JDDM
Journal of Dynamic Decision Making

Open Access

Original Research

System structure and cognitive ability as predictors of performance in dynamic system control tasks

Jan Hundertmark, Daniel V. Holt, Andreas Fischer, Nadia Said, and Helen Fischer

Department of Psychology, Heidelberg University, Heidelberg, Germany

In dynamic system control, cognitive mechanisms and abilities underlying performance may vary depending on the nature of the task. We therefore investigated the effects of system structure and its interaction with cognitive abilities on system control performance. A sample of 127 university students completed a series of different system control tasks that were manipulated in terms of system size and recurrent feedback, either with or without a cognitive load manipulation. Cognitive abilities assessed included reasoning ability, working memory capacity, and cognitive reflection. System size and recurrent feedback affected overall performance as expected. Overall, the results support that cognitive ability is a good predictor of performance in dynamic system control tasks but predictiveness is reduced when the system structure contains recurrent feedback. We discuss this finding from a cognitive processing perspective as well as its implications for individual differences research in dynamic systems.

Keywords: dynamic system control, complex problem solving, reasoning ability, working memory, cognitive reflection

It is a central question in problem solving and decision making research which task properties and situational factors determine the difficulty of a problem and how these demands interact with the abilities of a problem solver. On the most general level, intelligence is useful for many types of problems and indeed, problem solving ability is often considered a defining aspect of general intelligence (e.g., Sternberg, 1982). However, while in some problem domains the value of cognitive abilities is well established, in other domains it does not help much and occasionally even has adverse effects (e.g., Wiley & Jarosz, 2012). In dynamic system control paradigms intelligence has generally been shown to be beneficial (Stadler, Becker, Gödker, Leutner, & Greiff, 2015), but it is still largely an open question in which way different aspects of dynamic systems (e.g., the number of variables or types of functional relations) contribute to problem difficulty and why some dynamic systems show high correlations with cognitive abilities while others do not. We therefore investigated the main effects of two characteristics of dynamic systems, system size and presence of oscillatory eigendynamics, and how they moderate the influence of cognitive abilities on control performance. Additionally, we assessed the effects of cognitive load. Taken together, we cover three groups of determinants of performance in dynamic system control tasks (as classified by Funke, 1991): (a) system characteristics, (b) personal factors, and (c) context factors. Systematically combining this range of factors in a single study allowed us to analyze their interaction, in particular, how system characteristics moderate the effect of cognitive abilities and context factors in determining task performance.

To investigate these questions, we employed a computer-simulated microworld paradigm. In microworld tasks participants interact with computer-simulated dynamic systems of varying size and complexity (Kluge, 2008). Systems are usually presented with a semantic framing such as managing a business, operating a complex machine, or carrying out chemistry experiments. The semantic framing may or may not give cues about the internal structure of the system. The task goal usually consists of exploring and successfully controlling the system to reach a target state. Systems used in research vary widely in terms of complexity, realism, and prior knowledge required for successful control. The core idea of the microworld paradigm is to mimic essential characteristics of dynamic systems in the real world in a controlled laboratory environment (Brehmer & Dörner, 1993; Gray, 2002).

System characteristics

Early research on semantic aspects of complex problem solving investigated the extent to which prior knowledge could be applied to a given problem. This line of research demonstrated that misleading semantics are a huge impediment to successful system control (Beckmann, 1994) and that prior knowledge accounts for a large proportion of performance in some common microworld tasks (Wittmann & Süß, 1999). Driven by the desire to create psychometrically reliable assessment procedures, a more recent wave of research introduced semantically lean systems with highly reduced complexity, an approach termed "minimal complex systems" (e.g., Greiff, Wüstenberg, & Funke 2012). It emphasizes formal aspects of problem difficulty by describing systems in a linear structural equation framework. The main determinant of difficulty is assumed to be the number of variables and system relations. Studies using this approach report item difficulties roughly corresponding to this construction principle (e.g., Greiff et al., 2012; Wüstenberg, Greiff, & Funke, 2012), but the relation between specific system characteristics and difficulty is usually not analyzed in detail. Building on Berry and Broadbent's (1984, 1987, 1988) seminal sugar factory and person interaction tasks, which are conceptually similar to minimal complex systems (cf. Fischer et al., 2015), we focused on the formal system characteristics *system size* and presence or absence of *oscillatory eigendynamics* (OED).

Corresponding author: Jan Hundertmark, Center for Psychosocial Medicine, Heidelberg University Hospital, Im Neuenheimer Feld 672, 69120 Heidelberg, Germany. E-Mail: jan.hundertmark@med.uni-heidelberg.de

While system size may seem an obvious determinant of difficulty, surprisingly few studies have systematically investigated its effect in a controlled experimental design (e.g., Funke, 1985). Although Berry and Broadbent used small and large systems (e.g., 1984, 1987), they never compared the difficulty of these size variations in the same study. We operationalize system size as the number of variables and relations within a system. We expect large systems to be more difficult, as the increased number of target variables and relations makes system exploration and control cognitively more demanding.

Dynamic change over time is another crucial property of complex problems (Dörner, 1980, 1983). One frequently encountered type of dynamics in system control tasks is a form of recurrent feedback termed *eigendynamics*, in which an output variable feeds back on itself. The feedback can be implemented either as a constant multiplier, leading to exponential growth or decay, or a negative sign of the feedback term. The latter may result in an oscillatory pattern with the output variable autonomously jumping between two values from one turn to another. The underlying equation is still linear, although the system's behavior is not. In the present study, we applied the same OED as Berry and Broadbent (1984): We either included or excluded system relations with an output variable negatively feeding back on itself, in the form of $Y_{t+1} = 2 \times X_t - Y_t$, with $Y_{t+1} =$ the new output, $X_t =$ the input given by the participant and $Y_t =$ the previous trial's output.

OED are quite common to many real-world scenarios containing negative feedback mechanisms, e.g., predator-prey systems or economic boom-bust-cycles. Using a cold store control scenario, Dörner (1996) and Güss (2010) have shown that systems with oscillatory behavior caused by negative feedback are indeed difficult to control, possibly due to the limited utility of simple exploration strategies such as the systematic variation of isolated variables to discover contingencies (e.g., Chen & Klahr, 1999). Oscillation due to negative feedback may be more difficult to discern than simple time-based oscillation, e.g., based on a sine function, as they can be irregular and change with different inputs. We therefore expect a main effect of OED on task difficulty.

As the structure of systems containing OED is apparently difficult to discern and verbalize, they have been labeled "non-salient" by Berry and Broadbent (1988). This term stems from implicit learning research, which postulates two distinct learning systems (e.g., Berry & Broadbent, 1988, 1995; Reber, 1989; Sun, Slusarz, & Terry, 2005): an explicit system responsible for forming a conceptual representation and an implicit system that stores events and contingencies in the form of subsymbolic associative links. In this approach's language, "salient" relations are amenable to explicit, analytic reasoning, while implicit, automatic learning processes are more suited for acquiring knowledge about "non-salient" relations. What makes system features more or less salient may depend on a range of factors, such as whether they have an immediate effect or are time-delayed, whether random noise makes the system more intransparent or to what extent system structure matches participants' expectations (see Funke, 2003, for an overview). OED have been used as one paradigmatic manipulation to reduce a system's salience (e.g., Berry & Broadbent, 1984). As the meaning of "salience" is only loosely specified, we focus on the specific system characteristic of OED.

Cognitive abilities

Personal factors relevant for dynamic system control may include a broad range of characteristics from cognitive ability to motivation and personality (Funke, 1991). Here, we investigate the aspect of cognitive abilities. While initially evidence was mixed (Stadler et al., 2015), by now it can be considered a well-established finding that intelligence (often operationalized as reasoning ability) is a good predictor of performance for many dynamic system control tasks. In a recent meta-analysis, Stadler et al. (2015) report a mean effect size of Hedge's $g = .43$ for the relation of intelligence and performance in a set of 62 studies. However, except for the attenuation of effect sizes due to measurement error, little is known about moderating factors and boundary conditions of this relation (Stadler et al., 2015).

We expect that systems including OED are not only harder to control but also that reasoning ability is less predictive for performance in this case. This may seem counter-intuitive, as superior intelligence and reasoning ability are generally associated with excelling at difficult tasks. However, reasoning is not a void process, but adds value to existing knowledge by transforming and recombining it according to the rules of logic. Therefore, without explicit knowledge about the problem at hand, reasoning processes lack the "raw material" to operate on (Goode & Beckman, 2010). If we combine this insight with the observation by Berry and Broadbent (1984) that OED restricts the amount of explicit system knowledge acquired, it follows that reasoning cannot unfold its full potential in this case. This interpretation is in line with the Elshout-Raaheim hypothesis, according to which the utility of reasoning may be limited by the amount of knowledge available (Leutner, 2002).

Studies in which explicit information about system structure is provided consistently found that reasoning ability and control performance are correlated (e.g., Putz-Osterloh & Lüer, 1981; Kröner, Plass, & Leutner, 2005; Wüstenberg et al., 2012). However, the most convincing line of evidence for the moderating effect of structural knowledge stems from Goode and Beckmann (2010; also Goode, 2011). In these studies, the amount of structural knowledge available to participants was experimentally manipulated. Goode and Beckmann (2010) observed a notable difference in the correlation of intelligence and control performance depending on the amount of information provided. Due to a relatively small sample in combination with a conservative analysis strategy, this difference was not statistically significant. In a later study using a larger sample the pattern of correlations was replicated and clearly reached statistical significance (Goode, 2011).

System size in contrast should not play a major role for the effects of reasoning provided that structural system knowledge can be acquired. Again, this is supported by the results reported in Goode (2011), as modifying system complexity by adding variables and relations did not result in an interaction of intelligence and complexity for predicting performance. Larger systems may be more difficult to control, but the cognitive processes required do not fundamentally differ from those required for controlling smaller systems. We therefore expect no effect of system size on the predictiveness of reasoning for control performance. The validity of this analysis is of course contingent on the absence of artificial restrictions by ceiling or floor effects, but there were no indications for such restrictions in Goode and Beckman (2010) or Goode (2011).

We further included cognitive reflection (Frederick, 2005) in our study, due to its good predictiveness for various judgment and decision making tasks (e.g., Toplak, West, & Stanovich, 2014; Weber & Johnson, 2009). As cognitive reflection is a reasoning-related disposition, we expect a pattern similar to reasoning ability, i.e., a main effect on control performance and interactions with OED. Additionally, we investigated the effects of working memory on control performance. Although reasoning and working memory are highly correlated, we expect that the predictors are not completely exchangeable. Gonzalez, Thomas, and Vanyukov (2005) found that both constructs were good predictors of performance in the "Water Purification Plant" scenario and showed statistically separable unique contributions to performance. However, we expect the effect of working memory on performance to be moderated less by OED and more by system size and concurrent dual tasking (see below).

Context factors

Context factors neither relate to structure or semantics of the system to be controlled, nor directly to characteristics of the person working on the task (cf. Funke, 1991). They can, for example, include to what extent additional information about the system is provided (e.g., causal relation diagrams) or the goals given to participants (e.g., understanding systems structure versus reaching given control goals). In the present study, we investigated the effect of concurrent cognitive load on task performance as a relevant context factor.

To this end, we introduced a dual task manipulation using a concurrent 2-back working memory task (cf. Kirchner, 1958). A comparable manipulation using a random letter generation task has previously been used with variants of the person interaction task by Hayes and Broadbent (1988). They hypothesized that dual tasking should interfere with the working-memory intense selective processing in the salient condition more strongly than the nearly automatic unselective learning process in the non-salient condition. Contrary to expectations, Hayes and Broadbent did not find such a selective impairment of learning in the "salient" condition under dual tasking, although response times were slowed down significantly. Dual tasking only had an effect when learned responses had to be adapted for transfer to a modified second task. The authors suggest that the secondary task might not have been demanding enough to impair performance in the original system control task. However, another possibility is the very small sample ($N = 18$), resulting in low statistical power. For more robust evidence on this question, we included a dual task manipulation using a concurrent 2-back working memory task. We follow Hayes and Broadbent's original hypothesis and expect dual tasks conditions not only to be more difficult, but to specifically impair the selective learning processes necessary to successfully control the stable, non-oscillatory systems.

Summary

In this comprehensive study we aim to analyze three types of performance determinants in dynamic system control and their interactions. First, we quantify the relative effect of system size and oscillatory eigendynamics (OED) system relations on control performance. Second, we analyze the predictive validity of reasoning ability and working memory capacity for control performance, particularly the interaction of these predictors with system size and the presence of OED. Third, we study the effect of a cognitive load manipulation on control performance, again with a view toward its interactions with system characteristics.

Method

Participants

One hundred and twenty-eight university students volunteered to participate in the study. One participant did not complete the system control tasks and was excluded from analysis. Of the remaining participants, 103 were female, age ranged from 18 to 35 years with a median of 21 years, all were native German speakers. The experiment took about 90 minutes on average. Participants received either €12 or course credit as compensation. For multivariate and repeated-measures analyses missing values were imputed using the expected maximization procedure (2.7% of data for the system control tasks).

Design

Each participant completed eight dynamic system control tasks and several tests of cognitive ability. In the systems control tasks, three experimental factors were manipulated within-subjects (two levels each: system size, presence of OED, cognitive load) in a fully crossed design. Serial order of conditions was balanced using a Latin square design. The cognitive load manipulation was applied block-wise, i.e., either to tasks one to four or to tasks five to eight. As an exploratory intervention, we gave half of the participants a brief instruction encouraging either explicit, rule-based exploration or an intuitive strategy. Cognitive abilities measured included working memory, cognitive reflection and reasoning. Using a within-subjects design with 127 participants yields 97% power to detect medium-sized effects at $\alpha = .05$ (according to Cohen, 1988).

Materials

We designed four different basic types of dynamic system control scenarios in two parallel versions for a total of eight tasks. All scenarios were semantically framed as experiments in a biology laboratory where different substances (input variables) with fictitious labels, e.g., "Dilarin" or "Berophal", could be added to cell cultures to produce different cell characteristics (output variables), e.g., nutrient requirement or temperature sensitivity, see Fig. 1. The scenarios were turn-based, i.e., participants first changed the value of input variables using increment and decrement buttons (12 steps per variable) and then clicked a button to proceed to the next turn. The value of input variables remained stable unless manipulated by the participant, the value of output variables was determined by a set of simple linear equations (cf. Funke, 2001) with a small random component (see Table 1 for equations). Values of system variables were capped at predefined minimum/maximum values to prevent participants from maneuvering systems into irrecoverable states. Each scenario consisted of an exploration phase of 1.5 minutes followed by two control phases with different target values for 20 turns (or at most 2 minutes). Successful system control required participants to first experiment with different input values and their effects on the output variables during the exploration phase. In the subsequent control phases, they had to apply their knowledge and manipulate the input variables to reach given target values.

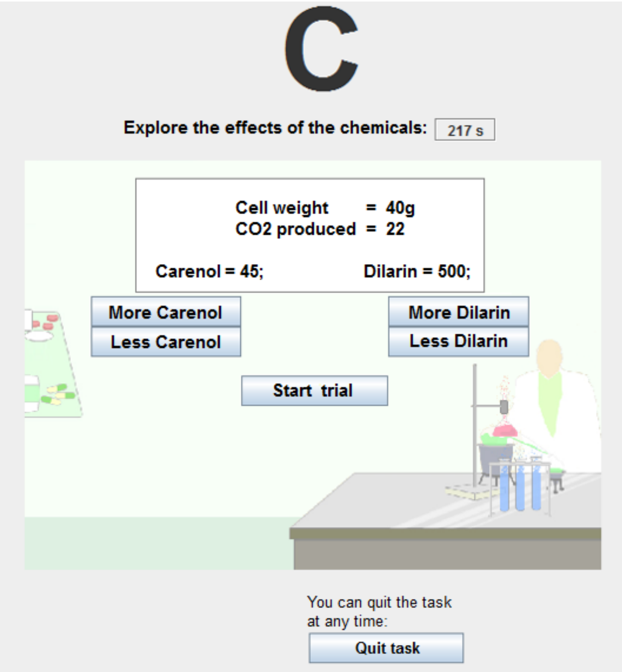

Figure 1. Task environment for a 2×2 mixed system (STA/OED, see Table 1) during the exploration phase with dual-tasking.

System size and *presence of OED* were experimentally manipulated with two levels each. System size was either small with one input and output variable (1×1 systems) or large with two input and output variables (2×2 systems). The OED factor was manipulated by either excluding or including OED in the system (cf. Berry & Broadbent, 1984, 1988). In the large systems, the OED was implemented for one of the output variables only. We refer to output variables excluding OED as stable (STA), because their values remain constant without the participant's intervention (except for a small random term). Taken together, the factors size and OED resulted in four basic system types for which two parallel versions each were constructed using different labels and numerical ranges (see Table 1). The structure of the small OED system was identical to Berry and Broadbent's (1984) tasks.

We employed a 2-back parallel task to create a constant but not overwhelming load on working memory in half of the system control tasks. Participants saw a sequence of large random letters on the top of the screen. Each letter was presented for 2.5 seconds, followed by a 500 ms inter-stimulus interval. Every time the current letter was the same as the second last letter back in the sequence, participants had to press the space key. We configured the task in such a way that a positive response was required in 30% of the trials. On errors, i.e., a false positive or a missed response (after 2500 ms), an acoustic beep was sounded.

Taken together, every participant completed eight scenarios: small and large systems including and excluding OED, once with and once without a parallel dual task (a $2 \times 2 \times 2$ fully crossed within-subjects design controlled for effects of task order).

Scoring of control performance

The control score was calculated by determining the proportion of turns during a control phase in which all variables of a system were within the target range. We chose the target range so that perfect control was in principle possible for every turn despite the random fluctuations. Scores were averaged over the two control phases for each system control task.

Cognitive Tests

We assessed working memory capacity using an adapted version of the Memory Updating (MU) task described in Lewandowsky, Oberauer, Yang, and Ecker. (2010). The task requires participants to simultaneously encode a set of three to five digits and sequentially apply simple arithmetic operations on them. Participants need to replace the memorized numbers with the results of the operation and recall them in a subsequent retrieval phase. In three validation experiments, the authors obtained high internal consistencies (average $\alpha = .87$) and showed that MU was the best single predictor of general working memory capacity in a battery of commonly used WM tests. The correlation with intelligence was found to be $r = .67$.

As an indicator of general reasoning ability, we used a short form of the Raven Advanced Progressive Matrices Test (Raven, Court, & Raven, 1985) developed by Arthur and Day (1994). In the present study we administered the short form with a time limit of 10 minutes. The original APM has been argued to be one of the purest available measure of analytical (fluid) intelligence (e.g., Raven, 1989; Carpenter, Just, & Shell, 1990). The short form shows an internal consistency of $\alpha = .72$, its retest-reliability is $r_{tt} = .75$, and it is strongly correlated with the APM long version, $r = .90$ (Arthur & Day, 1994).

The original Cognitive Reflection Test (CRT; Frederick, 2005) is a three-item questionnaire measuring the tendency to override a prepotent but incorrect response alternative and to engage in further reflection that leads to the correct response. The three questions are designed to make an intuitive yet erroneous answer spring to mind. For instance, the first question is "A bat and a ball cost \$1.10. The bat costs \$1.00 more than the ball. How much does the ball cost?" The correct answer (5 ct) requires the suppression of the

Table 1. Basic equations used in the system control tasks.

	Oscillatory eigendynamics	
System size	Absent (STA)	Present (OED)
1×1	$Y = X - 2 + R$	$Y = 2X - Y' + R$
2×2	$Y_1 = 1.8 \times X_1 - 0.45 \times X_2 + R$ $Y_2 = 0.8 \times X_2 + 0.45 \times X_1 + R$	$Y_1 = 1.8 \times X_1 - 0.45 \times X_2 + R$ $Y_2 = 1.3 \times X_2 + 0.95 \times X_1 - Y_2' + R$

Note. X = input value; Y = current trial's output value; Y' = preceding trial's output value; R = random noise. Equations adapated from Berry and Broadbent (1984, 1987, 1988).

impulsive answer (10 ct). The CRT was designed to assess a cognitive style related to readiness to engage in deliberate reflection, as postulated by dual process theories (see Stanovich & West, 2000; also Evans & Frankish, 2009). It has been shown that the CRT is closely related to measures of fluid intelligence (Frederick, 2005; Toplak et al., 2014) and particularly numerical reasoning ability (e.g., Campitelli & Gerrans, 2014). We used the expanded 7-item version as proposed by Toplak et al. (2014). Its correlation with the original CRT is $r = .86$ and its internal consistency is $\alpha = .72$.

Procedure

After participants gave written informed consent they received a short oral instruction explaining the tasks and the set-up of the experiment. System control tasks were presented first, followed by the assessment of cognitive predictor variables. As an exploratory manipulation, we varied the instruction type by presenting one of two different task descriptions to participants. In the *rule-based instruction* condition, we instructed participants to "carefully observe the experiments' results and try to form a rule in order to predict them accurately". In contrast, in the *intuition-based instruction* condition, we encouraged them to "just take the presented results in and [...] not try to calculate or form a rule" and to instead "observe the results attentively and use [their] intuition". This was repeated before every block for both conditions. The instructions aimed at eliciting a more selective (explicit) or unselective (implicit) learning mode, respectively. Past research has shown similar wordings to affect participants' approach to learning in dynamic system control tasks (cf. Berry & Broadbent, 1988; Gebauer & Mackintosh, 2007).

After completing all system control tasks, we employed a manipulation check and asked participants to rate in which way they processed the tasks on a one-item nine-level Likert scale ranging from *entirely intuitive* to *entirely rule-based*. Furthermore, all participants completed a computer-based Serial Reaction Time task (Robertson, 2007), which was intended as a measure of implicit learning ability. Due to technical problems data from this task were unusable and had to be excluded from analysis.

Results

Exploration

The median exploration time per task was 80.9 seconds (IQR = 95.8) with a median of 26 exploration turns (IQR = 43). Exploration was completed more quickly for the small systems (median 68.1 and 75.1 seconds for STA and OED) than for the large systems (median 95.9 and 96.5 seconds for STA/STA and STA/OED). The median number of exploration turns was comparable (24 and 25 turns versus 28.5 and 25 turns). Dual tasking had no effect on exploration time (median 78.23 seconds with dual tasking, 83.7 seconds without), Wilcoxon $W(127) = 4300$, $p = .57$, but a detrimental effect on the number of exploration turns (20 turns with dual tasking, 33 turns without), Wilcoxon $W(127) = 7001.5$, $p < .001$.

System characteristics and context factors

The effect of system characteristics and context factors on control performance was analyzed using a four-factor mixed ANOVA with *system size* (small or large) and *OED* (present or absent) as system characteristics, which were varied within-subjects. The context factors were *dual tasking* (present or absent, within-subjects) and *instruction* (rule-based or intuition-based, between-subjects). To reduce the inflation of Type I errors in multifactorial designs, we only report main effects and interactions for which hypotheses had been formulated. Fig. 2 illustrates the characteristic behavior of systems with STA (stable) or OED (oscillatory) dynamics. The mean control performance scores for the four different system types were .69 ($SD = .16$) for STA, .27 ($SD = .08$) for OED, .24 ($SD = .12$) for STA/STA, and .09 ($SD = .05$) for STA/OED. As displayed in Fig. 3, system size showed a strong main effect, $F(1,\ 125) = 1673.06$, $p < .001$, $\eta_g^2 = .55$, as did OED, $F(1,\ 125) = 870.55$, $p < .001$, $\eta_g^2 = .52$. Both factors interacted, $F(1, 125) = 310.77$, $p < .001$, $\eta_g^2 = .19$, indicating that the effect of OED partially depended on system size. Comparing performance for the two target variables within the large mixed system (STA/OED) replicated the pattern of the separate STA and OED systems, $F(1, 126) = 197.28$, $p < .001$, $\eta_g^2 = .38$.

The context factor dual tasking exerted a small but statistically significant main effect on performance in the expected direction, $F(1,\ 125) = 5.55$, $p = .02$, $\eta_g^2 = .01$. Contrary to expectation, it did not interact

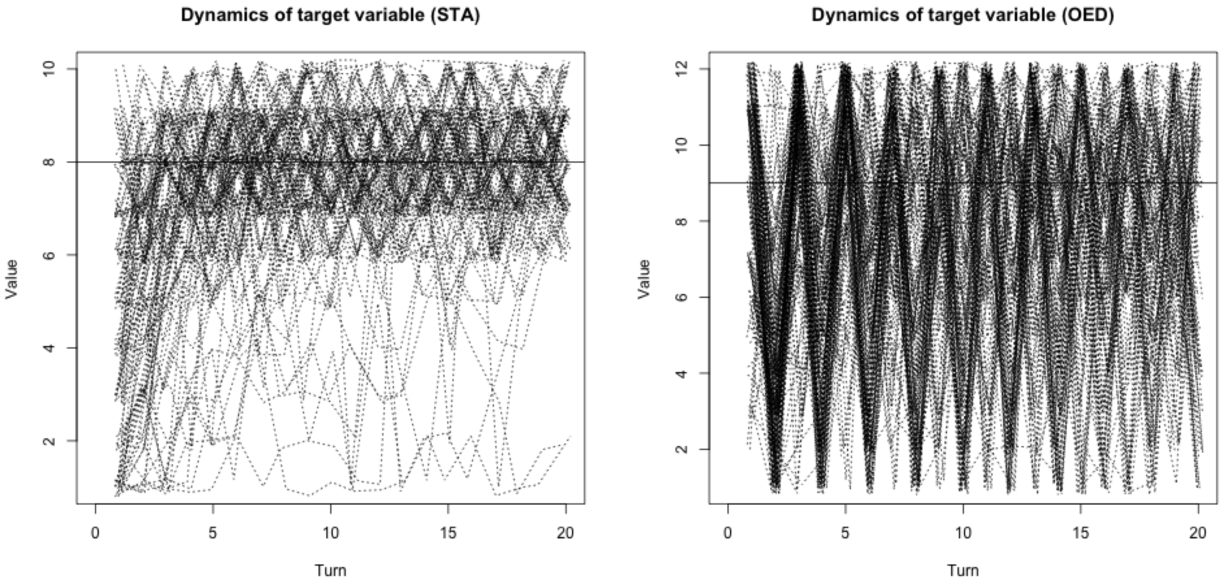

Figure 2. Dynamics of a 1×1 STA (stable) and a 1×1 OED (oscillatory) system showing the development of the target variable over 20 control turns. The horizontal line indicates the given target value. Each dotted line represents the output values of one participant.

with OED, $F(1,\,125) = 0.45$, $p = .50$. Mean control performance was .33 ($SD = .09$) without dual tasking and .31 ($SD = .08$) with dual tasking. Different instructions (encouraging rule formation or an intuitive approach) also showed a small but statistically significant effect on performance, $F(1,\,125) = 6.87$, $p = .01$, $\eta_g^2 = .01$ and no interaction with OED, $F(1,\,125) = 1.94$, $p = .17$. Mean control performance with rule-based instructions was .34 ($SD = .06$) and .31 ($SD = .08$) with intuition-based instructions. The self-rated processing style was not affected by the type of instruction given, $t(125) = 0.15$, $p = .88$.

Cognitive abilities

A regression analysis for predicting overall system control (averaged over all tasks) with the cognitive ability variables APM, CRT, and MU showed that in total 24.5% of performance variance could be explained by these predictors, $F(3,\,123) = 13.29$, $p < .001$. CRT was the strongest overall predictor, $\beta = .41$, $p < .001$, followed by APM, $\beta = .20$, $p = .05$, while MU did not significantly contribute, $\beta = -.11$, $p = .25$.

Table 2 lists the bivariate correlations between individual predictors and the different system types, supporting that CRT was a good predictor throughout, while MU was comparatively weak. To test whether the predictiveness of cognitive variables interacts with the presence or absence of OED in the tasks as hypothesized, we conducted William's tests for comparing dependent correlation coefficients for the small STA and OED systems. CRT showed the expected difference, $t(127) = 2.85$, $p < .01$, with a lower correlation in the OED condition, but APM and MU did not, $t(127) = 1.64$, $p = .10$, and $t(127) = 0.08$, $p = 0.93$. Combining all cognitive variables into a single general ability score by averaging z-standardized scores revealed that this overall ability variable also interacted with the absence of presence of OED, $t(127) = 2.03$, $p = .04$.

Correlations between cognitive ability variables and control performance may be attenuated by low reliabilities of the system control tasks. Cronbach's α for the two small STA system was only .42, and .28 for the two small OED systems. We therefore repeated the William's tests applying a one-sided correction for attenuation to the control performance scores before comparing correlation coefficients. Results support the initial analysis, even accentuating the interaction effects. CRT showed the expected difference, $t(127) = 6.94$, $p < .01$, and with correction for attenuation so did APM, $t(127) = 3.94$, $p < .001$, while there still was no effect for MU, $t(127) = 0.86$, $p = .39$. For the combined general ability score this analysis also yielded a significant effect, $t(127) = 4.83$, $p < .001$.

For the large systems, correlations of performance with cognitive abilities were not significantly different between those including or excluding OED, t's $< .40$, p's $> .69$. However, the analysis of whole systems may mask differences between the two target variables in the mixed system (STA/OED). We therefore conducted the comparisons of correlations of cognitive predictors and control performance just for the STA and OED variables within the mixed system (see Table 2). Similar to the results for the independent STA and OED systems, we found that the variable involving OED showed significantly lower correlations with two of the three cognitive predictors, $t(127) = 2.16$, $p = .03$ for the CRT and $t(127) = 2.09$, $p = .04$ for MU. For APM correlations did not significantly differ, $t(127) = .41$, $p = .68$. Again, these results were accentuated when correcting correlations for attenuation due to low reliabilities of the system control tasks (Cronbach's $\alpha = .41$ for STA variables and .17 for OED

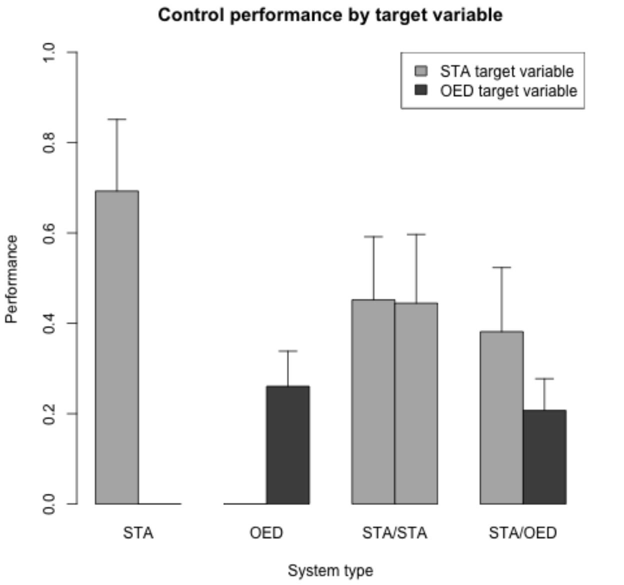

Figure 3. Performance by target variable, averaged over parallel task versions. Error indicators represent standard deviations. The value of each target variable is either controlled by a salient (light gray) or a non-salient relation (dark gray). (STA = 1×1 stable system, OED = 1×1 system containing oscillatory eigendynamics, STA/STA = 2×2 system with two stable target variables, STA/OED = 2×2 mixed system with one stable and one oscillatory target variable, see Table 1).

variables).

To investigate whether dual tasking moderates the predictiveness of working memory in this task, we compared the correlation of MU and performance between dual tasking conditions. For all four system types, the correlation coefficients did not differ with and without dual tasking, t's < 1.46, p's > 0.14.

Further analyses

Performance in the 2-back secondary task was generally low with an average hit rate of .34 (SD = .20), although consistent (Cronbach's α = .86). A 2×2 ANOVA showed a strong effect of system size on secondary task performance, $F(1, 126) = 68.00$, $p < .001$, but no effect of OED, $F(1, 126) = 0.06$, $p = .80$, and no interaction, $F(1, 126) = 0.01$, $p = .93$. This suggests that the larger systems were more working-memory demanding, thereby reducing cognitive resources for the secondary task.

In addition to the effect on system control performance reported above, we also observed a clear effect of dual tasking on response latency: Without dual tasking participants took an average of 2.20 (SD = 0.92) seconds per control turn and 2.77 (SD = 1.22) with dual tasking, $F(1, 126) = 38.40$, $p < .001$.

Discussion

We observed that manipulating the presence of oscillatory eigendynamics (OED) and system size changed difficulty as expected, while manipulating cognitive load and the instructions only had a small effect on control performance. Furthermore, we found that OED not only make system control more difficult, but that they can also reduce the effect of cognitive abilities on control performance.

Regarding system characteristics, we found that OED apparently were difficult to discern and control for most participants, in line with the results of Berry and Broadbent (1988). The small OED system was about as difficult as a stable system twice the size (STA/STA). What makes this finding particularly striking is that the mathematical change to the system structure was minimal, just an additional negative term in the linear equation. The difficulty pattern was replicated for the different target variables in the large mixed systems (STA/OED). The target variables behaved very similar to the small STA and OED systems, with the OED variable being much harder to control. These results show that operationalizing system complexity merely in terms of number of variables and relations does not fully cover complexity from a cognitive perspective. The emergent dynamic complexity of the system as a whole seems to be just as important, if not more so (e.g., Brehmer & Dörner, 1993; Gonzalez et al., 2005).

In their seminal work, Berry and Broadbent (1984) and Reber (1967) referred to systems which are easy, respectively difficult, to explore and control using deliberate reasoning strategies as “salient” or “non-salient”. However, we think that the effects of dynamic complexity produced by negative feedback go beyond Berry and Broadbent's suggestion that low salience simply makes it less likely that participants focus their exploration on the relevant parts of the system. Even when non-salient relations are detected and perhaps even partially understood (e.g., that there is oscillation), the system still may be more difficult to explore and control. Simple exploration strategies such as control-of-variables (Chen &Klahr, 1999) are harder to apply due to the prior system state's influence and the resulting instable system behavior. Furthermore, for the same reason it is difficult to derive the correct control interventions even if the system structure is understood.

We replicated the finding that cognitive ability is a good predictor of control performance (Stadler et al., 2015). Considering specific abilities, we found cognitive reflection to be the strongest overall predictor, followed by reasoning ability, while working memory capacity was a comparably weak predictor. This result is somewhat surprising, given the conceptual overlap between reasoning and working memory. As the measure of working memory employed, memory updating, is a well-established and reliable indicator, one possible explanation is that the relatively simple systems used in this study do not pose high working memory demands. This explanation is supported by the fact that the concurrent working memory load only had a small effect on performance.

Table 2. Correlations of reasoning ability, cognitive reflection, and working memory with control performance.

	System type				Target variable in STA/OED	
	STA	OED	STA/STA	STA/OED	STA	OED
Reasoning ability	**.33**	.17	**.23**	**.27**	**.32**	**.18**
Working memory	.16	.09	.17	**.20**	**.30**	.08
Cognitive reflection	**.45**	.16	**.28**	**.31**	**.39**	**.18**

Note. $N = 127$. Correlations for the target variables in the 2x2 mixed system (STA/OED) are shown separately. Correlations with $p < .05$ shown in bold. Coefficents above .23 are significant at $p < .01$, above .29 at $p < .001$.

Beyond their overall effects, cognitive abilities interacted with specific task characteristics. As expected, the predictors most closely related to abstract reasoning (APM, CRT) interacted with the presence or absence of OED. Specifically, these predictors were less correlated with performance in the small systems including OED. We found the same pattern in the large mixed system (STA/OED) when both target variables were analyzed separately. Working memory capacity, in contrast, did not show an interaction with the presence of OED, possibly due to its generally low predictiveness in this paradigm. These results also hold when statistically controlling for the low measurement reliability of the systems including OED. The only predictor interacting with system size was cognitive reflection, a statistically significant, but very small effect.

The interaction of cognitive abilities and system characteristics is in line with previous findings by Goode (2011), who showed that reasoning ability is less predictive for highly complex systems. The explanation given by Goode (2011, also Goode & Beckmann, 2010) is that reasoning ability can only unfold its effect if structural knowledge is acquired. However, as Berry and Broadbent (1984) have shown before, the presence of OED dramatically reduces the amount of structural knowledge acquired. Consequently, reasoning ability should be less predictive in systems including OED. Given that our results confirm this hypothesis, we conclude that, somewhat paradoxically, reasoning ability may be more helpful for relatively simple dynamic problems with an obvious structure. However, this result was obtained under laboratory conditions with a strict time limit and may be different when further opportunity for exploration or additional information sources are available.

Another conceivable criticism is that control performance in the OED systems is simply less reliable in psychometric terms and correlations with other constructs are therefore limited. We calculated corrections for attenuation as one approach to rule out this possibility. Furthermore, this criticism is based on the assumption that there is a stable trait or ability reflected in performance, which does not need to be the case. Alternatively, the performance scores can be considered a formative measure, i.e., they directly represent the degree of successful system control, which is the criterion to be predicted.

An alternative candidate for an ability underlying performance in tasks with OED would have been implicit learning ability, as suggested by the observation that implicit learning takes place in these systems (Berry & Broadbent, 1984). Although our measure of implicit learning ability was unusable for technical reasons, it is uncertain whether it would have added much explanatory value as a predictor. In a study using a relatively complex dynamic system, Danner et al. (2011) showed that the latent correlation (corrected for measurement error) between implicit learning ability and control performance was just $r = .26$ compared to $r = .86$ for intelligence. Furthermore, implicit learning as a unitary ability is not undisputed (Gebauer & Mackintosh, 2007) and its reliability seems to be generally low (Reber & Allen, 2000). Moreover, the time restrictions in our study and the tasks' superficial similarity despite their structural differences may have prevented implicit, instance-based learning (cf. Kaufman, 2011). The correlations between reasoning and system control performance in our study suggest that mainly explicit, deliberate learning was required. This interpretation is supported by studies that similarly found such correlations in explicit learning conditions but not in implicit learning conditions (shown for intelligence by Gebauer & Mackintosh, 2007, and for working memory capacity by Unsworth & Engle, 2005).

Supporting earlier findings by Hayes and Broadbent (1988), our results show that dual tasking slowed participants down, but only had a negligible effect on control performance. While in some reasoning tasks cognitive load affects both accuracy and response latency (e.g., Gilhooly, Logie, Wetherick, & Wynn, 1993), a dissociation of the two is also sometimes observed (e.g., Baddeley & Hitch, 1974). Our findings imply that in the present task it is possible to compensate experimentally reduced mental capacity by proceeding more slowly, and that participants seem to give priority to accuracy over speed. This pattern of results was the same for STA and OED systems. If performance in OED conditions was purely based on implicit learning, it should have been less affected by dual tasking. However, this was not the case, further supporting the interpretation that explicit learning may have been relevant in all conditions.

In summary, the present study demonstrates that the presence of oscillatory eigendynamics in a system has a strong effect on difficulty and can act as a moderator on the effect of reasoning and cognitive reflection on control performance. System size has an effect on difficulty, but shows only limited interaction with cognitive abilities. Furthermore, we found that analyzing target variables in the mixed (STA/OED) large systems separately mirror the pattern from comparing the separate small STA and OED systems. We therefore recommend the separate analysis of system parts for future cognitive research in dynamic system control. Our results may also be informative for the psychometric application of dynamic system control tasks, as they contribute towards a more differentiated understanding of the effects of system characteristics and cognitive abilities on task performance.

Acknowledgements: This research was partially supported by grant no. Fu 173/14 of the German Research Foundation (DFG). The authors would like to thank Katharina Berger and Antje Spiertz for assistance with data collection and three reviewers for their helpful comments on the manuscript.

Declaration of conflicting interests: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. DH and AF are co-editors of JDDM.

Author contributions: All authors contributed to the design of the study. AF, DH, and NS prepared the materials. JH coordinated the data collection. JH and DH analyzed the data and wrote the manuscript. All authors commented on and approved of the final version of the manuscript.

Supplementary material: Supplementary material available online.

Handling editor: Florian Kutzner

Citation: Hundertmark, J., Holt, D. V., Fischer, A., Said, N., & Fischer, H. (2015). System structure and cognitive ability as predictors of performance in dynamic system control tasks. *Journal of Dynamic Decision Making, 1*, 5. doi: 10.11588/jddm.2015.1.26416

Received: 10 December 2015
Accepted: 29 January 2016
Published: 9 February 2016

References

Arthur, W., & Day, D. V. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement, 54*, 394-403. doi: 10.1177/0013164494054002013

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower, *The psychology of learning and motivation, Vol. 8*, (pp. 47-89). New York, NY: Academic Press.

Beckmann, J. F. (1994). *Lernen und komplexes Problemlösen: ein Beitrag zur Konstruktvalidierung von Lerntests [Learning and complex problems solving: a contribution to the construct validation of tests of learning potential]*. Bonn: Holos.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology, 36A*, 209-231. doi: 10.1080/14640748408402156

Berry, D. C., & Broadbent, D. E. (1987). The combination of explicit and implicit learning processes in task control. *Psychological Research, 49*, 7-15. doi: 10.1007/BF00309197

Berry, D. C., & Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology, 79, 251-272*. doi: 10.1111/j.2044-8295.1988.tb02286.x

Berry, D. C., & Broadbent, D. E. (1995). Implicit learning in the control of complex systems. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective*. (pp. 131-150). Hillsdale, NJ: Erlbaum.

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior, 9*, 171-184. doi: 10.1016/0747-5632(93)90005-D

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory & Cognition, 42*, 434-447. doi: 10.3758/s13421-013-0367-9

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*, 404-431. doi: 10.1037/0033-295X.97.3.404

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development, 70*, 1098-1120. doi: 10.1111/1467-8624.00081

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Danner, D., Hagemann, D., Holt, D.V., Bechthold, M., Schankin, A., Wüstenberg, S., & Funke, J. (2011). Measuring performance in a Complex Problem Solving task: Reliability and validity of the Tailorshop simulation. *Journal of Individual Differences, 32*, 225-233. doi: 10.1027/1614-0001/a000055

Dörner, D. (1980). On the difficulties people have in dealing with complexity. *Simulation & Gaming, 11*, 87-106. doi: 10.1177/104687818001100108

Dörner, D. (1983). Heuristics and cognition in complex systems. In R. Groner, M. Groner, & W. F. Bischof (Eds.), *Methods of heuristics*. Hillsdale, NJ: Erlbaum.

Dörner, D. (1996). *The logic of failure: Recognizing and avoiding error in complex situations*. New York, NY: Perseus.

Evans, J., & Frankish, K. (2009). *In two minds: dual processes and beyond*. Oxford: University Press.

Fischer, A., Greiff, S., Wüstenberg, S., Fleischer, J., Buchwald, F., & Funke, J. (2015). Assessing analytic and interactive aspects of problem solving competency. *Learning and Individual Differences, 39*, 172-179. doi:10.1016/j.lindif.2015.02.008

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives, 19*, 25-42. doi: 10.1257/089533005775196732

Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle [Control of dynamic systems by generating and applying subjective causal models]. *Zeitschrift für Psychologie, 193*, 443-465.

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Erlbaum.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning, 7*, 69-89. doi: 10.1080/13546780042000046

Funke, J. (2003). *Problemlösendes Denken [Problem solving thinking]*. Stuttgart: Kohlhammer.

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

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition, 21*, 115-124. doi: 10.3758/BF03211170

Gonzalez, C., Thomas, R. P., & Vanyukov, P. (2005). The relationships between cognitive ability and dynamic decision making. *Intelligence, 33*, 169-186. doi: 10.1016/j.intell.2004.10.002

Goode, N. (2011). *Determinants of the control of dynamic systems: The role of structural knowledge* (Doctoral thesis, University of Sydney, Sydney, Australia). Retrieved from http://ses.library.usyd.edu.au/handle/2123/8967

Goode, N., & Beckmann, J. F. (2010). You need to know: There is a causal relationship between structural knowledge and control performance in complex problem solving tasks. *Intelligence, 38*, 345-352. doi: 10.1016/j.intell.2010.01.001

Gray, W. D. (2002). Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research. *Cognitive Science Quarterly, 2*, 205-227.

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic Problem Solving: A New Assessment Perspective. *Applied Psychological Measurement, 36*, 189-213. doi: 10.1177/0146621612439620

Güss, C. D. (2010). Fire and ice: Testing a model on culture and complex problem solving. Journal of Cross-Cultural Psychology, 1-20. doi: 10.1177/0022022110383320

Hayes, N. A., & Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition, 28*, 249-276. doi: 10.1016/0010-0277(88)90015-7

Kaufman, S. B. (2011). Intelligence and the cognitive unconscious. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 442-467). New York: Cambridge University Press.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*, 352-358. doi: 10.1037/h0043688

Kluge, A. (2008). Performance assessments with microworlds and their difficulty. *Applied Psychological Measurement, 32*, 156-180. doi: 10.1177/0146621607300015

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

Leutner, D. (2002). The fuzzy relationship of intelligence and problem solving in computer simulations. *Computers in Human Behavior, 18*, 685-697. doi: 10.1016/S0747-5632(02)00024-9

Lewandowsky, S., Oberauer, K., Yang, L.-X., & Ecker, U. K. H. (2010). A working memory test battery for MATLAB. *Behavior Research Methods, 42*, 571-585. doi: 10.3758/BRM.42.2.571

Putz-Osterloh, W., & Lüer, G. (1981). The predictability of complex problem solving by performance on an intelligence test. *Zeitschrift für Experimentelle und Angewandte Psychologie, 28*, 309-324.

Raven, J. (1989). The Raven Progressive Matrices: A review of national norming studies and ethnic and socioeconomic variation within the United States. *Journal of Educational Measurement, 26*, 1-16. doi: 10.1111/j.1745-3984.1989.tb00314.x

Raven, J. C., Court, J. H., & Raven, J. (1985). *Manual for Raven's progressive matrices and vocabulary scales* (Revised ed.). London: Lewis.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855-863. doi: 10.1016/S0022-5371(67)80149-X

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology, 118*, 219-235. doi: 10.1037/0096-3445.118.3.219

Reber, A. S., & Allen, R. (2000). Individual differences in implicit learning: Implications for the evolution of consciousness. In R. G. Kunzendorf & B. Wallace (Eds.), *Advances in consciousness research* (pp. 227-247). Amsterdam: John Benjamins.

Robertson, E. M. (2007). The serial reaction time task: implicit motor skill learning? *The Journal of Neuroscience, 27*, 10073-10075. doi: 10.1523/JNEUROSCI.2747-07.2007

Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence, 53*, 92-101. doi: 10.1016/j.intell.2015.09.005

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*, 645-665. doi: 10.1017/S0140525X00003435

Sternberg, R. J. (1982). *Handbook of human intelligence*. Cambridge: Cambridge University Press.

Sun, R., Slusarz, P., & Terry, C. (2005). The Interaction of the Explicit and the Implicit in Skill Learning: A Dual-Process Approach. *Psychological Review, 112*, 159-192. doi: 10.1037/0033-295X.112.1.159

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*, 147-168. doi: 10.1080/13546783.2013.844729

Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition, 33*, 213-220. doi: 10.3758/BF03195310

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology, 60*, 53-85. doi: 10.1146/annurev.psych.60.110707.163633

Wiley, J., & Jarosz, A. F. (2012). Working memory capacity, attentional focus, and problem solving. *Current Directions in Psychological Science, 21*, 258-262. doi: 10.1177/0963721412447622

Wittmann, W. W., & Süß, 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* (pp. 77-108). Washington, DC: American Psychological Association.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving. More than reasoning? *Intelligence, 40*, 147-168. doi: 10.1080/13546783.2013.844729