Journal of Experimental Psychology: General

Diffusion Modeling and Intelligence: Drift Rates Show Both Domain-General and Domain-Specific Relations With Intelligence

Veronika Lerche, Mischa von Krause, Andreas Voss, Gidon T. Frischkorn, Anna-Lena Schubert, and Dirk Hagemann

Online First Publication, May 7, 2020. http://dx.doi.org/10.1037/xge0000774

American Psychological Association

Journal of Experimental Psychology: General

ISSN: 0096-3445

http://dx.doi.org/10.1037/xge0000774

Diffusion Modeling and Intelligence: Drift Rates Show Both Domain-General and Domain-Specific Relations With Intelligence

Veronika Lerche, Mischa von Krause, Andreas Voss, Gidon T. Frischkorn, Anna-Lena Schubert, and Dirk Hagemann
University of Heidelberg

Several previous studies reported relationships between speed of information processing as measured with the drift parameter of the diffusion model (Ratcliff, 1978) and general intelligence. Most of these studies utilized only few tasks and none of them used more complex tasks. In contrast, our study (N = 125) was based on a large battery of 18 different response time tasks that varied both in content (numeric, figural, and verbal) and complexity (fast tasks with mean RTs of ca. 600 ms vs. more complex tasks with mean RTs of ca. 3,000 ms). Structural equation models indicated a strong relationship between a domain-general drift factor and general intelligence. Beyond that, domain-specific speed of information processing factors were closely related to the respective domain scores of the intelligence test. Furthermore, speed of information processing in the more complex tasks explained additional variance in general intelligence. In addition to these theoretically relevant findings, our study also makes methodological contributions showing that there are meaningful interindividual differences in content specific drift rates and that not only fast tasks, but also more complex tasks can be modeled with the diffusion model.

Keywords: intelligence, diffusion model, mathematical models, reaction time (RT) methods, fast-dm

Supplemental materials: http://dx.doi.org/10.1037/xge0000774.supp

One of the processes that has often been discussed as basis of individual differences in intelligence is speed of information processing (Jensen, 2006). This notion is supported by consistent empirical results showing moderate relationships between general intelligence[1] and response times (RTs) from a broad range of cognitive tasks (Sheppard & Vernon, 2008). Regarding these relationships between intelligence and RTs, (at least) two important observations have been made in the last decades: (a) the relationship between RT and intelligence does not seem to be specific to content domains (verbal, figural, numeric; Levine, Preddy, & Thorndike, 1987; Neubauer & Bucik, 1996); (b) the slower responses within one task are more highly related to intelligence than the faster responses, resulting in the formulation of the *worst performance rule* (Larson & Alderton, 1990; for a review, see Coyle, 2003; for methodological considerations, see Frischkorn, Schubert, Neubauer, & Hagemann, 2016; for a meta-analysis, see Schubert, 2019). Thus, in brief, the relationship between intelligence and speed of information processing seems to depend on the speed of trials, but not or only to a small degree on the specific task content.

However, there are some methodological limitations of previous studies that examined the relationship between intelligence and speed of information processing. One of these limitations has been pointed out by Schmiedek, Oberauer, Wilhelm, Süß, and Wittmann (2007): Regarding the worst performance rule, they noted that previous studies employed different RT bands resulting in only restricted numbers of trials per band, thereby limiting the reliability of estimates. Instead of employing RT bands, Schmiedek et al. (2007) used a mathematical model that takes into account information about RT distributions, and thus has a considerably higher information usage—the diffusion model (Ratcliff, 1978; see Voss, Nagler, & Lerche, 2013, for a review).

The diffusion model is a stochastic model that is applicable to binary response time tasks and allows the separation of different, otherwise confounded, processes. One parameter of this model—

Veronika Lerche, Mischa von Krause, Andreas Voss, Gidon T. Frischkorn, Anna-Lena Schubert, and Dirk Hagemann, Department of Psychology, University of Heidelberg.

Gidon T. Frischkorn is now at the Department of Psychology, University of Zurich.

This research was supported by grants from the German Research Foundation to Andreas Voss (Grant VO1288/2-2) and to the Graduate School SMiP (GRK 2277; Statistical Modeling in Psychology). We further thank our students Anne Bülow, Sarah Hladik, Henrike Jungeblut, Julia Karl, Sontje Nordholt, Yannick Roos, and Moritz Spielberger for their help during data acquisition. Veronika Lerche and Mischa von Krause contributed equally to this project.

Correspondence concerning this article should be addressed to Veronika Lerche, Department of Psychology, University of Heidelberg, Hauptstrasse 47-51, D-69117 Heidelberg, Germany. E-mail: veronika.lerche@psychologie.uni-heidelberg.de

[1] In this article, we use the term *general intelligence* to denote a general factor that statistically emerges in intelligence tests (in the sense of sampling theories, e.g., Kovacs & Conway, 2016). Our use of the term general intelligence does not imply that we assume this factor to be a causal factor. In fact, our study does not have the aim of providing any inferences regarding the question of causality.

drift rate—is supposed to provide a pure measure of speed of information processing, with other processes (such as speed of motoric response execution, or speed-accuracy settings) "partialled out." It is a known property of the diffusion model that changes in drift rate have a larger influence on the tail than on the leading edge of RT distributions. More specifically, Ratcliff and McKoon (2008) report that changes in the .9 quantile of RT distributions are typically four times as large as changes in the .1 quantile. Changes in other parameters of the diffusion model—which measure processes such as speed-accuracy settings (threshold separation parameter) or the duration of encoding and motoric processes (nondecision time parameter)—on the other hand, do not have this asymmetric influence on fast versus slow RTs. In line with this reasoning, Schmiedek et al. (2007) found the drift rate (but not other diffusion model parameters) to be related to intelligence. In the following years, other studies also supported the notion that intelligence as measured by classical intelligence tests is associated with the drift rate (e.g., Ratcliff, Thapar, & McKoon, 2011; Schmiedek et al., 2007; Schmitz & Wilhelm, 2016; Schubert, Hagemann, Voss, Schankin, & Bergmann, 2015).

In contrast to drift rate, mean RTs are influenced by a number of different processes (e.g., how cautious individuals are and how fast they execute the motoric response). In fact, for these other processes, for which the diffusion model provides distinct measures, no consistent correlations with intelligence have been found. The only relationship that has been reported several times is a small negative correlation of intelligence with nondecision time, indicating that more intelligent people are faster in nondecisional processes, that is, in encoding and/or motoric processes (McKoon & Ratcliff, 2012; Schubert et al., 2015; Schulz-Zhecheva, Voelkle, Beauducel, Biscaldi, & Klein, 2016). In several other studies, however, this relationship between intelligence and nondecision time has not been found (e.g., Schmiedek et al., 2007; Schmitz & Wilhelm, 2016). Critically, previous studies that examined relationships between diffusion model parameters and intelligence are based on only limited numbers of tasks and they used different estimation approaches, which might account for inconsistencies in the findings.

To sum up, according to the literature distinct effects of speed of information processing on RT distributions account for the worst performance rule. Furthermore, whereas drift rate seems to be consistently related to intelligence, for the other diffusion model parameters the current state of research is inconsistent. We will now come back to the question of domain-specificity of mental speed. The diffusion model, which has proved useful for the examination of the worst performance rule, might also help to gain further insights into this finding.

Interestingly, previous studies did not find clear support for a three-factor structure (numeric, figural, verbal) in RT tasks, suggesting that there are no substantial domain-specific factors of speeds of information processing (Levine et al., 1987; Neubauer & Bucik, 1996). This observation is in contrast to findings from intelligence tests that assume a hierarchical structure of intelligence with both a general factor and domain-specific factors (e.g., verbal, numeric, figural; Jäger, Süß, & Beauducel, 1997). However, it might be difficult to draw definite conclusions from the mental speed studies by Levine, Preddy, and Thorndike (1987) and Neubauer and Bucik (1996) as they did not explicitly disentangle processing speed from other processes. The mental speed measures used in these studies might, thus, have been distorted and may therefore have been no valid indicators of actual speed of information processing. Notably, the studies did find a tendency for domain-specific correlations (i.e., higher correlations between intelligence and mental speed in the respective domains) although their data did not contain compelling evidence for a hierarchical factor structure of mental speed. Moreover, effects were not consistent and very small. Thus, we hypothesize that the measures of processing speed used might not have been pure enough to find clear support for domain-specificity. Using drift rate as a purer measure of cognitive speed provides a more powerful and fairer test for the question, whether cognitive speed has stable domain-specific components. The diffusion model literature, though, so far only reports one general drift rate factor, and Schmiedek et al. (2007) see their results as suggesting that "underlying mechanisms could be relatively task-independent" (p. 425). Notably, however, previous diffusion model studies only used a very restricted number of tasks per domain. Accordingly, the existing literature does not allow to draw clear inferences as to whether there is only one common speed of information processing or whether there are domain-specific speeds. It is further unclear whether domain-specific processing speeds (if they exist) are related to the respective intelligence test scores or just to general intelligence.

To sum up, we see two important research gaps that have not been addressed by previous studies analyzing the association of cognitive speed and intelligence with the diffusion model framework. These gaps originate from restrictions in the number and breadth of the employed tasks. First, whereas previous studies found clear evidence for an association of drift rate and general intelligence, results regarding the other diffusion model parameters are less clear-cut. Second, previous diffusion model studies did not vary task content systematically, so it remains an open question whether there are also domain-specific factors of cognitive speed, and whether such domain-specific speeds are related to the respective intelligence test scores.

Another perspective on the research aims listed above relates to the diffusion model as a diagnostic tool: Whereas, in the past, the diffusion model was mainly employed for the analysis of differences between groups or conditions, in recent years it has been proposed to use this methodology also for the analysis of *interindividual differences* in cognitive processes (e.g., Frischkorn & Schubert, 2018; Ratcliff & Childers, 2015; White, Curl, & Sloane, 2016). Our study allows for an examination of whether there are in fact meaningful content-domain specific interindividual differences in the processing of information.

One further important goal of the present study is the comparison of easy (perceptual) tasks versus complex tasks (requiring more complex mental operations). In the past, it was often recommended to apply the diffusion model only to tasks with mean trial RTs of up to 1.5 s (e.g., Ratcliff & Frank, 2012; Ratcliff & McKoon, 2008; Ratcliff, Thapar, Gomez, & McKoon, 2004). Following this rule of thumb, the previous studies that examined links between intelligence and drift rate used easy tasks that required no complex mental operations and thus allowed for very rapid responding. Interestingly, first studies indicate that the diffusion model might also be applicable to more complex tasks, requiring several seconds for response selection (Aschenbrenner, Balota, Gordon, Ratcliff, & Morris, 2016; Lerche, Christmann, & Voss, 2018; Lerche & Voss, 2019). These studies, however, only examined single tasks (e.g., a complex figural task in the studies by Lerche & Voss, 2019) and did not compare easy with more complex tasks. In the present study, we use a large number of both easy and more complex tasks and examine whether the goodness-of-fit of the diffusion model differs between data from easy versus complex tasks.

Furthermore, we test the criterion validity of drift rate in the more complex tasks, analyzing whether drift rate is related to intelligence not only in the fast, but also in the more complex tasks. In fact, for more complex conditions stronger associations of intelligence and mental speed have been reported (Sheppard & Vernon, 2008; see also Coyle, 2017; Marshalek, Lohman, & Snow, 1983). More precisely, the relationship between intelligence and mental speed increases from very simple tasks (RTs of about 300 ms) to moderately complex tasks (RTs around 500–900 ms), but decreases again if tasks get even more complex (RTs of more than 1,200 ms; Jensen, 2005; see also Lindley, Wilson, Smith, & Bathurst, 1995). Thus, there seems to be an inverted-U-shaped relationship between task complexity and the correlation between intelligence and mental speed. In our study, we examine "easy" tasks (around 600 ms; i.e., moderately complex tasks according to the definition by Jensen) and "complex" tasks (around 3,000 ms). Jensen (2005) states the hypothesis that one reason for the decrease from moderately complex to complex tasks is that individual differences in performance strategies play a more important role in complex tasks. Furthermore, Lindley, Wilson, Smith, and Bathurst (1995) point out that in their complex task participants had to repeatedly scan between different task elements resulting in supplemental motor time so that RT became a less accurate measure of processing speed. Notably, drift rate is a more specific measure of processing speed with some strategies (different speed-accuracy settings) or the duration of encoding processes partialled out. Jensen (2005) also mentions that complex tasks show more task-specific factors that can weaken the correlation between RT and *g*. As we use a large number of tasks, we can use a structural equation modeling (SEM) approach, which helps us to control for task specificities. Thus, the use of diffusion modeling and SEM provides us with more specific measures of mental speed and the relationship between mental speed and intelligence. Accordingly, in our study we assume a substantial relationship between drift rate and intelligence also for the more complex tasks.

In the following paragraphs, we first give a brief introduction to the diffusion model (for more detailed information, see Ratcliff, Smith, Brown, & McKoon, 2016; Voss, Nagler, et al., 2013; Wagenmakers, 2009). Next, we present a review of previous studies that examined relationships between intelligence and diffusion model parameters. In the subsequent section, we present theoretical underpinnings of the relationship between drift rate and intelligence. After that, we examine the question of whether the diffusion model is also applicable to more complex RT tasks. Finally, we present the method and results of our study.

Introduction to the Diffusion Model

The diffusion model (Ratcliff, 1978) is a mathematical model that is applicable to decision tasks with two response options. When a participant works on a trial of such a binary task (e.g., color discrimination task, see Voss, Rothermund, & Voss, 2004) she is assumed to accumulate information continuously until she reaches one of two thresholds (see Figure 1). The two thresholds represent either the two response options (response coding) or the response accuracy (accuracy coding; e.g., Figure 1). The distance between the thresholds, the so-called threshold separation (*a*) reflects how much information needs to be accumulated to reach a decision. If individuals are more cautious, they will accumulate more information before they decide for one option. In this case, a larger threshold separation will cause longer RTs and—at the same

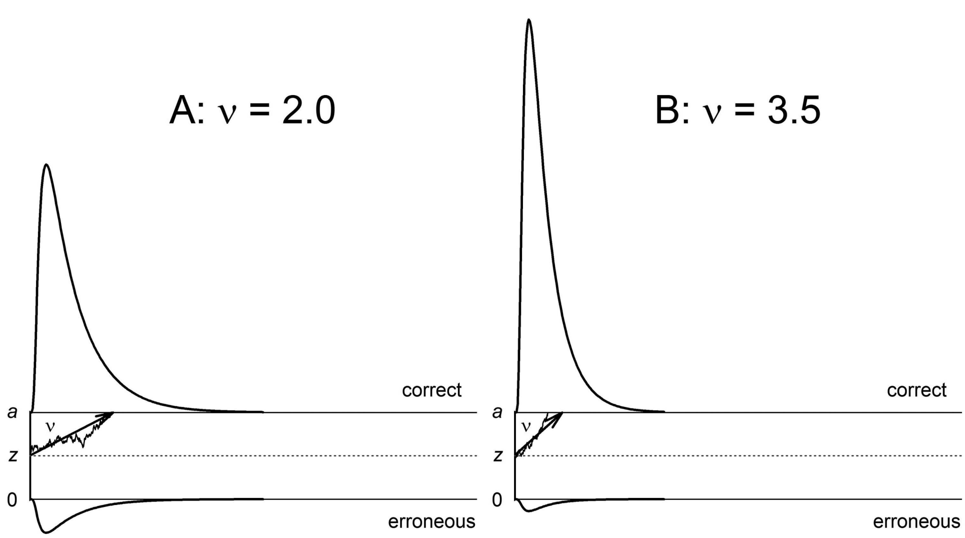

Figure 1. Illustration of the diffusion model. The most important model parameters are threshold separation (*a*), starting point *z* (here situated at the center between the two thresholds), nondecision time (t_0, not depicted in the figure) and drift rate ν. In Panel B, drift (ν = 3.5) is higher than in Panel A (ν = 2.0), which results in more accurate and faster responses.

time— higher accuracy because the decision processes will terminate at the wrong threshold more rarely.

Speed of information processing is denoted as drift (ν) and is illustrated by the arrows in Figure 1, with steeper arrows indicating faster accumulation of information. During information sampling, Gaussian noise is added constantly to the drift, reflecting random fluctuations in the decision process. Due to this noise, the accumulation process does not terminate after the same time and not always at the same threshold, even if the available information (i.e., the stimulus) is identical. The two panels of Figure 1 illustrate the influence of differences in drift on the RT distributions. It can be seen that if the drift is higher (Panel B) fewer errors are made resulting in a smaller distribution at the error threshold and a larger distribution at the correct response threshold. In addition, RT distributions for lower drift rates (Panel A) are more spread out than those for higher drift rates. Another diffusion model parameter is nondecision time (t_0) which subsumes the duration of all nondecision processes, such as encoding of information (preceding the decision process) and motoric response execution (succeeding the decision process). The last parameter of the basic diffusion model is starting point, which maps whether a decision is biased for one of the two response options.

Next to these four main model parameters, often three more parameters mapping intertrial variability of drift s_ν, starting point s_z (Ratcliff & Rouder, 1998) and of nondecision time s_{t0} (Ratcliff & Tuerlinckx, 2002) are estimated. However, the intertrial variability of drift and starting point cannot be estimated reliably and fixation of these parameters to zero can improve estimation of the main diffusion model parameters (Lerche & Voss, 2016; see also van Ravenzwaaij, Donkin, & Vandekerckhove, 2017).

Intelligence and Diffusion Modeling

It is well-known that intelligence shows a high stability over long time periods (e.g., Carroll, 1993; Larsen, Hartmann, & Nyborg, 2008). Accordingly, the rank-order stability of a diffusion model parameter is a prerequisite for it to be related to intelligence. Test-retest studies by Lerche and Voss (2017) provide first evidence that drift rates are rather time stable. More specifically, in Study 1, a lexical-decision task and a recognition memory task were completed at two sessions, separated by a 1-week interval. In a second study, participants worked on an associative priming task (again with a test-retest interval of one week). In all three tasks, drift showed acceptable test-retest correlations. The authors further conducted simulation studies based on the parameters estimated for the empirical data. Specifically, they simulated two data sets (reflecting the two sessions) based on identical parameter values. Interestingly, test-retest correlations of drift rates estimated from the real data were very similar to correlations based on simulated data. This suggests that the speed of information processing was very stable across measurements, and situation influences on drift rate are rather small.

A study by Schubert, Frischkorn, Hagemann, and Voss (2016) corroborates this idea. The authors conducted a test-retest study with a time interval of 8 months. They then used latent state-trait analyses to disentangle trait influences and situation influences. The most important finding was that drift rates had the highest consistencies, indicating that they were the most trait-like parameters. Accordingly, drift rate might be a good candidate for associations with intelligence, which is characterized by high temporal stability and great consistency (Danner, Hagemann, Schankin, Hager, & Funke, 2011).

In support of this hypothesis, in several studies relationships between general intelligence and drift rate have been reported (McKoon & Ratcliff, 2012; Ratcliff, Thapar, & McKoon, 2010; Ratcliff et al., 2011; Schmiedek et al., 2007; Schmitz & Wilhelm, 2016; Schubert et al., 2015; Schulz-Zhecheva et al., 2016). These studies measured drift rates from performance in different types of binary tasks. For example, Ratcliff, Thapar, and McKoon (2010) used a numerosity discrimination task, a recognition memory task, and a lexical-decision task. Intelligence was assessed by means of the vocabulary and matrix reasoning subtests of the Wechsler Adult Intelligence Scale. The authors observed substantial correlations between IQ (mean over the two scales) and drift rate as measured in the lexical decision (r = .53) and recognition memory task (r = .55). The correlation was smaller for the numerosity task (r = .24). As also alluded to by the authors this is not astonishing, as the subscales of the intelligence test that were administered did not address the numeric domain, but the verbal (vocabulary subtest) and figural domain (matrix reasoning subtest). Only small-to-moderate values were observed for the correlation of intelligence with threshold separation and nondecision time ($|r|_{max}$ = .33).

In a subsequent article, Ratcliff et al. (2011) reported correlations between IQ and diffusion model parameters from an item recognition memory task and an associative recognition memory task. Again, there were substantial correlations between the IQ scales and drift rate with r = .36–.68 for college age participants and r = .47–.67 for participants aged 60–74 years. For the oldest group (75- to 90-years-old), correlations were smaller (r = .18–.34), which was seen as partly attributable to floor effects and lower reliability of the vocabulary subtest. For threshold separation and nondecision time, an inconsistent pattern of mostly small correlations with IQ emerged across tasks and age groups. McKoon and Ratcliff (2012), who assessed participants of the same three age groups with the same two subtests of the Wechsler Intelligence Scale, also found IQ to be correlated with drift rates for associative recognition (rs between .24 and .68) and item recognition (rs between .49 and .68). In addition, nondecision times were negatively related to IQ, suggesting faster encoding and/or response execution of more intelligent participants.

Schubert et al. (2015) report results from three elementary cognitive tasks (Hick task, Sternberg memory scanning task, and Posner letter-matching task). Intelligence was assessed in this study with Raven's Advanced Progressive Matrices and with a shortened version of the knowledge test of the German Intelligenz-Struktur-Test 2000-R. In line with the results of the previously reported studies, the authors observed a correlation of r = .50 between the component score of drift rates from the different tasks (extracted from principal component analyses) and general intelligence. In addition, like in the study by McKoon and Ratcliff (2012), a negative relationship between intelligence and nondecision time emerged (r = −.42). Thus, the more intelligent individuals not only showed higher drift rates but also shorter nondecision times.

Schmiedek et al. (2007) used a larger number of different tasks: two lexical tasks, two numeric tasks, and four spatial tasks. For the assessment of intelligence, the authors employed tasks of the

Berlin Structure of Intelligence Test (BIS; Jäger et al., 1997). More specifically, three numeric, figural, and verbal tasks from the reasoning and psychometric speed operation scales were used. Based on structural equation modeling (SEM), the authors found that the latent factor of psychometric speed correlated highest with latent drift rate ($r = .59$), whereas the correlations were smaller for threshold separation ($r = -.42$) and nondecision time ($r = -.04$). Similarly, for reasoning the highest correlation emerged for drift rate ($r = .79$; threshold separation: $r = -.48$; nondecision time: $r = .25$).

Schmitz and Wilhelm (2016) also reported relationships of drift with intelligence. Using two different cognitive tasks and also employing SEM to link the drift rates to a measure of fluid intelligence (a figural sequence reasoning test from the BEFKI; Wilhelm, Schroeders, & Schipolowski, 2014) they found correlations with drift of $r = .15$ (nonsignificant) for visual search and of $r = .29$ for visual comparison. The authors did not report any significant correlations between fluid intelligence and the other diffusion model parameters.

Schulz-Zhecheva et al. (2016) tested a sample of participants aged 8 to 18 years with Cattell's Culture Fair Intelligence Test (CFT 20-R; Cattell & Cattell, 1960; Weiss, 2006) of fluid intelligence and measured diffusion model drift rates across four simple decision tasks. The latter consisted of deciding whether a number was odd or even, whether a number was smaller or larger than 50, whether an arrow pointed upward or downward and whether a line was shown in the upper or lower half of the screen. Once more, drift rate was by far the strongest correlate of fluid intelligence (*gf*; $r = .41$; nondecision time: $r = -.20$; threshold separation: $r = -.13$). The total *gf* factor variance explained by the diffusion model parameters was 19%.

In sum, drift rate seems to have a trait-like characteristic, showing moderate consistency across different tasks and temporal stability. Moreover, robust relationships between drift rates and intelligence have been reported across different studies and experimental tasks. In contrast, correlations of the other diffusion model parameters with intelligence are smaller and the pattern is less consistent. Apart from the relationship with drift rate, the finding that has been most often reported is a negative correlation between intelligence and nondecision time. However, this relationship only showed up in some of the studies.

From the previous diffusion model literature, no clear conclusions can be drawn regarding the existence of domain-specific drift rates. Whereas the findings by Schmiedek et al. (2007) speak in favor of task-independence of speed of information processing, other studies lend first support to the hypothesis that speed of information processing might differ between domains. For example, Ratcliff et al. (2010) who measured intelligence with a verbal and a figural test found a smaller correlation of intelligence with drift in a numeric task than in a verbal or a figural task. Furthermore, in the study by Schubert et al. (2016) drift rates showed smaller consistencies than typically observed in intelligence tests, suggesting that individual differences in drift rates also reflect task- and content-specific properties to a substantial degree. Importantly, a study that combines domain-specific intelligence assessment with a battery of various RT tasks that tackle these domains is still missing. It is an open question whether a domain-specific structure of speed of information processing can be found and if so, if such domain-specific drift rates correlate with the respective domain scores of an established intelligence measure. To address these questions, in our study, we put together a battery of 18 different binary RT tasks that address the three different domains of intelligence.

Relationship Between Drift Rate and Intelligence: Theoretical Considerations

As we described in the last section, empirical findings support the view that speed of information processing as measured by the drift rate of the diffusion model is related to intelligence. Next, we will outline why this relationship is theoretically plausible and why we assume that in more complex tasks relationships between drift rate and intelligence might be even stronger than in less complex tasks.

For illustration, let us consider the two mechanisms proposed by Salthouse (1996) to describe the assumed effect of age-related slowing on cognition, the *limited time mechanism* and the *simultaneity mechanism*. The *limited time mechanism* is supposed to be in effect when the time for solving a problem is limited and only little time is available for the higher-order integration of information, because earlier stages of information processing occupied too much time. The *simultaneity mechanism* assumes that, over time, information becomes less available in working memory. If older individuals need more time to process information, a greater amount of information will then be lost or at least fragmented by the time they start to integrate all processed information. Accordingly, we assume that individuals who have a reduced speed of information processing (i.e., a smaller drift rate) will suffer more from time constraints, as they have less time available for higher-order processing. Furthermore, for these individuals (in contrast to individuals with higher drift rates) more information will get lost during the accumulation process. The importance of temporal aspects in information-processing has also been stressed, for example, by the time-based resource-sharing (TBRS) model (Barrouillet, Bernardin, & Camos, 2004; Camos & Barrouillet, 2014). The model supports the view of a time-related decay of memory traces and regards the number of necessary memory retrievals and the time given to perform them as important factors influencing performance. More complex tasks will often require more memory retrievals than simple RT tasks (e.g., perceptual or recognition memory tasks), with time pressure kept constant between task types. Accordingly, more complex RT tasks might be more vulnerable to deficits in speed of accumulation of information. In other words, task-related differences in working memory demands might underlie higher relationships between more complex tasks and intelligence.

A similar idea is part of the *process overlap theory* (Conway & Kovacs, 2015; Kovacs & Conway, 2016; see also Kan, van der Maas, & Kievit, 2016), a recently proposed intelligence theory. According to this theory "executive/attentional processes" play an important role, underlying—among other—both the worst performance rule and the finding of higher relationships with intelligence for more complex tasks. Process overlap theory is considered a modern version of Thomson's sampling theory (Thomson, 1916). According to Thomson (1916), each mental test addresses a number of what has later often been called "bonds" (see Deary, Lawn, & Bartholomew, 2008, for a historical analysis). This account explains correlations of performance across tasks by an overlap of required psychological processes (in the intelligence literature also often referred to as *positive manifold*). Rather than assuming a causal general factor of intelligence, process overlap theory regards the *g* factor—that undoubtedly shows up in any factor analysis of cognitive ability test data—as an "emergent property" (Kovacs & Conway, 2016, p. 162).

In contrast to Thomson's theory, process overlap theory does not postulate an additive overlap of processes but assumes a bottleneck in form of multiplicatively linked "executive/attentional processes" (Kovacs & Conway, 2016; see Schubert & Rey-Mermet, 2019, for a critical discussion of the empirical testability of this hypothesis). Kovacs and Conway (2016) state that "*g* loadings depend on the involvement of executive processes seated primarily in the prefrontal cortex rather than on the number of processes measured" (p. 170) and define *complexity* as "the extent to which a test taps executive/attentional processes" (p. 164). Accordingly, they suppose the relationship between more complex tasks and intelligence is driven by the engagement of executive processes. Similarly, it is assumed that the slower trials in a task are more highly related to intelligence because they are indicators of failures in executive processes. We support this view of a common explanation of both these empirical observations. More specifically, we assume that the drift rate of the diffusion model might provide a methodological account for both observations. It has already been demonstrated that the drift rate provides an explanation for the worst performance rule (e.g., Schmiedek et al., 2007). So far, however, no study has examined relationships between intelligence and drift rate in more complex tasks. In our study, we examine complex tasks with RTs of about 3,000 ms, thus tasks for which according to Jensen (2005) relationships between mental speed and intelligence should be small because of higher influences of individual differences in strategies. As the diffusion model provides a more specific measure of mental speed (e.g., partialling out speed-accuracy settings), we assume that also for more complex tasks there should be a substantial relationship between mental speed (measured by means of the drift rate) and intelligence. This relationship might even be larger than for less complex tasks because of higher memory demands.

In short, we suppose that a higher speed of information processing helps to counteract time-related decay of memory. This might be particularly relevant for tasks with higher memory demands. In our study, we examine both fast tasks with little memory demands and more complex tasks with higher memory demands. As we will outline in the next section, we assume that the diffusion model is also applicable to such more complex tasks.

Diffusion Modeling for Fast Versus More Complex Tasks

In the past, the diffusion model has almost exclusively been applied to *fast* tasks. By this term, we refer here to tasks with a mean trial duration of below 1.5 s. The claim that the diffusion model is only applicable to such fast tasks has been repeatedly put forth (e.g., Ratcliff & Frank, 2012; Ratcliff & McKoon, 2008; Ratcliff, Thapar, et al., 2004) and has strongly influenced the choice of tasks for diffusion modeling for a long time. The reasoning underlying this restriction is that tasks with longer RTs were seen as more likely to violate basic assumptions of the diffusion model (such as the assumption that decisions are based on a single processing stage and that parameters remain constant over time within one trial). However, we question the idea that data from more complex tasks are more likely to violate assumptions of the diffusion model.

Let us first consider response time tasks that fulfill the 1.5 s rule, that is, typical RT tasks to which the diffusion model has been applied frequently, such as a color discrimination task. In this task, participants have to decide whether, for example, the color orange or blue prevails in a square filled with pixels of these two colors (e.g., Germar, Schlemmer, Krug, Voss, & Mojzisch, 2014; Voss et al., 2004). Participants are assumed to sample evidence from the perceptual dimension (here, color). In such perceptual tasks, it is plausible that participants continuously sample information (i.e., perceptions of color), until they are reasonable sure that one color prevails. However, the diffusion model has also often been applied to tasks in which a continuous sampling of information is less plausible. Imagine, for example, the lexical decision task (Ratcliff, Gomez, & McKoon, 2004). Here it is unclear, whether—during decision making—information of "wordiness" of a stimulus is accumulated with constant drift. Rather, different prelexical (e.g., bigram frequencies) and postlexical (e.g., similarity to existing words) processes could inform the decision with different impact, thus resulting in separate decision stages with different drift rates.

Because there is no way to assess the assumptions of the diffusion model analytically, the model has to be validated empirically, both regarding its general ability to fit empirical data and regarding the external validity of all model parameters. Such validation studies are essential for any cognitive model and any new type of task. One important tool in this regard are so-called selective influence studies that demonstrate that specific experimental manipulations with high face validity take impact on specific model parameters in a specific way. Importantly, such selective influence studies have shown comparably good validity of the diffusion model parameters for color discrimination (Voss et al., 2004) and recognition memory (Arnold, Bröder, & Bayen, 2015). Accordingly, even in the recognition memory task the model assumptions are apparently not seriously violated.

Imagine now a more complex task, for example, the complex figural task used in our study (see Figure 2, for an example stimulus). In each trial of this task, participants see several rectangles. Half of the rectangles are surrounded by a blue border and half of them by a red border. Participants have to estimate the total area of the blue-bordered rectangles and compare it to the total area of the red-bordered rectangles in order to assess which of these summed areas is larger. In studies by Lerche and Voss (2019), the variant of the complex figural task employed led to mean RTs of about 7 s per trial. Answers of participants to an open-framed question about their use of strategies revealed that a typical strategy is to sequentially pick pairs of rectangles and compare the two rectangles within one pair to each other (i.e., one red- and one blue-bordered rectangle). Apart from the high perceptual and spatial affordances (e.g., considering color of borders, and both width and height of rectangles at different positions on the screen), also memory processes are relevant. Participants need to remember which of the rectangles they have already compared and how large the differences were. Thus, this task can be partitioned into several subtasks. For example, each pair of rectangles could be seen as one subtask (with each of these subtasks consisting of further subtasks). Each subtask might be conceived of as

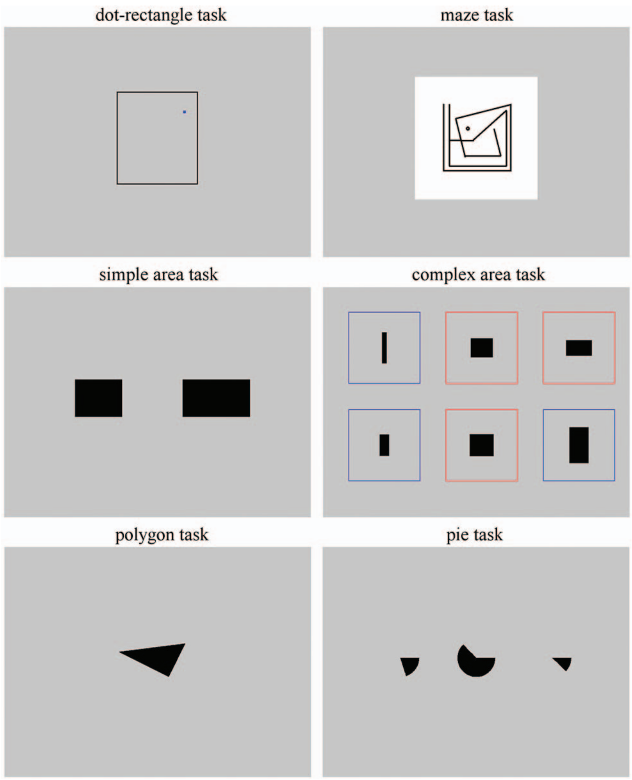

Figure 2. Example for stimuli from the fast figural tasks (left) and the slow figural tasks (right). See the online article for the color version of this figure.

having its own speed of information processing. Following the concept of the law of large numbers, with an increase in the number of subtasks, extreme values of drift rate in single subtasks might become less influential, allowing for an even better measurement of overall mental speed. Thus, we assume that the data of tasks such as the complex figural task can be modeled adequately by a constant drift (i.e., on average, information accrues toward the correct boundary) with Gaussian noise (reflecting nonsystematic influences).

Importantly, in selective influence studies based on the complex figural task, convergent and discriminant validity of the diffusion model parameters were comparable to what has been observed in the validation studies based on faster tasks (Lerche & Voss, 2019). Furthermore, in another study, data from a complex *verbal* task were entered into a diffusion model analysis (Lerche et al., 2018). In this task, participants had to assess the meaningfulness of sentences, which took 2.2 s on average. Results again demonstrated an excellent fit of the diffusion model. Thus, these first empirical findings support our claim that the diffusion model can also be applied to tasks with mean response times above 1.5 s. In the present study, we build upon these promising results and employ both fast and more complex tasks. We compare the model fit between these two types of tasks and examine the external validity (analyzing the relationship of drift rate with intelligence).

The Present Study

In the present study, an intelligence test battery and a battery of 18 binary RT tasks were administered to a sample of 125 participants. The RT tasks included both simple and complex tasks addressing three content domains (numeric, figural, and verbal). With our study, we pursued three main objectives: First, we aimed to replicate findings from previous studies showing that general intelligence correlates with drift rate measured across a variety of different tasks. That is, we expected a substantial relationship between general intelligence and the drift rates across tasks. Second, we wanted to examine whether there are domain-specific aspects of cognitive speed as measured by drift rates and—if so—whether these are related to the respective numeric, verbal, and figural aspects of intelligence, as measured by an intelligence test. Third, we aimed at further investigating the applicability of the diffusion model to more complex RT tasks, which require more time for response selection. Specifically, we compare model fit from nine fast and nine more complex tasks. We also examine how drift rates estimated from the more complex tasks specifically predict general intelligence.

Method

Participants

We determined the required minimum sample size for structural equation analyses with a power analysis following the procedure described by Kim (2005). According to this procedure, the proposed minimum sample size for a test of close model fit according to the root mean squared error of approximation (RMSEA) is 113 (df = 350, α = .05, β = .05). We recruited 125 participants for the study to ensure adequate power.[2]

We used different recruitment methods. The largest part of participants was recruited via a newspaper article. Others were hired via the participants' pool of the Psychological Institute of Heidelberg University in Germany using the software hroot (Bock, Baetge, & Nicklisch, 2014) or by means of fliers that were distributed at public places. We obtained informed consent from all participants. Participants were remunerated with 35€ after data collection was completed. In addition, all participants received feedback about their performance. Participants were between 18 and 65 years old (M = 36.0, SD = 14.3). Sixty-three percent were females. The percentage of students amounted to 50%.

Design and Procedure

The study consisted of three sessions. In the first session, participants had to work on an intelligence test.[3] In the second and third session, all RT tasks were administered (with nine of these tasks in each session). The order of tasks was identical for all participants and is provided in Table 1. Tasks of the three different

[2] Following suggestions of our reviewers, we kept the structural equation models simpler than in our original analysis plan. Most importantly, for the intelligence data, we used scale means rather than the single task scores, leading to a lower number of *df*s in our models.

[3] N = 11 participants had already participated in a previous study in which the same intelligence test was administered. These participants, therefore, only took part in the two PC assessments and received 25€.

Table 1
Overview of the 3 (Domain: Numeric vs. Verbal vs. Figural) × 2 (Speed: Fast vs. Slow) × 3 (Number of Tasks) = 18 RT Tasks

Domain	Fast	Slow
Numeric	• FN1: number discrimination task (2.2) number is greater vs. smaller than 500 • FN2: odd-even task (1.5) number is odd vs. even • FN3: simple inequation task (2.8) inequation is correct vs. wrong	• SN1: mean value computation task (1.8) 16 numbers with mean greater vs. smaller than 500 • SN2: equation task (2.5) equation is correct vs. wrong • SN3: complex inequation task (1.2) equation on left or right side is larger
Verbal	• FV1: word category task (2.6) word is adjective vs. noun • FV2: lexical decision task (1.1) letter combination is word vs. nonword • FV3: animacy task (1.7) noun is living vs. nonliving	• SV1: grammar task (1.4) sentence with grammatical error in possessive pronoun vs. noun • SV2: statement task (2.3) statement is correct vs. wrong • SV3: semantic category task (2.9) several nouns with one vs. two nouns not belonging to the superordinate category
Figural	• FF1: dot-rectangle task (1.9) dot within vs. outside of rectangle • FF2: simple area task (2.4) rectangles with larger area on the left vs. right side • FF3: polygon task (1.3) polygon is triangle vs. rectangle	• SF1: maze task (2.1) maze solvable vs. insolvable • SF2: complex area task (1.6) six rectangles with larger total area of red vs. blue bordered rectangles • SF3: pie task (2.7) three pie slices making more vs. less of a total pie

Note. The first letter indicates the task complexity (F = fast, S = slow); the second letter denotes the domain (N = numeric, V = verbal, F = figural). The numbers in parentheses indicate the time point of assessment (session and number in sequence).

domains and fast and slow tasks were presented alternatingly. After the third and the sixth task within each session, participants took a break of 3 min.

Each of the 18 tasks started with four practice trials. In these trials, participants received feedback about the correctness of their response (green checkmark vs. red cross for correct vs. erroneous responses, respectively; presentation duration: 1,500 ms). After the practice trials, 100 test trials (preceded by one warm-up trial) were administered. All tasks had a binary response format, with both responses correct in half of the trials. Simulation studies have shown that the diffusion model can provide reliable parameter estimates for about 100 or even fewer trials (Lerche, Voss, & Nagler, 2017). The practice and warm-up trials were discarded from subsequent analyses. The order of trials was determined randomly and was held constant for all participants. In each trial, participants had to press one of two keys (“A” or “L”). The key assignment was identical for all participants. Each trial started with the presentation of a fixation cross for 500 ms. Subsequently, the target was shown and remained on the screen until the participant responded. Participants were instructed always to respond as fast and accurately as possible. The next trial started after an inter-trial-interval of 500 ms.

The fast tasks took between 528 and 810 ms on average per trial (M = 655 ms) and the slow tasks took between 2,469 and 4,314 ms (M = 3,319 ms). The mean duration of assessment sessions was 71 min for Session 2 and 69 min for Session 3.

Intelligence Assessment

For the assessment of intelligence we used the BIS (Jäger et al., 1997) which relies on the bimodal Berlin intelligence structure model (Jäger, 1982). This model comprises operation-related and content-related components of general intelligence. Of interest to our study were the content-related components (numeric, figural, and verbal). The intelligence assessment was run in sessions of six participants at maximum and took on average 50 min.

Whereas Schmiedek et al. (2007) selected only nine tasks that were all taken from the reasoning and psychometric speed operations, we also used the memory tasks of the short scale BIS (Jäger et al., 1997), which resulted in a total of 12 tasks originating from three of the four operations tapped in the test (reasoning, psychometric speed, memory, and idea fluency). We excluded the tasks on idea fluency because they are more related to creativity than to the construct of intelligence (cf. Schmitz & Wilhelm, 2016). Consequently, verbal, numeric, and figural domains were represented by four tasks each. To keep the structural equation models as simple as possible, we used scale means as manifest variables for each of the three content domains.

Response-Time Tasks

The study consisted of 3 (domain: numeric vs. verbal vs. figural) × 2 (speed: fast vs. slow) × 3 (number of tasks) = 18 different RT tasks (see Table 1). In the following, we briefly describe the different tasks and materials.

Numeric tasks. The *fast numeric tasks* were the number discrimination task, the odd-even task, and the simple inequation task. In the *number discrimination task*, participants saw a number in each trial and had to assess whether this number was smaller or larger than 500. The numbers were randomly drawn from a uniform distribution ranging from 100 to 900 (excluding 500), with the restriction that half of the numbers were larger than 500 and that the mean deviation from 500 was identical for the numbers smaller and the numbers larger than 500. In the *odd-even task*, participants had to assess whether a presented number was odd or even. The numbers were randomly drawn from a uniform distribution ranging from 100 to 899 (i.e., a vector including 400 odd and 400 even numbers). In the *simple inequation task*, participants had to decide which of two numbers displayed left and right of the center of the screen was larger. The two simultaneously presented numbers were randomly drawn from a uniform distribution ranging from 1 to 20, with the restrictions that numbers were never identical and that the difference between the numbers did not exceed three.

The *slow numeric tasks* were the mean value computation task, the equation task and the complex inequation task. In the *mean value computation task*, 16 numbers were presented on the screen. Participants had to assess whether the mean of these numbers was smaller or larger than 500. The mean of the 16 simultaneously presented numbers of each trial was either 400 or 600, and the numbers were presented at random positions on the screen (overlapping of numbers was prevented). In the *equation task*, in each trial an equation was shown and participants had to assess whether the equation was correct or wrong. In half of the trials, a multiplication or division had to be performed, respectively. The erroneous equations were generated using several different principles. Specifically, for erroneous equations either the tens digit or the ones digit of the solution were set to incorrect values (e.g., $5 \cdot 7 = 25$ or $4 \cdot 12 = 40$, respectively), the operator was wrong (e.g., $11/3 = 33$), or the order of numerator and denominator was reversed (e.g., $8/64 = 8$). In the *complex inequation task*, participants had to decide which solution of two equations displayed on the left and right side of the screen was larger. The equations were sums and differences of two numbers (e.g., "9 – 6" vs. "19 – 17"). The two numbers were drawn randomly from a uniform distribution between 1 and 20, and the solutions of the sums and differences were in that range as well. The operations for the two equations were randomly determined and could be the same or different for the two equations. Furthermore, the difference between the solutions of the two equations was restricted to a maximum of three.

Verbal tasks. The *fast verbal tasks* were the word category task, the lexical decision task, and the animacy task. In the *word category task*, in each trial a word was presented and participants had to assess whether the word was an adjective or a noun. All words comprised of six letters and had one or two syllables. The words had frequency classes of 12 or above (according to the online dictionary project of the university of Leipzig, retrieved in May 2017, see http://wortschatz.uni-leipzig.de/de), which indicates that the German word "der" ("the") is used at least 2^{12} times as often as the selected stimuli. The mean frequency class of adjectives and nouns was identical ($M = 15$). Thus, all words had a low frequency in German language. In the *lexical decision task*, letter combinations were presented and participants had to assess whether or not these were German words. The stimuli were selected from a lexical decision study by Lerche and Voss (2017). The words were nouns consisting of one or two syllables and four to six letters. The words had a frequency class of 14 or 15 (retrieved in November 2014). The nonwords had been generated by replacement of vowels from valid word. Thus, all nonwords were pronounceable and had plausible bigram frequencies. In the *animacy task*, nouns were presented and participants had to classify these as living versus nonliving. The "living" stimuli could refer to humans, animals or plants. Two of the authors and two further independent raters classified the words unambiguously as living versus nonliving. The words consisted of one to three syllables, four to eight letters, and had frequency classes between 11 and 16 (retrieved in June 2017). The mean frequency class was identical for words classified as living or nonliving ($M = 13$).

The *slow verbal tasks* were the grammar task, the statement task, and the semantic category task. In the *grammar task*, participants read German sentences with grammatical errors and had to indicate whether the error was located in the possessive pronoun or in the noun. All sentences consisted of five words and had a very similar structure: They always started with a personal pronoun and further contained a predicate and an object with a possessive pronoun (e.g., "Er widerspricht seine Chef oft" = "He often contradicts his boss"; the error in the German statement is in the possessive pronoun that should read "seine*m*" instead of "seine"). In each trial, by changing one word—either the possessive pronoun or the object—the sentence could be corrected. The errors were generated using the wrong case (e.g., accusative instead of dative), the wrong gender, the wrong declension, or the wrong number.

In the *statement task*, four to six words were presented at different positions of the screen. The participants had to assess whether or not it was possible to create a true statement using all of the presented words. The words were distributed randomly across the screen. From each set of words one grammatically correct sentence could be composed. An example for a true statement is "ein Lastwagen ist sehr schwer" ("A truck is very heavy") and for a wrong statement is "reiche Menschen haben kein Geld" ("Rich people have no money").

In the *semantic category task*, five nouns were presented one above the other. There was one superordinate category to which most of the words (that is, three or four words) belonged. Either one or two words did not belong to this category. Participants had to indicate whether one or two words did not belong to this superordinate category. The selected words were members of the superordinate categories planets, seating furniture, fruit, tools, baking ingredients, medical specialists, geometric figures, grain, craftsmen, or organs reported by Scheithe and Bäuml (1995). Either three or four words belonged to the same category and one or two belonged to another superordinate category. For example, in one trial the words "Stuhl" (= chair), "Sonne" (= sun), "Sessel" (= armchair), "Sofa" (= sofa), and "Bank" (= bench) were shown. Here, the correct response was 1 because all words except one ("sun") belong to the same superordinate category "seating furniture." In another example, "Weizen" (= wheat), "Mond" (= Moon), "Jupiter" (= Jupiter), "Merkur" (= Mercury), and "Hirse" (= sorghum) were presented. In this case, the correct response was two, because two nouns ("wheat" and "sorghum") do not belong to the dominant category (planets). There are 10 different possibilities for the positioning of two minority category members among the five words and five possibilities for the positioning of one minority category member. Each possible positioning was used equally often.

Figural tasks. Example illustrations of the figural tasks are depicted in Figure 2. The *fast figural tasks* were the dot-rectangle task, the simple area task, and the polygon task. In the *dot-rectangle task*, a rectangle and a dot were shown. Participants had to indicate whether the dot was located within or outside of the rectangle. The rectangles varied in size while the dot was always of the same size. The form of the rectangle and the exact positioning of the dot were determined randomly. In the *simple area task*, two rectangles were shown side by side. Participants had to assess which of the two rectangles was larger. The edge lengths of the rectangles were determined randomly, with the area of the smaller rectangle always comprising 70% of the area of the larger rectangle. In the *polygon task*, polygons were shown and participants had to indicate whether the stimulus was a triangle or a quadrangle. The shapes of polygons were generated randomly.

The *slow figural tasks* were the maze task, the complex area task, and the pie task. In the *maze task*, mazes were presented with a dot positioned inside the maze. Participants had to assess whether or not it was possible to leave the labyrinth (starting from the position of the dot). The mazes were drawn manually with a graphics program. In the *complex area task* (cf. Lerche & Voss, 2019), in each trial six rectangles were shown. Three of them had a red border and three of them had a blue border. Participants had to compare the total area of all red-bordered rectangles with the total area of all blue-bordered rectangles and decide which area was larger. The larger area was always 1.3 times larger than the smaller area. The rectangles were generated randomly based on some restrictions (most importantly, the largest or smallest area was not indicative of the correct answer so that participants really had to assess the total area, see Lerche & Voss, 2019, for details). In the *pie task*, three pie slices were shown in each trial. Participants had to judge whether the three slices—if put together—add up to more or less than a full circle. Between trials, the slices summed up to either 95% or 105%, and each slice comprised between 5% and 95% of a full circle each. The combinations of slices were generated randomly with the restriction that from the summing of only two slices it was not possible to derive a correct answer.

Data Preparation

For all RT tasks, we discarded all responses faster than 300 ms. Furthermore, for each task, trials lying more than three interquartile ranges beneath the first or above the third quartile of the intraindividual logarithmized RT distributions were excluded (see also Tukey, 1977). The percentage of excluded trials was on average 1.3% per task and participant.

One participant interrupted accidentally the experimental program at the beginning of the penultimate task of the session, so that data from two tasks (mean value computation task and dot-rectangle task) are missing for this participant. Furthermore, separately for the different RT tasks, we removed the diffusion model parameter estimates of participants with inadequate model fit (i.e., fit <1% quantile of the simulated data, see below for details on the assessment of model fit; this resulted in an exclusion of 0.93% of the diffusion model parameter estimates). Next, we also excluded the diffusion model parameter estimates, mean RT, and accuracy for a specific person and task if the accuracy rate or mean RT for this specific task and person exceeded the Tukey criterion (i.e., distance from first or third quartile larger than three times the interquartile range; Tukey, 1977).[4] Finally, based on the estimated diffusion model parameters (v, a, t_0), accuracy rates, mean RTs and intelligence scale scores, we computed the Mahalanobis distances to detect multivariate outliers. Two of our participants exceeded the critical value of $\chi^2 = 140.89$ ($df = 93$, $p = .001$) and thus had to be excluded.

Parameter Estimation

We estimated the diffusion model parameters using the maximum likelihood optimization criterion implemented in *fast-dm-30* (Voss & Voss, 2007, 2008; Voss, Voss, & Lerche, 2015). Parameters were estimated separately for each participant and each task. Thresholds were associated with correct (upper threshold) and erroneous (lower threshold) responses. Accordingly, the starting point was centered between thresholds (z_r = 0.5). In addition, we fixed the intertrial variabilities of drift rate and starting point to zero. These two parameters cannot be estimated reliably from low trial numbers and the fixation of these parameters can even improve the estimation of the other model parameters (Lerche & Voss, 2016; see also van Ravenzwaaij et al., 2017). In sum, for each participant and each task we obtained estimates for threshold separation, drift rate, nondecision time, and the intertrial variability of nondecision time.

In order to examine the robustness of our results, we also conducted three additional types of parameter estimation. In the first, we associated the thresholds with the two response categories of the respective task (instead of correct and erroneous responses) and freely estimated the starting point. This way, we could check if accounting for a possible bias in starting point alters our results. With this estimation approach, we obtained two different drift rate estimates per task, one for each response category, and—after multiplying the drift rate for the category associated with the lower threshold by −1—computed the mean of the two drift rates as an overall estimate of drift per task. In our second additional estimation procedure, we examined whether practice effects might influence our pattern of results. Therefore, prior to parameter estimation, we excluded not only the four practice trials and the warm-up trial of each task, but also the subsequent 20 trials. Finally, we combined the two alternative estimation approaches obtaining parameter estimates with a freely estimated starting point while also excluding the 20 additional practice trials.

Some of the tasks employed in our study were similar to tasks that have already been used for diffusion model analyses: Specifically, lexical decision tasks (e.g., Dutilh, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2009; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008; Yap, Balota, Sibley, & Ratcliff, 2012), number discrimination (Ratcliff, 2014; Ratcliff, Thompson, & McKoon, 2015), odd-even tasks (Schmiedek et al., 2007; Schmitz & Voss, 2012), animacy discrimination tasks (Aschenbrenner et al., 2016; Spaniol, Madden, & Voss, 2006; Voss, Rothermund, Gast, & Wentura, 2013), and the complex area task (Lerche & Voss, 2019) have been analyzed with the diffusion model before. However, most tasks, in particular the slow RT tasks (with the exception of the complex area task), have not yet been examined by means of diffusion modeling. Thus, we were particularly interested in whether the model can fit data from all tasks (and especially from the slow tasks)

[4] To test the robustness of our main findings, in additional analyses we excluded univariate outliers in the diffusion model parameters (because we had obtained some extreme estimates, e.g., $t_0 \approx 0$, $a \approx 10$, $v > 10$). The pattern of results remained unchanged when we excluded these values.

reasonably well. Accordingly, we examined the model fit for all tasks (our procedure is reported in the Results section).

Structural Equation Modeling

Our structural equation modeling approach consisted of two main steps. First, we established a measurement model for drift rates and a model of the intelligence test scales, separately. Then, we combined these two models into one complete model. We used the *R* package *lavaan* (Rosseel, 2012) for the structural equation analyses. To deal with missing data we employed the full information maximum likelihood (FIML) estimator included in *lavaan*, which utilizes all available information.

We standardized all observed variables before they were entered into the structural equations to avoid estimation problems resulting from differing variances between the drift rates and the intelligence scale scores. As we were not interested in absolute values, fixing all means to zero is unproblematic. However, the analysis of correlations instead of covariances can lead to biased standard errors and fit indices (Cudeck, 1989). We accounted for this by fixing the model implied indicator variances to one, equal to the manifest indicator variances, as proposed by Cudeck (1989). For examination of model fit we used several fit indices: the χ^2 statistic, the comparative fit index (CFI), the RMSEA, and the Tucker-Lewis index (TLI). We used the cut-off criteria proposed by Hu and Bentler (1999) for evaluation of fit. Please note that due to the use of the FIML estimator, a mean structure was also estimated. We fixed all estimated indicator means to zero (as the variables were standardized), a fact that informs the degrees of freedom for all reported models.

We compared four different measurement models of drift rate. Because it was essential to keep the models as parsimonious as possible, we assumed parallel measurement of all factors by fixing all factor loadings to one and setting all residual variances of items loading onto the same factor equal (see Lord & Novick, 1968, Equations 3.3.1a and 3.3.1b, for the outline of a model of parallel measurement). The four models are shown in Figure 3. The first model (Model 1) assumed a general (*g*) factor of drift rate. This equals the assumption that the common variance in speed of information processing can be explained by a single, general factor contributing to all tasks. Model 2 did not include a *g* factor, but three uncorrelated domain factors. The idea behind this model is that there are different types of speed of information processing for figural, verbal and numeric tasks, and that these are unrelated to one another. In Model 3, we assumed a hierarchical structure of the factors: *g* was modeled as a higher-order factor and the domain factors as lower part of the factor hierarchy. The general factor is here interpreted as the common variance of the domain factors, which—in contrast to Model 2—are thought to be correlated. Thus, Model 3 assumes that speed of information processing has both a general component and domain-specific components.[5] Finally, in Model 4, we fit an extended version of Model 3 adding a factor that captures the specific variance of the slow tasks (M-1 approach; Eid, Lischetzke, Nussbeck, & Trierweiler, 2003). Here, the idea is that speed of information processing in the slower, more complex tasks shares specific common variance. This way, the interpretation of the *g* factor changes: It now comprises the domain-general shared variance of speed of information processing except for the variance solely shared by the slow tasks. As not all of the models are nested, we compare model fit based on AIC and BIC values.

For the BIS intelligence scales, we used a hierarchical model of domains and a superordinate *g* factor (Intelligence Model, see Figure 4). We employed scale means (instead of single item values) as single indicators for each domain (figural, numerical, verbal) to keep the model as simple as possible, fixing residual indicator (not: domain) variances to zero.[6] Domain factor variances were set equal for the three domains. We also fixed the unstandardized loadings of the indicators on *g* and on the domain factors to 1. While this assumption of perfect measurement and parallel structure is certainly an oversimplification, we made this decision because the BIS is an established instrument and the focus of this study is less on the structure of intelligence, but on the structure of speed of information processing and its relationship to intelligence. In the last step, we combined the best fitting model of drift rates and the BIS model (Combined Drift-Intelligence Model).

Although the focus of this work is on drift rate, we also fit the same model structures (Models 1 to 4, see Figure 3) to estimates of threshold separation (a), nondecision time (t_0) and mean logarithmized response times of correct responses. If a measurement model with acceptable fit emerged, we further tested the combined model (i.e., including the intelligence model). In the tables and plots, models are labeled accordingly (e.g., Drift Model 1 or RT Model 1). The data of our study is available on the Open Science Framework project page: https://osf.io/xpbwe/. On the project page, also an R Markdown file is available that allows an examination and replication of all the structural equation modeling analyses.

---

[5] In the literature on the structure of mental abilities, there is an ongoing debate on how hierarchical models compare with so-called bifactor models (see, e.g., Morgan, Hodge, Wells, & Watkins, 2015). The latter assume a structure of both uncorrelated domain factors and a *g* factor, also orthogonal to the other factors. Thus, bifactor models do not make the presumption that the common variance shared by all tasks is due to the variance shared between the domain factors. Empirically, bifactor models often tend to fit better, while at the same time being less understood from a substantive, theoretical perspective (Kan, van der Maas, & Levine, 2019). Bifactor models fit better because with all loadings estimated freely hierarchical models are more constrained: The hierarchical models assume that the proportions of indicator variance accounted for by the domain (residual) factors and the proportions accounted for by *g* are the same for all indicators within a domain (Gignac, 2016). In our modeling approach, we fixed all factor loadings to be equal within each factor, which leads to a case were hierarchical and bifactor models are mathematically equivalent, yielding identical fit indices and estimates of the corresponding variances. We decided to use a hierarchical model instead of a bifactor model because it can be interpreted more intuitively and because it is also the more common model of cognitive abilities found in the literature.

[6] Fixing the indicator variances to zero and using the domain factors as de-facto residuals was necessary to estimate the covariances between the drift domain residuals and the respective intelligence test components.

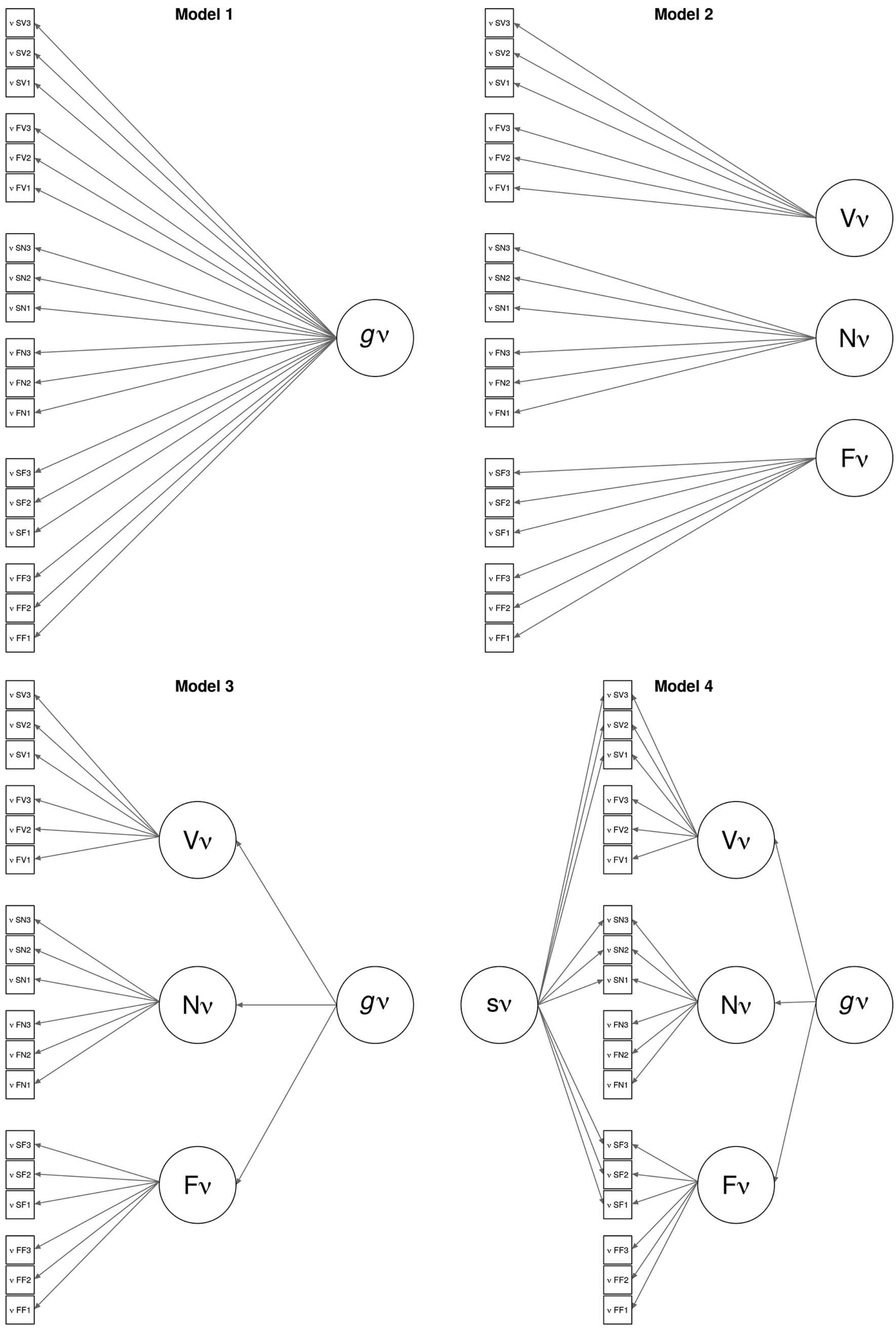

Figure 3 (opposite)

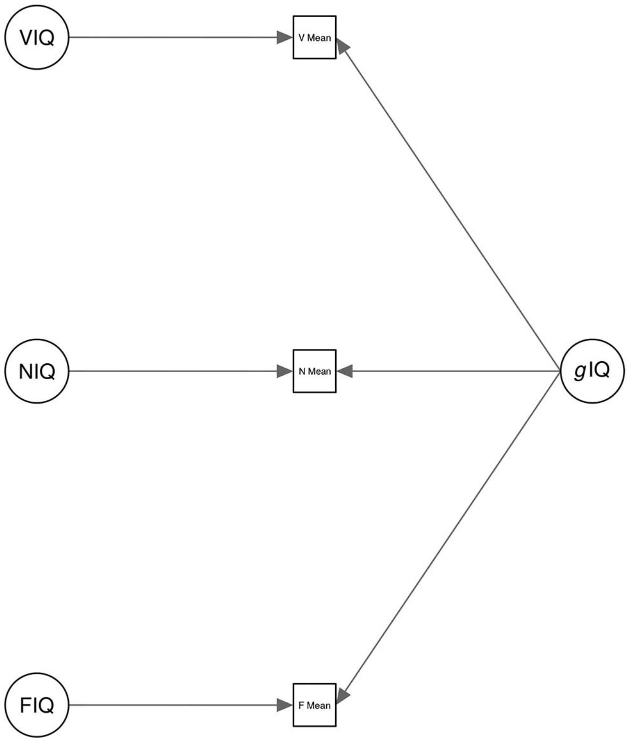

Figure 4. Intelligence Model. Scale means are used as indicators for verbal (VIQ), numeric (NIQ), and figural (FIQ) intelligence. *g*IQ = general intelligence. Indicator residuals are fixed to zero, domain factors serve as quasi-residuals, see Method.

Results

Figures A1 (fast tasks) and A2 (slow tasks) in the Appendix show boxplots of the response times for all 18 tasks. Appendices D to I report descriptive statistics of response times, accuracy rates, drift rates, threshold separations, nondecision times, and intelligence scores.

Fit of the Diffusion Model

Our analyses of model fit comprise two different approaches: First, we examined the fit values of the maximum likelihood optimization. For better interpretation of these values, we conducted simulation studies based on the estimated parameters to infer a criterion for the assessment of model fit (Voss, Nagler, et al., 2013). Second, we analyzed model fit by means of graphical illustrations comparing observed and estimated descriptive statistics.

In the maximum likelihood approach, parameter estimation is based on the maximization of the sum of logarithmized densities over all responses. Boxplots illustrating log-likelihood values for all tasks are given in Figure B1 (fast tasks) and Figure B2 (slow tasks) in the Appendix. Higher likelihood values indicate a better fit of data to the model. One problem with the interpretation of the log-likelihood values is that they depend on the parameter ranges of the specific task. For example, the RT distributions of slower tasks are more spread so that the sum of logarithmized densities is smaller (for an example illustration, see Figure 4 in Lerche & Voss, 2019). This makes it difficult to compare the performance of tasks with different parameter ranges.

To account for this, we conducted simulation studies. More specifically, for each task, we generated 1,000 random parameter sets from multivariate normal distributions, with means, variances, and covariances based on the distribution of estimated parameters. Thus, simulated parameter sets were similar to observed parameters. From each parameter set, we simulated one random data set (using *construct-samples,* which is part of the program *fast-dm*). Therefore, simulated data reflects the assumption that data is based on a diffusion process. Next, we reestimated parameters from simulated data using the same *fast-dm* settings as for the analyses of observed data (i.e., same number of estimated and fixed parameters, same optimization criterion). If the fit values for the real data are worse than those of the simulated data, the observed data probably do not result from a diffusion process only, and consequently, results from the diffusion model analyses might be invalid. Importantly, the distributions of log-likelihood values did not differ systematically between observed data and simulated data, suggesting an excellent model fit (see Figures B1 and B2).

We further defined a criterion to quantify the percentage of observed data sets with poor fit. Specifically, we computed the 1% quantile of the distribution of fit values from simulated data. Maximum likelihood values below this criterion are assumed to indicate poor model fit. This criterion is depicted as horizontal line in each plot. In addition, the plots give the percentage of data sets with fit values below this criterion. The percentages of suspicious fits are very low (at maximum 3.2%) and they are equal for the slow and fast tasks ($M = 1.1\%$). This suggests that the diffusion model fits equally well for the fast and slow RT tasks of our study.

We also examined the model fit graphically, in terms of the precision of predictions for accuracy rates and RT quartiles. Specifically, we constructed scatter plots for each type of task (Domain × Speed) that show the correspondence of different statistics (RT quantiles and accuracy rates) of observed data (*x*-axis) with the respective values predicted from the diffusion model results (*y*-axis; see Figures B3 and B4 in the Appendix

Figure 3. Drift Rate Models 1 to 4. The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. $g\nu$ = general drift rate factor; $V\nu$ = verbal drift rate factor; $N\nu$ = numeric drift rate factor; $F\nu$ = figural drift rate factor; $s\nu$ = method factor for drift rate in slow tasks. All unstandardized factor loadings are fixed to 1. Residuals are omitted from the plot for simplicity. We used the same model structures also for threshold separation, nondecision time, and mean logarithmized response times.

for the fast and slow tasks, respectively). In these figures, each point represents one participant in one task. The figures illustrate that the diffusion model fit the data very well as for all tasks the points are close to the diagonals (all correlations between the empiric and the respective estimated quartiles were larger than .97). Interestingly, the model fits at least as well for slow as for fast RT tasks. Thus, the graphical fit analyses are in accordance with the simulation-based analyses of maximum likelihood values.

The simulation studies and graphical analyses of model fit for the three alternative types of estimation (including estimates of starting point, excluding additional practice trials, and doing both) yielded similar results. The according plots are in the online supplemental material.

Structural Equation Modeling

We started by fitting the measurement models described above (Models 1 to 4, see Figure 3) to the drift rate estimates: Model 1, a *g* factor model; Model 2, a model of uncorrelated domains; Model 3, a hierarchical model of domains and a *g* factor; and Model 4, a model that further added a method factor for all slow decision tasks. Table 2 shows the fit indices for all drift rate models. Figures C1 to C4 in the Appendix show the results for Drift Models 1 to 4 and Appendix J to M report the parameter estimates for each of the four structural equation models, including the unstandardized solution, the corresponding standard errors and *p* values, and completely standardized estimates.

Model 4, the model containing a hierarchical structure of three content domain factors, a superordinate *g* factor, and a method factor for the slow tasks had the best fit in terms of AIC and BIC values (see Table 2) and also regarding the measures of absolute model fit, $\chi^2(df = 184) = 254.40$, CFI = .88; TLI = 0.90; RMSEA = 0.06. Accordingly, we decided to retain this model. It should be noted that the estimated residual variance of the figural drift factor did not differ significantly from zero und should therefore be interpreted accordingly. We kept it in the model in order to a) refrain from post hoc model adjustments and b) make possible replications easier to compare.

The intelligence model is illustrated in Figure C5 in the Appendix, Table 2 shows the fit, and Appendix N parameter estimates. As the fit was good, $\chi^2(df = 8) = 0.18$, CFI = 1.00; TLI = 1.03; RMSEA = 0.00, we used this model for the combined analyses.

Finally, we combined the best measurement model of drift rates (i.e., Model 4) and the Intelligence Model into a Combined Drift-Intelligence Model. We allowed freely estimated covariances between residual figural drift rate and residual figural BIS intelligence, residual numeric drift rate and residual numeric BIS intelligence, residual verbal drift rate and residual verbal BIS intelligence, and the superordinate *g* factor for drift rate and *g* BIS intelligence.[7] In addition, the covariance between the slow decision task factor and the *g* BIS intelligence factor was freely estimated, reflecting our hypothesis that speed of information processing in slow tasks might be especially closely related to general intelligence. Figure 5 shows the resulting model. Model fit was acceptable, $\chi^2(df = 241) = 406.49$; CFI = .82; TLI = 0.84; RMSEA = 0.07 (see Table 2). Table 3 shows the parameter estimates. All latent factors except the figural drift factor had variances significantly different from zero; the same was true for the covariances between them. The relative parts of the variances of the manifest indicators explained by the latent factors are reported in Table 4. Across all tasks, 20% of the variance of drift rates could be attributed to the *g* drift factor, while 3–16% were based on the domain-specific factors. For the complex tasks, an additional 10% of the variance was explained by the slow factor. Overall, the mean task specific and error variance was 63%.

The estimated correlation between figural intelligence and figural drift rate was .90. However, this value should not be overinterpreted because of the very low residual variance of figural drift rate, which did not differ significantly from zero. Numeric intelligence and numeric drift rate correlated with .74. The correlation between verbal intelligence and verbal drift rate was .50, while the correlation between domain general drift rate and general intelligence as measured by the BIS was .45. Finally, the method factor for slow decision tasks and the BIS *g* factor were also strongly correlated ($r = .68$). If the links of the *g* drift and slow drift factors to *g* BIS intelligence were modeled as a regression, the R^2 value of *g* BIS was .67. Thus, the domain general drift factor and the slow drift factor jointly explained two thirds of the variance in general intelligence.

We conducted several robustness checks to ensure our main findings would hold. First, we fit models with completely freely estimated factor loadings and residual indicator variances for both the best measurement model (Drift Model 4, freely estimated, see Figure C6 and Appendix O; see Table 2 for fit indices) and the combined drift-intelligence model (freely estimated, see Figure C7 and Appendix P; see Table 2 for fit indices). In terms of AIC and BIC values, the constrained Drift Model 4 was preferred to the freely estimated version. For the combined drift-intelligence model, AIC was lower for the free model, but the constrained model had the lower BIC value (i.e., better fit). Please note that the number of estimated parameters in the freely estimated models is very large for our sample size and the results should thus be interpreted with caution. In addition, estimation of the Combined Drift-Intelligence Model (freely estimated) yielded a nonpositive definite estimated covariance matrix.[8] Still, while the estimated unstandardized factor loadings in the freed models sometimes differed widely from unity and standard errors were much higher than in the constrained model, leading to statistically insignificant estimates, the main resulting covariances remained much the same. Namely, the estimated correlations between the factors in the freely estimated Combined Drift-Intelligence Model (compared with the constrained Combined Drift-Intelligence Model) were: .56 (.90) for the figural, .90 (.74) for the numeric, and .52 (.50) for verbal drift residual factors and their respective intelligence counterparts. A correlation of .42 (.45) was now found for the relation of *g* drift and *g* BIS and a correlation of .74 (.68) for the association of the slow factor and *g* BIS.

[7] We also fitted a Combined Drift-Intelligence Model freely estimating the covariances between all domain residuals. Only the theoretically implied covariances (figural drift <-> figural IQ, numeric drift <-> numeric IQ, verbal drift <-> verbal IQ) reached statistical significance, except for a negative correlation between verbal drift and figural intelligence ($r = -.34$, $p = .048$).

[8] This problem could be overcome by fixing the residual variance of the figural drift factor, that did not differ significantly from zero, to zero.

Table 2
Fit Indices of Drift Rate Models, Intelligence Model, and Combined Drift-Intelligence Model

Model	AIC	BIC	χ^2	df	CFI	TLI	RMSEA
Drift Model 1	5,773.69	5,776.50	350.71	188	.73	0.78	0.08
Drift Model 2	5,795.32	5,803.75	368.34	186	.69	0.75	0.09
Drift Model 3	5,711.05	5,722.30	282.07	185	.84	0.86	0.07
Drift Model 4	**5,685.38**	**5,699.44**	254.40	184	**.88**	**0.90**	**0.06**
Drift Model 4, freely estimated	5,688.59	5,772.96	207.61	159	.92	0.92	0.05
Intelligence Model	945.39	948.21	0.18	8	1.00	1.03	0.00
Combined Drift-Intelligence Model	6,507.19	6,538.12	406.49	241	.82	0.84	0.07
Combined Drift-Intelligence Model, freely estimated	6,496.67	6,603.53	341.97	214	.86	0.86	0.07

Note. Model 1 = *g* factor model; Model 2 = model of uncorrelated domains; Model 3 = hierarchical model of domains and a *g* factor; Model 4 = model 3 with additional method factor for all slow decision tasks; AIC = Akaike's information criterion; BIC = Bayesian information criterion; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean squared error of approximation. The best-fitting drift rate model among the four alternative models (Models 1 to 4) is in bold. In the freely estimated models, all loadings and residual variances were unconstrained.

Further evidence for the robustness of our results was provided by additional analyses based on different specifications of the diffusion models: Similar results emerged for the structural equation models when drift was estimated using the alternative diffusion model architectures that (a) also estimated the starting point, (b) excluded 20 additional practice trials, or (c) did both. Fit indices and parameter estimates for these models are given in the online supplemental material.

Table 5 shows the fit values for the measurement models of threshold separation, nondecision time, and mean logarithmized response times. Parameter estimates for all these models can be found in the online supplemental material. Of all the measurement models, only t_0 Models 1, 3, and 4 showed somewhat acceptable model fit (RMSEA < 0.08, CFI and TLI at least > 0.82), with Model 4 showing the lowest values in AIC and BIC. Thus, for nondecision time, a hierarchical model of domain factors, a superordinate g t_0 factor and a method factor for slow tasks provided the best fit. However, the residual variances for the figural and numerical domain factors did not reach statistical significance. Appendix Q shows the complete parameter estimates for this model. We also fit a combined model of nondecision time and the BIS intelligence scales (Combined t_0-Intelligence Model, see Table 5 for the fit measures). The model structure was identical to the Combined Drift-Intelligence Model. Table 6 shows the resulting estimates. The nondecision time domain factors were negatively correlated to the respective intelligence factor residuals, as were the g_{t0} factor and the slow$_{t0}$ factor to general intelligence.

Notably, none of our predefined models showed acceptable fit to the mean logarithmized response times. However, the relationship between response times and intelligence is of particular theoretical interest because response times are the measures of mental speed used in most previous studies. Therefore, we additionally conducted an exploratory principal components analysis to explore the covariance structure of response times in our sample. A parallel analysis (Horn, 1965) suggested the extraction of one general component that explained 58% of variance in response time variables. When added to the Intelligence (structural equation) Model as a manifest variable, the component scores explained 65% of the variance in gIQ ($\beta = .80$, $p < .001$; RMSEA = 0.00, CFI and TLI ≥1.00 for this model).

Discussion

Our study focused on the relationship between intelligence and drift rate—a measure of speed of information processing estimated in diffusion model analyses (Ratcliff, 1978). In contrast to previous studies that examined such relationships (e.g., Ratcliff et al., 2011; Schmiedek et al., 2007; Schmitz & Wilhelm, 2016; Schubert et al., 2015), we used a much larger set of RT tasks, and these tasks systematically addressed three content domains (verbal, numeric, and figural). More specifically, we employed six tasks for each of the three domains with half of the tasks of each domain being typical fast diffusion model tasks (mean RT of 660 ms), and the other half being more complex, slower tasks (mean RT of 3,320 ms). Thereby, our study is the first diffusion model study on intelligence that includes not only fast but also more complex RT tasks and uses a large number of tasks per content domain. This allowed us to examine three main substantial questions: First, we tested whether we can replicate the relationship between general intelligence and drift rate that has been found in previous diffusion model studies (e.g., Ratcliff et al., 2011; Schmiedek et al., 2007; Schmitz & Wilhelm, 2016; Schubert et al., 2015). Additionally, we also examined relationships of intelligence with mean RT and other diffusion model parameters. Second, we analyzed whether there are domain-specific aspects of speed of information processing and—if so—whether these domain-specific drift rate factors are related to the respective domains of the intelligence test BIS (Jäger et al., 1997).

In addition to these substantial questions, our study also allows the examination of two methodological issues. First, in the last years it has been proposed to use the diffusion model not only for the analysis of differences between groups or conditions (the typical application in most previous studies), but also for the examination of interindividual differences (e.g., Frischkorn & Schubert, 2018; Ratcliff & Childers, 2015; White et al., 2016). Our study is the first to allow a profound analysis of whether there are meaningful interindividual differences in the content-domain specific aspects of drift rates. Second, in the past, the diffusion model was typically only applied to fast RT tasks. Our study allows inferences about whether the diffusion model fits slower, more complex RT tasks similarly well as typical fast RT tasks. Furthermore, we could examine the external validity of drift rate in more complex tasks, analyzing the relationship with intelligence.

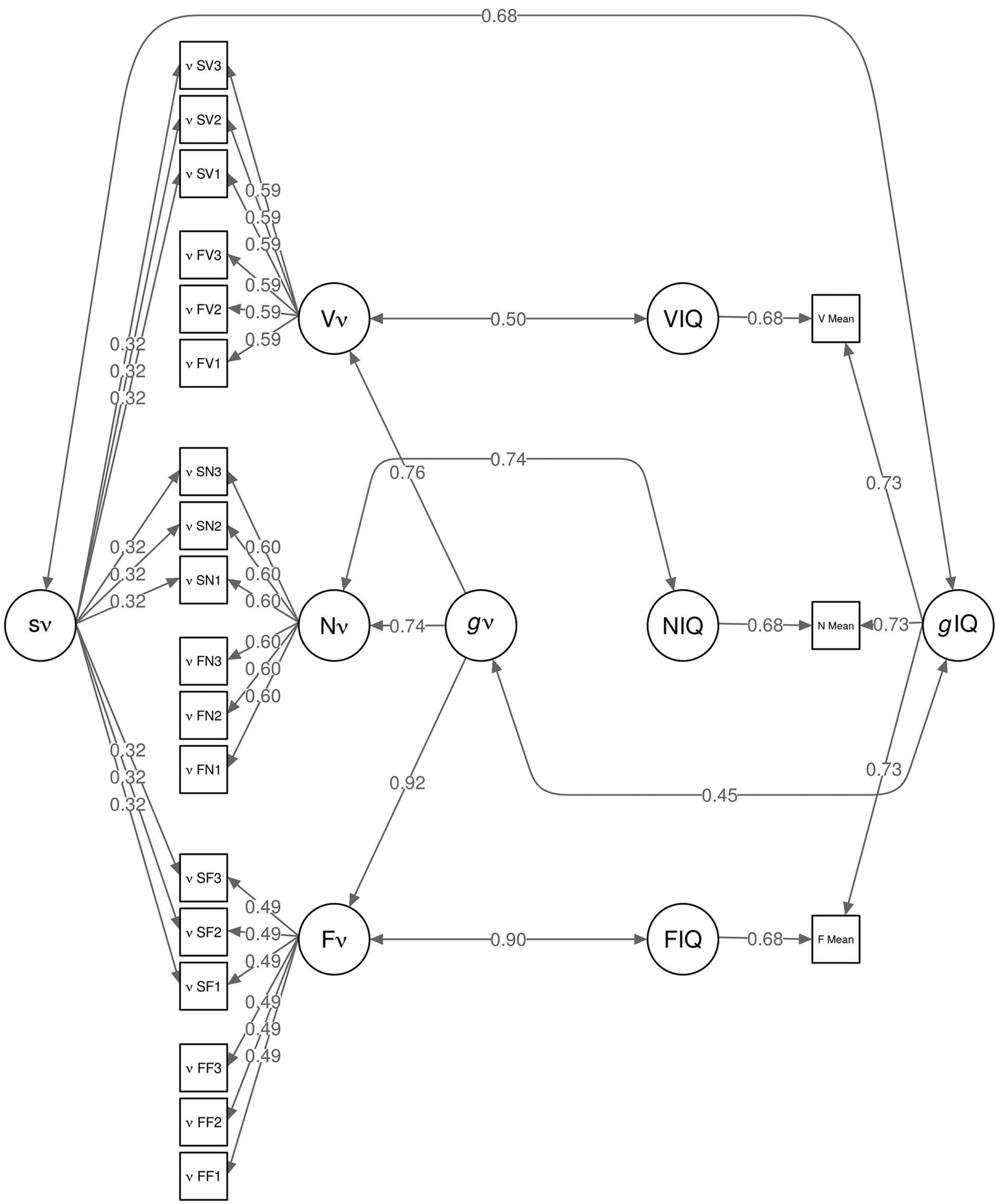

Figure 5 (opposite)

Table 3
Combined Drift-Intelligence Model

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
Loadings					
Fν on ν (each figural task)	1	0			0.487
Nν on ν (each numeric task)	1	0			0.603
Vν on ν (each verbal task)	1	0			0.591
sν on ν (each slow task)	1	0			0.322
gν on Fν	1	0			0.919
gν on Nν	1	0			0.742
gν on Vν	1	0			0.758
gIQ on F_Mean/on N_Mean/V_Mean	1	0			0.734
FIQ on F_Mean/NIQ on N_Mean/VIQ on V_Mean	1	0			0.679
Covariances					
gν with gIQ	0.148	0.035	[0.080, 0.216]	<.001	0.450
sν with gIQ	0.162	0.030	[0.102, 0.222]	<.001	0.684
Fν with FIQ	0.117	0.031	[0.057, 0.177]	<.001	0.899
Nν with NIQ	0.202	0.035	[0.134, 0.269]	<.001	0.736
Vν with VIQ	0.130	0.034	[0.063, 0.197]	<.001	0.497
Latent (residual) variances					
gν	0.200	0.025	[0.152, 0.248]	<.001	1
gIQ	0.539	0.039	[0.462, 0.617]	<.001	1
sν	0.104	0.023	[0.059, 0.149]	<.001	1
Fν	0.037	0.028	[−0.017, 0.091]	.182	0.156
Nν	0.163	0.032	[0.100, 0.227]	<.001	0.449
Vν	0.149	0.031	[0.089, 0.209]	<.001	0.426
FIQ/NIQ/VIQ	0.461	0.039	[0.383, 0.538]	<.001	0.461
Residual indicator variances					
ν (each fast figural task)	0.763	0.033	[0.698, 0.827]	<.001	0.763
ν (each fast numeric task)	0.637	0.031	[0.576, 0.697]	<.001	0.637
ν (each fast verbal task)	0.651	0.032	[0.589, 0.713]	<.001	0.651
ν (each slow figural task)	0.659	0.034	[0.593, 0.725]	<.001	0.659
ν (each slow numeric task)	0.533	0.034	[0.467, 0.599]	<.001	0.533
ν (each slow verbal task)	0.547	0.032	[0.486, 0.609]	<.001	0.547

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

Summary of Results

The presented structural equation analyses replicated findings of previous diffusion model studies in that we found a strong relationship between a general drift rate factor and general intelligence as measured by the BIS. As the general latent factor of drift rates in our study captured the shared variance of 18 different tasks, this provides strong support for the hypothesis that speed of information processing is closely linked to general intelligence. Furthermore, for two out of three content domains (verbal and numeric), we found significant domain-specific drift factors, indicating that there are domain-specific interindividual differences in mental speed that can be assessed with a diffusion model analysis. Strikingly, the three domain-specific latent factors accounted for roughly one third of the shared variance between tasks. Moreover, the domain-specific drift factors were closely related to the respective components of the standard intelligence test. Finally, fit of diffusion models was equally good for fast and more complex RT tasks and speed of information processing in the more complex tasks explained additional variance in general intelligence.

Domain-Specific Speeds of Information Processing

Our study is the first to reveal domain-specific drift factors, which we further found to be related to the respective domain scores of the intelligence test. The variance proportions explained

Figure 5. Combined Drift-Intelligence Model. The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. Completely standardized loadings are reported. Residuals are omitted from the plot for simplicity. The latent correlations between the drift domains and intelligence domains are between the drift domain residuals and the (quasi-residual) intelligence domain factors (see Method). *g*ν = general drift rate factor; Vν = verbal drift rate factor; Nν = numeric drift rate factor; Fν = figural drift rate factor; sν = method factor for drift rate in slow tasks. Scale means are used as indicators for verbal (VIQ), numeric (NIQ) and figural (FIQ) intelligence. *g*IQ = general intelligence. As the loadings of the drift domain factors are standardized on the different freely estimated variances of the domain factors, their standardized values differ although the unstandardized loadings are all fixed to 1.

Table 4
Percentage of Variance Explained by Latent Variables in Manifest Indicators in Combined Drift-Intelligence Model

Task type	*g* Factor	Slow factor	Domain factor	Residual
Fast figural	20.03		3.70	76.27
Slow figural	20.03	10.37	3.70	65.90
Fast numeric	20.03		16.30	63.67
Slow numeric	20.03	10.37	16.30	53.29
Fast verbal	20.03		14.85	65.12
Slow verbal	20.03	10.37	14.85	54.75

by the domain-specific drift factors for numeric and verbal drift are substantial (15% and 16%), challenging the view of only one general mental speed factor. Thereby, our study helps to reconcile research on mental speed with the literature that is based on standard intelligence testing. In the latter, a hierarchical structure with both a *g* factor and domain-specific factors is a very common assumption. Previous mental speed studies might have failed to reveal domain-specific factors due to measurement issues. Specifically, studies that did not employ the diffusion model might have examined a measure of mental speed that is confounded by other processes such as encoding speed, motoric speed, or speed-accuracy settings. The diffusion model has the great advantage of providing a more process-pure measure of mental speed. Furthermore, previous studies employing the diffusion model might have failed to find domain-specific drift rates because the number of tasks that had been used for each domain might have been too low.

Diffusion Modeling for Slower, More Complex RT Tasks?

In the past, it was assumed that the diffusion model is only applicable to fast RT tasks with mean trial RTs below 1.5 s (e.g., Ratcliff, Thapar, et al., 2004). However, first studies support the notion that the model might also be utilized for more complex tasks. Lerche and Voss (2019) conducted experimental validation studies (also often called "selective influence studies") based on a complex figural RT task, and Lerche, Christmann, and Voss (2018) examined model fit of a complex verbal task. The present study offers a unique possibility to compare model fit between easy and more complex tasks, because participants completed both nine complex tasks and nine fast tasks, which were—beside the differences in cognitive demands—very similar. Thus, we could compare model fit (in statistical terms and graphically) between fast and slow tasks and examine correlations with intelligence. Interestingly, the fit of the diffusion model was as good for the more complex as for the simpler tasks.

Furthermore, in our structural equation modeling analyses, a model that included an additional "slow drift factor" (i.e., a factor on which the drift rates of all slow tasks loaded) fitted data better than models without this factor. Furthermore, this slow drift factor was closely linked to general intelligence ($r = .68$). The explained variance (R^2) for drift rates from slow tasks was slightly higher than for drift from fast tasks, due to the latent slow factor that explained 10% of their variance. Thus, drift rates in the more complex tasks are closely related to intelligence, which provides evidence for a good criterion validity of drift rates in this kind of tasks.

The complex tasks that we employed in our study apparently differed in their demands in terms of, for example, memory (e.g., high demands in the "complex area task") or reasoning (e.g., high demands in the "word category task"). We did not manipulate or measure the specific demands in our study. However, it is notable that the diffusion model fit all of our complex tasks very well, thus, fit was independent of the specific task demands. In line with this finding are other recent studies that successfully applied sequential sampling models to tasks with high demands on memory or reasoning. One of them applied the diffusion model to a difficult recognition memory task (Aschenbrenner et al., 2016) and another one applied the linear ballistic accumulator model (Brown & Heathcote, 2008) to an inductive reasoning task (Hawkins, Hayes, & Heit, 2016).

Advantages of the Diffusion Model

Notably, the slow drift factor and the general drift factor together accounted for an impressive 67% of the variance of general

Table 5
Fit Indices of Threshold Separation (a), Nondecision Time (t_0), and RT Models

Model	AIC	BIC	χ^2	*df*	CFI	TLI	RMSEA
a Model 1	5,594.45	5,597.26	485.09	188	.67	0.73	0.11
a Model 2	5,813.55	5,821.99	700.20	186	.43	0.53	0.15
a Model 3	5,597.19	5,608.44	481.84	185	.67	0.73	0.11
a Model 4	**5,502.78**	**5,516.84**	385.42	184	**.78**	**0.82**	**0.09**
t_0 Model 1	5,610.96	5,613.77	316.75	188	.82	0.86	**0.07**
t_0 Model 2	5,791.36	5,799.80	493.15	186	.58	0.65	0.12
t_0 Model 3	5,607.52	5,618.77	307.31	185	.83	0.86	**0.07**
t_0 Model 4	**5,587.65**	**5,601.71**	285.44	184	**.86**	**0.88**	**0.07**
Combined t_0-Intelligence Model	6,457.09	6,488.03	390.73	241	.84	0.86	0.07
RT Model 1	4,887.05	4,889.87	801.89	188	.70	0.75	0.16
RT Model 2	5,076.96	5,085.40	987.80	186	.60	0.67	0.19
RT Model 3	4,794.67	4,805.92	703.50	185	.74	0.79	**0.15**
RT Model 4	**4,760.91**	**4,774.97**	667.75	184	**.76**	**0.80**	**0.15**

Note. Model 1 = *g* factor model; Model 2 = model of uncorrelated domains; Model 3 = hierarchical model of domains and a *g* factor; Model 4 = model 3 with additional method factor for all slow decision tasks; AIC = Akaike's information criterion; BIC = Bayesian information criterion; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean squared error of approximation. The best-fitting model among the four alternative models (Models 1 to 4) is always in bold.

Table 6
Combined t_0-Intelligence Model

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
Loadings					
Ft_0 on t_0 (each figural task)	1	0			0.540
Nt_0 on t_0 (each numeric task)	1	0			0.579
Vt_0 on t_0 (each verbal task)	1	0			0.614
st_0 on t_0 (each slow task)	1	0			0.275
gt_0 on Ft_0	1	0			1.016
gt_0 on Nt_0	1	0			0.948
gt_0 on Vt_0	1	0			0.894
gIQ on F_Mean/N_Mean/V_Mean	1	0			0.731
VIQ on V_Mean/NIQ on N_Mean/FIQ on F_Mean	1	0			0.682
Covariances					
gt_0 with gIQ	−0.266	0.031	[−0.327, −0.206]	<.001	−0.663
st_0 with gIQ	−0.023	0.025	[−0.071, 0.026]	.358	−0.112
Ft_0 with FIQ	−0.047	0.027	[−0.101, 0.007]	.086	−0.709
Nt_0 with NIQ	−0.103	0.030	[−0.161, −0.045]	.001	−0.819
Vt_0 with VIQ	−0.113	0.032	[−0.176, −0.051]	<.001	−0.604
Latent (residual) variances					
gt_0	0.301	0.021	[0.260, 0.343]	<.001	1
gIQ	0.535	0.041	[0.455, 0.615]	<.001	1
st_0	0.076	0.019	[0.039, 0.113]	<.001	1
Ft_0	−0.010	0.022	[−0.052, 0.033]	.657	−0.033
Nt_0	0.034	0.023	[−0.012, 0.080]	.146	0.101
Vt_0	0.076	0.026	[0.025, 0.127]	.003	0.201
FIQ/NIQ/VIQ	0.465	0.041	[0.385, 0.545]	<.001	1
Residual indicator variances					
t_0 (each fast figural task)	0.708	0.029	[0.651, 0.765]	<.001	0.708
t_0 (each fast numeric task)	0.665	0.029	[0.609, 0.721]	<.001	0.665
t_0 (each fast verbal task)	0.623	0.028	[0.567, 0.678]	<.001	0.623
t_0 (each slow figural task)	0.633	0.030	[0.574, 0.691]	<.001	0.633
t_0 (each slow numeric task)	0.589	0.030	[0.529, 0.649]	<.001	0.589
t_0 (each slow verbal task)	0.547	0.031	[0.486, 0.608]	<.001	0.547
F_Mean/N_Mean/V_Mean	0	0			

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

intelligence assessed by the BIS. It is striking that drift rate has such a close relation to intelligence in the present study. In our view, this strong relation—and the advantage of drift rate over mean RT—can be explained by two advantages of the diffusion model.

First, unlike mean RT, the drift provides a common metric that combines both RT and accuracy (Spaniol et al., 2006). Thus, when effects of cognitive ability spread over response latencies and accuracy (i.e., higher ability is negatively related to RT and positively related to accuracy of a task), a common metric is required that captures both effects. This is of special importance, when the main impact of cognitive ability is for one group of participants on speed and for others on accuracy.

Second, the diffusion model makes it possible to disentangle different processes of information processing. Most important, different—and conceptually independent—parameters map speed of information processing, speed-accuracy settings, and nondecision times. For example, participants might be faster or slower, because they are less or more cautious (i.e., error avoiding), respectively. Participants might also differ in the time needed for encoding or motoric responses (i.e., nondecision time parameter). For example, it has been consistently found that older participants are more cautious (i.e., higher threshold separations) and that they have higher nondecision times than younger participants (see Theisen, Lerche, von Krause, & Voss, 2019, for a meta-analysis). This example shows that the validity of pure RT as a measure for mental speed might be problematic (see Coyle, 2017, for a similar argument). In diffusion modeling, the response style (threshold separation) and nondecision time are removed analytically from the index for mental speed (drift). Therefore, drift rate is a more process-pure measure of mental speed than is mean RT, and is thus a better predictor for intelligence.

Are Relationships With Intelligence Specific for Drift Rate?

Importantly, in our structural equation analyses drift rates showed a clear pattern of correlations with intelligence, distinguishing between domain-general and domain-specific aspects, whereas the structural equation models of mean RT did not have a satisfactory fit. Similarly, previous studies that used chronometric tasks and varied the type of material (numeric, verbal, figural) failed to find clear support for domain-specific factors (Levine et al., 1987; Neubauer & Bucik, 1996). These studies examined behavioral variables which—as outlined in more detail in the previous section—are confounded with other processes involved in task execution such as speed-accuracy settings.

Apart from drift rate, for nondecision time, we also observed relationships with intelligence (fitting the same models as for drift rate resulted in a worse, but still acceptable, model fit). Higher scores in the intelligence test were associated with shorter nondecision times. Also in some previous studies, negative relationships between nondecision time and intelligence have been reported (McKoon & Ratcliff, 2012; Schubert et al., 2015; Schulz-Zhecheva et al., 2016), whereas in other studies no such relationship was found (e.g., Schmiedek et al., 2007; Schmitz & Wilhelm, 2016). Our study—which is based on a large number of RT tasks and might thus allow more solid inferences than previous studies—supports the view that there is also a relationship between nondecision time and intelligence (even though this relationship is smaller than for drift rate).

What does this relationship between intelligence and nondecision time indicate? It suggests that "intelligence" as measured by classical paper-and-pencil based intelligence tests is more than speed of information processing. In fact, as already mentioned previously, not only mean RTs in response time tasks, but also performance in paper-and-pencil-based intelligence tests like the BIS can be influenced by different processes. In intelligence tests, it is difficult to distinguish between the different processes that are involved in task completion, such as decision settings (i.e., whether individuals prefer speed or accuracy), motoric elements (e.g., how fast individuals write down their answers), encoding processes, and speed of information processing.[9] Thus, we suppose that nondecision time is related to the BIS because also the paper-and-pencil-based test measures to a certain extent nondecisional components. The nondecision time parameter of the diffusion model includes time needed for encoding and motoric processes. We hypothesize that the correlations with intelligence are probably mainly based on encoding processes rather than on motoric processes. It seems implausible that for motoric components a model with not only a general factor, but also domain-specific factors and a complex task factor emerges. In line with this argument, when the *Jensen box* is used—which allows a separation of the time needed for decision making (termed RT) from the time needed for finger movement (movement time)—RTs clearly increase with increasing task complexity, whereas movement times do not (Jensen, 1987, 2006; see also the differential–developmental model by Coyle, 2017). It is, however, highly plausible that encoding processes differ between domains. Furthermore, the complex task factor could be attributed to the fact that the stimuli in the more complex tasks consisted of more elements than the stimuli in the fast tasks (e.g., several numbers distributed over the screen in the mean value computation task in contrast to a single number presented in the center of the screen in the number discrimination task). Accordingly, more complex tasks pose higher demands on encoding than easier tasks. Importantly, by means of diffusion modeling, we get a purer measure of speed of information processing with the time needed for encoding and motoric components partialled out.

Limitations and Directions for Future Research

We want to make clear that we do not claim that mental speed is causally related to intelligence. In fact, a recent study based on an experimental approach did not find support for a causal link between mental speed (as measured by the drift rate of the diffusion model) and intelligence (Schubert, Hagemann, Frischkorn, & Herpertz, 2018). Rather, the authors suggest that structural properties of the brain may give rise to the association between mental speed and intelligence. The aim of our project was not to make any inferences regarding the question of causality.

Diffusion modeling allows for an examination of interesting research questions surrounding the *g* factor and other intelligence-related phenomena. One of these questions, which we addressed in our study, is the examination of whether there are domain-specific mental speeds. However, there are certainly further interesting research questions that could be examined by means of diffusion modeling in the future, for example the factor differentiation finding (e.g., Detterman & Daniel, 1989), which is regarded as one main feature of *g* (Kovacs & Conway, 2016).

Apart from the examination of further intelligence-related phenomena, it would also be important to explore relationships between drift rate and external criteria (e.g., grades at school/university, or job performance). Presently, we have no data on the predictive validity of drift rates for success in life; however, we think that future studies investigating this issue are important. Because our analyses revealed that in particular drift rate in more complex RT tasks showed strong relationships with intelligence, future research might focus on these more complex tasks.

In future studies, one might also examine whether the results that we observed in our study are moderated by the number of trials used in the RT tasks. Several diffusion model studies found that drift rate grows over time (Dutilh et al., 2009; Lerche & Voss, 2017; Petrov, Van Horn, & Ratcliff, 2011). Possibly, the 100 trials per task used in our study still give room for learning effects and relationships with intelligence might be even stronger or possibly smaller if higher trial numbers were employed, so that more trials could be discarded as practice trials.[10] A higher trial number would also increase reliability of estimates for drift (Lerche & Voss, 2017; Lerche et al., 2017). Further, in future studies we advise to use higher numbers of participants. The sample size of our study was relatively small for the application of structural equation modeling, leading to the use of very parsimonious parallel measurement models to ensure model convergence.

One aspect that is common to both the assessment of intelligence with the BIS and our computerized RT tasks (both "fast" and "slow" tasks) is the focus on speed. Chuderski (2013) showed that this focus on speed can have an important impact. He found that working memory capacity and fluid intelligence are isomorphic constructs when both are measured under time pressure. If, on the other hand, fluid intelligence is measured with no real time pressure, the relationship with working memory capacity decreases. The findings from the study by Chuderski (2013) suggest

[9] One notable exception is the explanatory model for performance in the Raven matrices by Carpenter, Just, and Shell (1990), in which different processes (incremental encoding, rule induction, goal management) were identified that contributed to the solution of the matrices. However, its application remains limited and its focus on Raven matrices forbids the generalization to other types of intelligence tests.

[10] Notably, our additional analyses in which we estimated parameters after exclusion of a larger number of practice trials did not result in a different pattern of results.

that relationships between drift rate in speeded RT tasks and intelligence measured under unspeeded conditions will probably be lower than the relationships we observed in our study which focused on speed. However, the difference in relationships between drift rate and speeded versus unspeeded intelligence tests would possibly be smaller than the differences between working memory capacity and speeded versus unspeeded fluid intelligence as measured by Chuderski (2013), because the isomorphic relation between working memory and fluid intelligence both assessed under speeded conditions might be partly attributable to nondecision time (e.g., speed of encoding). If the diffusion model is used, such influences can be "partialled out" so that we expect more similar relationships between speeded versus unspeeded intelligence testing and our performance measure (drift rate). It would be interesting to examine the size of the relationship between drift rate and unspeeded versus speeded intelligence testing in future research and compare it to the effect sizes found by Chuderski (2013).

One final aspect that we want to point out is that our findings do not lend support to an application of the diffusion model to all kinds of more complex, slower RT tasks. In tasks that require significantly more time than the approximately three seconds observed in our study, it becomes more likely that central assumptions of the diffusion model are seriously violated. In future studies it would be interesting to analyze tasks with substantially longer RTs (e.g., a matrices task with a mean RT of more than a minute; Partchev & De Boeck, 2012). Probably more important than the mean RT of a task are characteristics of the specific task. Even fast tasks can be poor candidates for diffusion modeling (e.g., because no continuous information uptake takes place). At the same time, even highly complex tasks that consist of many subtasks might be compatible with the diffusion model. In our study, the diffusion model provided a good fit for all employed tasks, and the relationships with intelligence speak in favor of the validity of the parameter drift rate. These tasks are interesting candidates for future diffusion model studies. If, however, researchers are interested in applying the diffusion model to any new tasks, these tasks (whether fast or slow) need to be carefully tested in terms of model fit and—even better—additionally with validation studies.

Conclusions

Prior research revealed relationships between general intelligence and the drift parameter of the diffusion model. This pattern proved to be robust in our structural equation modeling of a set of 18 binary RT tasks. Additionally, we expanded this research showing that there are content-domain specific (verbal, numeric, figural) aspects of cognitive speed, which are related to the respective components of a standard intelligence test. Moreover, slower, more complex tasks also proved to be closely linked to intelligence. Finally, we supply several more complex binary RT tasks that were fit well by the diffusion model and could thus be employed in future research projects.

Context of the Research

This research project is a cooperation of researchers from the departments of Quantitative Research Methods (VL, MVK, and AV) and Personality Research (GTF, ALS, and DH) of the Psychological Institute of Ruprecht-Karls-Universität Heidelberg. In this project, we could nicely combine the main expertise of the two labs, that is, diffusion modeling and intelligence research. In the preceding years, VL and AV have been contacted repeatedly by researchers who asked whether they could use the diffusion model also for more complex RT tasks. VL and AV conducted studies that provide first support for an extension to more complex tasks. Thereby arose the idea for a larger project, which includes numerous both fast and more complex RT tasks. GTF, ALS, and DH were always wondering whether there are domain-specific speeds of information processing but—because they usually additionally collect EEG data—they so far had refrained from running a study with such a large number of different RT tasks ($N = 18$). MVK is a PhD student who joined the team at the beginning of the recruitment for the study and has taken over an important role in the running of the study and the data analyses. He is currently examining the data further, focusing on age effects. One future research project will be the examination of relationships between drift rate in more complex tasks and external measures of performance (e.g., job performance).

References

Arnold, N. R., Bröder, A., & Bayen, U. J. (2015). Empirical validation of the diffusion model for recognition memory and a comparison of parameter-estimation methods. *Psychological Research, 79,* 882–898. http://dx.doi.org/10.1007/s00426-014-0608-y

Aschenbrenner, A. J., Balota, D. A., Gordon, B. A., Ratcliff, R., & Morris, J. C. (2016). A diffusion model analysis of episodic recognition in preclinical individuals with a family history for Alzheimer's disease: The adult children study. *Neuropsychology, 30,* 225–238. http://dx.doi.org/10.1037/neu0000222

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133,* 83–100. http://dx.doi.org/10.1037/0096-3445.133.1.83

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review, 71,* 117–120. http://dx.doi.org/10.1016/j.euroecorev.2014.07.003

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178. http://dx.doi.org/10.1016/j.cogpsych.2007.12.002

Camos, V., & Barrouillet, P. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: The executive and phonological loops. *Frontiers in Human Neuroscience, 8,* 900. http://dx.doi.org/10.3389/fnhum.2014.00900

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. http://dx.doi.org/10.1037/0033-295X.97.3.404

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511571312

Cattell, R. B., & Cattell, A. (1960). *Measuring intelligence with the culture fair tests*. Champaign, IL: Institute for Personality and Ability Testing.

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41,* 244–262. http://dx.doi.org/10.1016/j.intell.2013.04.003

Conway, A. R. A., & Kovacs, K. (2015). New and emerging models of human intelligence. *Wiley Interdisciplinary Reviews: Cognitive Science, 6,* 419–426. http://dx.doi.org/10.1002/wcs.1356

Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. *Intelligence, 31,* 567–587. http://dx.doi.org/10.1016/S0160-2896(03)00054-0

Coyle, T. (2017). A differential-developmental model (DDM): Mental speed, attention lapses, and general intelligence (g). *Journal of Intelligence, 5,* 25. http://dx.doi.org/10.3390/jintelligence5020025

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105,* 317–327. http://dx.doi.org/10.1037/0033-2909.105.2.317

Danner, D., Hagemann, D., Schankin, A., Hager, M., & Funke, J. (2011). Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence, 39*(5), 323–334. http://dx.doi.org/10.1016/j.intell.2011.06.004

Deary, I. J., Lawn, M., & Bartholomew, D. J. (2008). A conversation between Charles Spearman, Godfrey Thomson, and Edward L Thorndike: The International Examinations Inquiry Meetings 1931–1938. *History of Psychology, 11,* 122–142. http://dx.doi.org/10.1037/1093-4510.11.2.122

Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence, 13,* 349–359. http://dx.doi.org/10.1016/S0160-2896(89)80007-8

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review, 16,* 1026–1036. http://dx.doi.org/10.3758/16.6.1026

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods, 8,* 38–60. http://dx.doi.org/10.1037/1082-989X.8.1.38

Frischkorn, G. T., & Schubert, A.-L. (2018). Cognitive models in intelligence research: Advantages and recommendations for their application. *Journal of Intelligence, 6,* 34. http://dx.doi.org/10.3390/jintelligence6030034

Frischkorn, G. T., Schubert, A.-L., Neubauer, A. B., & Hagemann, D. (2016). The worst performance rule as moderation: New methods for worst performance analysis. *Journal of Intelligence, 4,* 9. http://dx.doi.org/10.3390/jintelligence4030009

Germar, M., Schlemmer, A., Krug, K., Voss, A., & Mojzisch, A. (2014). Social influence and perceptual decision making: A diffusion model analysis. *Personality and Social Psychology Bulletin, 40,* 217–231. http://dx.doi.org/10.1177/0146167213508985

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence, 55,* 57–68. http://dx.doi.org/10.1016/j.intell.2016.01.006

Hawkins, G. E., Hayes, B. K., & Heit, E. (2016). A dynamic model of reasoning and memory. *Journal of Experimental Psychology: General, 145,* 155–180. http://dx.doi.org/10.1037/xge0000113

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Jäger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen: Experimentell kontrollierte Weiterenwicklung eines deskriptiven Intelligenzstrukturmodells [Multimodal classification of intelligence achievement: Experimentally controlled, further development of a descriptive intelligence structure model]. *Diagnostica, 28,* 195–225.

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

Jensen, A. R. (1987). Process differences and individual differences in some cognitive tasks. *Intelligence, 11,* 107–136. http://dx.doi.org/10.1016/0160-2896(87)90001-8

Jensen, A. R. (2005). Mental chronometry and the unification of differential psychology. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence: Identifying the mechanisms of the mind* (pp. 26–50). New York, NY: Cambridge University Press.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam, the Netherlands: Elsevier.

Kan, K.-J., van der Maas, H. L. J., & Kievit, R. A. (2016). Process overlap theory: Strengths, limitations, and challenges. *Psychological Inquiry, 27,* 220–228. http://dx.doi.org/10.1080/1047840X.2016.1182000

Kan, K.-J., van der Maas, H. L. J., & Levine, S. Z. (2019). Extending psychometric network analysis: Empirical evidence against g in favor of mutualism? *Intelligence, 73,* 52–62. http://dx.doi.org/10.1016/j.intell.2018.12.004

Kim, K. H. (2005). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling, 12,* 368–390. http://dx.doi.org/10.1207/s15328007sem1203_2

Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry, 27,* 151–177. http://dx.doi.org/10.1080/1047840X.2016.1153946

Larsen, L., Hartmann, P., & Nyborg, H. (2008). The stability of general intelligence from early adulthood to middle-age. *Intelligence, 36,* 29–34. http://dx.doi.org/10.1016/j.intell.2007.01.001

Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence, 14,* 309–325. http://dx.doi.org/10.1016/0160-2896(90)90021-K

Lerche, V., Christmann, U., & Voss, A. (2018). Impact of context information on metaphor elaboration: A diffusion model study. *Experimental Psychology, 65,* 370–384. http://dx.doi.org/10.1027/1618-3169/a000422

Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology*. Advance online publication. http://dx.doi.org/10.3389/fpsyg.2016.01324

Lerche, V., & Voss, A. (2017). Retest reliability of the parameters of the Ratcliff diffusion model. *Psychological Research, 81,* 629–652. http://dx.doi.org/10.1007/s00426-016-0770-5

Lerche, V., & Voss, A. (2019). Experimental validation of the Diffusion Model based on a Slow Response Time Paradigm. *Psychological Research, 83,* 1194–1209. http://dx.doi.org/10.1007/s00426-017-0945-8

Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods, 49,* 513–537. http://dx.doi.org/10.3758/s13428-016-0740-2

Levine, G., Preddy, D., & Thorndike, R. L. (1987). Speed of information processing and level of cognitive ability. *Personality and Individual Differences, 8,* 599–607. http://dx.doi.org/10.1016/0191-8869(87)90057-2

Lindley, R. H., Wilson, S. M., Smith, W. R., & Bathurst, K. (1995). Reaction time (RT) and IQ: Shape of the task complexity function. *Personality and Individual Differences, 18,* 339–345. http://dx.doi.org/10.1016/0191-8869(94)00154-K

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

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the Radex and hierarchical models of intelligence. *Intelligence, 7,* 107–127. http://dx.doi.org/10.1016/0160-2896(83)90023-5

McKoon, G., & Ratcliff, R. (2012). Aging and IQ effects on associative recognition and priming in item recognition. *Journal of Memory and Language, 66,* 416–437. http://dx.doi.org/10.1016/j.jml.2011.12.001

Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research?: A comparison of fit in correlated factors, higher-order, and bi-factor models via Monte Carlo simulations. *Journal of Intelligence, 3,* 2–20. http://dx.doi.org/10.3390/jintelligence3010002

Neubauer, A. C., & Bucik, V. (1996). The mental speed—IQ relationship: Unitary or modular? *Intelligence, 22,* 23–48. http://dx.doi.org/10.1016/S0160-2896(96)90019-7

Partchev, I., & De Boeck, P. (2012). Can fast and slow intelligence be differentiated? *Intelligence, 40,* 23–32. http://dx.doi.org/10.1016/j.intell.2011.11.002

Petrov, A. A., Van Horn, N. M., & Ratcliff, R. (2011). Dissociable perceptual-learning mechanisms revealed by diffusion-model analysis. *Psychonomic Bulletin & Review, 18,* 490–497. http://dx.doi.org/10.3758/s13423-011-0079-8

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. http://dx.doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R. (2014). Measuring psychometric functions with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 870–888. http://dx.doi.org/10.1037/a0034954

Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision, 2015,* 237–279. http://dx.doi.org/10.1037/dec0000030

Ratcliff, R., & Frank, M. J. (2012). Reinforcement-based decision making in corticostriatal circuits: Mutual constraints by neurocomputational and diffusion models. *Neural Computation, 24,* 1186–1229. http://dx.doi.org/10.1162/NECO_a_00270

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182. http://dx.doi.org/10.1037/0033-295X.111.1.159

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922. http://dx.doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9,* 347–356. http://dx.doi.org/10.1111/1467-9280.00067

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20,* 260–281. http://dx.doi.org/10.1016/j.tics.2016.01.007

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19,* 278–289. http://dx.doi.org/10.1037/0882-7974.19.2.278

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology, 60,* 127–157. http://dx.doi.org/10.1016/j.cogpsych.2009.09.001

Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General, 140,* 464–487. http://dx.doi.org/10.1037/a0023810

Ratcliff, R., Thompson, C. A., & McKoon, G. (2015). Modeling individual differences in response time and accuracy in numeracy. *Cognition, 137,* 115–136. http://dx.doi.org/10.1016/j.cognition.2014.12.004

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9,* 438–481. http://dx.doi.org/10.3758/BF03196302

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48,* 36. http://dx.doi.org/10.18637/jss.v048.i02

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428. http://dx.doi.org/10.1037/0033-295X.103.3.403

Scheithe, K., & Bäuml, K.-H. (1995). Deutschsprachige Normen für Vertreter von 48 Kategorien [German-language norms for representatives of 48 conceptual categories]. *Sprache & Kognition, 14,* 39–43.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136,* 414–429. http://dx.doi.org/10.1037/0096-3445.136.3.414

Schmitz, F., & Voss, A. (2012). Decomposing task-switching costs with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 222–250. http://dx.doi.org/10.1037/a0026003

Schmitz, F., & Wilhelm, O. (2016). Modeling mental speed: Decomposing response time distributions in elementary cognitive tasks and correlations with working memory capacity and fluid intelligence. *Journal of Intelligence, 4,* 13. http://dx.doi.org/10.3390/jintelligence4040013

Schubert, A.-L. (2019). A meta-analysis of the worst performance rule. *Intelligence, 73,* 88–100. http://dx.doi.org/10.1016/j.intell.2019.02.003

Schubert, A.-L., Frischkorn, G. T., Hagemann, D., & Voss, A. (2016). Trait characteristics of diffusion model parameters. *Journal of Intelligence, 4,* 7. http://dx.doi.org/10.3390/jintelligence4030007

Schubert, A.-L., Hagemann, D., Frischkorn, G. T., & Herpertz, S. C. (2018). Faster, but not smarter: An experimental analysis of the relationship between mental speed and mental abilities. *Intelligence, 71,* 66–75. http://dx.doi.org/10.1016/j.intell.2018.10.005

Schubert, A.-L., Hagemann, D., Voss, A., Schankin, A., & Bergmann, K. (2015). Decomposing the relationship between mental speed and mental abilities. *Intelligence, 51,* 28–46. http://dx.doi.org/10.1016/j.intell.2015.05.002

Schubert, A.-L., & Rey-Mermet, A. (2019). Does process overlap theory replace the issues of general intelligence with the issues of attentional control? *Journal of Applied Research in Memory & Cognition, 8,* 277–283. http://dx.doi.org/10.1016/j.jarmac.2019.06.004

Schulz-Zhecheva, Y., Voelkle, M., Beauducel, A., Biscaldi, M., & Klein, C. (2016). Predicting fluid intelligence by components of reaction time distributions from simple choice reaction time tasks. *Journal of Intelligence, 4,* 8. http://dx.doi.org/10.3390/jintelligence4030008

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences, 44,* 535–551. http://dx.doi.org/10.1016/j.paid.2007.09.015

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 101–117. http://dx.doi.org/10.1037/0278-7393.32.1.101

Theisen, M., Lerche, V., von Krause, M., & Voss, A. (2019). *Age differences in diffusion model parameters: A meta-analysis*. Manuscript submitted for publication.

Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology, 8,* 271–281. http://dx.doi.org/10.1111/j.2044-8295.1916.tb00133.x

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison Wesley.

van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The EZ diffusion model provides a powerful test of simple empirical effects. *Psychonomic Bulletin & Review, 24,* 547–556. http://dx.doi.org/10.3758/s13423-016-1081-y

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental Psychology, 60,* 385–402. http://dx.doi.org/10.1027/1618-3169/a000218

Voss, A., Rothermund, K., Gast, A., & Wentura, D. (2013). Cognitive processes in associative and categorical priming: A diffusion model analysis. *Journal of Experimental Psychology: General, 142,* 536–559. http://dx.doi.org/10.1037/a0029459

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32,* 1206–1220. http://dx.doi.org/10.3758/BF03196893

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39,* 767–775. http://dx.doi.org/10.3758/BF03192967

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52,* 1–9. http://dx.doi.org/10.1016/j.jmp.2007.09.005

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology, 6,* 636. http://dx.doi.org/10.3389/fpsyg.2015.00336

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *The European Journal of Cognitive Psychology, 21,* 641–671. http://dx.doi.org/10.1080/09541440802205067

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58,* 140–159. http://dx.doi.org/10.1016/j.jml.2007.04.006

Weiss, R. H. (2006). *Grundintelligenztest Skala 2-Revidierte Fassung (CFT 20-R)* [General Intelligence Test Scale 2-Revised]. Goettingen, Germany: Hogrefe.

White, C. N., Curl, R. A., & Sloane, J. F. (2016). Using decision models to enhance investigations of individual differences in cognitive neuroscience. *Frontiers in Psychology, 7,* 81. http://dx.doi.org/10.3389/fpsyg.2016.00081

Wilhelm, O., Schroeders, U., & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8–10).* Göttingen, Germany: Hogrefe.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 53–79. http://dx.doi.org/10.1037/a0024177

Appendix A

Task Descriptives

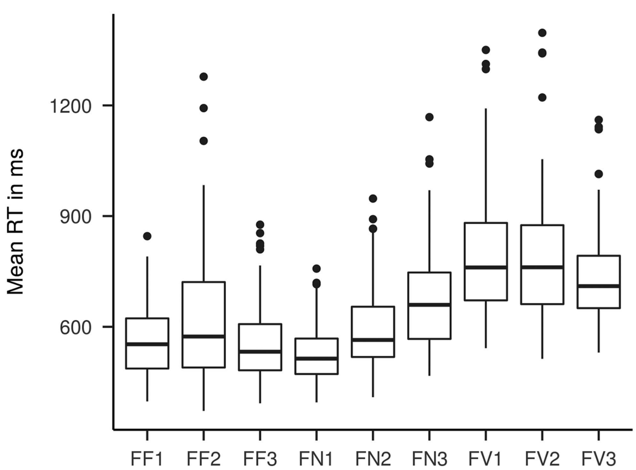

Figure A1. Boxplots of mean response times for all *fast* tasks. The first letter indicates the task complexity (F = fast); the second letter denotes the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. The boxplots display the first, second and third quartile. Outliers are values greater than 1.5 times the interquartile range from either end of the box.

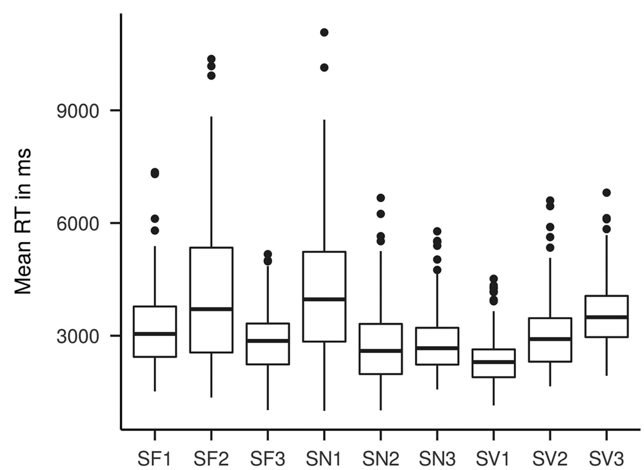

Figure A2. Boxplots of mean response times for all *slow* tasks. The first letter indicates the task complexity (S = slow); the second letter denotes the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. The boxplots display the first, second and third quartile. Outliers are values greater than 1.5 times the interquartile range from either end of the box.

(*Appendices continue*)

Appendix B

Diffusion Model Fit

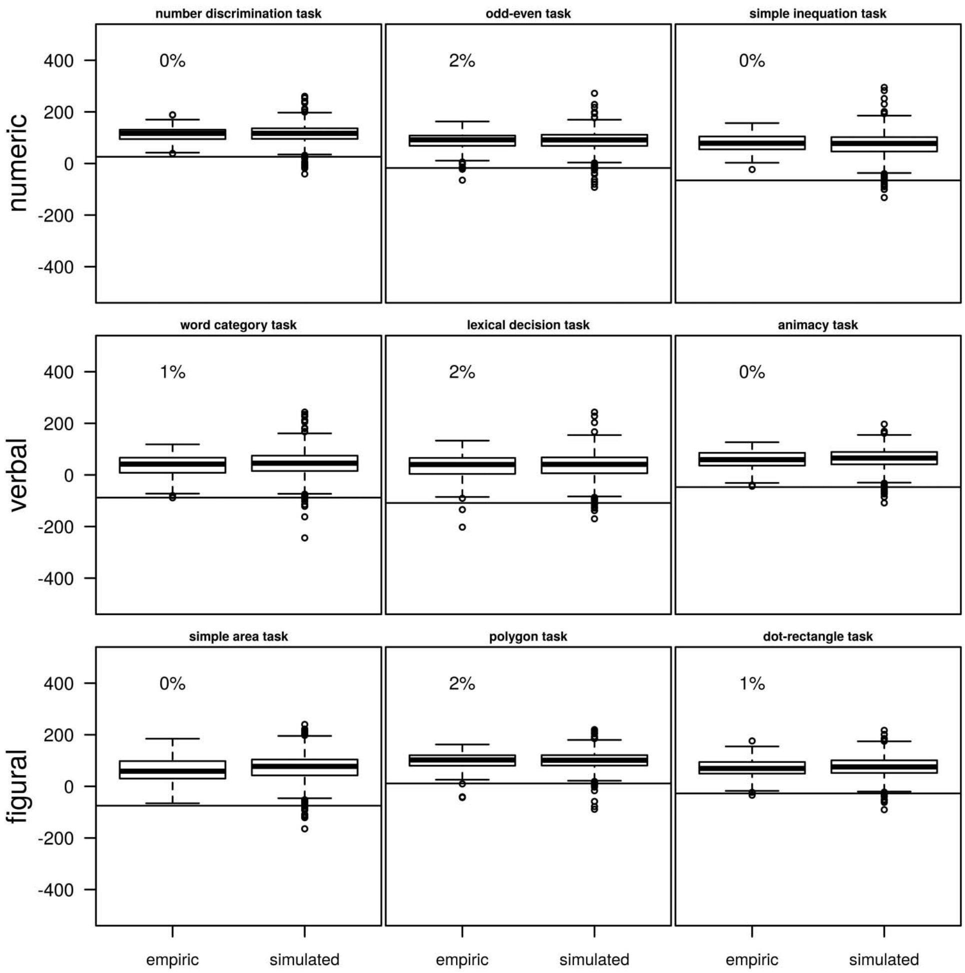

Figure B1. Model fit of all *fast* RT tasks. The boxplots show the maximum likelihood statistic (sum of logarithmized densities). Lower values indicate worse model fit. The horizontal line is the 1% percentile of fit values from 1,000 simulated data sets. For observed data, the percentage of fits that are worse than this critical value is also given.

(*Appendices continue*)

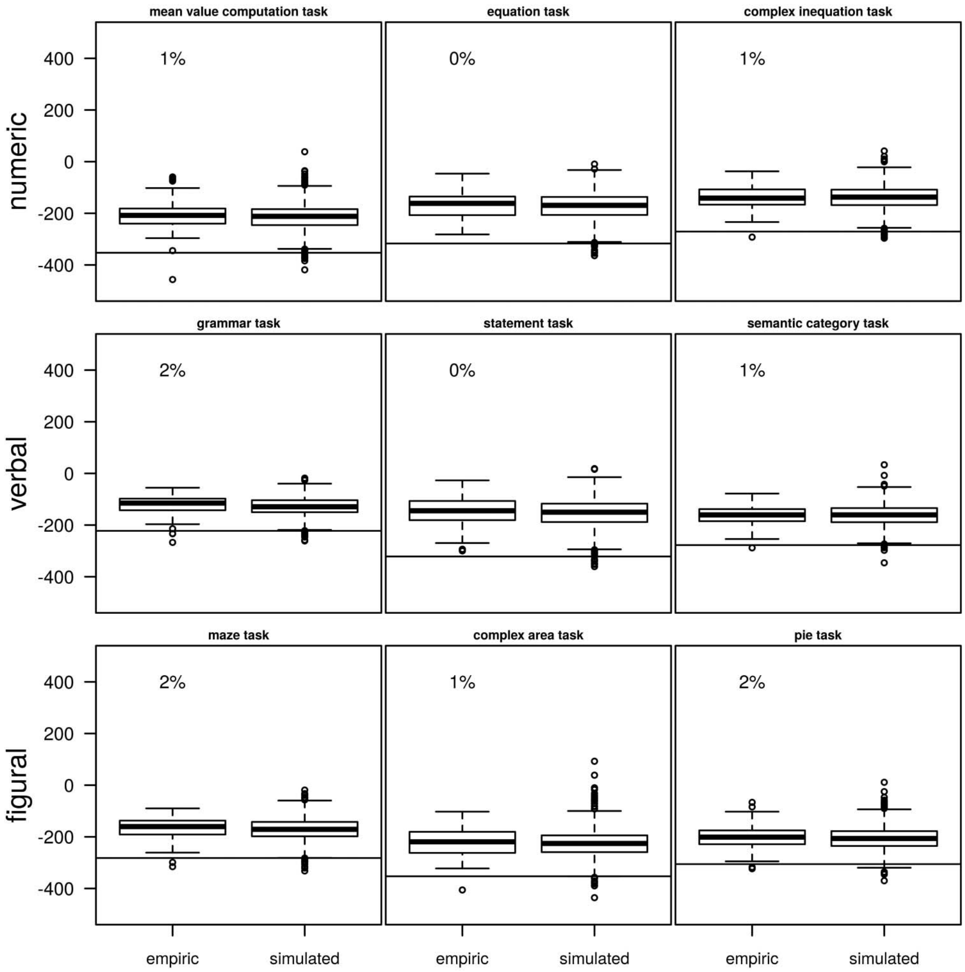

Figure B2. Model fit of all *slow* RT tasks. The boxplots show the maximum likelihood statistic (sum of logarithmized densities). Lower values indicate worse model fit. The horizontal line is the 1% percentile of fit values from 1,000 simulated data sets. For observed data, the percentage of fits that are worse than this critical value is also given.

(*Appendices continue*)

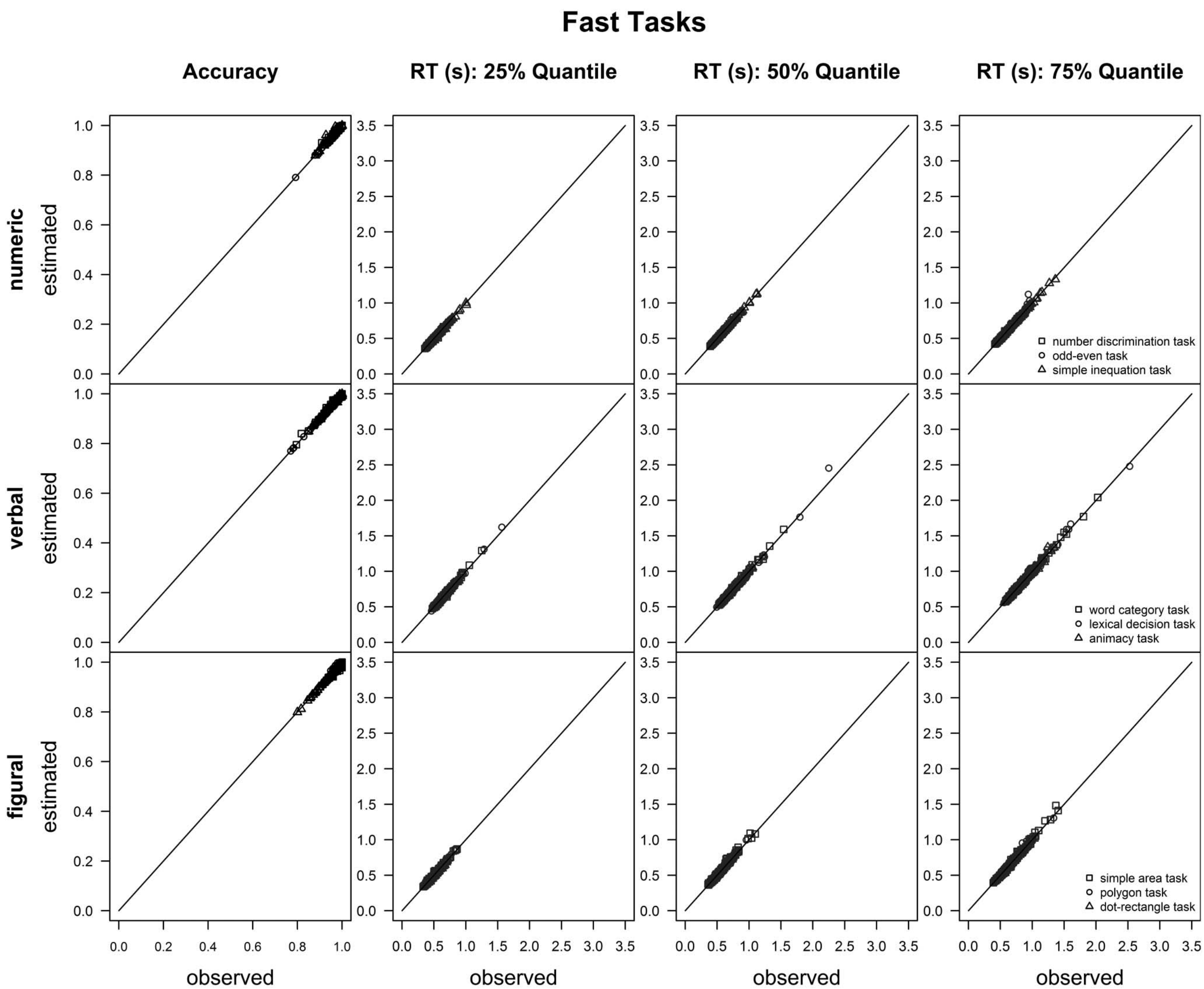

Figure B3. Model fit of the *fast* RT tasks based on the comparison of statistics (accuracy rate, first, second, and third RT quartile) of observed data and models' predictions. Each point represents one participant in one task. The diagonals indicate perfect model fit. One data point exceeding the scales of the third RT quartile plot was omitted.

(*Appendices continue*)

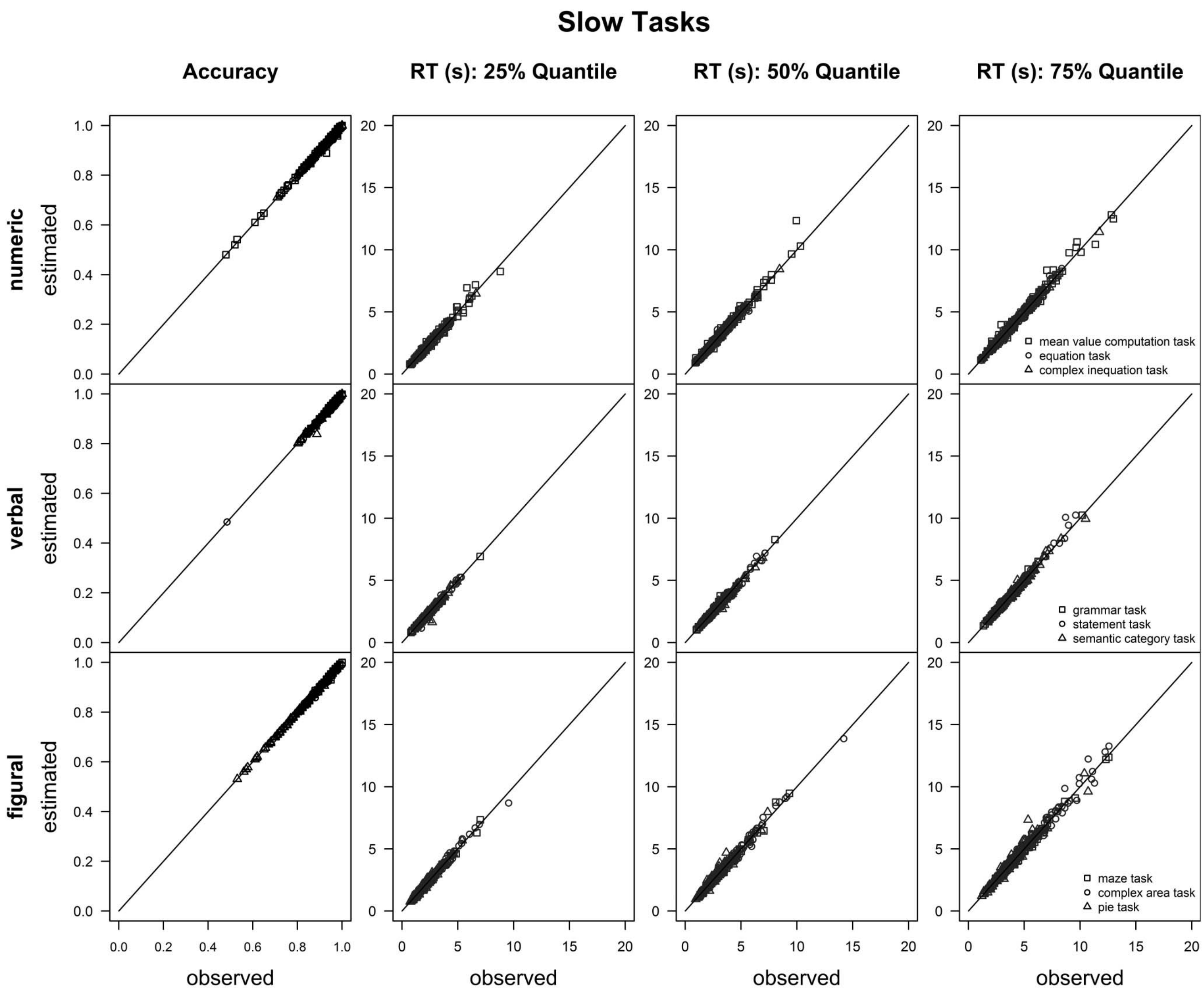

Figure B4. Model fit of the *slow* RT tasks based on the comparison of statistics (accuracy rate, first, second, and third RT quartile) of observed data and models' predictions. Each point represents one participant in one task. The diagonals indicate perfect model fit. Two data points exceeding the scales of the third RT quartile plot were omitted.

(*Appendices continue*)

Appendix C

Structural Equation Models

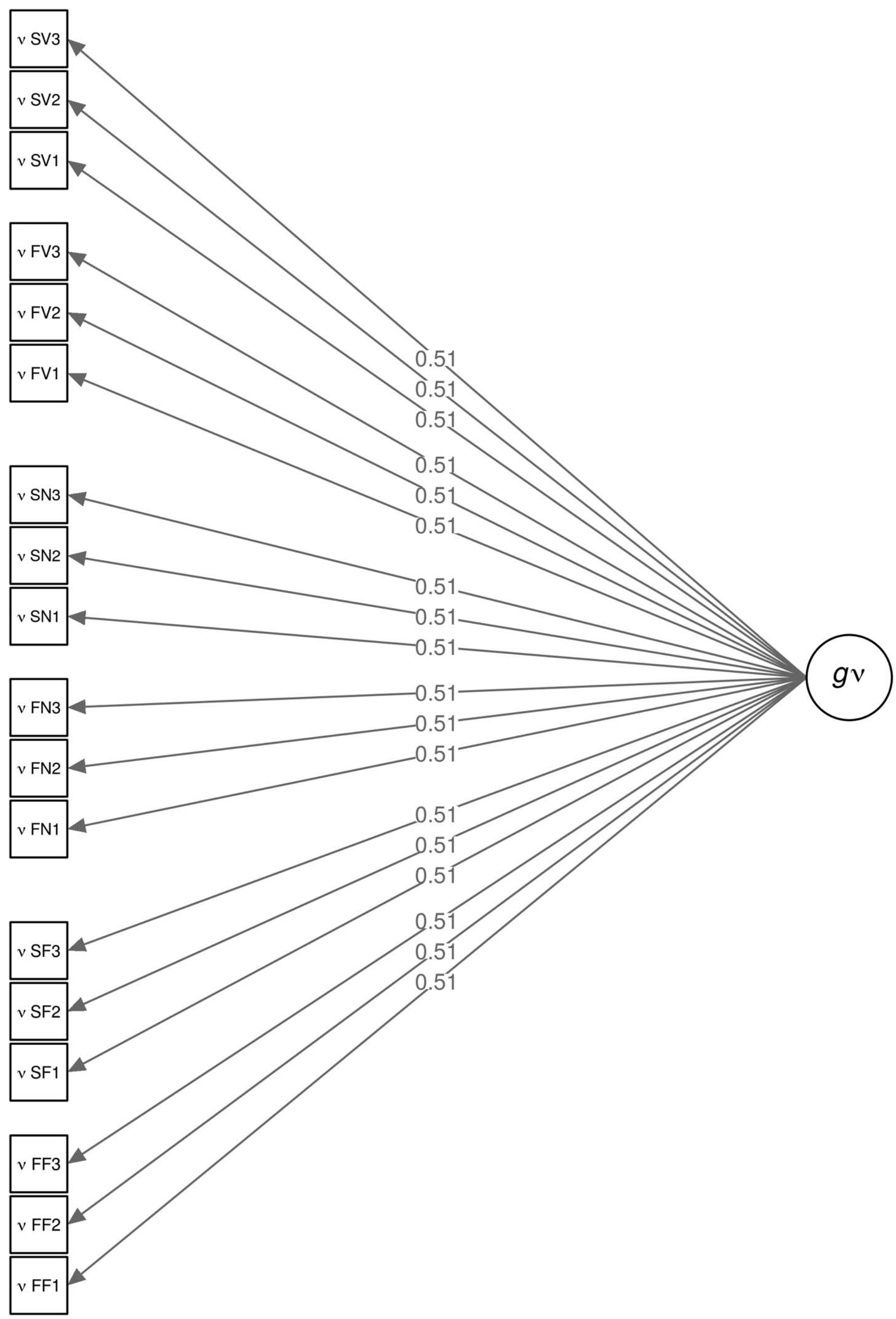

Figure C1. Drift Model 1. The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. Standardized loadings reported. Residuals are omitted from the plot for simplicity. $g\nu$ = general drift rate factor.

(*Appendices continue*)

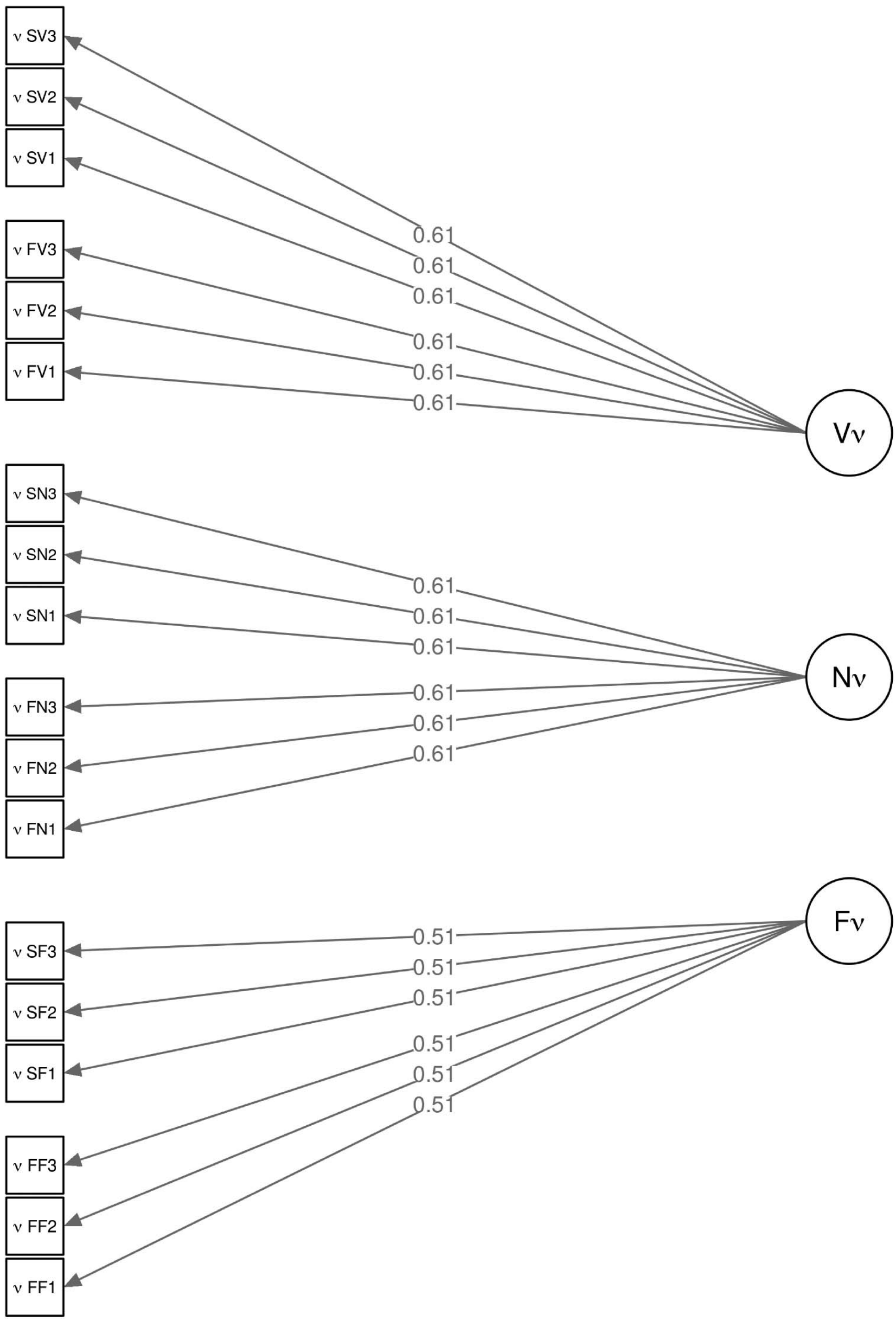

Figure C2. Drift Model 2. The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. Standardized loadings reported. Residuals are omitted from the plot for simplicity. Vν = verbal drift rate factor; Nν = numeric drift rate factor; Fν = figural drift rate factor.

(*Appendices continue*)

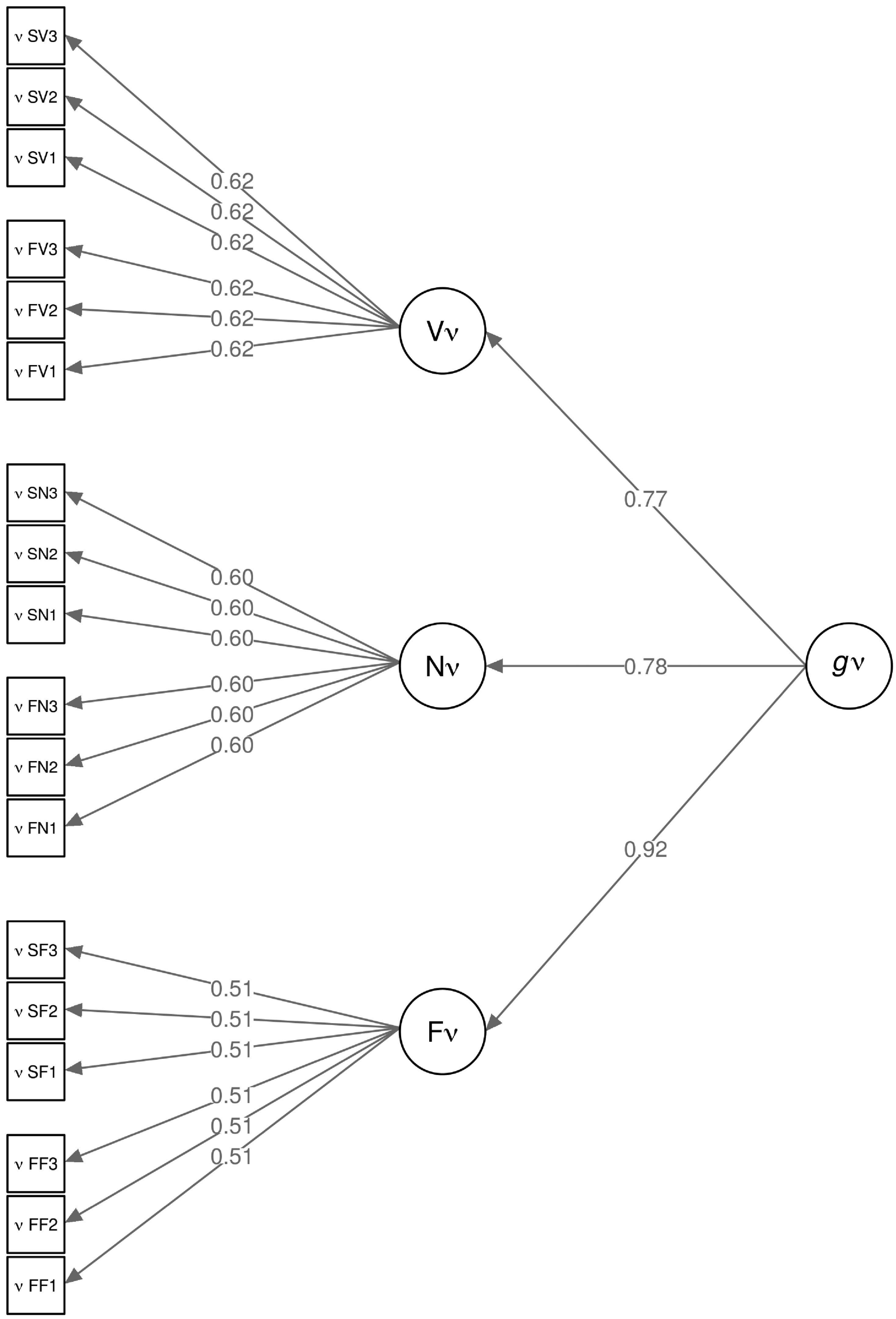

Figure C3. Drift Model 3. The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. Standardized loadings reported. Residuals are omitted from the plot for simplicity. gν = general drift rate factor; Vν = verbal drift rate factor; Nν = numeric drift rate factor; Fν = figural drift rate factor. As the loadings of the drift domain factors are standardized on the different freely estimated variances of the domain factors, their standardized values differ although the unstandardized loadings are all fixed to one.

(*Appendices continue*)

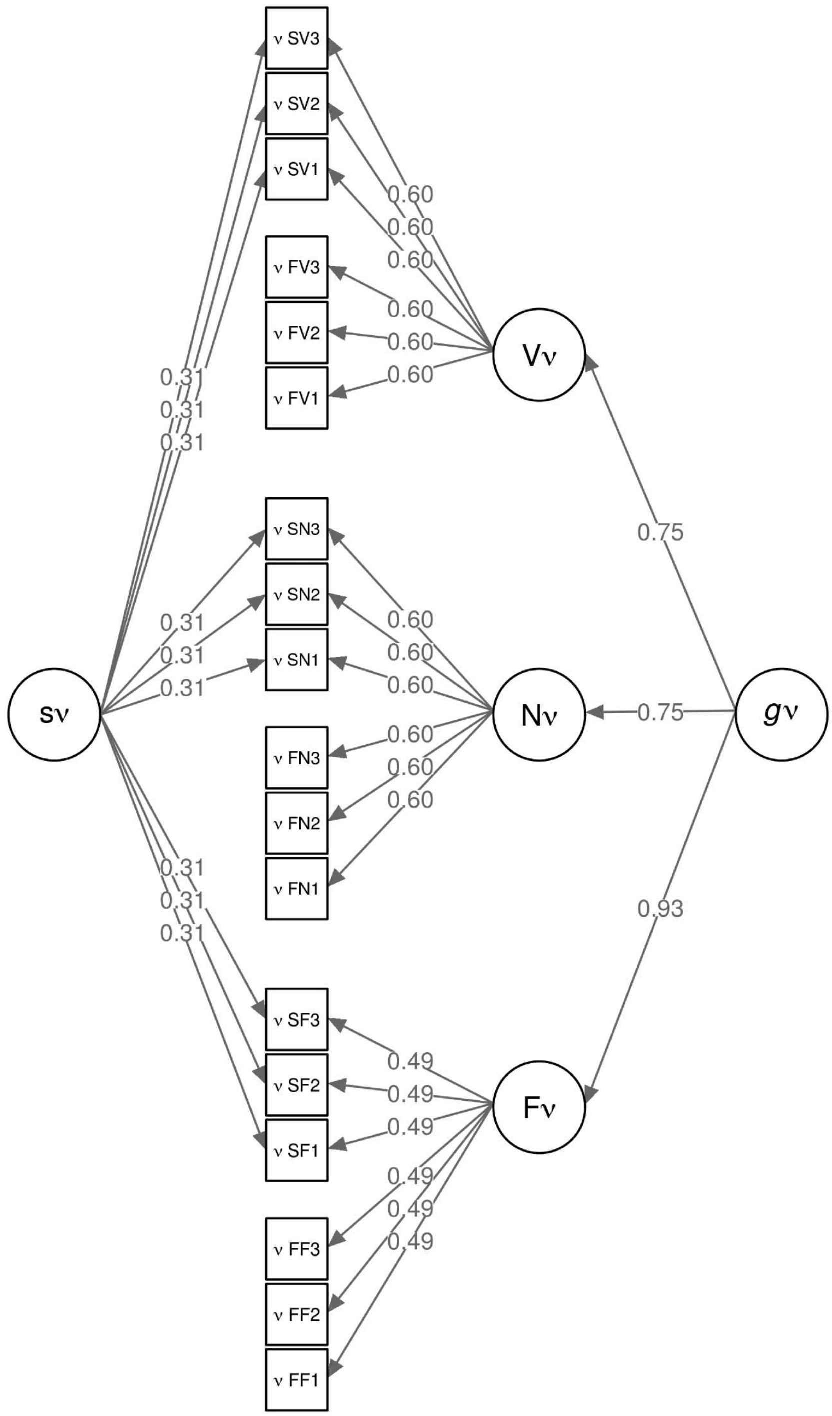

Figure C4. Drift Model 4. The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. Standardized loadings reported. Residuals are omitted from the plot for simplicity. gν = general drift rate factor; Vν = verbal drift rate factor; Nν = numeric drift rate factor; Fν = figural drift rate factor; sν = method factor for drift rate in slow tasks.

(*Appendices continue*)

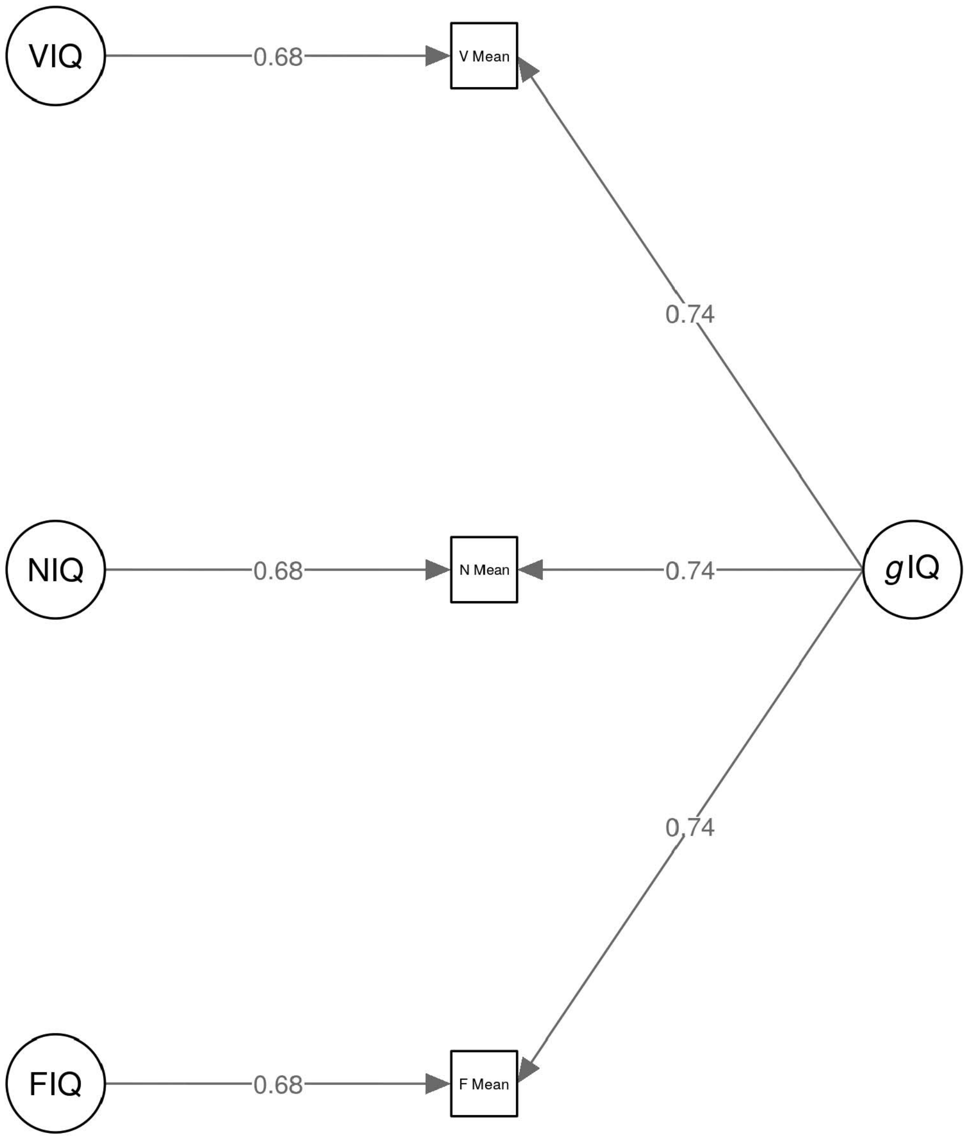

Figure C5. Intelligence Model. Scale means are used as indicators for verbal (VIQ), numeric (NIQ), and figural (FIQ) intelligence. gIQ = general intelligence. Completely standardized loadings are reported. Indicator residuals are fixed to zero, domain factors serve as quasi-residuals, see Method.

(*Appendices continue*)

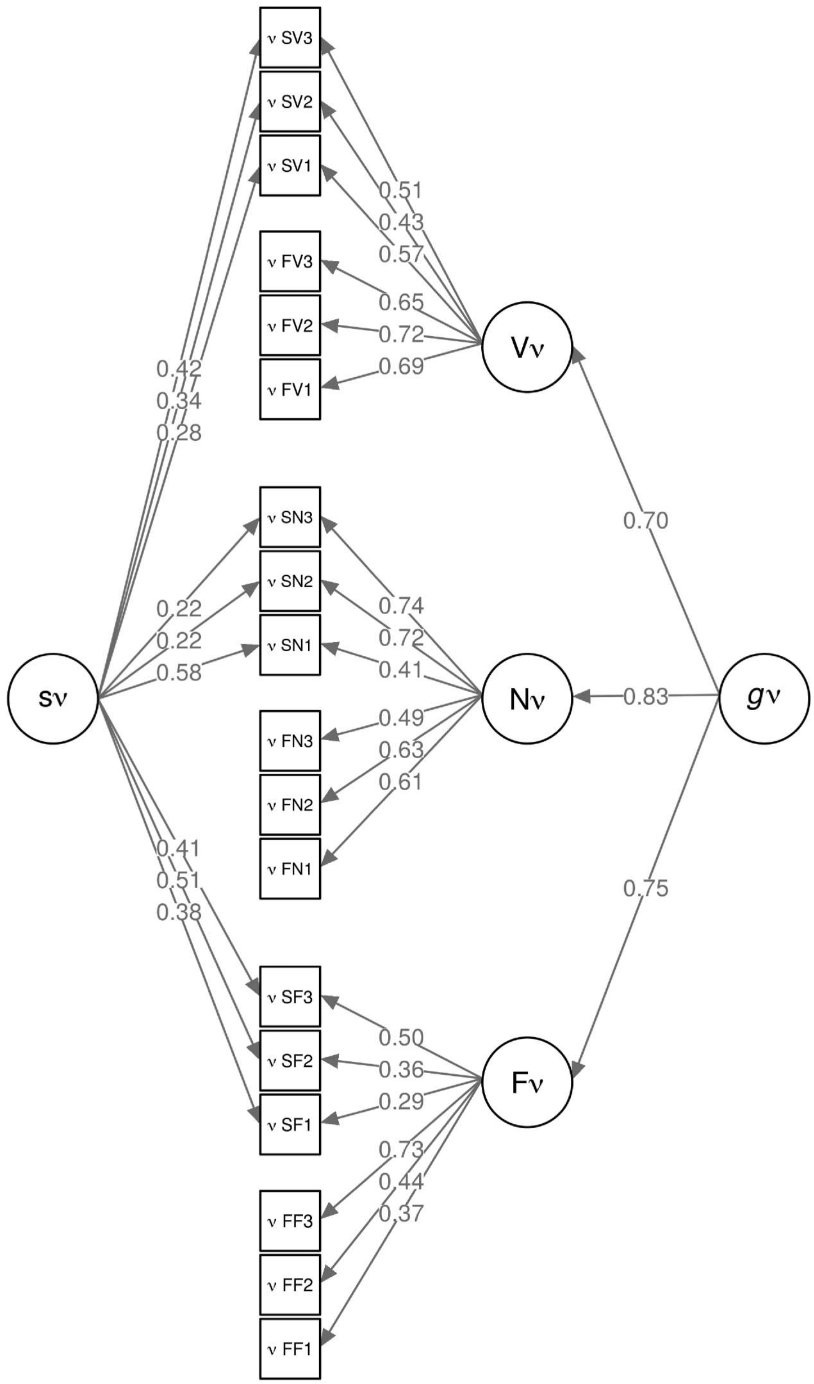

Figure C6. Drift Model 4 (freely estimated). The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. Standardized loadings reported. Residuals are omitted from the plot for simplicity. gν = general drift rate factor; Vν = verbal drift rate factor; Nν = numeric drift rate factor; Fν = figural drift rate factor; sν = method factor for drift rate in slow tasks.

(*Appendices continue*)

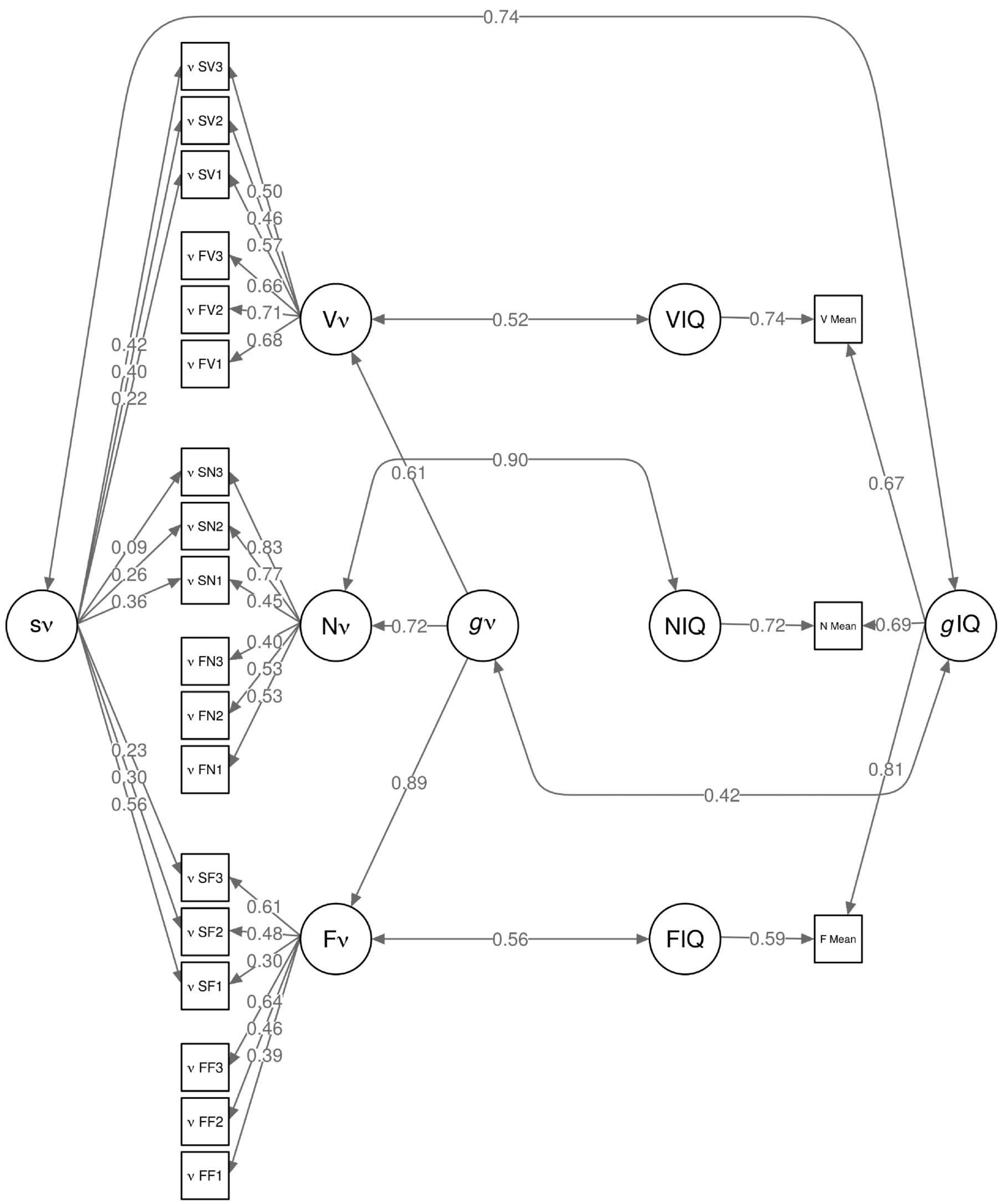

Figure C7. Combined Drift-Intelligence model (freely estimated). The first letter of the task indices denotes the type of task (F = fast, S = slow); the second letter indicates the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. Standardized loadings reported. Residuals are omitted from the plot for simplicity. The latent correlations between the drift domains and intelligence domains are between the drift domain residuals and the (quasi-residual) intelligence domain factors (see Method). gν = general drift rate factor; Vν = verbal drift rate factor; Nν = numeric drift rate factor; Fν = figural drift rate factor; sν = method factor for drift rate in slow tasks. Scale means are used as single indicators for verbal (VIQ), numeric (NIQ), and figural (FIQ) intelligence. gIQ = general intelligence.

(*Appendices continue*)

Appendix D

Descriptives of RT (in ms)

Task	*M*	*SD*	Minimum	Maximum
FF1	560	96	398	846
FF2	620	176	372	1,278
FF3	551	96	393	877
FN1	527	78	395	758
FN2	590	107	409	947
FN3	670	135	467	1,168
FV1	792	164	542	1,350
FV2	781	162	513	1,397
FV3	737	124	530	1,161
SF1	3,234	1,091	1,517	7,354
SF2	4,189	2,009	1,355	10,366
SF3	2,856	906	1,021	5,171
SN1	4,168	1,904	1,004	11,074
SN2	2,761	1,098	1,014	6,670
SN3	2,805	885	1,571	5,780
SV1	2,380	709	1,145	4,516
SV2	3,030	1,002	1,654	6,599
SV3	3,600	895	1,935	6,808

Note. The first letter indicates the task complexity (F = fast, S = slow); the second letter denotes the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. *SD* = standard deviation.

Appendix E

Descriptives of Accuracy Rate (in %)

Task	*M*	*SD*	Minimum	Maximum
FF1	93.65	2.88	84.54	97.00
FF2	98.68	1.60	93.00	100.00
FF3	97.71	1.90	91.58	100.00
FN1	98.03	2.26	89.00	100.00
FN2	97.68	2.03	91.00	100.00
FN3	97.17	2.74	88.00	100.00
FV1	96.22	3.76	79.55	100.00
FV2	95.11	3.97	78.35	100.00
FV3	97.18	2.41	87.00	100.00
SF1	95.53	2.91	87.00	100.00
SF2	86.69	6.50	69.00	100.00
SF3	80.47	9.10	53.06	97.00
SN1	90.76	8.11	61.00	100.00
SN2	91.16	5.48	72.00	98.00
SN3	93.51	3.71	82.00	100.00
SV1	96.36	2.39	88.00	100.00
SV2	95.11	2.61	85.86	99.00
SV3	94.24	4.77	80.21	100.00

Note. The first letter indicates the task complexity (F = fast, S = slow); the second letter denotes the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. *SD* = standard deviation.

(*Appendices continue*)

Appendix F

Descriptives of Drift Rate

Task	*M*	*SD*	Minimum	Maximum
FF1	3.16	0.73	1.79	6.42
FF2	3.26	1.02	1.43	7.16
FF3	4.27	0.96	2.38	8.01
FN1	4.97	1.82	2.41	16.50
FN2	3.95	0.97	2.12	8.52
FN3	3.97	1.39	2.00	12.23
FV1	2.81	0.88	1.37	6.25
FV2	2.68	0.78	1.12	4.83
FV3	3.21	0.89	1.54	6.61
SF1	0.94	0.20	0.52	1.61
SF2	0.58	0.17	0.17	0.97
SF3	0.50	0.18	0.09	1.02
SN1	0.70	0.22	0.15	1.30
SN2	0.80	0.25	0.39	1.48
SN3	1.08	0.33	0.57	2.15
SV1	1.17	0.20	0.64	1.79
SV2	1.03	0.29	0.54	1.99
SV3	0.90	0.23	0.39	1.63

Note. The first letter indicates the task complexity (F = fast, S = slow); the second letter denotes the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. *SD* = standard deviation.

Appendix G

Descriptives of Threshold Separation

Task	*M*	*SD*	Minimum	Maximum
FF1	0.91	0.21	0.46	1.71
FF2	1.53	0.53	0.66	3.61
FF3	1.16	0.61	0.63	5.52
FN1	1.47	1.31	0.44	10.00
FN2	1.20	0.51	0.62	3.90
FN3	1.36	1.03	0.50	10.00
FV1	1.52	0.73	0.53	5.76
FV2	1.33	0.44	0.55	2.62
FV3	1.35	0.55	0.66	5.61
SF1	3.75	1.44	1.73	10.00
SF2	3.71	1.37	1.45	8.05
SF3	3.06	0.81	1.36	5.10
SN1	4.00	1.53	1.21	10.00
SN2	3.25	0.92	1.13	6.35
SN3	2.85	0.92	1.52	6.79
SV1	3.08	0.84	1.71	7.07
SV2	3.19	0.87	1.35	5.14
SV3	3.69	1.23	1.75	10.00

Note. The first letter indicates the task complexity (F = fast, S = slow); the second letter denotes the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. *SD* = standard deviation.

(*Appendices continue*)

Appendix H

Descriptives of Nondecision Time (in ms)

Task	*M*	*SD*	Minimum	Maximum
FF1	423	65	273	587
FF2	359	66	242	592
FF3	411	56	236	555
FN1	388	67	135	539
FN2	427	57	313	678
FN3	499	96	192	789
FV1	513	76	226	850
FV2	527	74	367	749
FV3	520	65	333	732
SF1	1,286	495	137	2,969
SF2	1,480	918	63	5,874
SF3	913	397	230	2,657
SN1	1,628	1,207	0	5,794
SN2	844	309	36	2,097
SN3	1,501	422	628	2,983
SV1	1,092	348	366	2,525
SV2	1,448	420	910	3,746
SV3	1,635	413	68	3,280

Note. The first letter indicates the task complexity (F = fast, S = slow); the second letter denotes the domain (N = numeric, V = verbal, F = figural). See Table 1 for a brief description of all tasks. *SD* = standard deviation.

Appendix I

Descriptives of BIS Domain Scale Scores

Scale	*M*	*SD*	Minimum	Maximum
F_Mean	96.35	7.74	76.50	114.25
N_Mean	99.94	8.38	80.50	120.75
V_Mean	102.78	7.83	79.75	121.50

Note. V = verbal; N = numeric; F = figural; *SD* = standard deviation.

Appendix J

Drift Model 1 (*g* Factor)

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
			Loadings		
$g\nu$ on ν (each task)	1	0			0.509
			Latent (residual) variances		
$g\nu$	0.259	0.020	[0.219, 0.298]	<.001	1
			Residual indicator variances		
ν (each task)	0.741	0.020	[0.702, 0.781]	<.001	0.741

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

(*Appendices continue*)

Appendix K

Drift Model 2 (Uncorrelated Domains)

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
			Loadings		
Fν on ν (each figural task)	1	0			0.506
Nν on ν (each numeric task)	1	0			0.610
Vν on ν (each verbal task)	1	0			0.615
			Latent (residual) variances		
Fν	0.256	0.035	[0.188, 0.325]	<.001	1
Nν	0.371	0.033	[0.308, 0.435]	<.001	1
Vν	0.378	0.033	[0.314, 0.442]	<.001	1
			Residual indicator variances		
ν (each figural task)	0.744	0.035	[0.675, 0.812]	<.001	0.744
ν (each numeric task)	0.629	0.033	[0.565, 0.692]	<.001	0.629
ν (each verbal task)	0.622	0.033	[0.558, 0.686]	<.001	0.622

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

Appendix L

Drift Model 3 (Hierarchical Model of Domains and *g* Factor)

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
			Loadings		
Fν on ν (each figural task)	1	0			0.514
Nν on ν (each numeric task)	1	0			0.605
Vν on ν (each verbal task)	1	0			0.617
$g\nu$ on Fν	1	0			0.922
$g\nu$ on Nν	1	0			0.784
$g\nu$ on Vν	1	0			0.769
			Latent (residual) variances		
$g\nu$	0.225	0.024	[0.178, 0.271]	<.001	1
Fν	0.039	0.029	[−0.017, 0.096]	.171	0.149
Nν	0.141	0.033	[0.077, 0.206]	<.001	0.386
Vν	0.156	0.032	[0.092, 0.219]	<.001	0.409
			Residual indicator variances		
ν (each figural task)	0.736	0.032	[0.672, 0.800]	<.001	0.736
ν (each numeric task)	0.634	0.031	[0.573, 0.696]	<.001	0.634
ν (each verbal task)	0.620	0.031	[0.559, 0.680]	<.001	0.620

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

(*Appendices continue*)

Appendix M

Drift Model 4 (Hierarchical Model of Domains and *g* Factor and Slow Method Factor)

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
		Loadings			
sν on ν (each slow task)	1	0			0.308
Fν on ν (each figural task)	1	0			0.486
Nν on ν (each numeric task)	1	0			0.600
Vν on ν (each verbal task)	1	0			0.598
$g\nu$ on Fν	1	0			0.926
$g\nu$ on Nν	1	0			0.750
$g\nu$ on Vν	1	0			0.751
		Latent (residual) variances			
$g\nu$	0.202	0.025	[0.154, 0.251]	<.001	1
sν	0.095	0.022	[0.051, 0.138]	<.001	1
Fν	0.034	0.028	[−0.022, 0.089]	.235	0.142
Nν	0.158	0.033	[0.094, 0.222]	<.001	0.438
Vν	0.156	0.031	[0.095, 0.217]	<.001	0.435
		Residual indicator variances			
ν (each fast figural task)	0.764	0.034	[0.698, 0.830]	<.001	0.764
ν (each fast numeric task)	0.640	0.031	[0.579, 0.701]	<.001	0.640
ν (each fast verbal task)	0.642	0.032	[0.580, 0.704]	<.001	0.642
ν (each slow figural task)	0.670	0.034	[0.602, 0.737]	<.001	0.670
ν (each slow numeric task)	0.545	0.034	[0.479, 0.612]	<.001	0.545
ν (each slow verbal task)	0.547	0.032	[0.485, 0.610]	<.001	0.547

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

Appendix N

Intelligence Model

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
		Loadings			
*g*IQ on F_Mean/N_Mean/V_Mean	1	0			0.736
VIQ on V_Mean/NIQ on N_Mean/FIQ on F_Mean	1	0			0.677
		Latent (residual) variances			
*g*IQ	0.542	0.040	[0.465, 0.620]	<.001	1
FIQ/NIQ/VIQ	0.458	0.040	[0.380, 0.535]	<.001	1
V_Mean/N_Mean/F_Mean	0	0			

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

(*Appendices continue*)

Appendix O

Drift Model 4 (Hierarchical Model of Domains and *g* Factor and Slow Method Factor), Freely Estimated

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
			Loadings		
Fν on ν.FF1	1	0			0.365
on ν.FF2	1.213	0.685	[−0.128, 2.555]	.076	0.443
on ν.FF3	1.996	1.266	[−0.486, 4.477]	.115	0.729
on ν.SF1	0.793	0.624	[−0.430, 2.017]	.204	0.290
on ν.SF2	0.974	0.532	[−0.067, 2.016]	.067	0.356
on ν.SF3	1.364	0.802	[−0.207, 2.935]	.089	0.498
Nν on ν.FN1	1	0			0.610
on ν.FN2	1.035	0.144	[0.753, 1.318]	<.001	0.632
on ν.FN3	0.802	0.158	[0.492, 1.112]	<.001	0.489
on ν.SN1	0.673	0.188	[0.304, 1.042]	<.001	0.411
on ν.SN2	1.172	0.203	[0.774, 1.570]	<.001	0.715
on ν.SN3	1.206	0.217	[0.780, 1.632]	<.001	0.736
Vν on ν.FV1	1	0			0.690
on ν.FV2	1.045	0.126	[0.799, 1.291]	<.001	0.721
on ν.FV3	0.942	0.135	[0.678, 1.207]	<.001	0.650
on ν.SV1	0.828	0.123	[0.586, 1.070]	<.001	0.571
on ν.SV2	0.628	0.130	[0.372, 0.883]	<.001	0.433
on ν.SV3	0.741	0.136	[0.474, 1.008]	<.001	0.511
sν on ν.SF1	1	0			0.378
on ν.SF2	1.339	1.182	[−0.978, 3.656]	.257	0.507
on ν.SF3	1.080	0.997	[−0.875, 3.034]	.279	0.408
on ν.SN1	1.543	1.299	[−1.002, 4.088]	.235	0.584
on ν.SN2	0.587	0.673	[−0.733, 1.907]	.383	0.222
on ν.SN3	0.579	0.744	[−0.879, 2.038]	.436	0.219
on ν.SV1	0.749	0.501	[−0.233, 1.731]	.135	0.283
on ν.SV2	0.895	0.653	[−0.385, 2.175]	.170	0.339
on ν.SV3	1.099	0.654	[−0.182, 2.381]	.093	0.416
gν on Fν	1	0			0.748
*g*ν on Nν	1.860	1.370	[−0.825, 4.545]	.175	0.833
*g*ν on Vν	1.768	1.188	[−0.560, 4.096]	.137	0.700
			Latent (residual) variances		
*g*ν	0.075	0.100	[−0.121, 0.270]	.455	1
sν	0.143	0.214	[−0.276, 0.562]	.503	1
Fν	0.059	0.050	[−0.038, 0.156]	.235	0.441
Nν	0.114	0.071	[−0.026, 0.254]	.110	0.307
Vν	0.243	0.082	[0.081, 0.404]	.003	0.510
			Residual indicator variances		
ν.FF1	0.867	0.142	[0.589, 1.144]	<.001	0.867
ν.FF2	0.804	0.085	[0.637, 0.970]	<.001	0.804
ν.FF3	0.469	0.170	[0.136, 0.802]	.006	0.469
ν.FN1	0.628	0.090	[0.451, 0.804]	<.001	0.628
ν.FN2	0.601	0.094	[0.418, 0.784]	<.001	0.601
ν.FN3	0.760	0.074	[0.615, 0.906]	<.001	0.760
ν.FV1	0.524	0.083	[0.361, 0.687]	<.001	0.524
ν.FV2	0.480	0.086	[0.312, 0.648]	<.001	0.480
ν.FV3	0.577	0.082	[0.416, 0.738]	<.001	0.577
ν.SF1	0.773	0.158	[0.463, 1.083]	<.001	0.773
ν.SF2	0.617	0.096	[0.428, 0.806]	<.001	0.617
ν.SF3	0.585	0.090	[0.408, 0.762]	<.001	0.585
ν.SN1	0.491	0.098	[0.298, 0.684]	<.001	0.491
ν.SN2	0.439	0.071	[0.300, 0.578]	<.001	0.439
ν.SN3	0.411	0.073	[0.268, 0.553]	<.001	0.411
ν.SV1	0.594	0.079	[0.440, 0.748]	<.001	0.594
ν.SV2	0.698	0.082	[0.538, 0.858]	<.001	0.698
ν.SV3	0.566	0.094	[0.381, 0.750]	<.001	0.566

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

(*Appendices continue*)

Appendix P

Combined Drift-Intelligence Model, Freely Estimated

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
			Loadings		
F*ν* on *ν*.FF1	1	0			0.392
on *ν*.FF2	1.180				0.463
on *ν*.FF3	1.630				0.639
on *ν*.SF1	0.758				0.297
on *ν*.SF2	1.215				0.477
on *ν*.SF3	1.554				0.610
N*ν* on *ν*.FN1	1	0			0.526
on *ν*.FN2	1.011	0.187	[0.645, 1.377]	<.001	0.532
on *ν*.FN3	0.756	0.181	[0.401, 1.112]	<.001	0.398
on *ν*.SN1	0.860	0.202	[0.464, 1.257]	<.001	0.453
on *ν*.SN2	1.472	0.261	[0.960, 1.985]	<.001	0.775
on *ν*.SN3	1.572	0.252	[1.078, 2.066]	<.001	0.827
V*ν* on *ν*.FV1	1	0			0.679
on *ν*.FV2	1.043	0.123	[0.803, 1.284]	<.001	0.709
on *ν*.FV3	0.970	0.131	[0.714, 1.226]	<.001	0.659
on *ν*.SV1	0.846	0.118	[0.615, 1.076]	<.001	0.575
on *ν*.SV2	0.679	0.117	[0.450, 0.907]	<.001	0.461
on *ν*.SV3	0.740	0.120	[0.505, 0.976]	<.001	0.503
s*ν* on *ν*.SF1	1	0			0.564
on *ν*.SF2	0.537	0.230	[0.087, 0.988]	.019	0.303
on *ν*.SF3	0.399	0.191	[0.025, 0.773]	.036	0.225
on *ν*.SN1	0.641	0.219	[0.212, 1.070]	.003	0.362
on *ν*.SN2	0.469	0.236	[0.008, 0.931]	.046	0.265
on *ν*.SN3	0.151	0.188	[−0.218, 0.520]	.421	0.085
on *ν*.SV1	0.392	0.168	[0.063, 0.721]	.020	0.221
on *ν*.SV2	0.717	0.214	[0.297, 1.137]	.001	0.404
on *ν*.SV3	0.747	0.182	[0.391, 1.104]	<.001	0.421
gν on F*ν*	1	0			0.885
gν on N*ν*	1.091				0.720
gν on V*ν*	1.191				0.608
*g*IQ on F_Mean	1	0			0.808
*g*IQ on N_Mean	0.858	0.033	[0.794, 0.923]	<.001	0.693
*g*IQ on V_Mean	0.833				0.673
FIQ on F_Mean	1	0			0.590
NIQ on N_Mean	1	0			0.721
VIQ on V_Mean	1	0			0.740
			Covariances		
gν with gIQ	0.117				0.418
s*ν* with gIQ	0.336	0.062	[0.214, 0.458]	<.001	0.739
F*ν* with FIQ	0.060	0.035	[−0.008, 0.128]	.082	0.561
N*ν* with NIQ	0.237	0.038	[0.162, 0.312]	<.001	0.899
V*ν* with VIQ	0.208	0.046	[0.119, 0.298]	<.001	0.522
			Latent (residual) variances		
gν	0.121				1
*g*IQ	0.652	0.038	[0.578, 0.727]	<.001	1
s*ν*	0.318	0.127	[0.068, 0.567]	.013	1
F*ν*	0.033	0.017	[0.000, 0.067]	.053	0.217
N*ν*	0.134	0.036	[0.000, 0.067]	<.001	0.482
V*ν*	0.291	0.080	[0.134, 0.448]	<.001	0.630
FIQ	0.348	0.038	[0.273, 0.422]	<.001	1
NIQ	0.519	0.059	[0.404, 0.634]	<.001	1
VIQ	0.548	0.052	[0.446, 0.649]	<.001	1

(*Appendices continue*)

Appendix P (*continued*)

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
			Residual indicator variances		
ν.FF1	0.846				0.846
ν.FF2	0.786	0.067	[0.655, 0.916]	<.001	0.786
ν.FF3	0.591	0.097	[0.402, 0.780]	<.001	0.591
ν.FN1	0.723	0.085	[0.557, 0.890]	<.001	0.723
ν.FN2	0.717	0.075	[0.571, 0.863]	<.001	0.717
ν.FN3	0.842	0.064	[0.716, 0.967]	<.001	0.842
ν.FV1	0.538	0.080	[0.382, 0.695]	<.001	0.538
ν.FV2	0.497	0.077	[0.346, 0.649]	<.001	0.497
ν.FV3	0.566	0.079	[0.412, 0.720]	<.001	0.566
ν.SF1	0.594	0.102	[0.393, 0.794]	<.001	0.594
ν.SF2	0.681	0.076	[0.531, 0.831]	<.001	0.681
ν.SF3	0.578	0.061	[0.458, 0.697]	<.001	0.578
ν.SN1	0.664	0.078	[0.512, 0.817]	<.001	0.664
ν.SN2	0.330	0.054	[0.225, 0.435]	<.001	0.330
ν.SN3	0.309	0.051	[0.209, 0.409]	<.001	0.309
ν.SV1	0.621	0.076	[0.471, 0.771]	<.001	0.621
ν.SV2	0.624	0.080	[0.466, 0.782]	<.001	0.624
ν.SV3	0.570	0.082	[0.409, 0.731]	<.001	0.570
F_Mean/N_Mean/V_Mean	0	0			

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized. Caveat: unreliable estimates with some missing standard errors.

Appendix Q

Nondecision Time Model 4 (Hierarchical Model of Domains and *g* Factor and Slow Method Factor)

Parameter	Estimate	*SE*	95% CI	*p*	Std. Est.
			Loadings		
Ft_0 on t_0 (each figural task)	1	0			0.539
Nt_0 on t_0 (each numeric task)	1	0			0.582
Vt_0 on t_0 (each verbal task)	1	0			0.613
st_0 on t_0 (each slow task)	1	0			0.273
gt_0 on Ft_0	1	0			1.020
gt_0 on Nt_0	1	0			0.944
gt_0 on Vt_0	1	0			0.897
			Latent (residual) variances		
gt_0	0.302	0.021	[0.261, 0.344]	<.001	1
st_0	0.075	0.019	[0.038, 0.112]	<.001	1
Ft_0	−0.012	0.021	[−0.054, 0.031]	.592	−0.040
Nt_0	0.037	0.023	[−0.009, 0.083]	.117	0.108
Vt_0	0.074	0.026	[0.023, 0.124]	.004	0.196
			Residual indicator variances		
t_0 (each fast figural task)	0.709	0.029	[0.652, 0.767]	<.001	0.709
t_0 (each fast numeric task)	0.661	0.029	[0.605, 0.717]	<.001	0.661
t_0 (each fast verbal task)	0.624	0.028	[0.568, 0.680]	<.001	0.624
t_0 (each slow figural task)	0.635	0.030	[0.575, 0.694]	<.001	0.635
t_0 (each slow numeric task)	0.587	0.030	[0.527, 0.646]	<.001	0.587
t_0 (each slow verbal task)	0.550	0.031	[0.488, 0.611]	<.001	0.550

Note. Missing *p* values indicate fixed parameters. The standardized solution is completely standardized.

Received December 10, 2018
Revision received January 9, 2020
Accepted March 8, 2020 ■