large number of studies. As consequence, many correlational results suggesting the existence of a substantial relationship are available. Ackerman, Beier and Boyle (2005) report a metaanalytic investigation of 57 studies and suggest a correlation of .48. The inspection of the individual results reveals that this field of research shows a high degree of heterogeneity. There are rather low besides very high correlations. The results obtained by means of structural equation modeling are most impressive. Some studies even suggest near identity of working memory and intelligence with respect to individual differences. Typically, the relationship is investigated at the latent level in considering a number of (slightly differing) measures (e.g., Buehner, Krumm, & Pick, 2005; Colom, Abad, Rebollo, & Shih, 2005; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Colom & Shih, 2004; Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen & Christal, 1990).

The heterogeneity of results demands for an explanation. Actually, there is a number of potential explanations. For example, the difference between correlations observed at the manifest level on one hand and at the latent level on the other hand provides an

* Tel.: +49 69 798 23350; fax: +49 69 798 23847.
E-mail address: K.Schweizer@psych.uni-frankfurt.de.

0160-2896/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2006.11.004

explanation. Different degrees of similarity between the measures for the assessment of working memory on one hand and of intelligence on the other hand give rise to another explanation (Schweizer, 2005). Samples originating from different populations, which in varying degrees allow age to act as moderator can also be accepted as explanation since age was found to be an influential source (Salthouse, 2005). Furthermore, there is the impurity of measures as explanation. There may be different degrees of impurity. If impurity is given, one part of the observed relationship is due to the intended source of performance whereas the other part is due to another source. Impurity calls the interpretation of the result into question.

This paper concentrates on the so-called impurity problem with respect to the relationship between working memory and intelligence. It demonstrates that the impurity of measures can lead to an erroneous interpretation of results concerning the relationship. The identification of impurity requires the combination of differential and experimental methodologies as proposed by Cronbach (1957) some time ago. Furthermore, it is necessary to apply a special structural equation model, denoted fixed-links model (Schweizer, 2006a). In the following sections the impurity problem is presented in more detail, the methodology is described and applied to data obtained by means of measures of working memory and fluid intelligence. The investigation of data in considering the impurity problem ends up with the separation of common variance into parts due to the intended source of performance and due to other sources. If the part due to the other sources is negligible, impurity is not a problem.

1. The impurity problem

Measures representing cognitive concepts are normally not pure measures. In analyzing the processes contributing to performance in completing choice reaction time tasks Jensen (1982) came up with four processes: encoding, operation, binary decision and response. In many cognitive tasks the uptake of information is necessary for initiating the process of interest and the conduction of motor processes for terminating processing. However, neither one of these processes is essential for answering the research question. Furthermore, Van Zomeren and Brouwer (1994) argue that completing a test, which measures divided attention, additionally requires alertness, focused attention and sustained attention. This argument is pointing into the direction of the dilemma of attention research. The manifestation of attention includes the taking influence on some of the transformation processes, which constitute information processing. This is especially obvious in complex tasks which require a lot of executive control (Miyake, et al., 2000). As consequence, attention, which is a predictor of intelligence (Schweizer, Moosbrugger, & Goldhammer, 2005), typically contributes to correlations between measures representing transformation processes and measures of intelligence. It is the manifold of processes contributing to performance, which constitute the impurity problem of differential research since performance is almost always due to a number of different processes. Because of this problem the validation of measures is necessary. However, validation can only guarantee that the concept of interest, which the measure is expected to represent, is reflected by the major source of performance.

In experimental research the impurity problem is avoided by means of several provisions of which the most important one is experimental treatment (see Harris, 2003). The comparison of the results obtained for the various treatment levels enables the concentration on the relevant components of measurement. Although the experimental effect may be reflected by a small component of the measurement only, appropriate statistical methods enable its identification. The combination of treatments and statistical methods provides the opportunity to get a grip on the impurity problem. In sum, the experimental methodology including treatments and statistical methods allows the experimental researcher to overcome the impurity problem. The success of this methodology provides the blueprint for an analogical approach within differential research.

It needs to be added that impurity is not a problem which is restricted to ability research. In personality research impurity was already identified as an annoyance a long time ago. In this field of research the multitrait–multimethod approach (Campbell & Fiske, 1959) was proposed in order to identify impurity. The core of this approach is the multitrait–multimethod design. Furthermore, structural equation modeling has been applied for improving the quality of measurement. Since neither the multitrait–multimethod design nor structural equation modeling alone proved to be sufficient with respect to the impurity problem, this approach was combined with structural equation modeling (Kenny & Kashy, 1992). Although the multitrait–multimethod approach is intriguing, in the field of cognitive ability it is inappropriate since often a specific method of measurement is characteristic for the corresponding concept. As a consequence, a comparison of methods is not possible.

2. Experimental measures within differential research

It is interesting to note that there is already a tradition of experimental measures in differential research. There are tasks including several treatment levels, which have frequently been employed for investigating the relationship between cognitive processes and intelligence. It was a major characteristic of such treatment levels that they posed demands on the same cognitive processes. For example, there is the so-called choice reaction time task (Jensen & Munro, 1979), which has been especially often applied as measurement device. In this task the various treatment levels consist of different numbers of light bulbs, which are to be monitored. In the list-search task (Neisser, 1963) the target stimulus is included in lists differing in length, thus giving rise to different treatment levels. In the memory-scanning task (Sternberg, 1966) different numbers of letters have to be stored in short-term memory. The different amounts of information, which need to be stored and scanned, provide different treatment levels. Different numbers of ordering operations give rise to different treatment levels in the SWAPS task (Stankov, 2000). In all these tasks the investigation of the effects of the treatment levels usually reveal substantial differences in reaction time and accuracy whereas the results concerning the relationship of reaction time and accuracy on one hand and intelligence on the other hand are often neither impressive nor very consistent. For example, the increases from treatment level to treatment level observed in the choice reaction time task prove to be very small in a very large sample (Jensen, 1987).

3. The exchange test as (experimental) measure of working memory

The present work concentrates on cognitive processes associated with exchange operations as example. The execution of such operations can be ascribed to the central executive of working memory (Baddley, 1986). These operations are elementary and very easy to be performed if the items to be exchanged are simple enough. The exchange operations of interest are stimulated by the tasks of the Exchange Test (Schweizer, 1996). This test served as the measure of working memory within several studies which I conducted in my lab or the lab of a colleague. Furthermore, there are reports of applications by independent researchers (Neubauer, Stern, & Grabner, 2005; Stankov, Brine, & Bowman, 2005). This test possesses all the characteristics, which a measure of working memory should show according to Bayliss, Jarrold, Gunn and Baddeley (2003). These characteristics in combination with the simplicity of the cognitive operations, which are stimulated, render this test as ideal candidate for the investigation of the impurity problem.

In completing this test a few simple figures (lines, squares, cross, etc.) showing a specific ordering have to be exchanged mentally. Since the exchanges are restricted to neighboring figures, intermediary configurations, which must temporarily be retained, are generated. This means that the exchange operations are closely associated with storage operations. Exchange operations are also necessary in completing the Tower-of-Hanoi task. The difficulty characterizing this task is ascribed to the load on working memory (Carpenter, Just, & Shell, 1990). In completing the Raven problems the load on memory also proved to be of special importance (Unsworth & Engle, 2005). Analogically, in the Exchange Test a high number of exchange and storage operations can be assumed as source of error. Furthermore, there are processes which are independent of the number of figures to be exchanged. They include the initial generation of a mental representation and the response, which terminates processing. In sum, there are processes which are independent of the various treatment levels and processes which reflect the treatment levels.

The Exchange Test provides two types of performance measures, accuracy and reaction time. Furthermore, the error score does as well as accuracy because accuracy is defined as the difference between the number of trials and the number of errors. Since single exchange operations are so easy to be performed and specific strategies do not provide an advantage, storage problems are to be considered as the major source of error/lack of accuracy. Errors can occur directly or indirectly due to capacity limitation (Schweizer, 2000, 2001). Directly means the transgression of capacity limitation whereas indirectly indicates the failure to stay within the time limit. Mechanisms leading to time-dependent failure were described by several authors (Jensen, 1982; Salthouse, 1993). Therefore, errors, which occur indirectly, can be expected to show dependence on processing times whereas errors, which occur directly, should prove to be independent of processing times.

It is reasonable to expect an increase in reaction time from treatment level to treatment level since additional exchange and storage operations are stimulated. The execution of the additional operations requires the completion of the operations already stimulated by the previous treatment level. Although the test demands suggest a linear increase, there may be deviations due to either acceleration or slowing. Acceleration can happen because of the facilitation of processing (positive priming) and the slowing because of inhibition of processing (negative priming). Since the simple figures serving as material are changed from trial to trial, and speed is a basic property (Helmbold & Rammsayer, 2006), constancy or slowing is more likely than acceleration. Furthermore, the relationship between the exchange and storage operations on one hand and perceptual and motor operations on the other hand is to be considered. Since exchange operations presuppose the completed representation of the material, independence of the types of operations is likely. Moreover, motor operations, which terminate the trial, presuppose the completion of the exchange operations. So there is some reason for assuming that the corresponding processes are in agreement with the additive model (Sternberg, 1969).

4. Experimental measures as basis for the decomposition of variance

The experimental measures have so far not been combined with a statistical method, which is really suitable for the isolation of the experimental effect within the framework of differential psychology. The statistical method should decompose the variance into one part, which reflects the source stimulated by the experimental treatment, and another part, which reflects all the other sources. Formerly, there were comparisons of correlations associated with different treatment levels. However, very large samples were necessary for reaching the level of significance in such comparisons, and in many cases the consideration of a lot of individual results was required. Furthermore, difference scores and regression parameters were computed at the manifest level for getting grip on the concept of interest. However, the results obtained this way usually showed a low degree of reliability and, consequently, a high degree of heterogeneity. No wonder, there was a long-lasting debate about the usefulness of such measures (see Cronbach & Furby, 1970).

A major reason for the disappointing results was presumably the fact that the former investigations were conducted at the manifest level, which is the level of the observations. More favorable results can be expected at the latent level since the step from the manifest level to the latent level includes the elimination of error. Therefore, it is reasonable to investigate the effects of the treatment levels at the latent level. All the favorable results mentioned in the introduction section were obtained at the latent level. However, such an investigation cannot be performed properly by means of the conventional structural equation model (Jöreskog, 1973; Keesling, 1972; Wiley, 1973) since in this model the differences between the treatment levels are not retained. In the conventional model each one of the measures associated with one level either receives its own latent variable, or all the measures are linked to the same latent variable. Neither one of these alternatives is well suited for representing the systematic differences between the treatment levels. These alternatives do not guarantee the decomposition of variance according to the treatment levels.

5. The fixed-links model as means for the decomposition of variance

The model, which includes constraints of the loadings of the manifest on latent variables (Schweizer, 2006b), enables the appropriate representation of the systematic differences between the treatment levels. In a way these constraints can be considered as a differential weighting system for latent variables. Such constraints also characterize latent curve models (see McArdle, 1988; McArdle & Epstein, 1987; Meredith & Tisak, 1984, 1990). Latent curve models have been developed for the investigation of time-dependent change. The latent curve describes the change, which occurs within a definite time span, at the latent level. There are constrained loadings of manifest on latent variables, which determine the shape of the latent curve.

Although the treatment levels of an experimental investigation of cognitive processes are not associated with specific points in time, there is some degree of similarity to the notion of time-dependent change. The treatment levels usually include the stimulation of the same process and impose a specific temporal structure on the measurements. For example, the same process is stimulated once, two times, three times and so on, so that there are different numbers of repetitions. This means that the treatment levels represent different spans between the initiation and termination of processing. This consideration suggests similarity although the treatment levels are not arranged according to a specific sequence of points in time and although the material used for stimulation slightly changes.

A rather general approach is selected as outset in order to arrive at a formal representation of the systematic differences between the treatment levels, which applies to a very general audience. It is provided by the following model:

$$Y = g(\mathbf{X}) + e$$

where Y is the response variable, $\mathbf{X}$ the vector of explanatory variables (=latent factors) and e the residual. The function g is expected to represent the true effect of the explanatory variables due to stimulation. It is obvious that this model can be considered as a version of the true-score model (Novick, 1966; Lord and Novick, 1968) with the first summand as true score and the second summand as error score.

The simple regression model suggests the consideration of function g as the composite of two further functions a and b such that

$$g(\mathbf{X}) = a(\mathbf{X}) + b(\mathbf{X}).$$

The function giving the first summand represents the intercept and the function giving the second summand the slope. This equation includes very general descriptions of both the intercept and the slope so that each one of them can be specified in a variety of ways. Furthermore, it enables the simultaneous investigation of both the effects on the means and covariances (including variances). Information concerning the estimation of latent factors can readily be taken from Kline (2006).

The most simple specification of function a is achieved by including the sample mean of Y (M_y):

$$a(\mathbf{X}) = M_y$$

which is characteristic of regression analysis. Furthermore, there is the possibility to specify a as mean structure (Bentler & Yuan, 2000). Mean structures enable the consideration of the different means which are achieved as the result of the different treatment levels or the subdivision of the sample into groups. The means can be estimated and submitted to an investigation (e.g., Dolan et al., 2006). For example, in longitudinal research the mean course of development may be investigated by specifying a accordingly.

The appropriate specification of function b is especially important with respect to the formal representation of the expected systematic differences between the treatment levels. Since b is rather complex, the specification occurs in two steps. Firstly, it is necessary to specify b with respect to the explanatory variables (=latent factors) which contribute. In assuming that $k+1$ explanatory variables need to be considered, b is given by

$$b(\mathbf{X}) = b_0X_0 + b_1X_1 + \ldots + b_kX_k$$

where $b_i (i=0,\ldots, k)$ is a weight and $X_i (i=0,\ldots,k)$ an explanatory variable (=latent factor). Secondly, the effect of the experimental treatments needs to be specified appropriately. In following specifications proposed within the framework of latent curve analysis the polynomial function (McArdle, 1988; McArdle & Epstein, 1987) is selected for this purpose. Accordingly, the function b must be adapted to the different treatment levels, and the weights must be replaced by the corresponding constituents:

$$b(\mathbf{X})_j = j^0X_0 + j^1X_1 + \ldots + j^kX_k$$

where b associated with the index $j (j=1,\ldots, p)$ is the version of the function, which applies to the jth treatment level, and the weight j an integer which reflects the treatment level also (major parts of this function are denoted constituents in this paper instead of components in order to avoid confusion with the components of the model of measurement). It is useful to consider the constant, linear and quadratic constituents since higher-order constituents of the polynomial function are not very likely in small numbers of treatment levels. The weights of this equation make especially obvious that the constraints of the fixed-links model may be perceived as a differential weighting system for the latent variables.

Whereas the general model applies to both the means and covariances (including variances), the fixed-links model is restricted to covariance matrices. This means that the intercept of g is not considered and the observed scores are assumed to be standardized to a mean of zero ($\mathbf{y}$ instead of Y). The concentration on the fixed-links model requires the formal presentation of the parameters of the model of measurement in a way which is typical for structural equation modeling. Accordingly, the model of measurement associated with $\mathbf{y}$ is defined as

$$\mathbf{y} = \Lambda\eta + \varepsilon$$

where $\mathbf{y} = (y_1, \ldots, y_p)'$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_p{}')$ are the $p\times1$ vectors of observations and of error components, $\eta = (\eta_1, \ldots, \eta_q)'$ is the $q\times1$ vector of latent variables (=latent factors) and $\mathbf{\Lambda}$ the $p\times q$ matrix of loadings (=links relating latent to manifest variables) (Bollen, 1989, p. 18). Whereas in the conventional model of measurement most elements of $\mathbf{\Lambda}$ are constrained to zero and the other elements are free to be estimated, in the fixed-links model of measurement all (or almost all) the elements are constrained. The elements, which are constrained to zero in the conventional model, are also constrained to zero in the fixed-links model. However, both these models differ with respect to the elements of $\mathbf{\Lambda}$, which are free in the conventional model. In the fixed-links model these elements are assigned to the sets of numbers which as a whole must represent the effects of the treatment levels appropriately. Setting free the diagonal elements of the covariance matrix $\mathbf{\Phi}$ compensates for the restriction due

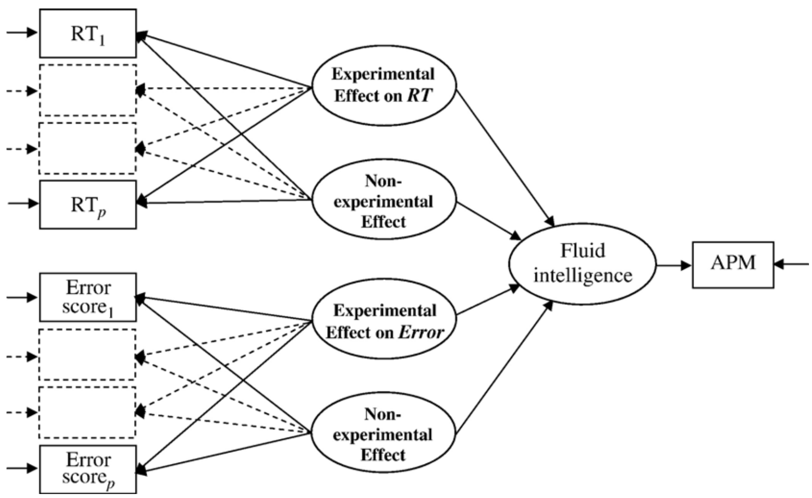

Fig. 1. Graphical representation of the model for the decomposition of variance into parts due to experimental and non-experimental sources with one dependent latent variable.

to the constraint of the elements of $\mathbf{\Lambda}$. This is obvious from the model of the covariance matrix $\mathbf{\Sigma}$:

$$\boldsymbol{\Sigma} = \Lambda\Phi\Lambda' + \Theta$$

(where $\mathbf{\Theta}$ is the diagonal matrix of error scores) since $\mathbf{\Lambda}$ and $\mathbf{\Phi}$ depend on each other. Both of them cannot be free at the same time.

6. The decomposition of variance into experimental and non-experimental parts

The precondition for the constraint of the elements of $\mathbf{\Lambda}$ is the availability of appropriate numbers. In latent curve models the numbers must guarantee the error-free representation of growth. There are two ways of making use of the polynomial function in latent curve and fixed-links models (see Schweizer, 2006a). In the first way, each one of the constituents (constant, linear, quadratic, etc.) is represented by means of one column of $\mathbf{\Lambda}$. This means that numbers giving rise to the corresponding curve are inserted into the columns of $\mathbf{\Lambda}$. Such numbers are included in the fifth equation of the previous section. An example constructed with respect to six treatment levels, constant, linear and quadratic constituents is given by

$$\Lambda = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix}.$$

In this case each constituent leads to one latent variable. As consequence, there is the decomposition of the variances of the manifest variables. Such decomposition is useful when the relationships between the latent variables representing different constituents and other latent variables need to be investigated. In the second way of employing the polynomial function the constituents are combined so that $\mathbf{\Lambda}$ includes one column only. This means that the weights of the polynomial function must be estimated outside of structural equation modeling and applied for obtaining the requested numbers. Since the investigation of impurity requires the decomposition of variance into at least two parts, only the first way is appropriate for this paper.

In the formal model each constituent is associated with one latent variable. Since the constituents of increase represent the effect of experimental treatment and the constituent of constancy is the referent of the effect of non-experimental sources, there are two types of latent variables, which guide the decomposition of variance. They guarantee that the experimental and non-experimental parts of variance result. In assuming one latent variable for each one of the experimental effects in reaction times and error scores and a further latent variable for each one of the non-experimental effects in reaction times and error scores the model included in Fig. 1 is achieved.

It additionally includes reaction times and error scores as manifest variables, which are assigned to the independent part of the model. The dependent part is composed of the latent variable representing intelligence and a measure of intelligence serving as manifest variable.

In small numbers of treatment levels and a restricted set of latent curves the constituents of the polynomial function are useful tools for representing data. Alternatively, the exponential function can be applied for this purpose (Meredith & Tisak, 1990). This function is especially useful in cases where an asymptotic curve underlies performance observed for the treatment levels. Since the different functions show different properties, it is reasonable to select the function with respect to the characteristics of the cognitive processes which need to be considered.

7. Aims of investigation

This paper serves several aims. Firstly, the relationship between working memory and intelligence shall be investigated in considering the impurity problem. This investigation can be expected to reveal whether the observed correlation is only due to the intended source or whether it is the result of several additional sources besides the intended source. Secondly, the decomposition of the variance into parts, which are due to experimental and non-experimental sources, is to be achieved. Although this aim is considered secondary according to its importance, its attainment is even the precondition for the attainment of the first aim. Thirdly, there is the demonstration of the benefits of combining differential and experimental methodologies as aim.

8. Method

8.1. Participants

The data were taken from three independent studies[1]. The sample of the first study included 124 participants (Schweizer & Koch, 2001a), the sample of the second study 104 participants (Schweizer & Koch, 2001b) and the sample of the third study 120 participants (Schweizer & Moosbrugger, 2004). Although the first and second studies were published in the same year, there was no person who participated in both studies. Because of incomplete data three participants were excluded so that the complete sample included 345 participants.

8.2. Measures

8.2.1. The measurement of fluid intelligence

Raven's (1962) Advanced Progressive Matrices (APM) served as the measure of fluid intelligence. Furthermore, this measure even showed a large loading on the *g* factor (Johnson, Bouchard, Krueger, & McGue, 2004).

8.2.2. The measurement of cognitive processes

The Exchange Test (Schweizer, 1996) was applied for measuring working memory. This test mainly stimulated exchange processes and storage processes. The participants had to perform exchanges of neighboring figures mentally until two lists of five simple figures corresponded. Furthermore, the participants had to count the number of exchanges. After the completion of the task the response button had to be pressed. Then the lists were removed from the screen, and the participant was asked to store the number of exchanges.

There were six treatment levels. The first treatment level required one exchange per task, the second treatment level two exchanges, the third and fourth treatment levels three exchanges, the fifth treatment level four exchanges and the sixth treatment level five exchanges. In order to have a monotonic increase of the number of exchanges, one of the two treatment levels requiring the same number of exchanges was eliminated. It was the third treatment level which was selected for elimination. After the elimination the treatment levels were renamed. The remaining levels were addressed as first to fifth levels.

8.3. Statistical analysis

In the first step descriptive statistics were computed for the error scores and reaction times of the treatment levels. In the second step APM, error scores and reaction times were correlated with each other. In the third step the models of measurement for error scores and reaction times constructed as fixed-links models were investigated. The fourth step served the investigation of the relationship between working memory and fluid intelligence by means of the complete fixed-links model. LISREL was applied for this purpose (Jöreskog & Sörbom, 2001).

Three latent variables were considered at the most. The three constituents of the polynomial function (constant, linear, quadratic) were used for the establishment of the links between the latent and manifest variables. The links were standardized so that

$$p\mathbf{I} = \mathrm{diag}(\Lambda'\Lambda)$$

where $\mathbf{\Lambda}$ was the $p \times q$ matrix of loadings, $\mathbf{I}$ the q×q identify matrix and p the number of manifest variables.

[1] The data were taken from studies which were already published because of two reasons: Firstly, a large sample size is necessary for obtaining accurate estimates. It takes a great deal of time to achieve such a sample size by computer-based testing. Secondly, a financial reward is necessary in order to attract motivated participants. Therefore, it would be very costly to have such a large sample.

Table 1
Means and standard deviations obtained by means of exchange test (N=345)

Statistic	Level				
	1	2	3	4	5
Reaction times					
Mean	1767	4776	8054	10929	13296
Standard deviation	701	2209	3575	5326	5206
Error scores					
Mean	0.31	1.31	2.70	4.07	6.00
Standard deviation	1.13	2.06	2.81	3.24	4.04

Standardization was considered necessary in order to obtain variances of the latent variables, which showed an acceptable size. In the absence of standardization the variances, which are estimated, usually differ considerably. The first column of $\mathbf{\Lambda}$ which is associated with the first latent variable leads to an acceptable size of variance whereas the last column yields a very small variance. The small size is due to the large numbers included in the last column of $\mathbf{\Lambda}$ (see example) since the numbers serving as constraints are included in the computation of the variances. Standardization assures that the variances can be compared among each other. However, it does not influence the level of significance.

Since the reaction times were computed on the basis of individual measurements obtained for correct responses, there were cases in which the number of individual measurements was too small to obtain a reliable estimate of the participant's reaction time. As consequence, there were so many missings in the highest treatment level that this level was excluded from correlational analysis and structural equation modeling. *Note*: the covariances of reaction times and error scores on one hand and intelligence on the other hand were recoded in order to obtain positive estimates in structural equation modeling.

9. Results

9.1. The effects of the treatment levels

Table 1 provides the means and standard deviations obtained for the error scores and reaction times.

The means of reaction times showed a monotonic increase from the first to fifth levels. In contrast, the standard deviations only increased from the first to fourth levels. The subsequent decrease presumably was due to the high number of missings resulting from the high proportion of errors in this treatment level. Therefore, only first to fourth level reaction times were included in the further investigations. In the error scores there were monotonic increases of both the means and standard deviations. Since there were twelve trials per treatment level, the mean of the fifth treatment level suggested that on average 50% of the responses were not correct.

9.2. The relationships among APM, reaction times and error scores

The correlations among APM, reaction times and error scores are given in Table 2. The first column provides the correlations of APM with reaction times and error scores. There was a decrease of the size of the correlations with reaction times from the first to fourth levels. In contrast, the sizes of the correlations with error scores increased from the first to fourth levels. On average the correlations with reaction times tended to surmount the correlations with error scores.

Very high correlations characterized the relationships among the reaction times. Most of the correlations among the error scores were of moderate size. The correlations of the pairs of one reaction time and one error score were low or even negative. Apparently, these

Table 2
Correlations of reaction times and error scores obtained by means of exchange test and APM scores (N=345)

Type	Level	APM	Reaction times				Error scores				
			1	2	3	4	1	2	3	4	5
APM		1.00									
RT	1	−.42	1.00								
	2	−.37	.74	1.00							
	3	−.27	.61	.75	1.00						
	4	−.25	.55	.68	.71	1.00					
Error	1	−.14	.11	.00	−.05	−.09	1.00				
	2	−.19	.05	−.04	−.07	−.08	.24	1.00			
	3	−.28	.09	−.02	−.09	−.13	.26	.41	1.00		
	4	−.30	.15	.06	.06	.10	.17	.23	.43	1.00	
	5	−.33	.08	−.04	−.11	−.08	.29	.29	.51	.47	1.00

Table 3
Fit statistics of fixed-links models of measurement for the reaction times

Model characteristic	χ^2	*df*	RMSEA	GFI	CFI	NNFI	AIC	Negative estimate
Models with latent variables linked to all the manifest variables								
Constant [a]	160.96	2	.481	.81	.77	.32	176.96	−67.11 (TD_{11})
Constant/linear [a]	23.93	1	.258	.97	.97	.83	41.93	−213.71 (PH_{11})
Constant/quadratic [a]	0.74	1	.000	1.00	1.00	1.00	18.74	−31.41 (TD_{11})
Constant/linear/quadratic [b]	1.46	1	.036	1.00	1.00	1.00	19.46	−68.25 (TD_{11})
Revised models with free loadings for first reaction time								
Constant	12.74	1	.185	.98	.99	.91	30.74	−741.86 (TD_{11})
Constant/linear [b]	0.47	1	.000	1.00	1.00	1.00	18.47	
Constant/quadratic	0.33	2	.000	1.00	1.00	1.01	16.33	
Constant/linear/Quadratic [c]	0.89	1	.000	1.00	1.00	1.00	18.89	

[a] In this model three correlations between error components were set free.
[b] In this model two correlations between error components were set free.
[c] In this model TD_{11} was fixed to zero in order to avoid a negative estimate.

correlations suggested independence of reaction times and error scores.

9.3. The fixed-links models of measurement

At first, the results referring to reaction times are considered. Models for confirmatory factor analysis, which included latent variables associated with constituents of the polynomial function, were applied to the reaction time data. These models included latent variables according to the following sets of constituents: the constant constituent or the constant and linear constituents or the constant and quadratic constituents or the constant, linear and quadratic constituents. The results are provided in Table 3.

Apparently, all the models with latent variables linked to all the manifest variables, for which the results were presented in the upper half of this Table, proved to be inappropriate. In all these models at least one estimate was negative, as it is obvious from the last column of this Table. Furthermore, in all these models the correlations between three pairs of error components had to be set free in order to achieve the reported degree of fit.

Since the first reaction time was always associated with a negative estimate, it was speculated that information processing due to the first treatment level could not really be compared with information processing due to the other treatment levels. In the first treatment level the result was obvious to the participants without performing an exchange operation. There was a high probability that the participants applied different strategies when performing according to the first treatment level and the other treatment levels. As consequence, it was decided to set the loadings of this reaction time on the corresponding latent variables free while keeping the other loadings constrained.

This decision led to a second set of models including all the combinations of the constituents of the sets described in the beginning of this section (see above). The results for the revised models are presented in the lower half of Table 3. There was again one model which showed an inappropriate estimate (see last column). In a second model (the model with latent variables for constant, linear and quadratic constituents) the negative estimate was eliminated by setting the corresponding parameter to zero. Of the remaining models there was one model which required additional correlations between error components. Only the model which included latent variables of the constant and quadratic constituents remained without restrictions and additionally showed a good degree of fit. Consequently, the best description of the data was achieved by a model which partly separated the first reaction time from the other reaction times.

The error scores were also investigated by means of models for confirmatory factor analysis, which included latent variables associated with constituents of the polynomial function. These models were rather similar to the models applied for investigating the reaction times: Again the models differed according to the constituents giving rise to latent variables: the constant constituent or the constant and linear constituents or the constant and quadratic constituents or the constant, linear and quadratic constituents. In applying all these models it was necessary to allow the error components of the second and third levels and the error components of the fourth and fifth levels to correlate with each other. The results obtained in investigating these models are provided in Table 4.

Only two of the four models led to appropriate results: the models including the latent variable of the constant constituent and the model including latent

Table 4
Fit statistics of fixed-links models of measurement for the error scores

Model[a] characteristic	χ^2	*df*	RMSEA	GFI	CFI	NNFI	AIC	Negative estimate
Constant	187.17	8	.255	.82	.61	.52	201.17	
Constant/linear	22.69	6	.090	.97	.96	.93	40.69	−0.11 (PH_{11})
Constant/quadratic	7.34	6	.025	.99	1.00	.99	25.34	
Constant/linear/quadratic	7.25	5	.036	.99	.99	.99	27.25	−0.29 (PH_{33})

[a] In each one of these model the correlations between two error components were set free.

variables of the constant and quadratic constituents. The other models led to negative estimates and, therefore, had to be excluded. Of the remaining models the model including latent variables of the constant and quadratic constituents showed the better results according to all the fit statistics. Furthermore, it was associated with the lowest AIC.

9.4. The structure of the prediction of fluid intelligence

Both the measurement models obtained for the reaction times and the error scores were combined into an overall model in order to predict fluid intelligence. The correlations of error components of one measurement model were retained (error components of second and third error scores and of the fourth and fifth error scores)[2]. Accordingly, there were four independent latent variables and one dependent latent variable. Each independent latent variable was directly linked to the dependent latent variable. The dependent latent variable was linked to Raven's (1962) Advanced Progressive Matrices as manifest variable. Since the manual of APM suggested a minimum reliability of .83, the corresponding error component was set to .31. Furthermore, it was necessary to set the correlation of the error component of the fourth treatment levels (reaction times and error scores) free. *Note*: in the following paragraphs the latent variables associated with specific constituents are addressed as corresponding constituents in order to facilitate communication.

This model showed a good degree of fit: $\chi^2=63.62$ ($df=32$, $p=.000$), RMSEA = .054, NFI = .95, NNFI = .97, CFI = .98, GFI = .96 and AGFI = .94. The variances of the independent latent variables were considerable (the variance of the constant constituent of reaction time: 312.05, the variance of the quadratic constituent of reaction time: 609.18, the variance of the constant constituent of error score: 0.48, the variance of the quadratic constituent of error score: 4.19). In both cases the variance due to increase surmounted the variance due to constancy. After the standardization[3] of the loadings with respect to variance, the elements of $\mathbf{\Lambda}$ provided interesting insight into the associations of manifest and latent variables. Table 5 includes these elements[4].

The first column refers to the constant constituent associated with the reaction times. Apparently, the first reaction time gave rise to the second to highest number of this column. Since this reaction time excluded repetitions, it could be assumed to be mainly due to the generation of the mental representation, the comparison and the motor response. Because of the special importance of the speed of perceptual processes and of motor response in other investigations the corresponding latent variable was denoted "PerMot Speed". In the second column there was an increase of the numbers from top to bottom. Therefore, the corresponding latent variable could be considered as the expression of cognitive speed associated with the exchange operations. It was denoted "Exchange Speed". The third column included the numbers referring to the constant constituent associated with error. This column showed the highest number with respect to the first error score. Since the corresponding task was not very demanding in the short run but in the long run, it was thought to reflect motivation and basic types of attention. The corresponding latent variable was denoted "Motivation/BasAttention". The numbers of the last column represented the quadratic constituent. Since the high numbers refer to the error scores of the fourth and fifth levels, this column could be assumed to be due to properties of working memory. This gave rise to the name "Working Memory".

[2] Since 36 correlations of errors components are possible, the percentage of error correlations is 5.5 that is only a bit above the chance level.

[3] This type of standardization is a routine procedure of every program for structural equation modeling. The variances of the manifest variables are set to one. This is a provision which assures that the coefficients are restricted to the range between +1 and −1.

[4] Please note that the numbers of this Table cannot be related to the numbers of the matrix provided on page 12 since there are different numbers of rows.

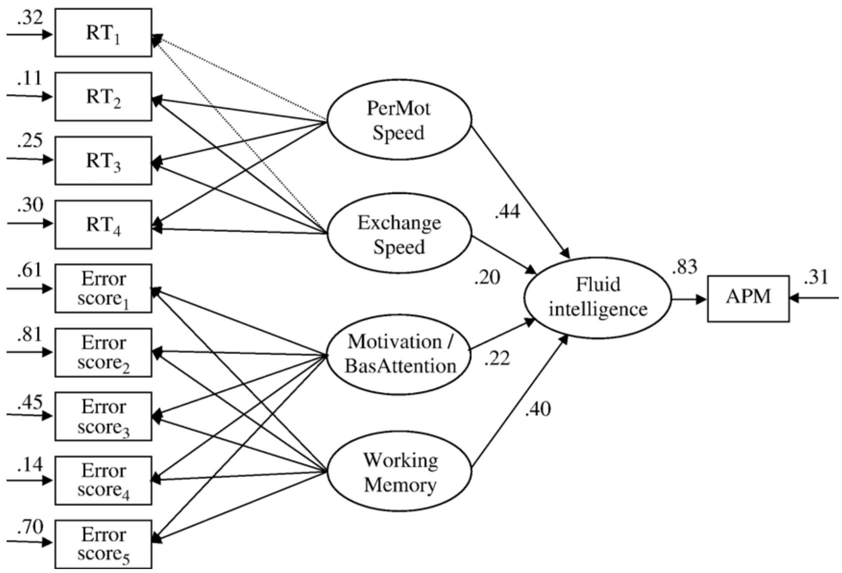

Fig. 2. Graphical representation of the fixed-links model with four independent latent variables and one dependent latent variable (dotted lines represent free links) (Note. The covariances of reaction times and error scores on one hand and intelligence on the other hand were recoded so that positive estimates of the gamma coefficients were obtained).

Fig. 2 provides the graphical representation of the model. This figure includes the estimates of the links between independent and dependent latent variables.

The latent variables of this Figure are termed according to the interpretations of the columns of Table 5 (see previous paragraph). Of special interest were the estimated λs concerning the first reaction time. The first reaction time showed the larger λ (.70) with respect to the constant constituent and the smaller λ (.36) with respect to the quadratic constituent. This relationship among the λs corresponded to expectations. The multiple correlation was .67. Apparently, 45% of the variance of fluid intelligence was explained by latent variables associated with reaction times and error scores.

All the latent variables substantially contributed to fluid intelligence: the latent variable of constant constituent (PerMot Speed) associated with reaction time $-\gamma=.44$ ($t=6.19$, $p<.01$), the latent variable of quadratic constituent (Exchange Speed) associated with reaction time $-\gamma=.20$ ($t=2.87$, $p<.01$), the latent variable of constant constituent (Motivation/BasAttention) associated with error score $-\gamma=.22$ ($t=2.33$, $p<.01$), and the latent variable of quadratic constituent (Working Memory) associated with error score $-\gamma=.40$ ($t=5.85$, $p<.01$). In the previous paragraph it is indicated that the four independent latent variables account for 45% of the variance of the dependent latent variable. In order to facilitate further reasoning, the amount of accounted variance is set to 100 (=45%). The two latent variables associated with experimental sources predict 45% of the accounted variance, and the two latent variables associated with non-experimental sources predict the remaining 55%. It was interesting to observe that neither the correlation between the latent variables associated with experimental sources ($r=-.11$, $t=-1.64$, n.s.) nor the correlation between the latent variables associated with non-experimental sources ($r=.17$, $t=1.69$, n.s.) reached the level of significance. Apparently, the latent variables of reaction times and error scores were independent of each other. This observation suggested that the errors were mainly due to capacity limitation and not due to time limitation.

Table 5
Standardized lambda coefficients of reaction times and error scores

Manifest variable	Latent variable			
	1	2	3	4
First reaction time	.70	.36	–	–
Second reaction time	.81	.48	–	–
Third reaction time	.50	.67	–	–
Fourth reaction time	.34	.80	–	–
First error score	–	–	.61	.13
Second error score	–	–	.33	.28
Third error score	–	–	.25	.48
Fourth error score	–	–	.21	.71
Fifth error score	–	–	.17	.91

10. Discussion

In the introductory section the impurity of measures was presented as possible cause of the extremely high correlations between measures of working memory and measures of intelligence. It gave rise to the present investigation based on data that were obtained by means of Exchange Test and Advanced Progressive Matrices. The results suggest that impurity is really a problem. There is approximately one half of variance, which is predicted by the Exchange Test and which can be ascribed to working memory, whereas the other half of variance does not originate from working memory although the Exchange Test also accounts for this half. The effect of impurity is more severe in reaction times than in error scores. All these observations signify that the impurity of measures is a problem, which should be taken seriously.

It may be argued that all these findings do not prove that the interpretations of the correlations addressed in the introductory section are really invalid because they are due to impurity. Although this argument is correct, it does not preclude the danger of erroneous interpretations. Therefore, precautions should be taken in order to avoid erroneous interpretations of results because of impurity. Furthermore, it may be objected that the Exchange Test is insufficient for representing such a complex structure like working memory. However, this latter argument does not count in the light of the finding that working memory is domain-general (Kane et al., 2004).

The insights provided by this investigation into the relationship between working memory and fluid intelligence are mainly due to the fixed-links model. The benefits of this model become especially obvious when it is compared with the conventional model. The fixed-links model enables the decomposition of variance with respect to the experimental treatments. Therefore, it needs to be highlighted that the constraints according to the experimental treatments in the fixed-links model provide the basis for substantiated interpretations of the latent variables. In contrast, in the conventional model the decomposition of variance is mainly the separation of true variance from error variance. Even if further decomposition is introduced there is no possibility to guide the decomposition *exactly* according to the experimental treatments.

The inclusion of linear and quadratic constituents into the matrix of loading may convey the impression that the statistical power of the model is low because of dependence between the constituents. In assuming that the constituents are moments, it may be argued that the quality of the higher-order moments depends on the quality of the lower-order moments. In this case errors can be assumed to be transferred from one level to the next level. Fortunately, the constituents can by no means be perceived as moments. They do not depend on each other, and they are estimated independently. For example, it is possible to have a quadratic constituent without having a constant or a linear constituent. Therefore, the inclusion of linear and quadratic constituents into the model does not impair the statistical power.

The constraint of all the loadings may give rise to concerns regarding the replicability of the results. These concerns are redundant since the constraints are not due to capitalization of chance. There were only a few alternatives which were constructed in a systematic way. Furthermore, the observation that the combination of constant and quadratic constituents serves best in reaction time data is not a singular result. A similar result was already achieved in choice reaction times (Schweizer, 2006a). A composite of constant and quadratic constituents served better than other types of composites in investigating choice reaction times.

Constraints can be interpreted in the same way as loadings. High constraints suggest a close association of the corresponding manifest and latent variables and low constraints a weak association. The relationships among the constraints are important. For example, in considering the quadratic constituent it becomes obvious that the corresponding latent variable is closely associated with a very few manifest variables only. In contrast, the consideration of the constant constituent reveals equal associations with all the manifest variables. Problems may arise when taking the perspective of the manifest variable. In this case it is recommendable to use standardized constraints in order to achieve a reasonable interpretation.

The results observed for the reaction times suggest three things. Firstly, the large non-experimental effect associated with reaction times can be regarded as evidence for the mental-speed approach (Neubauer, 1997). Obviously, there is a source of individual differences, which considerably contributes to fluid intelligence and is independent of exchange and storage processes. Secondly, the substantial contribution of the quadratic constituent makes obvious that repetition does not lead to the gradual merging of processes. Processing seems to show serial instead of parallel characteristics (Meyer, Yantis, Osman, & Smith, 1985). Apparently, processing due to the task demands is mainly controlled processing. Thirdly, the investigation reveals that the first reaction time differs from the other reaction times. This reaction time is not due to the whole set of processes leading to the other reaction times. Presumably, the participants do not really complete all the exchanges which are required by the tasks of the other treatment levels. After having noticed that two neighboring figures need to be exchanged the participants disrupt processing since they already know the answer to the question which is asked afterwards (What is the number of exchanges?). In this case performance is mainly due to perceptual processing (see Deary & Stough, 1996; Schweizer & Koch, 2003).

The increase of the error scores is well described by the curve resulting from the constant and quadratic constituents of the polynomial function. In the standardized solution it becomes apparent that the error score of the first treatment level is by far the best marker of the latent variable due to the constant constituent. Therefore, this latent variable receives its meaning mainly from this error score, which is presumably due to the ability to process routine tasks with high precision. Such continuous performance attention seems to correlate only moderately with intelligence (Rockstroh & Schweizer, 2001, 2004). In contrast, the quadratic constituent is likely to represent storage capacity since it receives the highest loadings from the errors scores of the upper levels. The quadratic constituent does presumably not reflect the exchange operations because of their simplicity. This observation supports the assumption of limited storage capacity as source of error (Schweizer, 2000, 2001). In this case errors resulting from the decay of information because of lack of time or of insufficient speed are not very likely (Jensen, 1982; Salthouse, 1993).

The complexity of the structure of the relationship between the independent and dependent latent variables is an exciting observation since it suggests a diversity of sources of fluid intelligence, which are associated with the demands of a rather simple cognitive test. Similar complexity can be observed in considering several other cognitive tests (Neubauer & Fink, 2003; Schweizer, 1998). The constant constituent of the reaction times and the quadratic constituent of the error scores are strong predictors of fluid intelligence whereas the quadratic constituent of the reaction time and the constant constituent of the error scores are weak predictors. This means that speed of routine perceptual and motor processing and the limitation of capacity are rather important sources of fluid intelligence and presumably also of general intelligence.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30−60.

Baddley, A. D. (1986). *Working memory.* Oxford: Clarendon Press.

Bayliss, D. M., Jarrold, C., Gunn, D., & Baddeley, A. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology. General*, *132*, 71−92.

Bentler, P. M., & Yuan, K. -H. (2000). On adding a mean structure to a covariance structure model. *Educational and Psychological Measurement*, *60*, 326−339.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Buehner, M., Krumm, S., & Pick, M. (2005). Reasoning = working memory ≠ attention. *Intelligence*, *33*, 251−272.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin*, *56*, 81−105.

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.

Colom, R., Abad, F. J., Rebollo, I., & Shih, P. (2005). Memory span and general intelligence: A latent-variable approach. *Intelligence*, *33*, 623−642.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence*, *32*, 277−296.

Colom, R., & Shih, P. (2004). Is working memory fractionated onto different components of intelligence? *Intelligence*, *32*, 431−444.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *30*, 116−127.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change" — or should we. *Psychological Bulletin*, *74*, 68−80.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time. Achievements, prospects, and problems. *American Psychologist*, *51*, 599−608.

Dolan, C. V., Colom, R., Abad, F. J., Wicherts, J. M., Hessen, D. J., & van de Sluis, S. (2006). Multi-group covariance and mean structure modeling of the relationship between the WAIS-III common factors and sex and educational attainment in Spain. *Intelligence*, *34*, 193−210.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology. General*, *128*, 309−331.

Harris, R. J. (2003). Traditional nomothetic approaches. In S. F. Davis (Ed.), *Handbook of research methods in experimental psychology* (pp. 41−65). Oxford: Blackwell.

Helmbold, N., & Rammsayer, T. (2006). Timing performance as a predictor of psychometric intelligence as measured by speed and power tests. *Journal of Individual Differences*, *27*, 20−37.

Jensen, A. R. (1982). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in research on intelligence*, *vol. 1* (pp. 93−132). 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. 103−175). Norwood, NJ: Ablex.

Jensen, A. R., & Munro, E. (1979). Reaction time, movement time and intelligence. *Intelligence*, *3*, 121−126.

Jöreskog, K. G. (1973). A general method for estimating the linear structural equation system. In A. S. Goldberg & O. D. Duncan (Eds.), *Structural Equation Models in the Social Sciences.* New York: Seminar Press.

Jöreskog, K. G., & Sörbom, D. (2001). LISREL (Version 8.5). Computer Software. Chicago: Scientific Software International.

Johnson, W., Bouchard, T. J., Krueger, R. F., & McGue, M. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, *32*, 95−107.

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 visuospatial memory span and reasoning. *Journal of Experimental Psychology. General*, *133*, 189−217.

Keesling, W. (1972). Maximum likelihood approaches to causal flow analysis. Ph.D. thesis, University of Chicago.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait–multimethod matrix by confirmatory factor analysis. *Psychological Bulletin*, *112*, 165−172.

Kline, R. B. (2006). *Structural equation modeling*, 2nd ed. New York: the Guilford Press.

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

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* London: Addison-Wesley.

McArdle, J. J. (1988). Growth curve analysis in contemporary psychological research. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 561−614). 2nd ed. New York: Plenum.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, *58*, 110−133.

Meredith, W., & Tisak, J. (1984, June). Tuckerizing curves. *Paper presented at the annual meeting of the Psychometric Society, Santa Barbara, CA*.

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

Meyer, D. E., Yantis, S., Osman, A., & Smith, J. E. K. (1985). Temporal properties of human information processing: Tests of discrete versus continuous models. *Cognitive Psychology*, *17*, 445−518.

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.

Neisser, U. (1963). Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology*, *76*, 376−385.

Neubauer, A. C. (1997). The mental speed approach to the assessment of intelligence. *Advances in Cognition and Educational Practice*, *4*, 149−173.

Neubauer, A. C., & Fink, A. (2003). Fluid intelligence and neural efficiency: Effects of task complexity and sex. *Personality and Individual Differences*, *35*, 811−827.

Neubauer, A. C., Stern, E., & Grabner, R. H. (2005). The neurophysiological correlates of human intelligence: Some recent qualifications of the neural efficiency hypothesis. *Paper presented at the 8th European Conference of Psychological Assessment, Budapest*.

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, *3*, 1−18.

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

Rockstroh, S., & Schweizer, K. (2001). The contribution of memory and attention processes to cognitive abilities. *Journal of General Psychology*, *128*, 30−42.

Rockstroh, S., & Schweizer, K. (2004). The effect of retest-practice on the speed–ability relationship. *European Psychologist*, *9*, 24−31.

Salthouse, T. A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology*, *84*, 171−199.

Salthouse, T. A. (2005). Developmental influences on adult intelligence: The Seattle Longitudinal Study. *Intelligence*, *33*, 551−554.

Schweizer, K. (1996). The speed–accuracy transition due to task complexity. *Intelligence*, *22*, 115−128.

Schweizer, K. (1998). Complexity of information processing and the speed–ability relationship. *Journal of General Psychology*, *125*, 89−102.

Schweizer, K. (2000). Cognitive mechanisms as sources of success and failure in intelligence testing. *Psychologische Beiträge*, *42*, 190−200.

Schweizer, K. (2001). On the role of mechanisms when the processing complexity is high. *European Psychologist*, *6*, 133−143.

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

Schweizer, K. (2006). The fixed-links model in combination with the polynomial function as a tool for investigating choice reaction time data. *Structural Equation Modeling*, *13*, 403−419.

Schweizer, K. (2006). The fixed-links model for investigating the effects of general and specific sources. *Methodology. European Journal of Research Methods for the Behavioral and Social Sciences*, *2*, 149−160.

Schweizer, K., & Koch, W. (2001). Kapazitätslimitierung und intellektuelle Leistungsfähigkeit [Capacity limitation and intellectual ability]. *Zeitschrift für Experimentelle Psychologie*, *48*, 1−19.

Schweizer, K., & Koch, W. (2001). A revision of Cattell's Investment Theory: Cognitive properties influencing learning. *Learning and Individual Differences*, *13*, 57−82.

Schweizer, K., & Koch, W. (2003). Perceptual processes and cognitive ability. *Intelligence*, *31*, 211−235.

Schweizer, K., & Moosbrugger, H. (2004). Attention and working memory as predictors of intelligence. *Intelligence*, *32*, 329−347.

Schweizer, K., Moosbrugger, H., & Goldhammer, F. (2005). The structure of the relationship between attention and intelligence. *Intelligence*, *6*, 589−611.

Stankov, L. (2000). Complexity, metacognition and fluid intelligence. *Intelligence*, *28*, 121−143.

Stankov, L., Birney, D., & Bowman, D. (2005). Fluid intelligence and test's complexity/difficulty: New evidence. *Paper presented at the 8th European Conference of Psychological Assessment, Budapest*.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652−654.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, *30*, 276−315.

Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: examining the correlation between operation span and Raven. *Intelligence*, *33*, 67−81.

Van Zomeren, A. H., & Brouwer, W. H. (1994). *Clinical neuropsychology of attention.* New York, NY: Oxford University Press.

Wiley, D. E. (1973). The identification problem for structural equation models with unmeasured variables. In A. S. Goldberg & O.D. Duncan (Eds.), *Structural Equation Models in the Social Sciences.* New York: Seminar Press.