Intelligence 39 (2011) 323–334

Contents lists available at ScienceDirect

Intelligence

Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning☆

Daniel Danner*, Dirk Hagemann, Andrea Schankin, Marieke Hager, Joachim Funke

Institute of Psychology, University of Heidelberg, Germany

ARTICLE INFO

Article history:
Received 31 January 2011
Received in revised form 6 May 2011
Accepted 3 June 2011
Available online 7 July 2011

Keywords:
Dynamic decision making
Complex problem solving
Implicit learning
Latent state-trait theory
Professional success

ABSTRACT

The present study investigated cognitive performance measures beyond IQ. In particular, we investigated the psychometric properties of dynamic decision making variables and implicit learning variables and their relation with general intelligence and professional success. $N=173$ employees from different companies and occupational groups completed two standard intelligence tests, two dynamic decision making tasks, and two implicit learning tasks at two measurement occasions each. We used structural equation models to test latent state-trait measurement models and the relation between constructs. The results suggest that dynamic decision making and implicit learning are substantially related with general intelligence. Furthermore, general intelligence is the best predictor for income, social status, and educational attainment. Dynamic decision making can predict supervisor ratings even beyond general intelligence.

1. Introduction

General intelligence is one of the most successful psychological constructs. Since Spearman's (1904) early investigations, there is a wealth of evidence for the reliability, stability, and validity of intelligence measures (Carroll, 1993). Furthermore, general intelligence is a powerful predictor of success in many domains of real life (Ng, Eby, Sorensen, & Feldman, 2005; Salgado et al., 2003; Schmidt & Hunter, 2004). Beside its undisputed usefulness, some researchers have suggested to use additional constructs for characterizing individuals' cognitive ability such as dynamic decision making and implicit learning (Dörner, 1980; Mackintosh, 1998).

The concept of dynamic decision making was developed by Dörner (1980, 1986) who proposed that situations in real life are complex and solving problems in real life requires managing complex information. He criticized that standard measures of general intelligence only assess whether individuals perform accurately and quickly in rather simple tasks but not whether they show intelligent behavior in complex tasks. Therefore, he suggested to measure performance in computer based scenarios that simulate complex, connected, dynamic, and non-transparent environments. Further on, he hypothesized that individual differences in dynamic decision making are unrelated to general intelligence but are substantially related to professional success.

Mackintosh (1998) suggested to consider another construct. He proposed that there are two independent mental systems: an explicit, hypothesis generating and testing system and an implicit, associative learning system. In particular, the explicit learning system is necessary for discovering regularities with intention and awareness (like in a numerical series task). The implicit learning system, on the other hand, detects contingencies without awareness or intention (like judging whether a sentence is grammatically

☆ This research was funded by German Research Foundation Grant DFG, Ha3044/7-1. We gratefully thank Andreas Neubauer, Anna-Lena Schubert, and Katharina Weskamp for conducting the assessment and Neil Patrick Harris and two anonymous reviewers for their helpful suggestions.

* Corresponding author at: University of Heidelberg, Institute of Psychology, Hauptstrasse 47-51, D-69117 Heidelberg, Germany. Tel.: +49 6221 547354; fax: +49 6221 547325.
E-mail address: daniel.danner@psychologie.uni-heidelberg.de (D. Danner).

0160-2896/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2011.06.004

right or wrong without being able to report the respective grammatical rule). Mackintosh suggested that standard intelligence tests capture individual differences in the explicit system but not individual differences in the implicit learning system. Therefore, he suggested to take individual differences in implicit learning into account. He hypothesized that these differences are independent from general intelligence measures but are nevertheless important predictors of educational and professional success.

Dörner and Mackintosh's proposals raise two interesting questions. Are there reliable individual differences in dynamic decision making and implicit learning which are independent from general intelligence? Can these differences predict real life performance beyond IQ? Investigating these issues will be the aim of the present study.

1.1. Previous findings

1.1.1. Dynamic decision making

Dörner's (1980, 1986) critique of standard intelligence tests laid the foundation for a field of research, which has been called *dynamic decision making* (Gonzalez, Vanyukov, & Martin, 2005) or *complex problem solving* (Funke, 2010). Over the years, several dynamic decision making tasks have been developed. For example, the Tailorshop scenario (Funke, 1983) simulates a fictional company where the participants have to control many variables like the number of workers or the costs for advertising to maximize their company value. Other tasks simulate a forestry (Wagener, 2001), a power plant (Wallach, 1998), or a space flight (Wirth & Funke, 2005) where the participants have to control several variables to reach a given goal state. Recently, dynamic decision making tasks have also been included in the Programme for International Student Assessment (PISA; Wirth & Klieme, 2003).

Over the years, there have been many studies investigating the relation between dynamic decision making and general intelligence. Whereas several studies found non-significant or only small correlations (for an overview see Kluwe, Misiak, & Haider, 1991), other studies reported significant standardized path coefficients between $\beta = 0.38$ and $\beta = 0.54$ from latent intelligence to latent dynamic decision making variables (Kröner, Plass, & Leutner, 2005; Rigas, Carling, & Brehmer, 2002; Wittmann & Hattrup, 2004). One study even found a correlation between a latent intelligence and a latent dynamic decision making variable of $r = 0.84$ (Wirth & Klieme, 2003).

There are only two studies that investigated the predictive validity of dynamic decision making measures. Wagener and Wittmann (2002) assessed a sample of $N = 35$ trainees and reported correlations between $r = 0.16$ and $r = 0.40$ between the performance in a dynamic decision making task and the performance in different assessment center tasks. However, the study did not report whether these relationships were incremental or due to an overlap between dynamic decision making and general intelligence. Kersting (2001) reported a correlation of $r = 0.37$ between the performance in a dynamic decision making task and supervisor ratings in a sample of $N = 73$ policemen. He further reported that this correlation remained significant after controlling for individual differences in general intelligence, $r = 0.29$, which points towards the incremental predictive validity of this dynamic decision making measure.

Taken together, these findings draw a rather heterogeneous picture of the relation between dynamic decision making and general intelligence and there is only preliminary evidence for the predictive validity of dynamic decision making variables.

1.1.2. Implicit learning

Mackintosh (1998) suggested to use artificial grammar learning tasks (Reber, 1967) to measure performance differences in implicit learning. In such a task, the participants are asked to learn a list of apparently arbitrary letter strings (like WNSNXS). Afterwards, they are told that these strings were constructed according to a complex rule system (a grammar) and they are asked to judge newly presented strings as grammatical or non-grammatical. Typically, the participants show above chance performance but are not able to report the grammar rules. Therefore, Reber (1967) suggested that the participants learned the grammar implicitly. Although Reber's interpretation released a long and fertile discussion about implicit learning processes, there have been only a few studies investigating the relation between performance in artificial grammar learning tasks and general intelligence.

Reber, Walkenfeld, and Hernstadt (1991) reported a correlation of $r = 0.25$ between the performance in an artificial grammar learning task and IQ, and Gebauer and Mackintosh (2007) reported respective correlations between $r = -0.03$ and $r = 0.17$ depending on the task and the instruction. To our knowledge, there is no published study investigating the relation between educational or professional success and the performance in an artificial grammar learning task. Thus, there is a paucity of evidence on the relation between implicit learning and general intelligence as well as on the relation between implicit learning and success in real life.

1.2. Some psychometric considerations

Previous studies that investigated the relation between general intelligence, dynamic decision making, and implicit learning treated the performance measures as trait-like variables. A trait may be defined as a variable that is stable over several measurement occasions, consistent across different situations, and consistent across different methods. However, the variance of a performance measure may capture additional factors beyond individual differences in a trait.

First, a performance measure may also be influenced by the specific measurement situation even in standardized experiments. For example, one person may be well rested whereas another person may already have worked several hours before testing. One person may be motivated to show maximum performance whereas another person may have gotten a stinging rebuke by his or her supervisor that day and may not be motivated to show performance at all. Because these effects may contribute unwanted variance, it may be beneficial to take this *occasion specificity* of performance variables into account.

Second, a performance measure may be influenced by the *specific method* that is used for the assessment. Hence, there may be individual differences in a performance measurement which are triggered by the method. For example, a verbal intelligence test may capture individual differences in general intelligence as well as individual differences in speech comprehension whereas a figural intelligence test may capture individual differences in general intelligence and visual thinking. Thus, individual differences in speech comprehension or visual thinking are method specific because they can only be assessed with verbal or figural test material. Similarly, a particular dynamic decision making task may measure performance differences, which are specific for this particular task but not for dynamic decision making in general.

Third, a performance measure may be influenced by *unsystematic measurement error*. For example, instructions may be ambiguous or persons may accidently make mistakes, which may result in a low reliability of performance measures. Because these effects may contribute unwanted variance, it seems worthwhile to investigate these factors with respect to dynamic decision making and implicit learning variables in greater detail.

These considerations have been formalized in Steyer et al.'s latent state-trait theory (Steyer, Schmitt, & Eid, 1999). In a nutshell, latent state-trait theory proposes that the measurement i of a variable Y can be decomposed into a trait ξ_i, a state residual ζ_i, a method residual η_i, and an unsystematic error residual ε_i, thus $Y_i = \xi_i + \zeta_i + \eta_i + \varepsilon_i$. Given the independence of these factors (Steyer et al., 1999), the variance of this measurement can be decomposed as $\sigma^2(Y_i) = \sigma^2(\xi_i) + \sigma^2(\zeta_i) + \sigma^2(\eta_i) + \sigma^2(\varepsilon_i)$, and the factor variances may be estimated with a structural equation model as shown in Fig. 1. As can be seen in this figure, the latent trait factor is defined as a variable that is consistent across several measurement occasions and methods, whereas the latent state residual and the method factor are specific for the individual measurement occasion and the assessment method, respectively. Hence, these models allow to separate the different contributions of the trait, the measurement occasion, and the measurement method to the manifest variables.

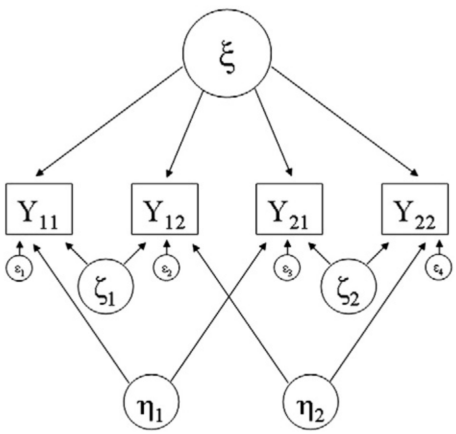

Fig. 1. Latent state-trait structural equation model. Y_{11} = variable at measurement occasion 1 with method 1, Y_{12} = variable at measurement occasion 1 with method 2, Y_{21} = variable at measurement occasion 2 with method 1, Y_{22} = variable at measurement occasion 2 with method 2, ξ = trait variable, ζ_1 = state residual 1, ζ_2 = state residual 2, η_1 = method residual 1, η_2 = method residual 2, ε_1 = error 1, ε_1 = error 2, ε_1 = error 3, ε_1 = error 4.

There have been many applications of latent state-trait models in different domains of personality research, which demonstrated substantial effects of the measurement occasion or the method on behavioral variables (e.g., Eid, Notz, Steyer, & Schwenkmezger, 1994; Schmitt & Steyer, 1993; Steyer, Schwenkmezger, & Auer, 1990; Yasuda, Lawrenz, Whitlock, Lubin, & Lei, 2004; Ziegler, Ehrlenspiel, & Brand, 2009) and physiological variables (e.g., Hagemann, Hewig, Seifert, Naumann, & Bartussek, 2005; Hermes et al., 2009). However, there have been no applications of latent state-trait models on performance variables yet, even if some findings suggest that it may be instructive to consider the occasion specificity and method specificity of these variables.

For example, in some studies the participants completed the same dynamic decision making task for several times (Süß, Kersting, & Oberauer, 1993; Wittmann & Hattrup, 2004) and the performance between subsequent task correlated only moderately (between $r = 0.37$ and $r = 0.62$). This points either towards a low reliability or towards a substantial occasion specificity of the variables. Moreover, Wirth and Klieme (2003) reported structural equation models, which implied a correlation of $r = 0.33$ between two dynamic decision making tasks ($r = 0.47$ when corrected for attenuation) and Gebauer and Mackintosh (2007) reported a correlation of $r = 0.15$ between two artificial grammar learning task ($r = 0.21$ when corrected for attenuation). These findings suggest a substantial method specificity of performance measures. Therefore, a further aim of the present study was to investigate the occasion specificity and the method specificity of dynamic decision making and implicit learning variables.

1.3. The present study

The present study investigated the psychometric properties of general intelligence, dynamic decision making, and implicit learning measures within the framework of latent state-trait theory. Therefore, each construct was measured with two methods at two measurement occasions. A further scope of this study was the relation between the respective trait variables and real life performance. We expected that general intelligence is a powerful predictor of professional success and we further expected that there are individual differences beyond IQ that are also able to predict professional success.

2. Method

2.1. Participants

There were $N = 173$ employees (113 females, 47 males, 13 not reported) completing the first measurement occasion and $N = 151$ completing the second measurement occasion. The participants were recruited via newspaper announcement from different branches and different companies around Heidelberg. The participants' jobs were rated according to the

International Standard Classification of Occupations (ISCO-88 COM). 6% rated themselves as legislators, senior officials, and managers, 25% as professionals, 11% as technical and associate professionals, 14% as clerks, 40% as service workers and shop and market sales workers, 1% as craft and related trade workers, 1% as plant and machine operators and assemblers, and 1% as elementary occupations. The participants' mean age was $M=43.34$ ($SD=11.22$).

2.2. Measures

2.2.1. Advanced progressive matrices (APM)

The APM (Raven, Court, & Raven, 1994) were used as an indicator for participants' general intelligence. A computer adapted version of the test was administered. According to the test manual, the number of solved items of the second set was taken as a performance indicator. These raw scores were transformed to *z*-scores for further analysis, because the APM and the Berlin Intelligence Structure Test were scaled differently.

2.2.2. Berlin intelligence structure test (BIS)

The short version of the BIS (Jäger, Süß, & Beauducel, 1997) was used as a second indicator of general intelligence. The BIS consists of a variety of tasks like an analogical reasoning task, a visual memory task, and a numerical series task (for an English description, see Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002). The test was administered and the raw scores were computed according to the test manual. We did not compute IQ scores because there is no adult normative sample for the BIS. For further analysis the raw scores were transformed to *z*-scores.

2.2.3. Artificial grammar learning tasks

Implicit learning was measured with two artificial grammar learning tasks (Reber, 1967). The procedure and the stimuli were adopted from Gebauer and Mackintosh (2007). The artificial grammar learning tasks consisted of a learning phase and a testing phase. In the *learning phase,* 30 letter strings were presented and the participants were instructed to memorize them. Each string was presented individually for 3 s on a 17 in. screen of a personal computer (e.g., WNSNXS). The participants were asked to repeat the strings correctly by pressing the respective letters on the keyboard. When a string was repeated correctly, the feedback "correct" was given and the next string occurred. When a string was repeated incorrectly, the feedback "false" was given and the string was displayed again until repeated correctly. After a participant repeated ten strings correctly, these ten strings were simultaneously displayed for 90 s on the screen and the participant was asked to repeat them silently. After a participant repeated all 30 strings correctly the learning phase was finished and the participant was informed that all strings in the learning phase were constructed according to a complex rule system. In the *testing phase,* 80 new strings were presented (see Appendix A). There were 40 grammatical strings that were constructed according to the same rule system as the strings in the learning phase (e.g., WNSWWW). In addition, there were 40 non-grammatical strings that contained one letter at a position that violated the rule system (e.g., NTSWWN). The participants were instructed to judge the letter strings as grammatical or non-grammatical. To judge a string as grammatical, the participants had to press the A-key of the keyboard, to judge a string as non-grammatical, the L-key. The order of presentation of the strings was fixed across the participants in a random order. The percentage of correct judgments in the testing phase was taken as the performance indicator. The stimuli for the first artificial grammar learning task were constructed according to Fig. 2. The stimuli for the second artificial grammar learning task were constructed according to Fig. 3.

2.2.4. Tailorshop

The Tailorshop simulation (Funke, 1983) was used as a dynamic decision making task. The Tailorshop is a computer based scenario and requests the participants to lead a fictional company which produces and sells shirts for twelve simulated months. Several variables can be manipulated like the number of workers, the expenses for advertising etc. (see Fig. 4). The state of a variable in a given month influences the state of the same and other variables in the following month

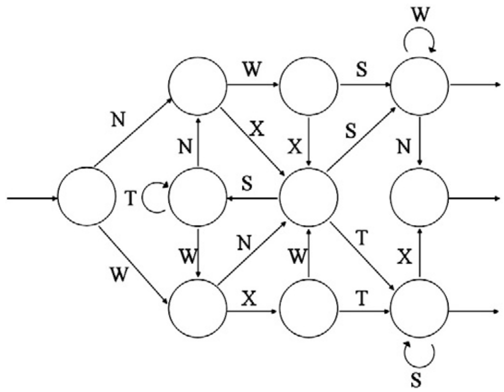

Fig. 2. Grammar 1 that was used in the first artificial grammar learning task.

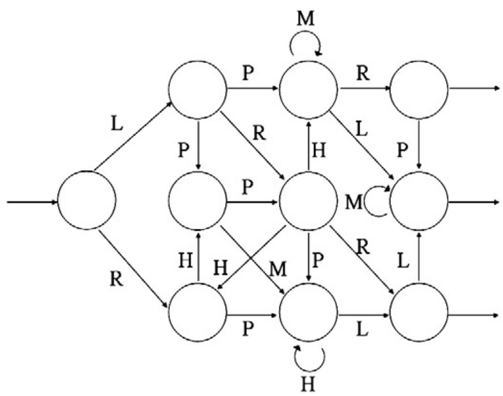

Fig. 3. Grammar 2 that was used in the second artificial grammar learning task.

but the participants do not know how the variables are connected (for a more detailed description see Funke, 1983, 2010). The participants completed a training phase, a knowledge test, and a control phase. In the training phase the participants controlled the system for six simulated months and were instructed to find out as much as possible about the scenario. The knowledge test consisted of twelve questions that measured how much the participants learned about the Tailorshop so far. In the control phase the participants were instructed to maximize their company value during twelve simulated months. For the purpose of the present study only data from the control phase were analyzed. The percentage of months with an increase in the company value between the second and the twelfth month was taken as the performance indicator, because Danner et al. (2011) have shown that this is a reliable and valid performance indicator.

2.2.5. Heidelberg finite state automaton (HFA)

The HFA (Wirth & Funke, 2005) was taken as a second indicator for dynamic decision making. The scenario is computer based and simulates a space flight where the participants can control a space ship and a vehicle with a user interface (see Fig. 5). The scenario consists of a training phase, a knowledge test, and a control phase. During the 15 minute training phase the participants were instructed to find out how to control the space ship and the vehicle. The knowledge test consists of 16 items and measures how much the participants have learned about the system so far. The control phase consists of 22 items where a target state is given which the participants have to reach by controlling the system (e.g., landing the space ship on a specified planet). For the purpose of the present study, only data from the control phase were analyzed. The percentage of correctly solved items was taken as the performance indicator.

2.2.6. Professional success

The participants' professional success was measured with two instruments. *Objective professional success* was measured by the participants' income (thirteen categories), self-rated social status (seven categories), and the participants' highest educational attainment (nine categories). To adjust for different scaling, the three variables were z-transformed ($M = 0$, $SD = 1$) for further analysis. In addition, professional success was measured by *supervisor ratings* with five items (e.g., “The employee demonstrates competence in all job-related tasks”) on a six-point Likert scale.

2.3. Procedure

There were two measurement occasions. The first measurement occasion started in July 2009 (till September 2009) and consisted of session 1 and session 2. Both sessions took place within one week for each participant. The second measurement occasion started in December 2009 (till February 2010) and consisted of session 3 and session 4, which also took place within one week. The participants were assessed in small groups of not more than four persons. Each session took approximately 2.5 h.

The participants completed the same tasks at both measurement occasions. During session 1 (and session 3) the participants completed an artificial grammar learning task with grammar 1, the APM, and the Heidelberg Finite State Automaton. During session 2 (and session 4), the participants completed an artificial grammar learning task with grammar 2, the short version of the BIS, and the Tailorshop simulation. After the first session, each participant received an envelope with a questionnaire for his or her supervisor. During the third session, the participants additionally completed a questionnaire about their professional success.

2.4. Statistical analysis

To investigate the relations between the variables, we used structural equation models. The parameters of the models were estimated using the maximum likelihood

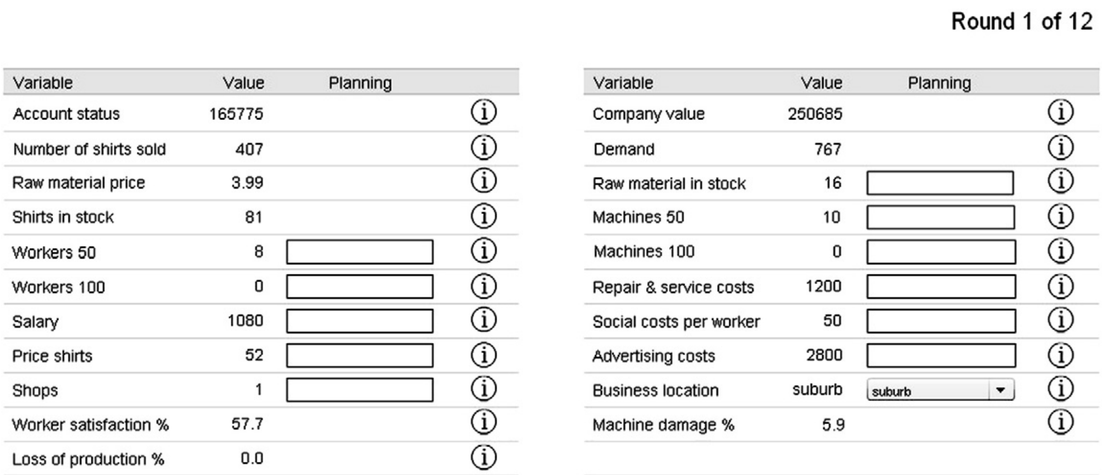

Fig. 4. Screenshot of the graphical user interface of the Tailorshop (labels translated).

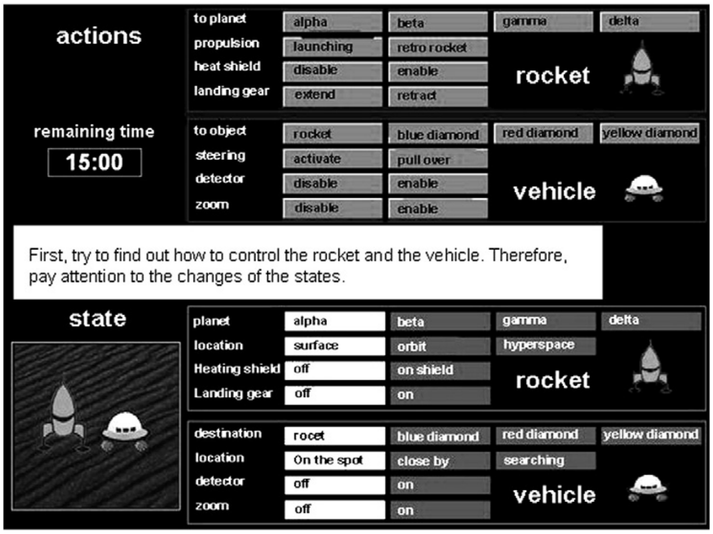

Fig. 5. Screenshot of the graphical user interface of the Heidelberg Finite State Automaton (labels translated).

algorithm implemented in Amos 18 (Arbuckle, 2006). In a first step, we investigated latent state-trait measurement models separately for intelligence, dynamic decision making, and implicit learning. In a second step, we investigated the correlation between the latent trait variables. In a third step, we performed a latent regression analysis to investigate relations between the constructs in greater detail.

3. Results

3.1. Raw scores

The raw scores of the measurements are reported in Table 1. The number of solved items in the Advanced Progressive Matrices at the first measurement occasion was $M=21.64$ (SD $=5.80$), which corresponds to an IQ of $M=100.62$ ($SD=22.55$). There are no normative samples for the Berlin Intelligence Structure Test, the Tailorshop, the Heidelberg Finite State Automaton, or the artificial grammar learning tasks. However, the present scores are similar to previous results. The mean score of the BIS was $M=96.30$ ($SD=6.21$) at the first measurement occasion and $M=99.21$ ($SD=6.38$) at the second measurement occasion. According to Jäger et al. (1997), a mean score of $M=100$ corresponds to an average performance. In the present study, the participants solved $M=10.79$ ($SD=5.80$) HFA items at the first measurement occasion and M $=13.44$ ($SD=5.95$) HFA items at the second measurement occasion. This result is similar to Wirth and Klieme (2003), who reported that their participants solved $M=11$ HFA items on average. The judgment accuracy in the artificial grammar learning tasks varied between $M=61.58$ ($SD=7.11$) and $M=63.90$ ($SD=7.24$), which corresponds to the findings of Gebauer and Mackintosh, who reported mean accuracies between $M=59.16$ ($SD=8.59$) and $M=69.93$ ($SD=7.52$) for the same artificial grammar learning tasks that were used in the present study.

3.2. Measurement models

We used a basic latent state-trait model (Steyer et al., 1999) with a state residual ζ for each measurement occasion and a method factor η for each instrument to control for effects of the measurement occasion and method effects (see Fig. 1). All path coefficients were fixed to one and the variances of all latent variables were estimated. If a first estimation revealed negative or non-significant variances, then these variances were fixed to zero and the model was estimated again.

3.2.1. Intelligence

A first analysis of the basic model revealed a good model fit, $\chi^2(1)=0.30$, $p=0.569$, RMSEA $=0.00$, CFI $=1.00$. However,

Table 1
Mean and standard deviation of raw scores.

	Measurement occasion 1		Measurement occasion 2	
Task	*M*	*SD*	*M*	*SD*
APM	21.64	5.80	22.94	7.02
BIS	96.30	6.21	99.21	6.38
Tailorshop	2.68	3.21	3.15	3.77
HFA	10.79	5.80	13.44	5.95
AGL1	61.58	7.11	62.83	6.87
AGL2	63.90	7.24	62.20	7.70

Note. APM = number of solved items in the Advanced Progressive Matrices, BIS = scores in Berlin Intelligence Structure Test, Tailorshop = number of months with an increase in the company value, HFA = number of items solved in the Heidelberg Finite State Automaton, AGL1 = percent of correct judgments in the artificial grammar learning task with grammar 1, AGL2 = percent of correct judgments in the artificial grammar learning task with grammar 2.

the estimated variance for ζ_1 was negative ($\zeta_1 = -14.14$, $p = 0.016$), and the estimated variance for ζ_2 was not significant ($\zeta_2 = 9.60$, $p = 0.125$). Therefore, these parameters were set to zero and the model was estimated again. The modified model fitted the data well, $\chi^2(3) = 5.59$, $p = 0.133$, RMSEA = 0.07, CFI = 1.00, and the difference in the fit of the models was not significant, $\Delta\chi^2(2) = 4.29$, $p = 0.117$. Therefore, this model could be accepted. The estimated model parameters are reported in Table 2.

3.2.2. Dynamic decision making

The basic latent state-trait model fitted well with the data, $\chi^2(1) = 0.9$, $p = 0.335$, RMSEA = 0.00, CFI = 1.00. However, the latent state residuals were negative ($\zeta_1 = -49.20$, $p = 0.183$) or non-significant ($\zeta_1 = 48.30$, $p = 0.257$). The modified model without latent state residuals also fitted well with the data, $\chi^2(3) = 3.25$, $p = 0.355$, RMSEA = 0.02, CFI = 1.00; $\Delta\chi^2(2) = 2.35$, $p = 0.309$. Thus, this model could be accepted. The estimated model parameters are presented in Table 2.

3.2.3. Implicit learning

The basic latent state-trait model fitted well with the data, $\chi^2(1) = 0.13$, $p = 0.719$, RMSEA = 0.00, CFI = 1.00. However, the variances of the latent state residual and the latent method variables were non-significant ($\zeta_1 = 6.57$, $p = 0.128$; $\zeta_2 = 2.35$, $p = 0.585$; $\eta_1 = -0.06$, $p = 0.988$; $\eta_2 = -4.96$, $p = 0.250$). Therefore, these variances were set to zero. This modified model fitted the data well, $\chi^2(5) = 3.19$, $p = 0.671$, RMSEA = 0.00, CFI = 1.00; $\Delta\chi^2(4) = 3.06$, $p = 0.548$, and this model was accepted. The estimated model parameters are presented in Table 2.

3.2.4. LST parameters

Based on these estimates, several latent state-trait parameters may be computed such as coefficients of reliability, trait-specificity (also referred to as consistency), occasion-specificity, and method-specificity. These parameters have a range between zero and one, and a greater value indicates a greater specificity. The reliability coefficient of a measurement i reveals how great the proportion of systematic variance in this measurement is. It is computed as $[\sigma^2(\xi_i) + \sigma^2(\zeta_i) + \sigma^2(\eta_i)] / \sigma^2(Y_i)$. The trait-specificity coefficient of a measurement i reveals how great the proportion of trait differences in a measurement is. It may be computed as $\sigma^2(\xi_i)/\sigma^2(Y_i)$. The occasion-specificity coefficient of a measurement i indicates the effects of the situation and the interaction between the situation and the person on the measurement. It may be computed as $\sigma^2(\zeta_i)/\sigma^2(Y_i)$. The method-specificity coefficient of a measurement i reveals how great the proportion of individual differences is due to the method (e.g., task) used. This coefficient is computed as $\sigma^2(\eta_i)/\sigma^2(Y_i)$.

These parameters are presented in Table 3. As can be seen, the general intelligence measurements revealed great reliabilities, great trait-specificities, and low method-specificities. The Heidelberg Finite State Automaton measurements also showed great reliabilities, but smaller trait-specificities and greater method-specificities. The Tailorshop measurements revealed small reliabilities and small trait-specificities. All implicit learning measurements revealed very small reliabilities and trait-specificities. Since all measurement models fitted well without state residuals, the estimated occasion-specificity was zero for all measurements.

3.2.4. Professional success

Objective professional success was measured with three indicators at session 3. A measurement model with one latent success variable, equal path coefficients ($\beta = 1$), and a latent error variable for each manifest variable was specified. The model fitted the data well, $\chi^2(2) = 2.46$, $p = 0.293$, RMSEA = 0.04, CFI = 0.98. Therefore, this model was accepted. The composite reliability (Raykov, 1997) of the items' mean score was 0.71. The participants' *supervisor ratings* were measured with a five item questionnaire. A measurement model with one latent success variable, equal path coefficients ($\beta = 1$), and a latent error variable for each manifest variable fitted the data well, $\chi^2(9) = 11.93$, $p = 0.217$, RMSEA = 0.04, CFI = 0.99. Thus, this model was accepted. The composite reliability of the items' mean score was 0.95.

Table 2
Estimated variances for measurement models (p-values in brackets).

	Intelligence	Dynamic decision making	Implicit learning
ξ	0.73 (<0.001)	317.12 (<0.001)	14.87 (<0.001)
ζ_1	0 (fixed)	0 (fixed)	0 (fixed)
ζ_2	0 (fixed)	0 (fixed)	0 (fixed)
η_1	0.14 (0.015)	144.66 (0.046)	0 (fixed)
η_2	0.24 (<0.001)	257.37 (<0.001)	0 (fixed)
ε_1	0.14 (<0.001)	425.17 (<0.001)	35.92 (<0.001)
ε_2	0.18 (<0.001)	637.27 (<0.001)	33.38 (<0.001)
ε_3	0.11 (<0.001)	146.00 (<0.001)	35.29 (<0.001)
ε_4	0.06 (0.014)	145.21 (<0.001)	43.83 (<0.001)

Note. ξ = trait variable, ζ_1 = state residual 1, ζ_2 = state residual 2, η_1 = method residual 1, η_2 = method residual 2, ε_1 = error 1, ε_2 = error 2, ε_3 = error 3, ε_4 = error 4. The different scaling of the variables affects the magnitude of the variances estimates.

Table 3
Reliability, trait- and method-specificity of measurements.

Task	Measurement occasion	Reliability	Trait-specificity	Method-specificity
APM	1	0.86	0.72	0.14
APM	2	0.83	0.70	0.13
BIS	1	0.90	0.67	0.22
BIS	2	0.95	0.71	0.24
Tailorshop	1	0.52	0.36	0.16
Tailorshop	2	0.42	0.29	0.13
HFA	1	0.80	0.44	0.36
HFA	2	0.80	0.44	0.36
AGL1	1	0.29	0.29	0.00
AGL1	2	0.31	0.31	0.00
AGL2	1	0.30	0.30	0.00
AGL2	2	0.25	0.25	0.00

Note. APM = Advances Progressive Matrices, BIS = Berlin Intelligence Structure Test, HFA = Heidelberg Finite State Automaton, AGL1 = artificial grammar learning task with grammar 1, AGL2 = artificial grammar learning task with grammar 2.

3.3. Relations between intelligence, dynamic decision making, implicit learning, and professional success

We specified an omnibus model, which simultaneously tested all measurement models described above and allowed free correlations between the latent trait variables and the latent professional success variables. The specified model revealed a good model fit, $\chi^2(174) = 197.74$, $p = 0.105$, RMSEA = 0.03, CFI = 0.98 and thus was accepted. The correlations between the latent variables are shown in Table 4. As can be seen, there were significant and substantial correlations between all performance variables. The greatest correlation was between intelligence and dynamic decision making, $r = 0.86$, $p < 0.001$. There was also a correlation of $r = 0.78$, $p < 0.001$ between objective professional success and general intelligence. There were further substantial correlations between objective professional success and dynamic decision making, $r = 0.52$, $p < 0.001$, and between objective professional success and implicit learning, $r = 0.31$, $p = 0.030$. The only significant correlation with supervisor ratings was the correlation with dynamic decision making, $r = 0.25$, $p = 0.021$.

3.4. Prediction of objective professional success

To investigate the relation between performance variables and objective professional success in greater detail, we specified a latent regression model according to Fig. 6. As can be seen, dynamic decision making, implicit learning, and professional success were regressed on intelligence. The residuals of this regression are the proportions of trait variances which are independent from general intelligence. The dynamic decision making and implicit learning residuals were used to predict the proportion of construct variance in objective professional success that could not be explained by general intelligence.

The specified model revealed a good model fit, $\chi^2(95) = 114.44$, $p = 0.085$, RMSEA = 0.03, CFI = 0.98. The standardized path coefficients are shown in Fig. 6. As can be seen, dynamic decision making as well as implicit learning revealed trait variances, which were independent from general intelligence. In addition, general intelligence was the only significant predictor of objective professional success. Neither the path coefficient from the residual dynamic decision variable to the residual professional success variable, nor the path coefficient from the residual implicit learning variable to the residual professional success variable was significant. Therefore, these path coefficients were set to zero and the model was estimated again. The modified model also revealed a good model fit, $\chi^2(97) = 117.62$, $p = 0.076$, RMSEA = 0.04, CFI = 0.98; $\Delta\chi^2(2) = 3.18$, $p = 0.204$. Thus, this model was accepted.

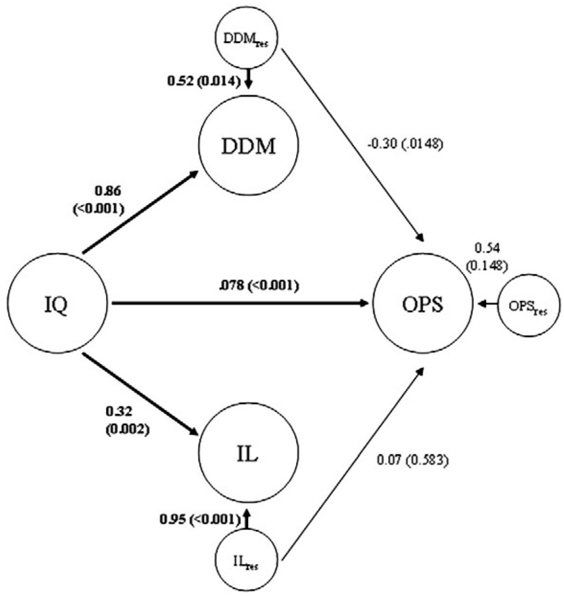

Fig. 6. Latent Regression Analysis with standardized path coefficients (*p*-values in brackets). IQ = latent general intelligence variable, DDM = latent dynamic decision making variable, IL = latent implicit learning variable, OPS = latent objective professional success variable, DDM_{res} = latent residual for dynamic decision making, IL_{res} = latent residual for implicit learning, OPS_{res} = latent residual for professional success.

3.5. Prediction of supervisor ratings

The relations between general intelligence, dynamic decision making, implicit learning, and supervisor ratings were investigated analogously to the analysis described above. The specified model fitted the data well, $\chi^2(126) = 125.86$, $p = 0.487$, RMSEA = 0.00, CFI = 1.00. The standardized path coefficients are shown in Fig. 7. As can be seen, dynamic decision making was the only significant predictor of participants' supervisor ratings. Neither the path coefficient from the general intelligence variable, nor the path coefficient from the residual implicit learning variable was significant. A modified model, which fixed these parameters to zero, revealed an adequate model fit, $\chi^2(128) = 126.00$, $p = 0.533$, RMSEA = 0.00, CFI = 1.00; $\Delta\chi^2(2) = 0.14$, $p = 0.932$. Therefore, this model was accepted.

Table 4
Correlation between latent success and latent trait variables (p-values in brackets).

	Intelligence	Dynamic decision making	Implicit learning	Objective professional success
Dynamic decision making	0.86 (<0.001)			
Implicit learning	0.32 (0.005)	0.26 (0.033)		
Objective professional success	0.78 (<0.001)	0.52 (<0.001)	0.31 (0.030)	
Supervisor ratings	0.03 (0.760)	0.25 (0.021)	−0.02 (0.871)	−0.07 (0.559)

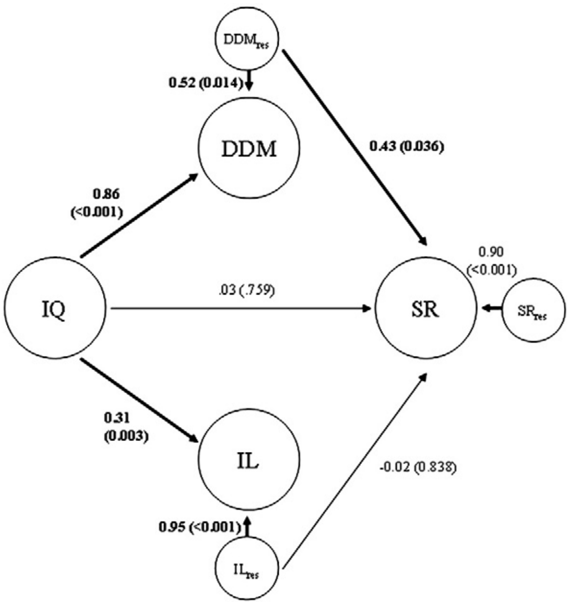

Fig. 7. Latent Regression Analysis with standardized path coefficients (*p*-values in brackets). IQ = latent general intelligence variable, DDM = latent dynamic decision making variable, IL = latent implicit learning variable, SR = latent supervisor rating variable, DDM$_{res}$ = latent residual for dynamic decision making, IL$_{res}$ = latent residual for implicit learning, SR$_{res}$ = latent residual for supervisor ratings.

4. Discussion

The present study investigated Dörner's (1980) and Mackintosh's (1998) hypotheses that dynamic decision making and implicit learning are cognitive abilities that are independent from general intelligence.

In a first step, we analyzed the psychometric properties of intelligence variables, dynamic decision making variables, and implicit learning variables within the framework of latent state-trait theory. All measurement models fitted well without latent state residuals. This indicates that the performance measures were not affected by situational factors such as individual differences in fatigue or individual differences in the form of the day. Furthermore, the general intelligence variables revealed high trait specificities and low method specificities, which indicate a high proportion of trait differences in these performance measures. The dynamic decision making and implicit learning variables, on the other hand, revealed lower trait specificities and greater method specificities, which suggests that these variables capture task specific performance differences as well. However, even if the trait specificities were small, the variances of the latent trait variables were still significant. This indicates that there are true individual trait differences in dynamic decision making and implicit learning.

In a second step, we analyzed the relations between these latent trait variables. The present results suggest that there are substantial relations between general intelligence, dynamic decision making, and implicit learning. In particular, there was a great correlation ($r = 0.86$) between the latent general intelligence variable and the latent dynamic decision making variable. This result goes in line with previous findings of Wirth and Klieme (2003), Wittmann and Hattrup (2004), and Kröner et al. (2005) who also reported great relations between measures of dynamic decision making and measures of general intelligence. Taken together, these findings contradict Dörner's hypothesis that dynamic decision making and general intelligence are independent variables.

The correlation between the latent implicit learning variable and the latent general intelligence variable was of medium size ($r = 0.32$). This goes in line with the findings of Reber et al. (1991) and Gebauer and Mackintosh (2007) who also reported low to medium correlations between measures of implicit learning and general intelligence. This finding does not support Mackintosh's hypothesis that implicit learning and general intelligence are independent constructs. However, general intelligence could only explain 10.24% of the implicit learning trait variance, which suggests that there are substantial individual differences in implicit learning beyond IQ.

Taken together, this pattern of result suggests that there are substantial relations between cognitive performance measures, which have been developed within very different domains. Measures of general intelligence have a long research tradition and were developed to measure persons' general mental ability. Measures of dynamic decision making arose in the domain of complex problem solving and were designed to explore persons' ability to deal with realistic problems. And measures of implicit learning were developed in the domain of cognitive psychology in order to study persons' ability in making intuitive decisions. The present findings suggests that these performance measures share a substantial proportion of common variance but also reveal variance proportions that are independent from each other. This fits well with hierarchical intelligence models like Carroll's (1993) three-stratum theory of cognitive abilities. In particular, Carroll suggested that the structure of human cognitive abilities may be explained by a hierarchical structure with three levels (three strata). On the lowest level (stratum 1) there are 64 different specific ability factors like reading comprehension, memory span, or general sound discrimination. According to Carroll, these specific abilities are not independent and therefore may be grouped together to eight more general ability factors (stratum 2), which are fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, broad auditory perception, broad retrieval ability, broad cognitive speediness, and processing speed. On the top of the hierarchy (stratum 3) there is a single general ability factor that explains the correlation between the stratum 2 factors. In Carroll's model there are no ability factors such as dynamic decision making or implicit learning. Accordingly, these constructs may be seen as supplementary aspects of human cognitive ability. However, the present results fit well with the concept of a hierarchical structure of human cognitive ability. In particular, the results of the structural equation models revealed that the overlap between the performance in the Tailorshop and the Heidelberg Finite Automaton may be explained by a more general dynamic decision making ability factor. In the same vein, the overlap between the different artificial grammar learning tasks could be explained by an implicit learning ability factor. Furthermore, there were substantial correlations between general intelligence, dynamic decision making and implicit learning that could be explained by one single general ability factor. Taken together, these results suggest that dynamic decision making and implicit learning may be supplementary abilities that fit well into a hierarchical concept of human cognitive ability. However, the present findings do not sufficiently allow to draw a conclusion on which stratum these ability factors may be located. Investigating this may be an interesting issue for future research.

In a third step, we analyzed whether dynamic decision making and implicit learning are powerful predictors of professional success beyond IQ. The zero correlation between objective professional success and supervisor ratings ($r = 0.07$) suggests that both variables capture different aspects of professional success. One reason for this may be that income, social status, and education attainment are rather profit-based indicators, whereas supervisor ratings may also capture social aspects. According to this, both aspects were analyzed separately.

There were substantial correlations between *objective professional success* and dynamic decision making ($r = 0.52$) as well as between objective professional success and implicit learning ($r = 0.31$). This suggests that both performance measures are able to predict objective professional success. However, when general intelligence was included as a predictor, then general intelligence remained the only significant predictor ($\beta = 0.78$). This finding is consistent with the literature and emphasizes the meaningfulness and usefulness of IQ measures (e.g., Schmidt & Hunter, 2004).

There was a substantial relation between the participants' *supervisor ratings* and dynamic decision making even when general intelligence was simultaneously considered ($\beta = 0.43$). This replicates findings of Kersting (2001) who also reported an incremental predictive value of dynamic decision making measures on participants' supervisor ratings. Furthermore, this result points towards the practical value of dynamic decision making measures and suggests that dynamic decision making measures may provide insights into aspects of professional success, which cannot be predicted by general intelligence. Therefore, Dörner's hypothesis that dynamic decision making has an incremental predictive value is partially supported. The relation between supervisor ratings and implicit learning was close to zero ($r = -0.02$) and not significant. Thus, this result may be seen as preliminary evidence against Mackintosh's hypothesis that implicit learning is a useful predictor of professional success. There was no significant correlation between supervisor ratings and general intelligence. At first sight, this finding is astonishing because there is a wealth of evidence for the relation between general intelligence and supervisor rating (e.g., Ng et al., 2005; Salgado et al., 2003; Schmidt & Hunter, 2004). However, the samples in these studies typically consist of employees within a single department or company whereas the sample in the present study consisted of employees of different companies and occupational groups. In particular, there may be a relation between general intelligence and supervisor ratings within single companies or occupational groups but not between. For example, a broker with an IQ of 130 may be rated as more successful than a broker with an IQ of 100 but a journalist with an IQ of 130 may still be rated as less successful than the broker with the IQ of 100.

4.1. Implications for assessment

The present results show that the APM as well as the Berlin Intelligence Structure Test yield measures with good trait specificities (0.67 to 0.72). Furthermore, there was a strong relation ($r = 0.78$) between general intelligence and objective professional success. Therefore, general intelligence tests seem to be a good choice for measuring cognitive ability.

There was also a relation between the dynamic decision making trait variable and objective professional success ($r = 0.52$) and between the dynamic decision making trait variable and supervisor ratings ($r = 0.25$). However, the performance measures of the Tailorshop simulation and the Heidelberger Finite State Automaton showed trait specificities between 0.29 and 0.44. This suggests that less than half of the variance in these performance measures is due to trait differences in dynamic decision making. Therefore, the trait-specificity of both tasks should be improved before they are used for an individual assessment. A more theory-orientated development of dynamic decision making tasks may help to reach this goal.

There was a relation of medium size between the implicit learning trait variable and objective professional success ($r = 0.31$). However, the latent regression analysis revealed that this relation was due to an overlap with general intelligence. This suggests that there is no incremental predictive value of implicit learning measures. The trait specificities of the artificial grammar learning measures were between 0.25 and 0.31. There was no method specificity of these variables, which suggests that the low trait specificity was due to unsystematic measurement error. Therefore, lengthening the test may help to enhance the trait-specificity. However, whether such an approach increases the reliability or rather causes fatigue effects is an open issue.

5. Conclusion

The present findings acknowledge the overall approval and usefulness of general intelligence measures. In addition, the results demonstrated that there are significant individual trait differences in cognitive performance beyond IQ. In particular, there was a large proportion of trait variance in implicit learning, which was independent from general intelligence and in addition, dynamic decision making revealed an incremental predictive validity. These findings make dynamic decision making as well as implicit learning attractive for the research of individual differences.

Appendix A

Table A1
Letter strings for grammar 1 sorted for different parts of the assessment.

Phase	Strings
Learning phase	WNSNXS NXSTWXT WNSTTWXT WXWSNXT NWXTS NXSNWXS WNTSSS NXSWXTX WXWSNWSN WNSNWXS NXSWXT NXSWWWW NWSWN WNSNWXTX NXSWNTX NXS WNSWWWW WXWSWN NXSTNWS WNSTWNSW WNSWXTX WNTSSX WNSNWSW WNTX NXSWNSW WNSNXTS NXTSSS NXSNWSN WNSNWSN NXSWNTSS
Testing phase (correct items)	WNSN NWSW NWSN NXSWW NWXSW NXTSX WNSWNS NXSWNT NWXTSS WNSWWW WNSWNT NXTSSX WXWTSX WNSNXT NWSWWN NXSNWS NXSNWXT NXTSSSX WNSTWNS NXSTNXS WNTSSSX WNSWXWT NXSNXTX WXWSWXT NWXSWNS NWXSNWS NXSTXWNT WNSTNWXT WXWSNWXS NXSTWNTX WXWSTWXT WXWSNWSW WXWSWXTX NXSTNWSN NWXSWXWS WXWTSSSS WNSTNXTX WXWSWNSW NWXSWXTX NXSNXSWN
Testing phase (incorrect items)	TXSWNT TWXTSX NTSWWN WWSWNS WNWWNT NWSXWN NWXSSW WXWTST TXWTSSX SWXSWNS WSSWWWN WSSWXTS NWWSWXT NXNTNXS WNTTSSX NWXWSSX NWXSNTS WNSNXXX WXWSWST WNSWXWN WXWSWNW XNSTWNTS TWXSWXWS TWSWWWWN NNXSWXWT WSSTTNXT WNNWNSWW WNNNWXSW NWXWWXTX WXWXWXTX WXWSXWXT WXWSNWSW NXSWWXWN WXWSWNWW WXWSNWNS

Table A2
Letter strings for grammar 2 sorted for different parts of the assessment.

Phase	Strings
Learning phase	LRHMMLM LRPHLLMM RHPHR RHPHMMLM LRHL LPMHLLMM LPPHLM RHPRLMMM LRHMRP RHPHMMRP LPPPLL RPHHHLLM RHPHL LPPRLMMM LPR LRHRPMMM LPPRL LPMMRPMM RHPHRP LPMHHLLM LPMMRP RHMHLLMM LPLM RPHHHHLL LRR LRRLMMM RHMHHL LPPRLMM RPLLMMM RHPHLMM
Testing phase (correct items)	LRPHHHL RHMHHHL LRHMLMMM LPRP LPRPMM LPRPMMM LPLMMMM RHPHMMML LPMR LPMRPM RPHHHLL LPPHMLM LPPHMMRP RPHL LRHLMM RHPHMML RHPHLMMM LPMLLMMM LPLMM LRHMML RPHLLMM LPMHHHLL LPPHLMMM RHPHLM LPPHMML RHMLLMM RPLLMMMM LPPPHLLM LPMMML LPMLLMM RHMHLLM RHPHMLMM LPRPMMMM LPLMMM RHPHMLM RHPRLMM LRPHHHLL RHPPHLLM
Testing phase (incorrect items)	RPRL LLRPMM RHHPHHL RHMHHPL LRPHMHLL LPLR LPPMRP LPHMMR RHPRLMH LPMHHPLL HHMLL LPLRMM RPPLLMM PPLLMMMM LRHMMHPM LPHHL LPMMHM LPPLRPM PHPHMMML LPPHLMHM LPPLL RPHHPL RPHHRLL MPPHMMRP LPPPHLRM LPLMP LPMMMP LPPHPML LPHMMMRP RHMHHLLP HRHLMM MHPPHLL LPPHPRP LPPMMRPM LPMLLMMP LLMHHL RMPPLLM LPPHMHM LPPLHHHL RHPPHLLL

Table A3
Correlations between the manifest variables.

	APM1	APM2	BIS1	BIS2	Tailor1	Tailor2	HFA1	HFA2	AGL1	AGL2	AGL3	AGL4	Income	Status	Education
APM2	0.83***														
BIS1	0.66***	0.65***													
BIS2	0.70***	0.69***	0.91***												
Tailor1	0.32***	0.29***	0.25**	0.28***											
Tailor2	0.33***	0.25**	0.30***	0.30***	0.48***										
HFA1	0.57***	0.54***	0.55***	0.54***	0.32***	0.41***									
HFA2	0.60***	0.59***	0.52***	0.56***	0.39***	0.43***	0.79***								
AGL1	0.24**	0.29***	0.16*	0.27**	0.11	0.04	0.19*	0.13							
AGL2	0.05	0.11	0.12	0.12	0.04	0.03	0.03	0.03	0.36***						
AGL3	0.05	0.08	0.13	0.10	0.10	0.07	0.20*	0.16*	0.27**	0.27***					
AGL4	0.12	0.16	0.16*	0.14	0.06	0.08	0.11	0.09	0.23**	0.32***	0.29***				
Income	0.23**	0.11	0.21**	0.22**	0.09	0.14	0.17*	0.21**	0.01	0.17*	0.13	−0.04			
Status	0.31***	0.23**	0.21**	0.29***	0.14	0.12	0.24**	0.23**	0.12	0.05	0.08	0.04	0.29***		
Education	0.43***	0.42***	0.47***	0.48***	0.11	0.05	0.18*	0.14	0.19*	0.16*	0.09	0.02	0.15	0.28*	
Supervisor	0.02	0.00	0.03	0.00	0.20*	0.12	0.14	0.18*	0.00	−0.08	0.04	0.03	0.01	−0.07	−0.07

Note. $^{*}p<0.050$, $^{**}p<0.010$, $^{***}p<0.001$, APM1 = Advances Progressive Matrices at measurement occasion 1, APM2 = Advances Progressive Matrices at measurement occasion 2, BIS1 = Berlin Intelligence Structure Test at measurement occasion 1, BIS2 = Berlin Intelligence Structure Test at measurement occasion 2, Tailor1 = Tailorshop at measurement occasion 1, Tailor2 = Tailorshop at measurement occasion 2, HFA1 = Heidelberg Finite State Automaton at measurement occasion 1, HFA2 = Heidelberg Finite State Automaton at measurement occasion 2, AGL1 = artificial grammar learning task with grammar 1 at measurement occasion 1, AGL2 = artificial grammar learning task with grammar 2 at measurement occasion 1, AGL3 = artificial grammar learning task with grammar 1 at measurement occasion 2, AGL4 = artificial grammar learning task with grammar 2 at measurement occasion 2, Income = participants' yearly income, Status = participants' self rated social status, Education = participants' educational level, Supervisor = participants' supervisor ratings, N varied between $N=173$ and $N=151$ due to dropouts between the first and the second measurement occasion.

References

Arbuckle, J. L. (2006). *Amos 7.0 user's guide.* Chicago: SPSS.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Danner, D., Hagemann, D., Holt, D., Bechtold, M., Schankin, A., Wüstenberg, S., & Funke, J. (2011). *Measuring performance in dynamic decision making: reliability and validity of the Tailorshop simulation.* : Journal of Individual Differences. doi:10.1027/1614-0001/a000055.

Dörner, D. (1980). On the difficulty people have in dealing with complexity. *Simulation & Gaming*, *11*, 87–106.

Dörner, D. (1986). Diagnostik der Operativen Intelligenz. *Diagnostica*, *32*, 290–308.

Eid, M., Notz, P., Steyer, R., & Schwenkmezger, P. (1994). Validating scales for the assessment of mood level and variability by latent state-trait analyses. *Personality and Individual Differences*, *16*, 63–76. doi:10.1016/0191-8869(94)90111-2.

Funke, J. (1983). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica*, *29*, 283–302.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*, 133–142. doi:10.1007/s10339-009-0345-0.

Gebauer, G. F., & Mackintosh, N. J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*, 34–54. doi:10.1037/0278-7393.33.1.34.

Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior*, *21*, 273–286. doi:10.1016/j.chb.2004.02.014.

Hagemann, D., Hewig, J., Seifert, J., Naumann, E., & Bartussek, D. (2005). The latent state-trait structure of resting EEG asymmetry: Replication and extension. *Psychophysiology*, *42*, 740–752. doi:10.1111/j.1469-8986.2005.00367.x.

Hermes, M., Hagemann, D., Britz, P., Lieser, S., Bertsch, K., & Naumann, E. (2009). Latent state-trait structure of cerebral blood flow in a resting state. *Biological Psychology*, *80*, 196–202. doi:10.1016/j.biopsycho.2008.09.003.

Jäger, A. O., Süß, H. -M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur — Test. Form 4.* Göttingen: Hogrefe.

Kersting, M. (2001). Zur Konstrukt- und Kriteriumsvalidität von Problemlöseszenarien anhand der Vorhersage von Vorgesetztenurteilen über die berufliche Bewährung. *Diagnostica*, *47*, 67–76. doi:10.1026//0012-1924.47.2.67.

Kluwe, R. H., Misiak, C., & Haider, H. (1991). The control of complex systems and performance in intelligence tests. In H. A. H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 227–244). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kröner, S., Plass, J. L., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, *33*, 347–368. doi:10.1016/j.intell.2005.03.002.

Mackintosh, N. J. (1998). *IQ and human intelligence.* Oxford: Oxford University Press.

Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success. A meta-analysis. *Personnel Psychology*, *58*, 367–408. doi:10.1111/j.1744-6570.2005.00515.x.

Raven, J. C., Court, J. H., & Raven, J. (1994). *Manual for Raven's progressive matrices and mill hill vocabulary scales. Advanced progressive matrices.* Oxford: Oxford Psychologists Press.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*, 173–184. doi:10.1177/01466216970212006.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior*, *6*, 855–863. doi:10.1016/s0022-5371(67)80149-x.

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *17*, 888–896. doi:10.1037/0278-7393.17.5.888.

Rigas, G., Carling, E., & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence*, *30*, 463–480. doi:10.1016/s0160-2896(02)00121-6.

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, *88*, 1068–1081. doi:10.1037/0021-9010.88.6.1068.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162–173. doi:10.1037/0022-3514.86.1.162.

Schmitt, M. J., & Steyer, R. (1993). A latent state-trait model (not only) for social desirability. *Personality and Individual Differences*, *14*, 519–529. doi:10.1016/0191-8869(93)90144-r.

Spearman, C. (1904). 'General intelligence', objectively determined and measured. *The American Journal of Psychology*, *15*, 201–293. doi:10.2307/1412107.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*, 389–408. doi:10.1002/(sici)1099-0984(199909/10)13:5<389::aid-per361>3.0.co;2-a.

Steyer, R., Schwenkmezger, P., & Auer, A. (1990). The emotional and cognitive components of trait anxiety: A latent state-trait model. *Personality and Individual Differences*, *11*, 125–134. doi:10.1016/0191-8869(90)90004-b.

Süß, H. -M., Kersting, M., & Oberauer, K. (1993). Zur Vorhersage von Steuerungsleistungen an computersimulierten Systemen durch Wissen und Intelligenz. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *14*, 189–203.

Süß, H. -M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability — and a little bit more. *Intelligence*, *30*, 261–288. doi:10.1016/s0160-2896(01)00100-3.

Wagener, D. (2001). *Psychologische Diagnostik mit komplexen Szenarios.* Taxonomie, Entwicklung, Evaluation. Lengerich: Pabst Science Publishers.

Wagener, D., & Wittmann, W. W. (2002). Personalarbeit mit dem komplexen Szenario FSYS: Validität und Potential von Verhaltensskalen. *Zeitschrift für Personalpsychologie*, *1*, 80–93. doi:10.1026//1617-6391.1.2.80.

Wallach, D. (1998). *Komplexe Regelungsprozesse.* Eine kognitionswissenschaftliche Analyse. Wiesbaden: Deutscher Universitäts-Verlag.

Wirth, J., & Funke, J. (2005). Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme. In E. Klieme, D. Leutner, & J. Wirth (Eds.), *Problemlösekompetenz von Schülerinnen und Schülern*. Wiesbaden: Verlag für Sozialwissenschaften.

Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, *10*, 329–345. doi:10.1080/0969594032000148172.

Wittmann, W., & Hattrup, K. (2004). The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioral Science*, *21*, 393–409 doi:10.1002/sres.653.

Yasuda, T., Lawrenz, C., Whitlock, R. V., Lubin, B., & Lei, P. -W. (2004). Assessment of intraindividual variability in positive and negative affect using latent state-trait model analyses. *Educational and Psychological Measurement*, *64*, 514–530. doi:10.1177/0013164403258445.

Ziegler, M., Ehrlenspiel, F., & Brand, R. (2009). Latent state-trait theory: An application in sport psychology. *Psychology of Sport and Exercise*, *10*, 344–349. doi:10.1016/j.psychsport.2008.12.004.