



## Reasoning=working memory≠attention

Markus Buehner<sup>a,\*</sup>, Stefan Krumm<sup>b</sup>, Marion Pick<sup>b</sup>

<sup>a</sup>*Department of Psychology, Ludwig-Maximilians-University, Munich, Leopoldstrasse 13, 80802 Munich*

<sup>b</sup>*Department of Psychology, Philipps-University, Marburg, Gutenbergstrasse 18, 35032 Marburg, Germany*

Available online 2 March 2005

---

### Abstract

The purpose of this study was to clarify the relationship between attention, components of working memory, and reasoning. Therefore, twenty working memory tests, two attention tests, and nine intelligence subtests were administered to 135 students. Using structural equation modeling, we were able to replicate a functional model of working memory proposed by Oberauer, Suess, Wilhelm, and Wittmann (2003) [Oberauer, K., Suess, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, 31, 167–193]. The study also revealed a weak to moderate relationship between the “selectivity aspect of attention” and working memory components as well as the finding that “supervision” was only moderately related to “storage in the context of processing” and to “coordination”. No significant path was found from attention to reasoning. Reasoning could be significantly predicted by “storage in the context of processing” and “coordination”. All in all, 95% of reasoning variance could be explained. Controlling for speed variance, the correlation between working memory components and intelligence did not decrease significantly.

© 2005 Elsevier Inc. All rights reserved.

*Keywords:* Intelligence; Reasoning; Working memory; Attention

---

### 1. Introduction

The term “working memory” was first used by Miller, Galanter, and Pribram (1960; cited according to Baddeley, 2001) in their book “Plans and the structure of behavior”. During the following years the

---

\* Corresponding author. Tel.: +49 89 21806257; fax: +49 89 2180523000.

E-mail address: [buehner@edu.uni-muenchen.de](mailto:buehner@edu.uni-muenchen.de) (M. Buehner).

working memory concept became more and more popular, especially in the prominent model of [Baddeley and Hitch \(1974\)](#). Since then, the working memory concept is one of the most important constructs in cognitive psychology (see [Miyake & Shah, 1999](#)).

### 1.1. Models of working memory

[Oberauer, Suess, Wilhelm, and Wittmann \(2000, 2003\)](#) assumed that working memory can be separated into two facets: a content facet and a functional facet. The content facet contains verbal/numerical material and figural/spatial material (see [Baddeley, 1986](#)), whereas the functional facet is separated into the components “storage in the context of processing,” “coordination,” and “supervision”.

[Oberauer et al. \(2003, p. 169\)](#) describe processing as “the transformation of information or the derivation of new information” and “storage in the context of processing” as “retention of briefly presented new information over a period of time in which the information is no longer present”. A characteristic task for “storage in the context of processing” is a dual task, where participants have to remember words, then perform another task and finally recall the remembered words (see [Fig. 1](#)). This factor is similar to the factors “updating” and “working memory capacity” by [Miyake, Friedman, Emerson, Witzki, Howerter, and Wager \(2000\)](#) and by [Engle, Tuholski, Laughlin, and Conway \(1999\)](#).

“Coordination” is “the ability to build new relations between elements and to integrate relations into structures” ([Oberauer et al., 2003, p. 169](#)). One task applied by [Oberauer et al. \(2003\)](#) to assess “coordination” is flight control, where the participants have to monitor moving objects (planes) and ensure that no crashes occurred (see [Fig. 2](#)). To avoid crashes participants had to stop planes by pressing the space bar, then pick one plane by using the left mouse button, and finally change the direction of this plane with the arrow keys.

According to [Oberauer et al. \(2003, p. 169\)](#), “supervision” involves “the monitoring of ongoing cognitive processes and actions, the selective activation of relevant representations and procedures, and the suppression of irrelevant, distracting ones”. A task representing supervision is “switching numerical”. Numbers appear in a clockwise order in one of four cells of a  $2 \times 2$  matrix (see [Fig. 3](#)). Participants have to switch from one decision rule (odd versus even) to the other (above 500 versus below 500). The “supervision factor” of [Oberauer et al. \(2000, 2003\)](#) is related to the factor “task set switching” in [Miyake et al. \(2000\)](#). The latter function is similar to the central executive in the working memory model of [Baddeley and Hitch \(1974\)](#) and similar to the definition of selective attention.

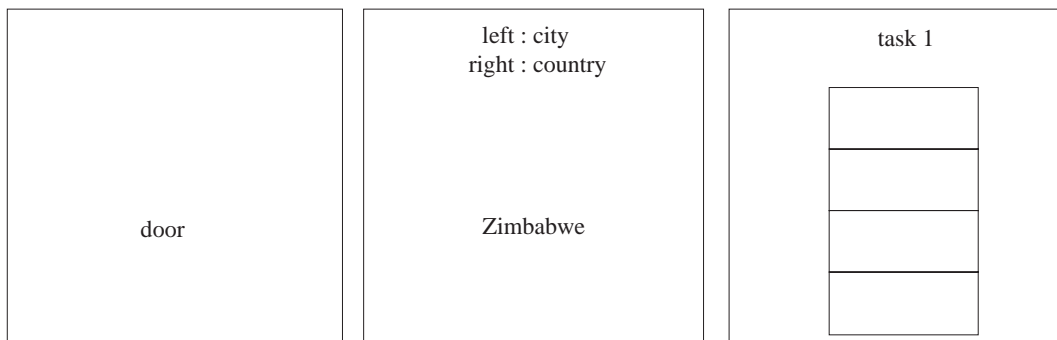


Fig. 1. Verbal dual task: first a sequence of nouns is presented, then a series of choice-reaction time task has to be performed, and finally the nouns have to be remembered in the correct sequence.

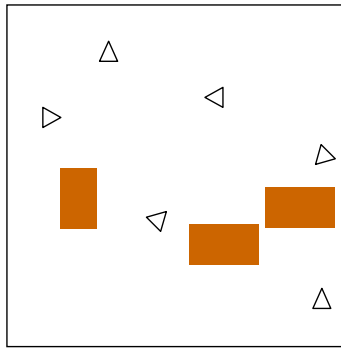


Fig. 2. Flight control: participants had to actively avoid crashes by stopping plane movement and redirecting planes.

Oberauer et al. (2000) confirmed a facet model using 23 working memory tests mostly taken from previous literature. However, in this study the separation of “storage in the context of processing” and “coordination” was not successful. In a more recent study, Oberauer et al. (2003) were able to distinguish between the functions of “coordination” and “storage in the context of processing”. They used 20 new distinct tasks for the assessment of the postulated facets. The study revealed high correlations between the factors “storage in the context of processing” and “coordination” ( $r=0.78$  to  $0.80$ ). It seemed to be difficult to clearly separate these two factors. The correlation between “supervision” and the other working memory factors have only been small, indicating that “supervision” was not a central component of working memory.

1.2. Intelligence and working memory

Many Studies have been conducted to explore the relationship between working memory and intelligence. Kyllonen and Christal (1990) found a factor correlation between a subset of working memory tasks and reasoning of  $r=0.74$  and  $r=0.93$  in different studies. According to Wittmann and Suess (1999), the only drawback of their study was that some of their working memory tasks were conceptually highly similar to tasks used in psychometric intelligence research. Wittmann and Suess (1999) revealed that “g” and working memory shared 53.5% of common variance. Intelligence was measured with the Berlin–Structure–Intelligence Test (BIS, Jaeger, Suess, & Beauducel, 1997).

<table border="1"> <tr> <td>left : odd right : even</td> <td>left : odd right : even</td> </tr> <tr> <td>812</td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td>left : high right : low</td> <td>left : high right : low</td> </tr> </table>	left : odd right : even	left : odd right : even	812				left : high right : low	left : high right : low	<table border="1"> <tr> <td>left : odd right : even</td> <td>left : odd right : even</td> </tr> <tr> <td></td> <td>433</td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td>left : high right : low</td> <td>left : high right : low</td> </tr> </table>	left : odd right : even	left : odd right : even		433			left : high right : low	left : high right : low	<table border="1"> <tr> <td>left : odd right : even</td> <td>left : odd right : even</td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td>157</td> </tr> <tr> <td>left : high right : low</td> <td>left : high right : low</td> </tr> </table>	left : odd right : even	left : odd right : even				157	left : high right : low	left : high right : low	<table border="1"> <tr> <td>left : odd right : even</td> <td>left : odd right : even</td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td>652</td> <td></td> </tr> <tr> <td>left : high right : low</td> <td>left : high right : low</td> </tr> </table>	left : odd right : even	left : odd right : even			652		left : high right : low	left : high right : low
left : odd right : even	left : odd right : even																																		
812																																			
left : high right : low	left : high right : low																																		
left : odd right : even	left : odd right : even																																		
	433																																		
left : high right : low	left : high right : low																																		
left : odd right : even	left : odd right : even																																		
	157																																		
left : high right : low	left : high right : low																																		
left : odd right : even	left : odd right : even																																		
652																																			
left : high right : low	left : high right : low																																		

Fig. 3. Switching numerical: a complete trial with 4 tasks is presented; at the first and the second task the decision rule odd versus even had to be applied, whereas at the following two tasks the decision rule high (above 500) versus low (below 500) had to be applied.

Wittmann and Suess (1999) also found evidence for differential validity of three working memory components predicting reasoning (BIS-K): “supervision”:  $a=0.22$ ; “verbal/numerical working memory”:  $a=0.55$ ; and “spatial working memory”:  $a=0.47$ . In their study, they used working memory tasks, which were quite different to reasoning tasks. Colom, Flores-Mendoza, and Rebollo (2003) found a relationship between factor scores of working memory and intelligence ranging from  $r=0.69$  to  $r=0.71$  measured with the SPM by Raven (CEPA, 1993) and the letter series of the Primary Mental Ability Test (Thurstone, 1938). In a more recent study Engle et al. (1999) found a high loading ( $a=0.59$ ) of working memory on fluid intelligence measured with the CFT (Cattell, 1973) and the SPM (Raven, Court, & Raven, 1977). This result was confirmed by Conway, Cowan, Bunting, Theriault, and Minkoff (2002). Colom, Rebollo, Palacios, Juan-Espinosa, and Kyllonen (2004) found that working memory is almost perfectly predicted by “g”. This finding is extensively unique and in contrast to Ackerman, Beier, and Boyle (2002) as well as in contrast to other studies (Conway et al., 2002; Schweizer & Moosbrugger, 2004; Suess, Oberauer, Wittmann, Wilhelm, & Schulze, 2002). Ackerman et al. (2002, p. 569) summarize: “In sum, the general finding is that WM performance is positively and significantly related to tasks of reasoning or fluid intelligence but neither correlations nor path coefficients are of the magnitude (i.e., 0.80 s) reported by Kyllonen and Christal (1990)”. So the crucial question remains: Is working memory really perfectly predicted by “g”? Besides the magnitude of the coefficients another question is of importance and remained unaddressed: Which components of a differentiated working memory model predict reasoning?

Suess et al. (2002, p. 262) summarized crucial shortcomings of previous working memory research: (1) working memory was often measured using a small unsystematic set of tasks; (2) working memory capacity was conceptualised as an undifferentiated construct, although most theories of working memory contain different facets; (3) studies relating working memory to intelligence constructs usually focused on a single mental ability; and (4) some working memory tasks were indistinguishable from common reasoning tasks.

To overcome these shortcomings, Suess et al. (2002)<sup>1</sup> conducted an analysis with carefully selected test. The authors used a more differentiated model of working memory similar to Oberauer et al. (2000). They found evidence that working memory and “g” measured with the BIS (Jaeger et al., 1997) are strongly related. Oberauer et al. (2003) provided a theoretically well-founded model of working memory. This model allows the separation of “storage in the context of processing” and “coordination”. Therefore, this model was chosen for the current study to evaluate the impact of working memory facets on reasoning ability.

### 1.3. Attention and reasoning

Wittmann and Suess (1999) stated that, beyond working memory, speed could be a promising predictor of reasoning. Speed is a very broad construct containing different factors (see Carroll, 1993). Jensen (1980) and Eysenck (1982) speculated about the relationship between speed and intelligence. They assumed that speed or efficiency of neural transmission in the brain affects performance on elementary cognitive tasks as well as on reasoning tasks. A construct, which is related to speed, is attention. Schweizer (1995) mentioned that divided attention is also a candidate for predicting reasoning. Schweizer, Zimmermann, and Koch (2000) pointed out that moderate correlations between

<sup>1</sup> Suess et al. (2002) and Wittmann and Suess (1999) used the same data set.

divided attention tasks and intelligence were found in literature (cited according to Necka, 1996; Roberts, Beh, & Stankov, 1988; Roberts, Beh, Spilsbury, & Stankov, 1991; Schweizer et al., 2000; Stankov, Roberts, & Spilsbury, 1994) whenever the task demands were higher. They have shown that several measures of attention all requiring the selection of relevant stimuli correlated from  $r=0.21$  to 0.58 with intelligence.

#### *1.4. Models of attention*

Despite of the large amount of previous research, it is difficult to find a commonly accepted model of attention. Some factor analytic models could not be replicated (Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991; O'Donnell, MacGregor, Dabrowski, Oestreicher, & Romero, 1994; Strauss, Thompson, Adams, Redline, & Burant, 2000) or extracted factors did not correspond to models of cognitive psychology (see model of Shum, McFarland, & Bain, 1990). Especially the use of complex attention tests was criticized (Kremen, Seidman, Faraone, Pepple, & Tsuang, 1992). The authors of the Test Battery for Attentional Performance (TAP, Zimmermann & Fimm, 2002) intended to measure several attention concepts with simple and specific tests. According to Sturm (2002) as well as Van Zomeren and Brouwer (1994) these tests cover the “selectivity aspect of attention”.

To overcome the shortcomings mentioned above and to evaluate the connection between intelligence and working memory, the facet model postulated by Oberauer et al. (2003) was adopted in this study. This model is highly discriminating and has strong theoretical background. To measure the “selectivity aspect of attention” proposed by van Zomeren and Brouwer (1994), TAP tests were applied. The classification of attention by Van Zomeren and Brouwer (1994) converged with commonly accepted concepts of attention. The TAP tests provide specific tests to assess this classification. To assess reasoning ability, the Intelligence Structure Test 2000 R (I-S-T 2000 R) by Amthauer, Brocke, Liepmann, and Beauducel (2001) was administered.

#### *1.5. Goal of the study*

The main goal of the study was to clarify the relationship between working memory, attention, and reasoning. Therefore, as a necessary first step, a replication of the factor structure of working memory proposed by Oberauer et al. (2003) was conducted. In a second step the model of Oberauer et al. (2003) was extended by the factor “selectivity aspect of attention” in order to clarify the relationship between attention and reasoning.

## **2. Method**

### *2.1. Participants*

The test battery was administered to 135 students of the Philipps-University, Marburg. Six participants did not complete all the administered tests and were therefore excluded. To rule out the possibility that outliers distorted the relationships between the tests, participants, whose results were three standard deviations above or below the mean score of any test, were removed from the data set, resulting in 5 excluded participants. Statistical analyses were conducted with 124 students. Participation

was voluntary. Afterwards, students received a performance feedback as well as credit points for their participation. The mean age was 22 years ( $SD=3.0$ ,  $Range=18-31$ ), and 75.8% of the students were female. Subjects could be categorized into the following academic units: psychology 62.1%, education 10.5%, economics 10.5%, and others 16.9%. The average amount of academic training measured in semesters was 3.6 semesters. ( $SD=3.26$ ,  $Range=1-16$ ).

## 2.2. Measures

The selected working memory tests – provided by the test battery programmed by Oberauer et al. (2003) – covered the full working memory model of Oberauer et al. (2003). Two different TAP tests were conducted to measure the “selectivity aspect of attention” according to Sturm (2002): Go/No-Go and the Divided Attention Test. To measure reasoning ability, the Intelligence Structure Test 2000-R (Amthauer et al., 2001) was administered. All these tests were presented in a computer-based format. The test trials of working memory and attention tasks were not started until the participants had solved all practice trials correctly.

## 2.3. Working memory tests

### 2.3.1. Processing tasks

Eight binary choice reaction time tasks (CRTs) were conducted. For each two CRTs, the same type of stimulus material was used, but for each task the decision criteria differed. The stimulus appeared in a rectangular frame in the middle of the screen. Participants were asked to respond as quickly and correctly as possible by pressing the keys labelled “right” or “left” on the keyboard. The tasks were organized into one practice block of 15 trials and five test blocks of 16 trials each. Feedback was given after each block. The pseudorandom sequence of stimuli was the same for all participants. Nouns were used as stimuli for the verbal CRTs. CRT categories required a distinction between animal versus plant terms, CRT syllables a distinction between one versus two syllables. Numerical CRTs consisted of three-digit numbers. The odd–even CRTs required a decision whether the presented number was odd or even. In large–small CRTs, a quick decision had to be made, whether the presented number was above or below 500. Arrows were applied as stimuli at the first two spatial CRTs. The arrows’ locations varied within a frame and pointed into different directions. Participants had to react to the different directions of the arrows in the CRT up–down task; in the CRT above–below task they had to react to the location of arrows within the frame (upper or lower half of the frame). Within the other two spatial tasks,  $3 \times 3$  matrices were presented which were partially filled. In the CRT symmetry task, the participants had to decide whether the matrices were symmetrical or not, in the CRT parts task, whether the matrices consisted of either one or two separated parts.

To build scores, false responses were eliminated (the mean percentage of errors within one CRT task varied from 4.0 to 8.62%) as well as reaction times below 200 ms and times exceeding the individual’s mean by three standard deviations. Then the log-transformed reaction times were aggregated within blocks.

### 2.3.2. Storage in the context of processing tasks

The “storage in the context of processing” component of the working memory model was assessed by dual tasks. One processing and one storage task were combined for each trial. The procedure was as follows: First, the materials to be remembered were presented one immediately after another (1-s inter-



stimulus interval). Second, participants had to perform a series of CRTs described above, which were unrelated to the material to be remembered. The CRTs lasted for 5 s (no matter how many trials the participants had performed within this time) to keep the time between learning and recall constant and to measure the recall independent of the processing speed. Finally, the participants were asked to recall the memory set (see Fig. 1).

The materials to be remembered were either nouns, digits, patterns ( $3 \times 3$  matrix, partially filled), or spatial locations of dots. The stimuli always had to be recalled in the correct order. For dual tasks with verbal material nouns had to be recalled and, in between, CRT categories task had to be performed. The number of nouns to be remembered increased from 3 to 7. Numerical dual tasks combined CRT odd–even tasks and a series of digits to be remembered. Three items were administered for each memory load, whereas memory loads varied from 4 to 8 digits. Also, two spatial dual tasks were applied. The first one combined CRT pattern symmetry with a task where the spatial location of dots presented (within a rectangle frame) had to be remembered. In the course of the second spatial task, participants had to remember several partially filled  $3 \times 3$  patterns and perform CRT arrows up–down tasks. The spatial dual tasks consisted of memory loads varying from 2 to 4, each level represented by five items.

Two scores were obtained from these dual tasks: the number of elements correctly remembered (memory performance) and the log-transformed reaction times for the CRT subtasks. Since the correlations between these two subtask scores were low, and since it is common practice to evaluate storage and processing tasks according to memory performance only (e.g., Daneman & Carpenter, 1980), the analyses were based only on the dual tasks' memory scores.

### 2.3.3. Coordination tasks

The “coordination” component of the working memory model was measured by monitoring tasks. Changing relations between several independently changing objects had to be monitored. Participants were instructed to detect certain critical relations. In order to compute and to continuously update the relations between the objects, simultaneous access to them was required.<sup>2</sup>

The verbal monitoring task consisted of a  $3 \times 3$  matrix with a word in each of the nine cells. One randomly chosen word was replaced every 2 s. The space bar had to be pressed whenever three rhyming words were presented in either the horizontal, vertical, or diagonal line. During one trial, 2 to 5 target rows appeared within 10 to 20 replacements. In the numerical monitoring task, three-digit numbers were presented in each of the 9 cells. Rows with equal last digits had to be detected. One randomly chosen number changed every 1.5 s. After each trial, feedback about hits, misses, and false alarms was presented. Scores were obtained by subtracting false alarms from hits.

“Flight control” was the first spatial monitoring task. A number of airplanes (ranging from 5 to 9 during the 15 items) represented by triangles moved across the screen in various directions with 4 different speeds. Mountains (clusters of brown squares) were located on the screen. Unpredictably, airplanes appeared on the border of the screen. Their flight direction maintained the same until they left the screen. The instruction was to monitor that no plane crashed either with another plane or a mountain. Plane movement could be stopped by pressing the space bar, then one plane had to be chosen by mouse click and redirected. Traffic started again after pressing the space bar. The participants were told that they started with 100 credit points at each trial. Each crash would cost 10 points and each movement stop 3 points. The goal was to avoid

---

<sup>2</sup> Only the no-memory versions of the coordination task were applied. Oberauer et al. (2003) used both no-memory and memory coordination tasks.

crashes and to stop the planes as seldom and as briefly as possible. Duration of movement stops was also measured. Without interruption each trial lasted about 12 s. Feedback was given after each trial regarding the number of crashes, the remaining points, and the cumulative duration of movement stops. Scores were obtained by counting the number of crashes (see Fig. 2).

“Finding squares,” the second spatial coordination task, consisted of 8 to 12 red dots randomly located within a  $10 \times 10$  matrix. Two randomly chosen dots changed their position every 1.5 s. Twenty items were presented. Participants had to press the space bar whenever four dots formed a square. Position and size of the square were not relevant. Scores were obtained by subtracting false alarms from hits.

#### 2.3.4. *Supervision tasks*

Supervision was measured by combining two CRT tasks using the same kind of stimulus material as described above. The stimuli (words, numbers, arrows, patterns) appeared in a clockwise order in one of four cells of a  $2 \times 2$  matrix. Participants had to switch from one decision rule (e.g., plant versus animal) in the upper two cells to the second decision rule (one syllable versus two syllables) in the lower two cells. About 50% switch trials and about 50% no-switch trials were provided by this design. There was one practice block and six test blocks with 16 trials each for the four switching tasks (see Fig. 3).

Two indicators of executive processes were derived from the switching tasks and the CRTs: specific switching costs (see model 1 described below) and general switching costs (see models 2 and 2a described below). *Specific switching costs* were derived from the supervision task only. As can be seen in Fig. 3 no switching of decision rules is required between reacting to the stimulus in the upper left square and the stimulus in the upper right square. The same holds true for the two lower squares. In contrast to this, switching is required between reacting to the stimulus in the upper right square and the lower right square. The same holds true for the lower left and the upper left square. Specific switching costs were defined as the difference between the log-transformed switching and no-switching reaction times within the switching task. To obtain *general switching costs* only no-switching reaction times of the supervision tasks were applied. Additionally, reaction times of the relevant basic processing tasks (CRTs) were obtained. General switching costs were defined as the difference between log-transformed no-switching RTs and baseline RTs from the two single CRTs. Again, false responses (the mean percentage of errors within one switching task varied from 3.16 to 3.87%) as well as reaction times below 200 ms and times exceeding the individual's mean by three standard deviations were eliminated to build scores.

### 2.4. *Selectivity aspect of attention*

#### 2.4.1. *Go/No-Go test (condition 2)*

According to Zimmermann and Fimm (2002), this test was intended to measure selective attention. The test contains five squares ( $3 \times 3$  cm) each showing different patterns. The participants had to memorize two target squares. Hereafter, in the middle of the screen, single squares appeared separately one after another. The subjects were instructed to respond to two target squares out of five. The test consists of 60 successive trials (24 targets and 34 non-targets). The median of the reaction time was assessed.

#### 2.4.2. *Divided-attention test (series 4)*

This test was intended to measure divided attention and the focus concept as a part of selective attention. Participants had to deal with two simultaneous tasks (one visual and one acoustic). The visual task consisted of a matrix of  $4 \times 4$  dots (size:  $10 \times 10$  cm). Seven little x's were superimposed randomly



over the 4×4 dots. When four x's formed a square, the subjects had to react as quickly as possible by pressing a button. In the acoustic task, the subjects had to react to series of alternating high (2000 Hz) and low (1000 Hz) tones. Whenever the same tone occurred twice, the subjects had to react as quickly as possible by pressing a button. The task contained 15 visual and 15 acoustic targets out of 85 visual non-targets and 185 acoustic non-targets. Since both scores did not correlate highly with each other, the median of the reaction time responding to squares and the median of the reaction time responding to tones was assessed (2 separate scores).

## 2.5. Reasoning

### 2.5.1. I-S-T 2000 R (basic module; *Amthauer et al., 2001*)

The computer-based basic module of the I-S-T 2000 R was administered to measure reasoning ability. Participants had to perform nine subtests with 20 tasks each. Only a limited period of time was given to complete each subtest. First, three subtests to measure verbal reasoning were applied: completing sentences, verbal analogies, and similarities. Within the “completing sentences” subtest, participants had to choose the appropriate phrase out of five alternatives to complete an unfinished sentence correctly. In the subtest “verbal analogies,” three words were presented. A relation existed between the first and the second given word. Between the third given word and one word of five alternatives, a similar relation could be applied. Participants had to choose the correct word. When performing the subtest “similarities,” participants were shown six words. The goal was to select two that had a common generic term. Then, three subtests to measure numerical reasoning were administered: calculations, number series, and signs. In the subtest “calculations,” 20 arithmetical problems had to be solved. “Number series” required participants to continue a line of numbers according to a certain rule. No multiple-choice alternatives were given for “calculations” and “number series”. In the subtest “signs,” mathematical equations were presented with the operators left out. The operators had to be added correctly. Finally, three subtests to measure figural reasoning were administered: figures, cubes, and matrices. The “figures” subtest consisted of geometrical figures cut into pieces. A decision had to be made, on which geometrical figure was resembled by the pieces. In the “cubes” subtest, participants were presented five cubes with different signs on their sides. Also, critical cubes were shown, which could be transferred into one out of the five reference cubes when turned. Appropriate cubes had to be allocated. The last subtest “matrices” consisted of a 2×2 matrix with three different figures. The rule underlying the placement of the three figures had to be detected and the correct missing figure had to be chosen according to that rule. Scores were built by aggregating correct answers for verbal (VIQ), numerical (NIQ), and figural (FIQ) subtests, as well as for the whole test (reasoning). *Beauducel, Brocke, and Liepmann (2001)* could show that the I-S-T 2000 R is a theoretically well-founded measure of reasoning.

## 2.6. Procedure

Participants were tested in groups of about 2 to 5 in a laboratory. Each participant took part in two sessions lasting 3 to 3.5 h each, separated by 1–2 weeks. The first session began with the attention tasks of the TAP: Go/No Go and Divided Attention. Then working memory tasks were administered: Switching followed immediately after the corresponding CRTs, followed by the corresponding dual tasks. Afterwards all coordination tasks were conducted. During the second session, the I-S-T 2000 R was administered following another test not mentioned in this part of the study. All tests were administered in German.

## 2.7. Statistical analysis

### 2.7.1. Confirmatory factor analyses

Confirmatory factor analyses (maximum likelihood) were conducted with AMOS 5.0. All test scores were *z*-transformed. Moreover, all scores (e.g., reaction times) were recorded in a way that high scores expressed high performance. The assumption of multivariate normality could not be confirmed by the Mardia-Test (multivariate kurtosis=84.894, *c. r.*=17.615,  $p < 0.001$ ). The assessment of normality for each task revealed that only the distribution of the task “finding squares” showed a severe violation of normality (negatively skewed). All other tasks met the standards recommended by West, Finch, and Curran (1995), namely skewness  $< 2$  and kurtosis  $< 7$ . Therefore, the Bollen–Stine bootstrapping (Byrne, 2001, p. 284) method was conducted to obtain a corrected  $\chi^2$ -value. In order to replicate the model of Oberauer et al. (2003), it was decided to maintain the “finding squares” task in all analyses, knowing that this procedure is somewhat critical when assessing construct reliability.

The assessment of the global goodness-of-fit was based on the Standardized Root Mean Square Residual (SRMR) and the Root Mean Squared Error of Approximation (RMSEA) as recommended by Hu and Bentler (1999). The following rules (Hu & Bentler, 1999, p. 27) were used to assess the global-fit between the tested model and the data: RMSEA  $\leq 0.06$  and SRMR  $\leq 0.11$ . Additionally, the Comparative-Fit-Index was provided. According to Hu and Bentler (1999), a cut-off value of about 0.95 is appropriate.

### 2.7.2. Tested models

According to the working memory model of Oberauer et al. (2003), a model (model 1) with three factors was tested first: “supervision” with *specific switching costs*, “storage in the context of processing,” and “coordination” (see also Fig. 4). Second, a facet model (Oberauer et al., 2003) was investigated (model 2a, (see also Fig. 5)) with three functional factors: “supervision” with *general switching costs*, “storage in the context of processing,” “coordination,” and two content factors: verbal/numerical, figural/spatial. Additionally, a model similar to model 2a was tested, but without the content factors (model 2b, (see also Fig. 6)). In a third step, an extended model (model 3, (see also Fig. 7)) with attention was examined. This model contains model 2b (Oberauer et al., 2003 without content factors) and in addition a factor which covered the “selectivity aspect of attention”. In the next step, a regression analysis was conducted using a latent variable approach (model 4, (see also Fig. 8)). The independent variables were the working memory factors (“supervision,” “coordination,” “storage in the context of processing”) and the factor attention (“selectivity aspect of attention”); the dependent variable was the factor reasoning built with the verbal, numerical, and figural facet of the I-S-T 2000 R. In a last step, a model (model 5, (see also Fig. 9)) with five factors (“reasoning,” “supervision” with general switching costs, “storage in the context of processing,” “coordination,” and the “selectivity aspect of attention”) and a “g”-factor was tested. This model is in line with Carroll (1993). It was taken into account that reasoning, “storage in the context of processing,” and “coordination” were highly related, and correlated errors of these constructs were specified.

### 2.7.3. Multiple regression analysis

In the second part of the analysis, *z*-standardized scores were built for each working memory component, the attention factor, and reasoning. Therefore, a mean score of the relevant *z*-standardized task scores was built for each of the three working memory components, the “selectivity aspect of attention” and reasoning. These five scores were used to conduct a multiple regression analysis with reasoning as the dependent variable.

Table 1  
Means, standard deviations, and reliability estimates of the sample and related norm samples

Tests	<i>M</i> (sample)	<i>SD</i> (sample)	<i>r</i> <sub>tt</sub> (sample)	<i>M</i> (norm)	<i>SD</i> (norm)	<i>r</i> <sub>tt</sub> (norm)
<i>Attention tasks</i>						
Div. attention squares (DAS)	760	83	– <sup>a</sup>	826	213	0.99 <sup>b</sup>
Div. attention tones (DAT)	552	79	– <sup>a</sup>	534	118	0.99 <sup>b</sup>
Go/No-Go (GN)	502	61	– <sup>a</sup>	536	71	0.99 <sup>b</sup>
<i>Working memory tasks</i>						
CRT categories	612	72	0.92 <sup>b</sup>	703	171	0.96 <sup>c</sup>
CRT syllables	600	129	0.96 <sup>b</sup>	638	160	0.97 <sup>c</sup>
CRT odd–even	543	97	0.95 <sup>b</sup>	554	84	0.94 <sup>c</sup>
CRT large–small	497	67	0.95 <sup>b</sup>	507	59	0.94 <sup>c</sup>
CRT up–down	420	72	0.93 <sup>b</sup>	468	60	0.93 <sup>c</sup>
CRT above–below	458	56	0.92 <sup>b</sup>	431	64	0.92 <sup>c</sup>
CRT parts	628	134	0.95 <sup>b</sup>	673	116	0.96 <sup>c</sup>
CRT symmetry	846	344	0.95 <sup>b</sup>	825	236	0.93 <sup>c</sup>
Specific SC verbal (SSCV)	259	162	0.84 <sup>b</sup>	246	186	0.78 <sup>c</sup>
Specific SC number (SSCN)	274	173	0.86 <sup>b</sup>	332	198	0.79 <sup>c</sup>
Specific SC arrows (SSCA)	295	178	0.81 <sup>b</sup>	299	205	0.84 <sup>c</sup>
Specific SC pattern (SSCP)	244	137	0.76 <sup>b</sup>	247	252	0.77 <sup>c</sup>
General SC verbal (GSCV)	219	120	0.60 <sup>b</sup>	241	175	0.83 <sup>c</sup>
General SC number (GSCN)	108	87	0.79 <sup>b</sup>	102	97	0.83 <sup>c</sup>
General SC arrows (GSCA)	166	110	0.59 <sup>b</sup>	145	120	0.79 <sup>c</sup>
General SC pattern (GSCP)	32	123	0.42 <sup>b</sup>	23	93	0.66 <sup>c</sup>
Dual task verbal (DTV)	3.8	0.7	0.80 <sup>c</sup>	3.6	0.7	0.89 <sup>c</sup>
Dual task numerical (DTN)	5.1	0.7	0.78 <sup>c</sup>	5.1	0.5	0.83 <sup>c</sup>
Dual task dots (DTD)	1.8	0.5	0.72 <sup>c</sup>	1.6	0.5	0.86 <sup>c</sup>
Dual task patterns (DTP)	2.3	0.4	0.77 <sup>c</sup>	2.2	0.7	0.91 <sup>c</sup>
Monitoring verbal (MV)	1.3	0.6	0.74 <sup>c</sup>	1.4	0.7	0.76 <sup>c</sup>
Monitoring number (MN)	1.7	0.4	0.64 <sup>c</sup>	1.7	0.5	0.59 <sup>c</sup>
Flight control (FQ)	1.9	0.9	0.85 <sup>c</sup>	1.0	0.5	0.86 <sup>c</sup>
Finding squares (FS)	1.8	2.9	0.34 <sup>c</sup>	1.3	0.7	0.73 <sup>c</sup>
<i>Reasoning tasks</i>						
Completing sentences	14.10	2.56	0.53 <sup>c</sup>	13.72	3.43	0.63 <sup>c</sup>
Verbal analogies	13.42	2.65	0.58 <sup>c</sup>	13.07	2.95	0.68 <sup>c</sup>
Similarities	13.69	2.42	0.53 <sup>c</sup>	12.56	3.16	0.71 <sup>c</sup>
Calculations	13.81	3.69	0.79 <sup>c</sup>	14.17	3.85	0.80 <sup>c</sup>
Number series	15.17	4.41	0.89 <sup>c</sup>	14.05	4.89	0.90 <sup>c</sup>
Signs	14.41	3.76	0.83 <sup>c</sup>	14.07	3.89	0.84 <sup>c</sup>
Figures	13.35	3.18	0.64 <sup>c</sup>	12.20	3.83	0.76 <sup>c</sup>
Cubes	11.11	4.36	0.84 <sup>c</sup>	11.47	4.17	0.79 <sup>c</sup>
Matrices	11.74	2.73	0.54 <sup>c</sup>	11.23	3.43	0.66 <sup>c</sup>

CRT=choice reaction time. Means and *SD* are given for untransformed times in milliseconds (ms) for CRT, time differences for switching variables, mean raw scores for the other tasks; reliability always refers to the scores used in the analyses. For divided attention and Go/No-Go tasks the median of reaction times (ms) is given. Related norms: attention (Zimmermann & Fimm, 2002); working memory (Oberauer et al., 2003); reasoning (Amthauer et al., 2001).

<sup>a</sup> Could not be calculated.

<sup>b</sup> Split-half reliability.

<sup>c</sup> Cronbach alpha.

### 3. Results

#### 3.1. Descriptive statistics and reliability

In Table 1 and Table 2, means, standard deviations, and reliability estimates of the subtests and factors are provided. Additionally in Table 1 the means and standard deviations found by Oberauer et al. (2003) are shown to point out differences of both samples. This might be important for further interpretations. Reliability estimates for the working memory tests are sufficient except for the tasks “finding squares,” “switching pattern,” and “switching arrows”. However, in comparison to Oberauer et al. (2003), the standard deviation and, consequently, the reliabilities were reduced. The same holds true for the matrices subtest of the I-S-T 2000 R (Amthauer et al., 2001). The I-S-T 2000 R total score of the present sample ( $SW=110$ ,  $SD=8$ ) was one standard deviation above the norm ( $N=2020$ ). Thus, the standard deviation is reduced in comparison to the norm (about 20%).

To rule out the possibility that correlations between different tests changed over practice, the corrected item–total correlation of each test block with the relevant subtest was calculated. Additionally, the correlation within blocks of each subtest was investigated. Finally, the correlation of each test block with reasoning was computed. None of the three analyses revealed major or systematic changes of the test blocks’ correlations within one subtest except for the practice blocks which were not entered into the calculation of test scores.

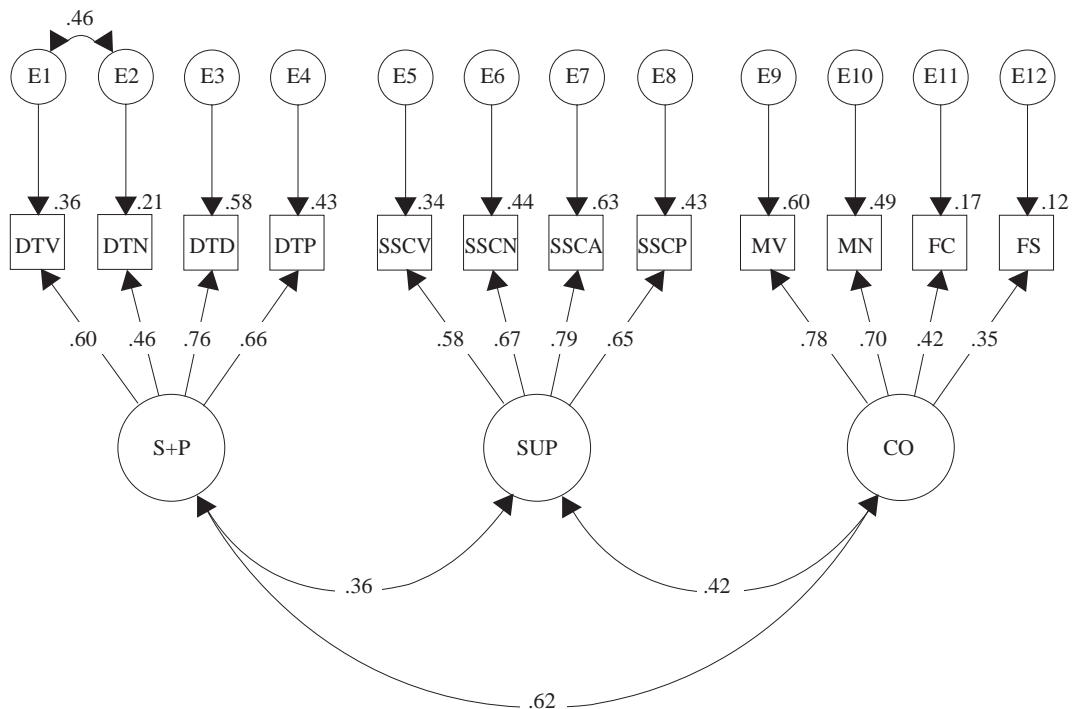


Fig. 4. Standardized solution of model 1 as reported by Oberauer et al. (2003). For a description of the used abbreviations see Tables 1 and 2. Unstandardized estimates (standard error; critical ratio): S+P↔SUP: 0.11 (0.04; 2.61); S+P↔CO: 0.27 (0.07; 3.70); SUP↔CO: 0.17 (0.06; 3.03).

3.2. Tested models

3.2.1. Model 1 (working memory, functional facets, specific switching costs)

Model 1 (see Fig. 4) revealed a good overall model-fit ( $\chi^2 [50]=62.786$ , n. s., RMSEA=0.046 [0.000, 0.078], SRMR=0.067). The loadings on the functional factors were all moderate to high (“storage in context of processing”:  $a=0.46$  to  $a=0.76$ ,  $p<0.001$ ; “supervision”:  $a=0.58$  to  $a=0.79$ ,  $p<0.001$ ; “coordination”:  $a=0.35$  to  $a=0.78$ ,  $p<0.001$ ). The functional factors correlated moderately to high: “coordination” and “supervision”:  $r=0.42$  ( $p<0.01$ ); “storage in the context of processing” and “supervision”:  $r=0.36$  ( $p<0.05$ ); “storage in the context of processing” and “coordination”:  $r=0.62$  ( $p<0.001$ ). The significantly correlated errors between E1 (dual task verbal) and E2 (dual task numerical) might have occurred since we specified no content facets (verbal/numerical). This holds true for all following models. Oberauer et al. (2003) showed that this significant correlation diminishes when content facets are specified.

3.2.2. Model 2a (working memory, content facets, functional facets, general switching costs)

Model 2a (see Fig. 5) could not be confirmed since negative variances occurred. The model was rejected and not further considered (Table 2).

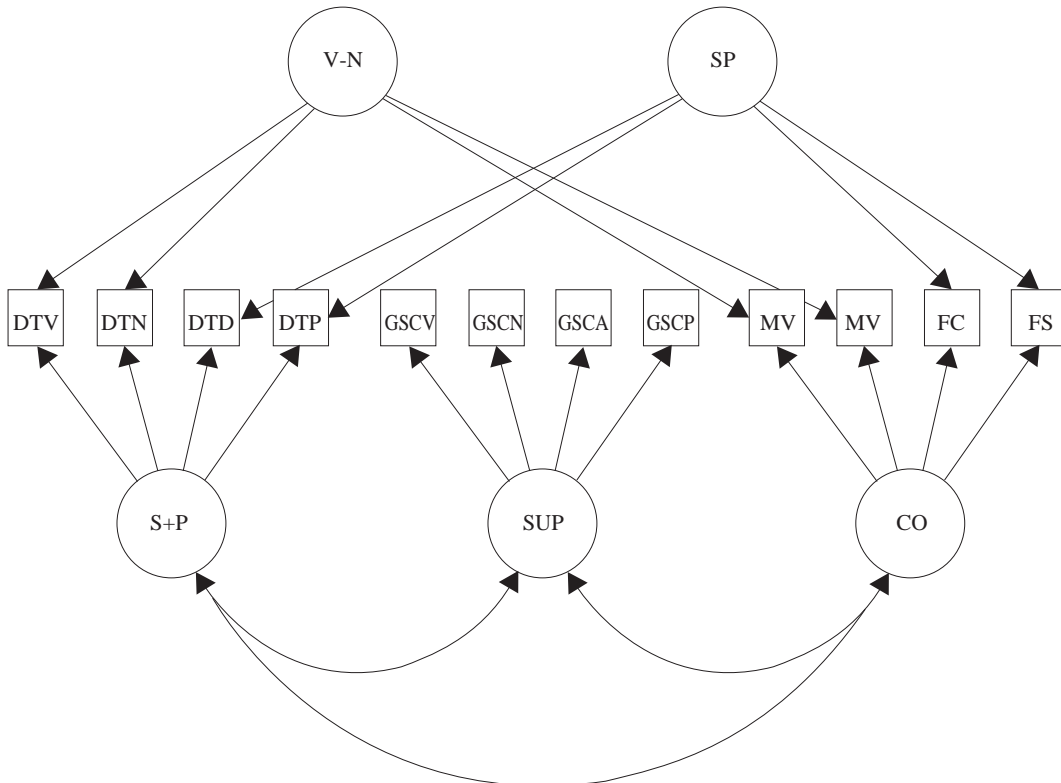


Fig. 5. Model 2a with general switching costs and two content factors. For a description of the used abbreviations see Tables 1 and 2.

Table 2  
Reliability of test batteries and construct reliability

Tests	bat <sup>r</sup> tt	constr <sup>r</sup> tt
<i>TAP</i>		
Selectivity aspect of attention (SEL)	— <sup>a</sup>	0.75
<i>Working memory</i>		
Storage in the context of processing (S+P)	0.90 <sup>b</sup>	0.74
Supervision (SUP); (model 1)	0.92 <sup>b</sup>	0.78
Supervision (SUP); (model 2a)	0.78 <sup>b</sup>	0.66
Coordination (CO)	0.87 <sup>b</sup>	0.74
<i>IST-2000-R</i>		
Reasoning (INT)	0.86 <sup>b</sup>	0.67

TAP=test battery for attentional performance. Reliabilities of test batteries have been calculated by a formula for test battery reliability by [Lienert and Raatz \(1998, p. 330\)](#). Reliabilities for latent variables as applied in structural equation modeling have been calculated according to [Hancock and Mueller \(2001\)](#).

<sup>a</sup> Could not be calculated.

<sup>b</sup> Cronbach alpha.

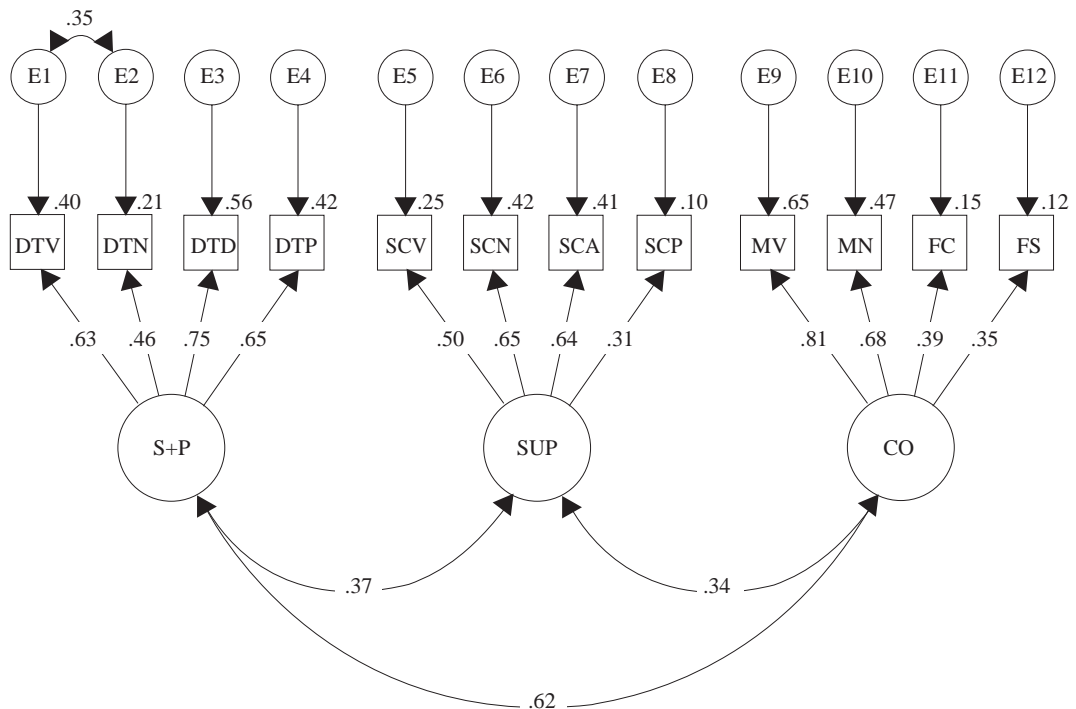


Fig. 6. Standardized solution of model 2b with general switching costs similar to [Oberauer et al. \(2003\)](#). For a description of the used abbreviations see [Tables 1 and 2](#). Unstandardized estimates (standard error; critical ratio): S+P↔SUP: 0.12 (.05; 2.36); S+P↔CO: 0.34 (0.08; 4.12); SUP↔CO: 0.14 (0.06; 2.32).



3.2.3. Model 2b (working memory, functional facets, general switching costs)

This model, slightly modified compared to the model of Oberauer et al. (2003) since the content facets have been deleted, showed a very good overall model fit ( $\chi^2 [50]=54.163$ , n. s., CFI=0.98, RMSEA=0.026 [0.000, 0.065], SRMR=0.0600). The loadings on the functional factors were all moderate to high (“storage in context of processing”:  $a=0.46$  to  $a=0.75$ ,  $p<0.001$ ; “supervision”:  $a=0.31$  to  $a=0.65$ ,  $p<0.05$ ; “coordination”:  $a=0.35$  to  $a=0.81$ ,  $p<0.001$ ). The functional factors correlated moderately to high: “coordination” and “supervision”:  $r=0.34$  ( $p<0.05$ ); “storage in context of processing” and “supervision”:  $r=0.37$  ( $p<0.05$ ); “storage in context of processing” and “coordination”:  $r=0.62$  ( $p<0.001$ ).

3.2.4. Model 3 (attention and working memory)

The model (see Fig. 7) revealed an excellent model fit ( $\chi^2 [83]=89.931$ , n. s., CFI=0.98, RMSEA=0.026 [0.000, 0.058], SRMR=0.0603). Low to moderate significant correlations between the factors “selectivity aspect of attention” and “storage in the context of processing” ( $r=0.34$ ,  $p<0.05$ ), “selectivity aspect of attention” and “supervision” ( $a=0.36$ ,  $p<0.05$ ), as well as “selectivity aspect of attention” and “coordination” ( $a=0.50$ ,  $p<0.001$ ) could be found.

3.2.5. Model 4 (attention, working memory and intelligence)

The model 4 (see Fig. 8) revealed a very good overall model fit ( $\chi^2 [126]=144.239$ , n. s., CFI=0.96, RMSEA=0.034 [0.000, 0.058], SRMR=0.0626). All in all, 95% of the reasoning variance could be explained. Two paths from “storage in the context of processing” ( $a=0.67$ ,  $p<0.001$ ) and “coordination” ( $a=0.40$ ,  $p<0.01$ ) to reasoning reached significance. The factors “supervision” and “selectivity aspect of

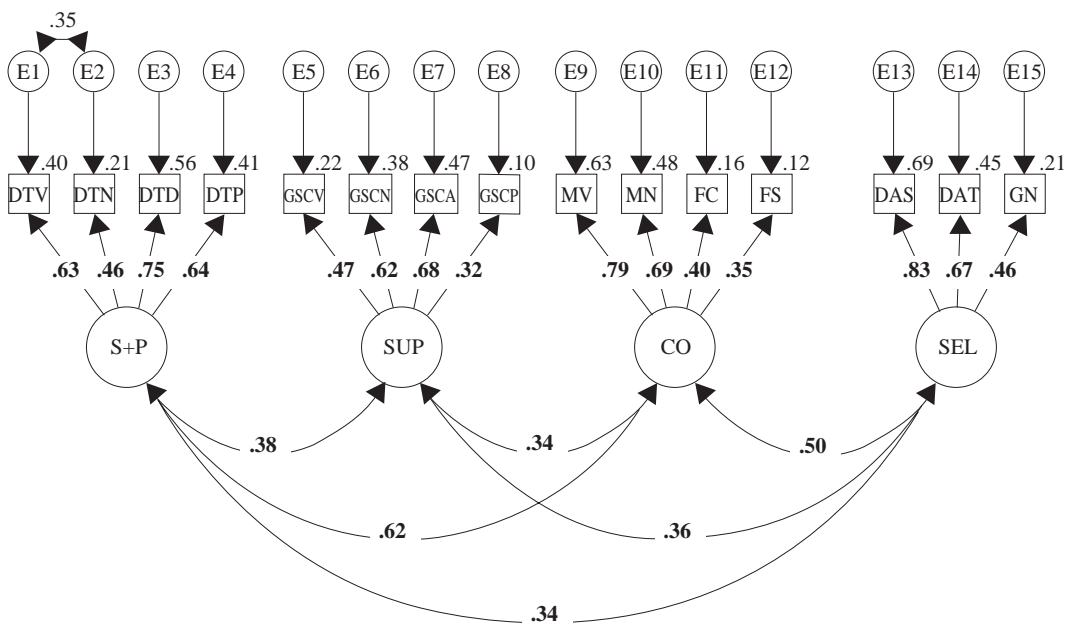


Fig. 7. Standardized solution of model 3. For a description of the used abbreviations see Tables 1 and 2. Unstandardized estimates (standard error; critical ratio): S+P $\leftrightarrow$ SUP: 0.11 (0.05; 2.34); S+P $\leftrightarrow$ CO: 0.29 (0.08; 3.88); S+P $\leftrightarrow$ SEL: 0.17 (0.07; 2.54); SUP $\leftrightarrow$ CO: 0.12 (0.06; 2.22); SUP $\leftrightarrow$ SEL: 0.14 (0.06; 2.35); CO $\leftrightarrow$ SEL: 0.33 (0.09; 3.74).

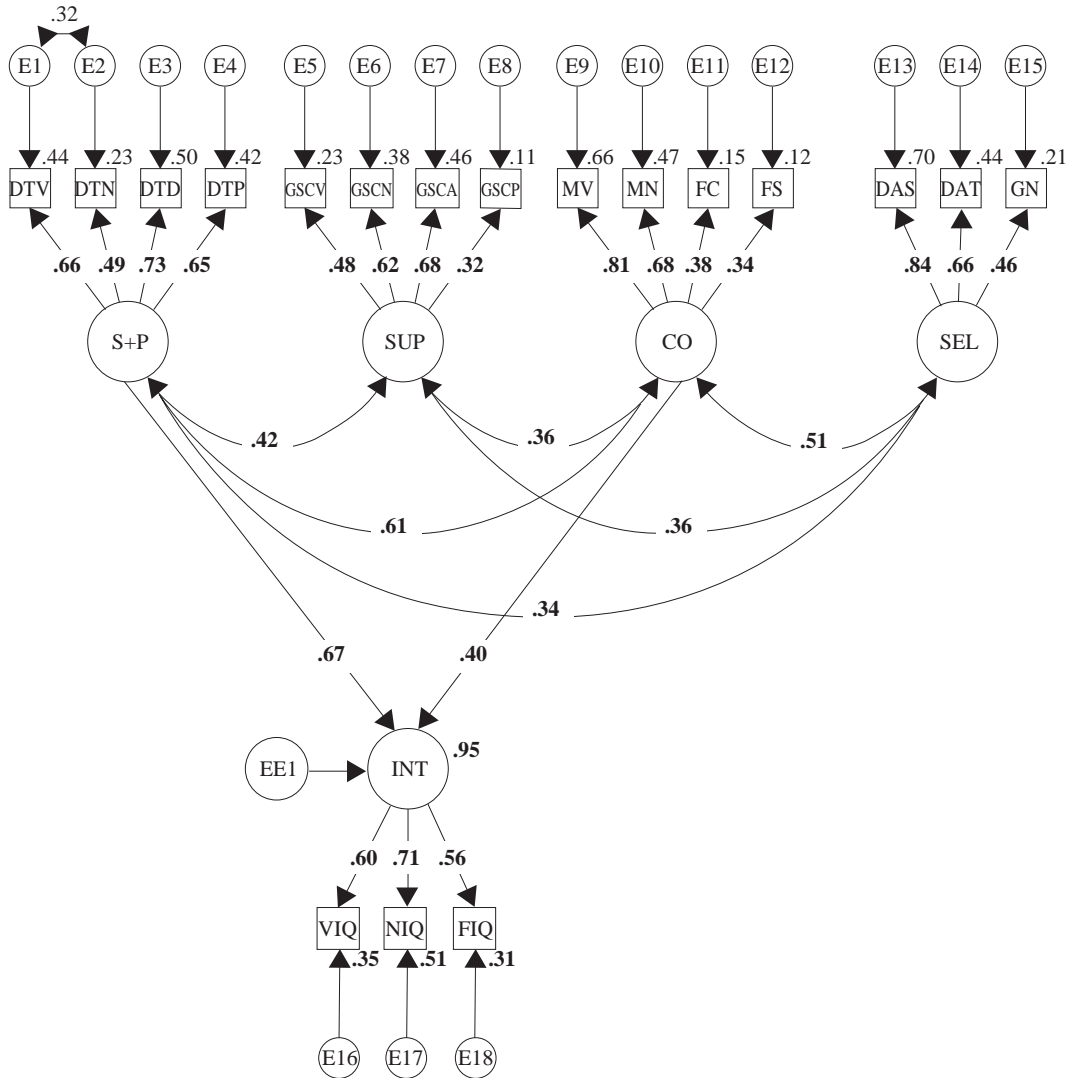


Fig. 8. Standardized solution of model 4. For a description of the used abbreviations see Tables 1 and 2. VIQ=verbal intelligence; NIQ=numerical intelligence; FIQ=figural intelligence. Unstandardized regression weight estimates (standard error; critical ratio): S+P→INT: 0.63 (0.17; 3.80); CO→INT: 0.30 (0.11; 2.72).

attention” did not add any significant explanatory power beyond “storage in the context of processing” and “coordination”.

### 3.2.6. Model 5 (g-factor model)

Model 5 (see Fig. 9) also showed an excellent model-fit ( $\chi^2$  [126]=144.901, n. s., CFI=0.96, RMSEA=0.035 [0.000, 0.058], SRMR=0.0638). As expected, reasoning ( $a=0.79$ ,  $p<0.001$ ) had the highest loading on “g,” followed by “coordination” ( $a=0.73$ ,  $p<0.001$ ), “attention” ( $a=0.63$ ,  $p<0.001$ ), “storage in the context of processing” ( $a=0.58$ ,  $p<0.01$ ), and by “supervision” ( $a=0.55$ ,  $p<0.01$ ). The

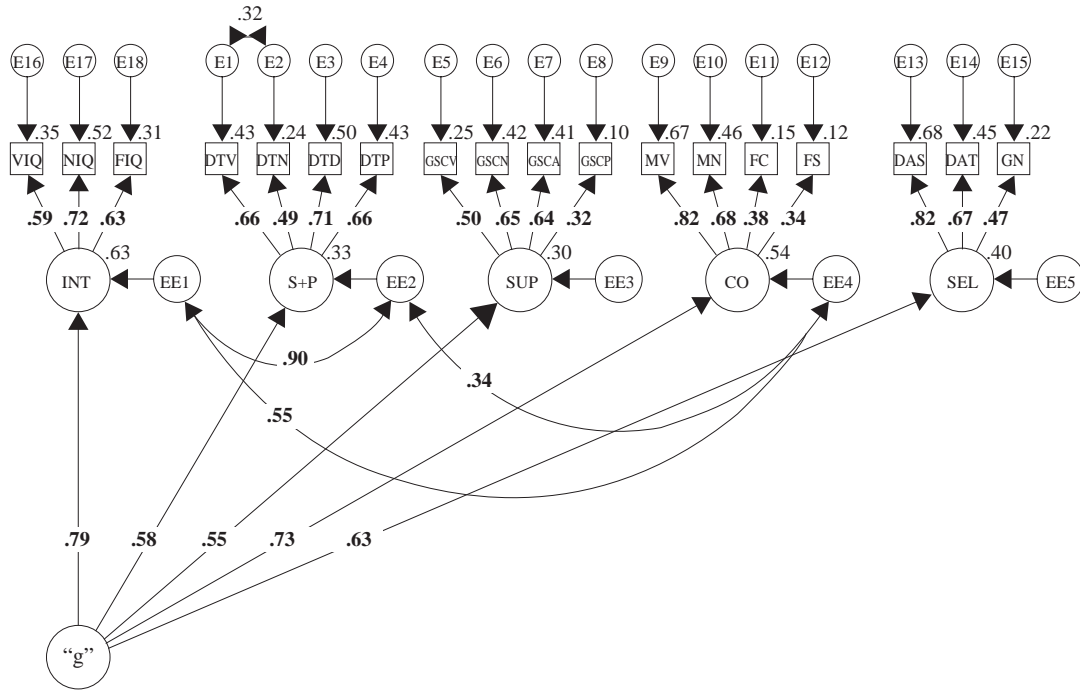


Fig. 9. Standardized solution of model 5 (“g”-factor model). For a description of the used abbreviations see Tables 1 and 2. VIQ=verbal intelligence; NIQ=numerical intelligence; FIQ=figural intelligence. Unstandardized regression weight estimates (standard error; critical ratio): “g”→INT: 0.47 (0.12; 3.85); “g”→S+P: 0.36 (0.12; 3.85); “g”→SUP: 0.27 (0.09; 3.07); “g”→CO: 0.60 (0.14; 4.17); “g”→SEL: 0.52 (0.13; 3.90).

changed path loadings in contrast to model 4 might be due to the correlated error variances of the latent constructs reasoning, “storage in the context of processing,” and “coordination”. The loadings of these constructs represent the unique contribution to “g”. The error correlation was necessary because of the large amount of common variance of these three constructs. The attempt was made to specify the correlated error variances as a latent variable (working memory bimodal model), but this failed due to negative variances. This might have occurred because of a small sample size.

### 3.3. Regression analysis

With the z-standardized scores of the working memory components and attention, 49% of the reasoning variance could be explained (correlations and beta weights, see Table 3). The same amount of explained variance could be achieved, when entering factor scores instead of z-standardized scores into the regression analysis. In contrast to the structural equation model (model 4) the working memory component “supervision” turned out to be a significant predictor of reasoning as well ( $\beta=0.13$ ;  $p<0.05$ ). This might be due to variance components which were not included in the factor scores (specific task variance). Compared to the other two working memory components, the beta-weight of “supervision” is considerably lower.

To explore if the relationship between “reasoning” and the working memory components (“supervision”:  $r=0.31$ ,  $p<0.001$ ; “coordination”:  $r=0.51$ ,  $p<0.001$ , “storage in the context of

Table 3

Regression analysis of the z-standardized scores (beta-weights, correlations, and partial correlations)

z-standardized scores	Beta	<i>T</i>	<i>p</i>	<i>r<sub>simple</sub></i>	<i>r<sub>partial</sub></i>
Selectivity aspect of attention	0.03	0.04	n.s.	0.31	0.04
Storage in the context of processing	0.37	6.06	<0.001	0.63	0.49
Coordination	0.22	3.32	<0.01	0.51	0.29
Supervision	0.13	2.13	<0.05	0.31	0.19
Processing speed	0.06	0.95	n.s.	0.40	0.09

Regression method=Enter. The dependant variable was reasoning.  $r_{simple}$ =Bivariate correlation between the corresponding z-standardized score and reasoning.  $r_{partial}$ =Correlation between the corresponding z-standardized score and reasoning when all other scores were partialled out.

processing”:  $r=0.63$ ,  $p<0.001$ ) as well as the “selectivity aspect of attention” ( $r=0.31$ ,  $p<0.001$ ) can be restored to processing speed, a z-standardized score of choice reaction times (mean score of eight tasks) was partialled out. The partial correlations remained rather unaffected except for the “selectivity aspect of attention” and reasoning, where a significant decrease of the correlation occurred (from  $r=0.31$  to 0.14,  $p<0.05$ , one-sided).

After controlling for “attention,” “supervision,” and “choice reaction times,” simultaneously the correlation between “storage in the context of processing” and “coordination” decreased significantly (from  $r=0.44$  to 0.31,  $p<0.05$ , one-sided), but still remained significant ( $p<0.001$ ) (Table 3).

#### 4. Discussion

The present study replicated the finding that “storage in the context of processing” is the best predictor of reasoning. Furthermore, it was revealed that “coordination” is also a significant predictor of reasoning. “Supervision” and the “selectivity aspect of attention” had only little or no impact on reasoning. Moreover, the variance explained by the “selectivity aspect of attention” predicting reasoning can be reduced to speed variance. The variance explained by working memory components (especially “storage in the context of processing” and “coordination”) on reasoning was 95% regarding the latent factors. However, using a multiple regression analysis, the explained variance was reduced to 49%. This might be due to the moderate construct reliability of all constructs and therefore, a big correction of attenuation occurred.

Furthermore, this study replicated the functional facets of the model proposed by Oberauer et al. (2003), but not the content facets. Nevertheless, the well-known fact was replicated that the factor “supervision/speed” correlated only moderately with the factors “storage in the context of processing” and “coordination” (Oberauer et al., 2000, 2003). The correlation between “selectivity aspect of attention” and “coordination” was moderate to high.

One goal of this study was to replicate the model proposed by Oberauer et al. (2003). The excellent global-fit confirmed the structure of working memory found by Oberauer et al. (2003). However, the content factors could not be confirmed. This might be due to the reduced standard deviations and (consequently) lower reliabilities of some working memory tasks. The present study also replicated two highly correlated functional facets “coordination” and “storage in the context of processing”. As expected, the correlations between these functional facets and “supervision” were found to be moderate (see Oberauer et al., 2003).

Another goal was to clarify the connection between working memory components and concepts of attention. All in all, there was no strong relationship between working memory and the “selectivity aspect of attention”. The highest correlation was found between the “selectivity aspect of attention” and “coordination” ( $r=0.50$ ). This might be due to the similarity of the divided attention task (squares) and the motor task “finding squares”. Both tasks used rapidly changing objects. This has to be clarified within experimental settings.

The main goal of the study was to clarify the relationship between a facet model of working memory, attention, and reasoning. It was replicated that “storage in the context of processing” is the best predictor of reasoning (e.g., Conway et al., 2002; Engle et al., 1999; Suess et al., 2002). Surprisingly, “coordination” was the best predictor of “g”. Thus, it seems to be likely that working memory components had a differential validity concerning reasoning and “g”. It has to be noted, that in the present study “g” was mainly built by working memory components. This “g” measure is not comparable with the highly elaborated model suggested by Carroll (1993). Thus, the present result has to be interpreted carefully and should be replicated with a broader selection of “g” components. There is much evidence that, besides “storage in the context of processing,” “coordination” is a key facet of working memory in predicting reasoning and “g”. Using structural equation models “supervision” did not significantly add explanatory power to the prediction of reasoning (see Suess et al., 2002). In the regression analysis a significant prediction could be achieved. As mentioned above, this might be due to specific task variance. In spite of this, “supervision” does not play the key role in predicting reasoning.

Besides the differential validity of working memory in predicting reasoning, another important question was addressed: Is working memory really perfectly predicted by “g?” as reported by Colom et al. (2004) but challenged by Ackerman et al. (2002) and other studies (Conway et al., 2002; Schweizer & Moosbrugger, 2004; Suess et al., 2002). The present study revealed evidence for both arguments: the path coefficients between “storage in the context of processing” and reasoning respectively “g” were similar to coefficients reported by Ackerman et al. (2002). Thus, these findings are in line with their argumentation. Nevertheless, reasoning can almost perfectly be predicted by working memory when adding the “coordination” component of working memory! This supports findings by Colom et al. (2004). The present study might suggest that a high correlation between working memory and reasoning does only occur when – in addition to “storage in the context of processing” – “coordination” is assessed. Most likely Colom et al. (2004) used tasks which require both: “storage in the context of processing” and “coordination”.

Attention had no significant impact on reasoning. Moreover, the variance explained by the “selectivity aspect of attention” predicting reasoning can be reduced to speed variance. In contrast to this, the common variance between working memory components and reasoning is not mediated by speed variance.

The variance explained by working memory components on reasoning was 95% regarding the latent factors. The high amount of explained variance might be due to the correction of attenuation, which is performed using structural equation modeling. The correction of attenuation refers to construct reliability. The lower the loadings on the latent variables, the higher is the correction. The construct reliability of the latent factors was on the lower or underneath the lower bound recommended by Hancock and Mueller (2001). This might explain the strong attenuation. However, using multiple regression analyses, the explained variance was reduced to 49%. This amount of variance is in line with previous research (e.g., Colom et al., 2003; Wittmann & Suess, 1999).

The present study has some possible limitations. First, it has to be mentioned that the participants were mainly female. This obviously limits the generality of the results. In contrast to other studies the information to be held in memory during the task “storage in the context of processing” was irrelevant to the decision task. Often working memory load is related to the decision task. Again, this affects the comparability between this and other studies. Another very important limitation is the fact that to a certain extent some of the applied reasoning tasks obviously measure crystallized intelligence, especially the verbal tasks. Future research should clarify the role of working memory in predicting  $g_f$  and  $g_c$ . In the present study the variance of the relevant measures is partly reduced due to the participants’ high performance. Therefore, in more heterogeneous samples even higher correlations can be expected.

All in all, the model proposed by Oberauer et al. (2003) showed a differential validity in predicting reasoning and “g”. It provided and provides a useful basis to predict a large amount of reasoning variance. Finally, the questions addressed by Suess et al. (2002), “What is critical for successful performance in reasoning tasks?”, could be answered: It is “storage in the context of processing” and “coordination!”

## Acknowledgement

We thank the anonymous reviewers and Earl Hunt for their valuable comments and suggestions.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at [doi:10.1016/j.intell.2005.01.002](https://doi.org/10.1016/j.intell.2005.01.002).

## References

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, *131*, 567–589.
- Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R)—Handanweisung* [Intelligence-Structure-Test 2000 R (I-S-T 2000 R)—Handbook]. Goettingen: Hogrefe.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A. D. (2001). Is working memory still working? *American Psychologist*, *56*, 849–864.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–90). San Diego: Academic Press.
- Beauducel, A., Brocke, B., & Liepmann, D. (2001). Perspectives on fluid and crystallized intelligence: Facets for verbal, numerical, and figural intelligence. *Personality and Individual Differences*, *30*, 977–994.
- Byrne, B. (2001). *Structural equation modelling with AMOS, basic concepts, applications, and programming*. London: Erlbaum.
- Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. New York: Cambridge University Press.
- Cattell, R. B. (1973). *Measuring intelligence with the culture fair test*. Champaign, IL: Institute for Personality and Ability Testing.
- Centro editor de psicologia aplicada (1993). *Testes das Matrizes Progressivas de Raven—Escala Geral* [Standard Progressive Matrices]. Adapted by Francisco Campos; Revised by Suzana Ezequiel da Cunha. RJ: CEPA.



- Colom, R., Flores-Mendoza, C., & Rebollo, I. (2003). Working memory and intelligence. *Personality and Individual Differences*, *34*, 33–39.
- 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.
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.
- 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*, *3*, 309–331.
- Eysenck, H. J. (1982). *A model for intelligence*. Heidelberg: Springer.
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Srboom (Eds.), *Structural equation modeling: Present and future—Festschrift in honor of Karl Jreskog* (pp. 195–216). Lincolnwood, IL: Scientific Software International, Inc.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
- Jaeger, A. O., Suess, H. -M., & Beauducel, A. (1997). *Test fuer das Berliner Intelligenz-Strukturmodell (BIS 4)* [Berlin–Structure–Intelligence Test (BIS 4)]. Goettingen: Hogrefe.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: The Free Press.
- Kremen, W. S., Seidman, L. J., Faraone, S. V., Pepple, J. R., & Tsuang, M. T. (1992). Attention/information-processing factors in psychotic disorders. Replication and extension of recent neuropsychological findings. *The Journal of Nervous and Mental Disease*, *180*, 89–93.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, *14*, 389–433.
- Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse* [Construction and analyses of tests]. Weinheim: Psychologie Verlags Union.
- Mirsky, A. F., Anthony, B. J., Duncan, C. D., Ahearn, M. B., & Kellam, S. G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychology Review*, *2*, 109–145.
- 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.
- Miyake, A., & Shah, P. (1999). *Models of working memory. Mechanisms of active maintenance and executive control*. Cambridge: University Press.
- Necka, E. (1996). The attentive mind: intelligence in relation to selective attention, sustained attention, and dual task performance. *Polish Psychological Bulletin*, *27*, 3–24.
- Oberauer, K., Suess, H. -M., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity—facets of a cognitive ability construct. *Personality and Individual Differences*, *29*, 1017–1045.
- Oberauer, K., Suess, H. -M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167–193.
- O’Donell, J. P., MacGregor, L. A., Dabrowski, J. J., Oestreicher, J. M., & Romero, J. J. (1994). Construct Validity of neuropsychological tests of conceptual and attentional abilities. *Journal of Clinical Neuropsychology*, *4*, 596–600.
- Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices*. London: H.K. Lewis.
- Roberts, R. D., Beh, H. C., Spilbury, G., & Stankov, L. (1991). Evidence for an attentional model of human intelligence using the competing task paradigm. *Personality and Individual Differences*, *12*, 445–455.
- Roberts, R. D., Beh, H. C., & Stankov, L. (1988). Hick’s Law, competing tasks, and intelligence. *Intelligence*, *12*, 101–120.
- Schweizer, K. (1995). Hypotheses on biological and cognitive sources of general intelligence. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *16*, 67–81.
- Schweizer, K., & Moosbrugger, H. (2004). Attention and working memory as predictors of intelligence. *Intelligence*, *32*, 329–347.
- Schweizer, K., Zimmermann, P., & Koch, W. (2000). Sustained attention, intelligence and the crucial role of perceptual processes. *Learning and Individual Differences*, *2*, 271–286.

- Shum, D. H. K., McFarland, K. A., & Bain, J. D. (1990). Construct validity of eight tests of attention: Comparison of normal and closed head injured samples. *The Clinical Neuropsychologist*, 4, 151–162.
- Stankov, L., Roberts, R., & Spilsbury, G. (1994). Attention and speed of test-taking in intelligence and aging. *Personality and Individual Differences*, 17, 273–284.
- Strauss, M. E., Thompson, P., Adams, N. L., Redline, S., & Burant, C. (2000). Evaluation of a model of attention with confirmatory factor analysis. *Neuropsychology*, 2, 201–208.
- Sturm, W. (2002). *Diagnostik von Aufmerksamkeitsstoerungen in der Neurologie* [Diagnostic of attention deficits in neurology]. *Aktuelle Neurologie*, 29, 25–29.
- Suess, H. -M., Oberauer, K., Wittman, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence*, 30, 261–288.
- Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, 1.
- Van Zomeren, A. H., & Brouwer, W. H. (1994). *Clinical neuropsychology of attention*. New York: Oxford Press.
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Thousand Oaks, CA: Sage Publications, Inc.
- Wittmann, W. W., & Suess, H. -M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 77–108). Washington, DC: American Psychological Association.
- Zimmermann, P., & Fimm, B. (2002). *Testbatterie zur Aufmerksamkeitspruefung (TAP), Version 1.7. Handbuch* [Test Battery for Attentional Performance (TAP) Version 1.7. Handbook]. Wuerselen: Vera Fimm/Psychologische Testsysteme.