IRT Branching Models for Individual Differences in Dual-Processing Theory of Reading Comprehension

A DISSERTATION SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Bowen Liu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor
Mark L. Davison

August 2020

## Acknowledgments

The year 2019 had been challenging for my family, my friends, and my relatives who live in Wuhan, China. They went through a lot during the COVID-19 outbreak. I want to show my profound gratitude for those who helped my family get through the tough times and support the communities back to normal.

I would like to express my sincere gratitude to my advisor, Mark L. Davison for his patience, motivation and professional guidance. He has been supportive since the day I became a QME student. Mark supported me, not only by providing a research assistantship over these years, but also academically through the path to learn about IRT. He helped me come up with the thesis topic and guided me over the research. I would like to thank Professor Michael Rodriguez for his evaluation and measurement courses, which aroused my interest and brought me to this major. I would also like to thank Ernest Davenport for his thoughtful feedback and his courses on factor analysis and advanced regression. Thank you to Xiaoou Li for her input provided in the prospectus meeting.

Numerous people helped me during my study. I would like to thank all my friends and the staff in 250 Education Sciences Building who helped me during my graduate education. As an international student, I appreciate those who helped me during my OPT and H1B applications.

I am grateful to my parents and my brother who give me unconditional love and support throughout my life. I would like to thank my husband who encourages and inspires me during my graduate study.

## Abstract

Dual process theories of cognition, achievement, and social cognition can be modeled with three-dimensional branching models. In the reading assessment context, this research studied the question of whether it is possible to derive three dimensions (speed of reading, accuracy of a fast reading process, and accuracy of a slow reading process) that are both interpretable and empirically distinct. In the reading context, the fast process is called “automaticity,” a level of reading that may be necessary for the facile application of reading in further education and careers. Unless automaticity is attained, the effort of the reading process itself can interfere with comprehension and learning of reading content.

Answering the questions about the distinctness of dual process theory dimensions required addressing a missing data problem found in earlier research on dual process theories of intelligence and achievement. This study introduces the polytomous scoring method in dealing with missing data in branching models, specifically Partchev and De Boeck’s (2012) proposed branching model differentiating fast and slow intelligence. In the area of reading achievement, rather than intelligence, this study investigated the effect of missing values caused by a branching model, and it proposes a polytomous scoring method as an alternative way to deal with missing values. The current study used a computer-administered reading comprehension assessment, which recorded students’ responses and response time for each item. In this study, both dichotomous and polytomous coding are discussed in the methodology section. All dichotomous variables were fitted with unidimensional and multidimensional 2PL models. Polytomous variables were fitted with unidimensional and multidimensional graded response models The results

were compared with Partchev and De Boeck's (2012) research. The current study found the same result as did Partchev and De Boeck's (2012) research on model comparison in which three-dimensional models fit better than two dimensional models. However, the theta correlations of three dimensions were different from their research. Raw scores were computed and used for validity evidence to aid the interpretation of latent dimensions.

## Table of Contents

.

## List of Tables

## List of Figures

# Chapter 1: Introduction

## 1.1 Overview

In psychology, dual-process theories have been widely discussed in cognition and social cognition. Researchers mainly focus on automatic processing and cognitive ability (e.g., Bargh, 1990; Brewin & Beaton, 2000; Monsell & Driver, 2000). Prior studies suggest that there may be two distinct cognitive systems, automatic and controlled, underlying thinking and reasoning. The automatic process is fast and automatic in nature, whereas, the deliberate process involves slow reaction, strategy, and use of central working memory. The two processes are thought to be represented in two types of variables, one representing the speed of response and one representing accuracy. Individual differences in dual processing can be described by branching models with three constructs: one that accounts for individual differences in the choice of a fast versus a slow process, a second that accounts for individual differences in the accuracy of the fast process, and a third that accounts for individual differences in the accuracy of the slow process. The branching model is a meta-theoretic outline that can be filled-in various ways.

With the development of computer-based testing, recording responses and response time has become accessible and common. In past decades, psychometricians placed increasing attention on analyzing response time. They formulated models and theories to analyze response time and explain the relationship between speed and accuracy (van der Linden 2007, Partchev & De Boeck, 2012, Semmes, Davison, & Close, 2011, De Boeck, Chen, & Davison, 2017). The branching model in Partchev and De Boeck is based on a dual-process theory. The branching model has a speed dimension and two ability dimensions, reflecting the automatic process and the deliberate process. However, the branching model and the method of item scoring cause large amounts of missing values.

Missing data interferes with the valid interpretation of test behaviors and the performance of examinees. A large amount of missing values also leads to sample size issues and problems of parameter estimation. Item response models were usually developed under the assumption of no missing data (Zhang & Walker, 2008). Therefore, researchers need to appropriately handle missing data because the appropriateness and relative effectiveness of IRT models may be sensitive to the amount of missing data. If it is not appropriately considered, missing data can lead to biased estimates of item parameters and a persons' ability. Missing data may make it difficult, if not impossible, to distinguish between the three separate dimensions of speed, fast accuracy, and slow accuracy.

As analyzing data with missing responses is an unavoidable problem, a number of methods have been developed for dealing with missing data. Smits, Mellenbergh, and Vorst (2002) mentioned that "No single method emerged as really superior". The studies (De Ayala, Plake, and Impara, 2001; Finch, 2008; Leite & Beretvas, 2010; Ludlow and O'Leary, 1999; Schafer, Khare, & Ezzati-Rice, 1993; Tabachnick & Fidell, 2001) showed that none of the missing data treatments is perfect for one kind of missing responses. Those treatments are "sensitive" to all aspects of the dataset, which are types of missing responses and data (dichotomous or polytomous), percent of missingness, test length, and sample size. Multiple imputations, EM algorithms, and response function methods provide more robust results of parameter estimation than single imputation, deletion, ignored and scored as incorrect or fractional correct. Only one study (Moustaki & O'Muircheartaigh, 2000) employed the method of scoring missing values as an additional category. However, their study did not investigate the impact of missingness. In considering which missing data treatments should be used, it's important to understand the dataset and research questions to be answered before deciding on a specific technique for handling missing data.

## 1.2 Purpose of Study

The issue dealt with in this study is the problem of distinguishing the three separate dimensions in dual process models such as that in Partchev and De Boeck's (2012) research on differentiating fast and slow intelligence In addressing this question, however, one must first address the issue of missing data that arises in past approaches to the problem.. This study will investigate the effect of missing values caused by a branching model, and it will propose a polytomous scoring method as an alternative way to deal with missing values. The current study will use a computer-administered reading comprehension assessment, which recorded students' responses and response time for each item.

This paper explores a polytomous item scoring method to deal with missing values on a new computer administered reading comprehension assessment in the context of an IRT branching model. The current study reviews previous research on effects of missing values and missing data treatments. Then a polytomous scoring method to deal with missing values is proposed. To demonstrate the polytomous scoring method, a comparative study was conducted similar to Partchev and De Boeck's (2012) study of differentiating fast and slow intelligence. Their study is based on a two-level branching model for dichotomously scored items. Their model and method of item scoring cause large amounts of missing values. The current study employed their dichotomously scored branching model and but also investigated a polytomously scored branching model. The results of the current study are compared with Partchev and De Boeck's (2012) research. This paper begins by reviewing the previous research on missing data treatments and an overview of Partchev and De Boeck's (2012) research. The sample and analysis method are described in the second section. Results are presented in the third section and conclusions are drawn in the final section.

The first research question is how well will model fitting techniques work with large amounts of missing values? In this study, a new approach designed to reduce missing values is proposed. The second question is how well will this new approach perform in estimating person parameters on speed and accuracy dimensions? The question will be addressed with real and simulated data.

# Chapter 2: Literature Review

## 2.1 Dual-Process Theory

In the past decade, dual-process theories have become influential in several areas of psychology. Atkinson and Shiffrin (1968) proposed the idea of two-divisions of memory and information process. Later, Schneider and Shiffrin (1977) presented the theory of information processing based on the *automatic process* and *controlled process*. They stated that *automatic process* involves only a small amount of cognitive processing and any new automatic process needs consistent training, whereas, the *controlled process* is a temporary sequence of actions that require cognitive processing, such as analytic thinking and search for long- or short-term memories. In recent research, dual-process theories have been widely discussed and proposed in many fields, including learning (Sun, Slusarz, & Terry, 2005), reasoning (Reber, 1989), social cognition (Smith & Collins, 2009; Smith & DeCoster, 2000) and decision making (Kahneman & Frederick, 2002; Bernheim & Rangel, 2004; Evans, 2008). The theories vary in different fields, but they have in common two different processing modes in completing cognitive tasks. The two processes can be distinguished as consciousness vs unconsciousness, fast vs slow, high capacity vs low capacity, and automatic vs controlled. Most theories posit that people have the capacity for both processes and that there are individual differences in the extent to which people rely on one process rather than the other. Different authors assigned various names for these two processes. Table 1 lists the names that were used in previous research.

Table 1. Different Names used in Dual-Process Theories

| References | Process 1 | Process 2 |
|---|---|---|
| Schneider and Shiffrin (1977) | Automatic | Controlled |
| Stanovich (1999) | System 1 | System 2 |
| Fodor (1983) | Input | Higher cognition |
| Epstein (1994) | Experiential | Rational |
| Chaiken (1980) | Heuristic | Systematic |
| Reber (1989) | Implicit | Explicit |
| Evans (2008) | Heuristic | Analytic |
| Sloman (1996) | Associative | Rule-based |
| Wason & Evans (1976) | Type 1 | Type 2 |
| Reyna (2004) | Intuitive | Analytic |
| Strack & Deutsch (2004) | Reflective | Impulsive |
| Betsch (2008) | Automatic | Deliberate |

There are four features that distinguish these two processes: consciousness, evolution, functions and individual differences. Dual-process theories are latent in the problem of consciousness. Freud stated that the human mind includes two minds: conscious and unconscious. Dual-process theories of cognition relate to consciousness. Process 1 and process 2 are differentiated by cognitive unconsciousness and cognitive consciousness respectively. Process 2 involves thinking and level of control which is supported by process 1 under unconsciousness (Evans & Over, 1996; Stanovich, 2004). The unconscious may control people's first reaction or behavior without being noticed, and process 2 is used for explanation or behavior adjustment. Schneider and Shiffrin (1977) stated a similar concept by stating that cognitive unconsciousness, lower-order cognition, involves perception and cognitive consciousness needs analytic thinking. The automatic process is usually transitioned to the controlled process by consistent training. In dual process theories of reading (e.g. Laberge & Samuels, 1974), the controlled process precedes the unconscious process called "automaticity."

Some researchers believe that process 1 occurred earlier than process 2 in human evolution (Evans & Over, 1996; Stanovich, 1999; Reber, 1989). Other animals also have process 1 cognition, however, process 2 is related to language, reflective consciousness, and high-order consciousness, which other animals may not have. As for functional differences, researchers asserted that process 1 is fast and automatic with prior knowledge and belief (Evans, 2008; Stanovich, 1999), whereas, process 2 is slow, controlled and rule-based (Sloman, 2002). Process 2 requires logical reasoning, a higher thinking mode, that can infer rules from people's minds. Individual differences in dual-process theories state that process 2 is correlated with general intelligence, such as analytic reasoning, working memory and decision making (Reber, 1989; Stanovich, 1999; Barrett et al., 2004). Working memory and analytical thinking skills can be used to measure people's performance level in cognitive tasks and cognitive capacity.

Dual-process theories have been widely applied to reasoning, social cognition, and decision making. Wason and Evans (1975) card selection tasks showed that responses were not always consistent with logic. It could be influenced by prior knowledge or, belief bias. Later on, Evans (1989) proposed heuristic-analytic theory and stated that the heuristic process is fast, automatic and belief-based, whereas analytic reasoning is slow, sequential and rule-based. The heuristic process provides initial responses and can be intervened by analytic reasoning. Various studies have shown that participants tend to draw inference from their belief and prior knowledge, which influence the reasoning process (De Neys et al., 2005; Markovits et al., 1998; Verschueren et al., 2005). Dual-process theories of social cognition were proposed in the early 1980s and mainly discussed issues about the unconscious processing of social information, such as stereotyping and person perception (Chaiken, 1980; Chaiken & Trope, 1999; Wilson, 2002). There are two cognitive systems: system 1 has existed since evolution and is shared in common with other animals' cognition, whereas, system 2 is rule-based processing and distinctively human. Recently, more and more researchers focus on decision making in dual-process theories. The Default-interventionist view of dual-process theories in decision making (Evans & Stanovich, 2013; Kahneman, 2011) is similar to the heuristic-analytic theory of reasoning (Evans, 2008). The heuristic process provides fast and intuitive judgment and can be corrected by the analytical process to make reflective decision-making. The analytical process takes more time than intuitive judgment.

In conclusion, dual-process theories are a family of theories applied to reason, social cognition and decision making. Although dual-process theories are discussed in a wide range of areas, they all agree that two cognitive processes exist in humans when conducting cognitive tasks. Various researchers have proposed different names to label the two cognitive processes.

### 2.2 Response Time and Accuracy Models

Recently, psychometricians have drawn attention to dual-process theories. The two cognitive processes, automatic and controlled, are two dimensions in measurement models. Automatic and controlled processes reflect speed and response accuracy. Psychometricians are interested in models that measure speed and accuracy relations. With the development of computerized testing, collecting response time becomes easier and more accurate. The availability of response time has led to psychometricians incorporating both response accuracy and response time on tests of intelligence (De Boeck, Chen, & Davison, 2017; De Boeck & Partchev, 2012; Partchev & De Boeck, 2012) and achievement (Davison, Semmes, Huang, & Close, 2012; Davison et al, 2012; Maris & van der Maas, 2012; van der Linden, 2007). In recent years, different models have been proposed to measure the relationship between response time and accuracy, such as hierarchical modeling (van der Linden, 2007; Glas & van der Linden, 2010; Klein Entink, Fox, & van der Linden, 2009), IRT models (Thissen, 1983; Furneaux,1961) and IRT models of categorized response times (De Boeck & Partchev, 2012, Partchev & De Boeck, 2012). In this section, hierarchical models and IRT models that consider both response time and responses will be briefly discussed.

### *2.2.1 Hierarchical Models*

Van der Linden (2007) proposed a two-level hierarchical model jointly modeling response time and responses. The first level is for individual examinees and a set of fixed items, which assumes that each examinee takes the test at a fixed speed. It assumes that examinees took the test with the same speed throughout the test and excluded the effects of fatigue, learning and so on. The second level is for the population of examinees and items and in which item and person parameters are treated as random. The relation between speed and ability across people is measured at the second level. The author used the lognormal model for response times:

$$ln\, t_{ij} = -\tau_j + \beta_i + \epsilon_{ij}\,, \qquad (1)$$

where $t_{ij}$ denotes the response time of person *j* on item *i,* $\tau_j$ is the speed parameter for person *j*. $\beta_i$ is a time-intensity parameter for item *i*. A larger value means that item *i* needs more time to complete. $\epsilon_{ij}$ is a normally distributed variable with variance equal to $\alpha_i^{-2}$. $\alpha_i$ is the time discrimination, which reflects how well an item differentiates fast and slow examinees. The larger $\alpha_i$, the smaller the variance of log response time on item *i* and the better it differentiates people who are fast and slow.

For item responses, van der Linden (2007) used the three-parameter logistic model:

$$P\left(X_{ij} = 1\middle|\theta_j\right) = c_i + (1 - c_i)\frac{exp[a_i(\theta_j - b_i)]}{1 + exp[a_i(\theta_j - b_i)]}, \qquad (2)$$

where $\theta_j$ is the ability parameter for person *j*, $a_i$, $b_i$ and $c_i$ are item discrimination, difficulty and guessing parameters for item *i.* Therefore, the lognormal model of response time in van der Linden (2007) can be expressed as

$$f\left(t_{ij}; \tau_j, \alpha_j, \beta_j\right) = \frac{\alpha_j}{t_{ij}\sqrt{2\pi}} exp\{-\frac{1}{2}[\alpha_j(ln\, t_{ij} - (\beta_j - \tau_j)]\} \qquad (3)$$

Several models are extensions of van der Linden’s hierarchical model. One extension that uses the Box-Cox transformations on response times was proposed by Entink, van der Linden, and Fox (2009), which solved the issue of violating the normality caused by the log transformation. Van der Linden’s model assumes conditional independence between response and response time for an individual. Ranger and Ortner (2012) allow dependence between response and response time on the same item. Their model assumes that response time differs across items, which means that examinees might answer an item faster with a lower accuracy rate. This model allows a more flexible assumption on response times. Different from van der Linden's hierarchical model (2007), in which a covariance structure is assumed on $\theta$ and $\tau$ at the second level, Molenaar, Tuerlincks and van der Mass (2015) proposed a cross-relation term in their response time model. The response time model is

$$ln\, t_{ij} = \beta_i + \varphi_i \tau_j - \varphi_i \rho_d \theta_{jd} + \varepsilon_{ij} \tag{4}$$

Here, $\varphi_i$ is the time discrimination parameter on item *I* and $\rho_d$ is a slope parameter in the regression of the log response times on the latent ability variable. $\varphi_i \rho_d \theta_{jd}$ is the cross-relation function and assumes that item *i* measures the *d*th dimension. The authors claimed that the cross-relation function improves the measurement precision of $\theta$.

*2.2.2 IRT Models*

Thissen (1983) proposed a two-parameter logistic (2PL) model for both response time and responses. For response times, the log transformation was used for a positively skewed distribution of response time. The log transformation for response time used in his research is given by:

$$log(T_{ij}) = \mu + \tau_j + \beta_i - \gamma[a_i(\theta_j - b_i)] + \varepsilon_{ij},\ \varepsilon_{ij} \sim N(0, \sigma^2)\,. \tag{5}$$

The log response time includes mean log response time $\mu$, time-intensity for each item *i* ($\beta_i$), a speed parameter for person *j* ($\tau_j$), and the log-odds of success in the item response model, which is expressed as $a_i(\theta_j - b_i)$. $a_i$ and $b_i$ are the item discrimination and item difficulty parameters for item *i*. $\gamma$ is the regression coefficient that adjusts the speed-accuracy tradeoff. If $\gamma > 0$, response time is expected to be longer when the log odds of answering an item correctly is smaller.

Roskam (1987) built a one-parameter logistic (1PL) response model by adding the processing time to the ability parameter. Hence, the model is given by:

$$P\left(X_{ij} = 1 \middle| \theta_j\right) = \left\{1 + exp[-(\theta_j + ln\, t_{ij} - b_i]\right\}^{-1} \tag{6}$$

This model suggests that as the difficulty of item $i$ increases, the response time could be longer. Also, increasing in time implies an increase in the probability of answering item $i$ correctly.

Verhelst, Verstralen and Jansen (1997) proposed a similar model to Roskam's 1PL model (1987). The model has a shape parameter $\pi_i$ and is given by

$$P\left(X_{ij} = 1 \middle| \theta_j\right) = \left\{1 + exp[-(\theta_j + ln\, t_{ij} - b_i]\right\}^{-\pi_i} \tag{7}$$

Verhelst et al. assume that examinees select the amount of time to answer an item, and the response time follows a two-parameter gamma distribution with rate considering the shape parameter of response time $\pi_i$. When $\pi_i = 1$, the distribution reduces to the 1PL model.

Wang and Hanson (2005) introduced a three-parameter logistic (3PL) model instead of the 1PL model. The response function is given by

$$P\left(X_{ij} = 1 \middle| \theta_j\right) = c_i + (1 - c_i)\{1 + exp\,[-a_i(\theta_i - \rho_j d_i / t_{ij} - b_i)]\}^{-1} \tag{8}$$

The model replaces ln $t_{ij}$ by $-\ \rho_j d_i / t_{ij}$, where $\rho_j$ and $d_i$ are the slowness parameters for person $j$ and item $i$.

This section summarized the response time and accuracy models that are most commonly discussed in the research. There are some models not mentioned here, such as the diffusion model and race model. These models are not frequently used in empirical analysis because of the complexity of the parameter estimation. Small data with the estimation of at least seven parameters may result in high error variance of parameter estimation and sensitivity to starting values. In summary, the response time models described above have some similarities. Authors work on different methods of transformation to adjust for the speed-accuracy tradeoff and treat response time as a continuous variable.

### 2.3 De Boeck's Branching Model of Speed and Accuracy

Partchev and De Boeck (2012) investigated the issue of fast and slow intelligence by fitting psychometric branching models on a Raven-like matrices test and a verbal analogies test. The two intelligence tests both contain multiple-choice items and were administered in a computerized but non-adaptive format. Students were given adequate time to finish the tests.

The psychometric branching model is based on three latent traits: speed, fast accuracy, and slow accuracy, and item parameters corresponding to each of these. The underlying assumption is that a person first activates a fast or a slow process, and the process activated then determines which ability (fast accuracy or slow accuracy) will govern the probability that the person correctly answers the item. Their branching theory of intelligence suggests that there are two intelligence dimensions, and they need to be measured separately.

In their study, within-person and within-item median splits were used to distinguish fast and slow responses. The within-person median split is based on each individual's distribution of response times over items, whereas, the within-item median split is based on each item's distribution of response times over persons.. In the within-person median split, a person's response for a given item is fast if the response time is less than the median response time for that person over all items; otherwise, it is slow. Similarly, in the item median split method, the person's response time is based on each item. If the response time is less than the item median response time for that item over all persons, it is considered as fast. In this research, only the item median method will be used.

Figure 1 shows a diagram of the branching model. Level 1, based on the speed dimension, distinguishes fast and slow based on a median split, and level 2 corresponds to accuracy based on student's correct and incorrect answers within fast or slow branches. This model leads to three binary sub-responses variables shown below.

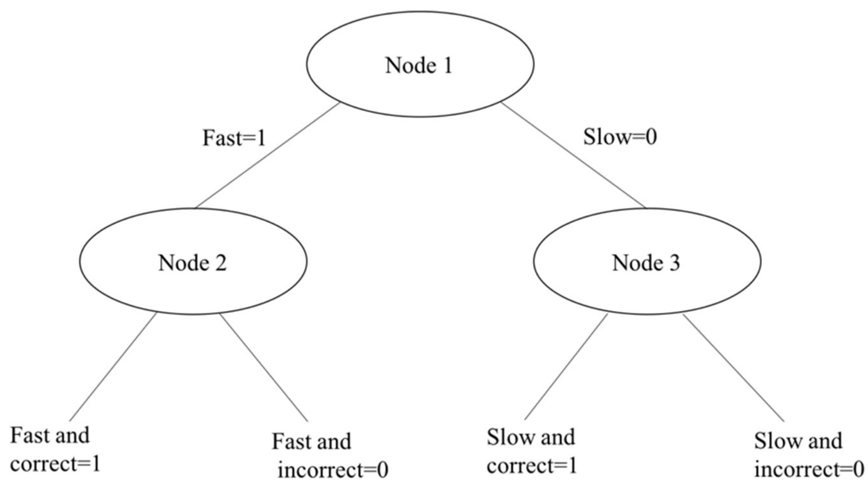



*Figure 1. Structure of Partchev and De Boeck's (2012) Branching model on Differentiating Fast and Slow Intelligence*

The first binary sub-response is the speed variable corresponding to Node 1 in Figure 1, which differentiates fast and slow. The first response variable for item *j* (*j* = 1, …, *J*) is

$$X_{1j} = 1 \text{ if the response is fast} \qquad (9)$$

$$= 0 \text{ if the response is slow.}$$

The second and third sub-responses are based on the first sub-response. The second sub-response, corresponding to Node 2 is the fast accuracy dimension, which is for correct vs. incorrect under the fast branch. The second response variable is

$$X_{2j} = 1 \text{ if the response is fast and correct} \quad (10)$$

$$= 0 \text{ if the response is fast and incorrect}$$

$$= \text{missing if the response is slow.}$$

Similarly, the third sub-response corresponding to Node 3 is the slow accuracy dimension, which indicates correct vs. incorrect under the slow branch. The third response variable is

$$X_{3j} = 1 \text{ if the response is slow and correct} \quad (11)$$

$$= 0 \text{ if the response is slow and incorrect}$$

$$= \text{missing if the response is fast.}$$

The branching model includes three latent dimensions. $\theta_1$ underlies the response variable $x_{1j}$ and is the person's propensity to activate the fast process rather than the slow process. Students with high scores on the dimension are more likely to give a fast rather than a slow response. The second dimension $\theta_2$underlies the response variable $x_{2j}$ and is the person's propensity to give an accurate response given a fast process. The third dimension $\theta_3$underlyies the response variable $x_{3j}$ and is the person's propensity to give an accurate response given a slow process.

The branching model then includes three sub-models, one for each response variable. For the first response variable $x_{1j}$, the model is

$$p_{1j(x_{1j}=1)} = \frac{exp\ (\theta_1 - d_{1j})}{1 + exp\ (\theta_1 - d_{1j})} \quad (12)$$

where $d_{1j}$ is a parameter reflecting the time demands of the item. $p_{1j}(x_{1j} = 1)$ is the person's probability to respond quickly given the time demands of the item. The second submodel, for the second response variable, is a conditional probability, conditioned on $x_{1j} = 1$:

$$p_{2j}(x_{2j} = 1|x_{1j} = 1) = \frac{exp\ (\theta_2 - b_{2j})}{1 + exp\ (\theta_2 - b_{2j})} \quad (13)$$

where $b_{2j}$ is the difficulty of the item when answered with a fast process. The third submodel, for the third response variable $x_{3j}$, is also a conditional probability, conditioned on $x_{1j} = 0$:

$$p_{3j}(x_{3j} = 1|x_{1j} = 0) = \frac{exp\ (\theta_3 - b_{3j})}{1 + exp\ (\theta_3 - b_{3j})} \quad (14)$$

where $b_{3j}$ is the difficulty of item $j$ given that it is answered with a deliberate response process. All three of these models have only one item parameter per item.

The intelligence findings of De Boeck and colleagues provide only limited support for the model (De Boeck & Partchev, 2012; Partchev & De Boeck, 2012). They studied items from two intelligence tests. The first was a set of verbal analogies items. The second was a set of matrix items similar to those on Raven's Progressive Matrices For both the verbal analogies and the matrix items, the AIC and BIC favored their three-dimensional model over a two-dimensional alternative with a speed dimension and a single, overall accuracy dimension. However, they found only modest evidence for distinct fast intelligence and slow intelligence dimensions. In support of their theory, there was only a modest correlation over items between item parameters on the fast and slow intelligence dimensions. However, the model-based estimate of the correlation between the fast and slow intelligence person locations were 0.879 and 0.869 for verbal and matrix items respectively for the within item median split of fast and slow. The correlations between speed and accuracy were more or less uncorrelated in the verbal test, but negatively correlated in the matrix test. In short, the data suggested that items had different locations along the fast and slow intelligence dimensions, but the locations of people were very similar across the two dimensions. The study by De Boeck et al. (2017), a study of mathematics achievement items, also found support for the hypothesis of different item locations, but similar person locations, across fast and slow mathematics achievement dimensions. The authors have been unable to find strong evidence for distinct individual differences dimensions of fast and slow intelligence. Could this failure be attributed to the large amount of missing data in their method?

This section summarized the branching models of speed and accuracy proposed by De Boeck (2012). The multidimensional branching models reflects two-dimension of accuracy, which are a person's ability to answer items correctly via fast/slow process and one dimension of speed. For any person-item combination, only two of the three sub-responses are observed. The slow accuracy sub-response is missing if the first sub-response is fast. The fast accuracy sub-response is missing when the first response is slow. Partchev and De Boeck (2012) stated that missing values in this study would not affect the results of a maximum likelihood-based model estimation because the missing values are missing at random (MAR). That is, missing values in fast accuracy and slow accuracy sub-responses depend on the speed dimension underlying the first sub response variable but they do not rely on the fast accuracy or the slow accuracy variable itself.

As stated above, Partchev and De Boeck's (2012) branching model leads to a large number of missing values. The item median split would cause 50% missing values in each dichotomously scored variable. Even though the missing values in this study are MAR, the estimations of item parameters and person abilities might be affected by the sparse data condition.

## 2.4 Missing data

### *2.4.1 Patterns and Types*

There are three patterns of missing data: univariate, monotone, and arbitrary. Suppose there are *J* variables, denoted as $Y_1, Y_2, \ldots, Y_J$. A dataset with a monotone pattern is one such that when $Y_p$ is missing, then $Y_{p+1}, Y_{p+2}, \ldots, Y_J$ are missing as well. This pattern often occurs in a longitudinal dataset in which participants drop out at one point. The second common pattern is the univariate pattern where only one variable contains missing values. If the missing value exists for any variable and any participant in a random way, it's called an arbitrary pattern. In practice, the pattern of missing data will be an arbitrary pattern in most datasets, missing data that have a monotone pattern or a univariate pattern are easier to impute than an arbitrary pattern (Dong and Peng, 2013).

Missing data are categorized into three types: Missing Completely at Random (MCAR), Missing at Random (MAR), and Missing Not at Random (MNAR). In some IRT research, researchers divided missing data into five categories (Mislevy & Wu, 1988): random assignment, target testing, adaptive testing, not-reached items, and omitted items.

MCAR only occurs randomly across all observations. It assumes that the probability of missing responses does not depend on observed or unobserved variables. For example, random sampling from a population for a survey can cause MCAR. The missing responses are ignorable because the observed data are a random sample from the complete data. Ignoring missing data that are MCAR will not introduce bias but will have larger effects on sampling error as the sample size decreases.

MCAR involves a strong assumption. A weaker assumption is MAR, which assumes that the probability of an observation being missing $Y_p$ depends on $Y_{p+1}$, $Y_{p+2}$,…, $Y_J$ but does not depend on $Y_p$ itself. Missing data that are MAR can be ignorable when estimating a person ability by using direct-likelihood or Bayesian inferences (Mislevy & Wu, 1996). The third type of missing data is MNAR, which means that the probability of missing responses is associated with the variable itself. Thus, the missingness is nonignorable and problematic.

From the perspective of IRT application, types of missing data were categorized into five types (Mislevy & Wu, 1988), but the types involve assumptions similar to MCAR, MAR, and MNAR. Random assignment is defined so that an examinee is randomly administered one test form to minimize carryover effects. Therefore, random assignment satisfies the assumption of MCAR in which the missingness is random and does not depend on the observed variable itself or other variables.

Targeted testing and adaptive testing have some similarities, in which the difficulty of the test is based on the examinee's background information. The missingness of targeted testing is according to the administrator-determined procedure. For example, an easy test form can be administered to third grade and a difficult form to fifth grade. As the probability of missingness does not depend on any observed or unobserved variables or itself, the targeted testing satisfies MCAR. The adaptive testing selects items based on a person's preceding responses to maximize the information for each person (van der Linden and Glas, 2010). The probability of missing responses depends on a person's responses to the previous items b t not the item itself. Hence, the adaptive testing satisfies MAR.

The most frequent and problematic missing responses are not-reached items and omitted items that appear frequently in educational measurement and in which examinees did not have time to answer the last few items or they reached the item and did not answer it. Mislevy (2013) states that not-reached items are MCAR when the persons' chances of responding correctly would not differ if there is no time limit. Whereas, the omitted items are nonignorable. In the treatment of nonresponse, the nature of the missing data mechanism plays an important role. Rose, Davier, and Xu (2010) mentioned that MCAR and MAR do not hamper likelihood-based or Bayesian inference. The missing data mechanism can be considered to be ignorable and does not need to be considered.

Missing responses happen in different ways, and the type of missing data is not conspicuous. Moreover, different missing data treatment applies to different types of missing data. Therefore, deciding the type of missing responses would be essential for analyzing data with missing responses. Lord (1980) said that not-reached responses only appear at the end of the test. For omitted responses, if the examinee has a chance to answer an item correctly that is larger than 1/A, where A is the number of options in an item, this examinee should not omit it. Sijtsma and Van der Ark (2003) stated that whether a missing response is ignorable or nonignorable, researchers have to depend on the pattern of missing responses in the data matrix. If there is no relationship between missing responses and other observed variables, the missing responses may be defined as the MCAR type. When a relationship to other observed variables is found, researchers may use these variables as covariates in multivariate analyses to impute scores. However, it is impossible to verify whether the MAR condition holds based on observed data. When a more complex pattern of relationships of missing responses is found, missing responses may be considered nonignorable. However, missing data will likely never be purely MCAR, MAR, or MNAR in real-world analyses. Also, in practical applications of IRT, item responses may not be observed from all examinees to all items. Hence, categorizing the missing responses into different types requires knowing as much as possible about the sample used for testing.

### *2.4.2 Previous Studies of Missing Data Treatments*

There is a vast literature on missing data treatments and the impact of different treatments. The percentage of missingness, test length, and sample size are also investigated in those studies. The most frequent missing data treatments discussed in previous research include: ignore, score missing values as incorrect, deletion, single imputation, multiple imputation, EM algorithm, etc. This section only focused on the studies related to IRT analysis.

#### *Ignore and Scored as Incorrect or Fractional Corrected*

Some studies have shown that traditional methods of ignoring missing data, scoring missing data as incorrect (IN) or fractional correct (FR) are not optimal approaches.

Lord (1974) applied a 3PL model on empirical data to show that omitted items should not be treated as incorrect when one is interested in accurately estimating either examinee ability or item parameters, which would lead to higher bias in estimating item discrimination and item difficulty. Later on, Lord (1980) did not recommend ignoring omitted responses because the examinees would obtain much higher $\hat{\theta}$ than they actually have. De Ayala, Plake, and Impara (2001) investigated traditional methods of ignoring omitted items and treating omitted items as incorrect on the estimation of person abilities with the 3PL model. They concluded that omitted items should not be scored as incorrect because the accuracy of estimated theta decreased when the number of omitted items increased. Comparing the results for scored as incorrect to ignoring omitted responses, ignoring omits showed more accurate ability estimates under MLE. Also, MLE ability estimation was not affected by 10.3% or fewer omitted items.

Ludlow and O'Leary (1999) investigated not-reached and omitted items in large-scale data sets. Their results noted that treating not-reached items as incorrect can cause higher item difficulties than they should have as well as inflated goodness-of-fit statistics. Finch (2008) investigated the accuracy of estimation for item parameters with the 3PL model. The author applied missing data treatments of missing as incorrect, fractionally correct and ignored, EM, multiple imputation and response function for six conditions, which were sample sizes of 500 and 1,000 with 5%, 15%, 30% MAR or MNAR. The results indicated that parameter estimation bias in the MNAR condition was greater than MAR for all missing data treatments. Moreover, treating missing values as incorrect led to the greatest parameter estimation bias. The guessing parameter was underestimated and had a lower standard error in the MNAR condition. FR method with a fraction of 1/5 had much lower bias than scored as incorrect. It should be noted the authors did not include a condition with the high rate of missingness (50%) in the dual processing theory studies (e.g. Partchev & De Boeck, 2012).

### *Deletion*

The easiest approach to missing data treatment in data analysis is deletion. If the data are MCAR, pairwise deletion produces consistent estimates of the parameters. If the data are MAR, the pairwise deletion may yield biased estimates (Glasser, 1964). Case deletion is not a preferred method for dealing with missingness, except perhaps if the amount of missing data is very small, and the data are MCAR (Tabachnick & Fidell, 2001). Zhang and Walker (2008) investigated listwise and pairwise deletion, single imputation, and multiple imputations with missing values that are MCAR. They found that with low missing data rates (15% and 30%), the estimation accuracy of a person's latent traits obtained with pairwise deletion and hot-deck imputation was very close to that obtained from examinees without missing data. When missing data rates increased to 50%, higher rejection rates of person-model fit were obtained.

### *Multiple Imputation*

Finch (2008) stated that multiple imputation was associated with less bias and smaller standard errors than the EM algorithm and response function methods. In terms of percent of missingness, the estimation bias increased as the percent of missing data increased. Under the condition of adaptive testing, Harmes, Parshall, and Kromery (2003) found that MI techniques performed better for item calibration, which had a lower bias in terms of item parameters. They also found that in general, longer proportional test lengths, which increased from 26 items to 52 items, resulted in smaller amounts of bias across item parameter estimates. Variations in sample size had a much more pronounced effect on the levels of bias. As the number of examinees increased, the amount of bias generally decreased, and the variability across the distribution of bias and magnitude of outliers decreased. Other researchers have shown that imputing ordinal data using the MI model yielded acceptable results when as much as 30% of the data were missing (Leite & Beretvas, 2010; Schafer, Khare, & Ezzati-Rice, 1993).

*EM Algorithm*

EM was superior to listwise deletion, pairwise deletion, and mean imputation in terms of the bias, root means square error, and confidence interval coverage for coefficient alpha.

In extending this research, Bernaards and Sijtsma (2000) found that EM algorithm techniques accurately estimate factor loadings for ordinal data. They stated that the EM approach was the optimal method for Likert data. In the MNAR condition, EM was also typically associated with the largest underestimation of item difficulty of the methods studied here. EM was associated with somewhat greater bias than MI. This result may be due in large part to the fact that the EM approach relies on an assumption of multivariate normality that clearly does not apply to dichotomous item responses. In terms of sample size, the EM algorithm did not perform well with smaller sample sizes that have large amounts of missing responses.

*Score as additional Category*

Moustaki and O'Muircheartaigh (2000)[1] discussed missing values in attitude scales with categorical items, which treats nonresponses as a separate response category and fit the nominal response model. They investigated the relationship between attitude and the probability of not responding to one item.

## 2.5 Synopsis

Smits, Mellenbergh and Vorst (2002) mentioned that "No single method emerged as really superior". The studies reviewed above showed that none of the missing data treatments is perfect for one kind of missing responses. Those treatments are "sensitive" toseveral aspects of the dataset, which are types of missing responses and data (dichotomous or polytomous), percent of missingness, test length, and sample size. Multiple imputation and EM algorithm methods provide more robust results of parameter estimation than single imputation, deletion, ignored and scored as incorrect or fractional correct. Only one study used the method of scoring missing values as an additional category. However, their study did not investigate the impact of missingness. In considering which missing data treatments should be used, it's important to understand the dataset and research questions to be answered before deciding on a specific technique for handling missing data.

[1] The review of Moustaki and O'Muircheartaigh's (2000) study is only based on their abstract.

## 2.6 Pilot Simulation Study

### *2.6.1 Data Generation*

A small simulation study was used to investigate whether the analysis of De Boeck and colleagues could recover the structure described in their theory if it existed. The generating parameters were based on the parameters of Form 3.1 on MOCCA, a 40 item, multiple-choice measure of reading achievement designed for 3$^{rd}$ – 5$^{th}$ graders. A simulation with *n*=500 cases and 40 dichotomous items was conducted. The data were generated and analyzed by R programming.

The independent variables were the means of item difficulty and the means of item discriminations. The means of item parameters from Form 3.1 were utilized in this simulation study. Table 2 shows the conditions of item parameters used for data generation. For dimension 1, 40 item discrimination parameters were uniformly distributed from 1-1.5 and item difficulties were simulated with $N(0,1)$, which provided nearly equal responses of 1 and 0. The items parameters for dimension 2 were simulated with a normal distribution of $\mu_{a2}$ =1.8, $\mu_{b2}$ =-0.3 and variance equal 1. Item parameters were simulated for dimension 3 were generated from a normal distribution with $\mu_{a3}$ =1.7, $\mu_{b3}$ =-0.5 and variances equal 1.

Table 2. Conditions of Item parameters

| | Item Discrimination | Item Difficulty |
|---|---|---|
| Dimension 1 | Random assign 1-1.5 | $\mu_{b1}$ =0 |
| Dimension 2 | $\mu_{a2}$ =1.8 | $\mu_{b2}$ =-0.3 |
| Dimension 3 | $\mu_{a3}$ =1.7 | $\mu_{b3}$ =-0.5 |

For each simulated respondent, $\theta_1$, $\theta_2$ and $\theta_3$ were drawn from a multivariate normal distribution with mean vector $\mu = 0$ and covariance matrix $\sum$. The diagonal elements of $\sum$ were 1.00. The off-diagonal values were set as 0.33 and 0 as the correlation between $\theta_1$ and $\theta_2$, and the correlation between $\theta_2$ and $\theta_3$.

Table 3. Correlations between dimensions

| | Dimension 1 | Dimension 2 | Dimension 3 |
|---|---|---|---|
| Dimension 1 | 1 | 0.33 | 0.33 |
| Dimension 2 | 0.33 | 1 | 0 |
| Dimension 3 | 0.33 | 0 | 1 |

For each item, there is a speed variable, a fast accuracy variable and a slow accuracy variable. For dimension 1, 50% fast responses and 50% slow responses will be randomly drawn based on the parameters. The model for dimension 1 is the 2PL model:

$$p_{1j}(x_{1j} = 1) = \frac{exp\ (a_{1j}(\theta_1 - b_{1j}))}{1 + exp\ (a_{1j}(\theta_1 - b_{1j}))} \tag{15}$$

where $p_{1j}(x_{1j} = 1)$ is the probability of a person responding quickly on item *j* (*j* = 1,…, *j*). $a_{1j}$ and $b_{1j}$ are the item discrimination and item difficulty of item *j* for the first dimension. The item “difficulty” reflects the time demands of the item.

Responses for dimension 2 and dimension 3 still depend on dimension 1. If the simulated response for item 1 is fast, then the response for dimension 2 will be generated based on dimension 2 item parameters and theta parameter. The response for item 1 on dimension 3 will be missing. The model for dimension 2 is conditioned on $x_{1j} = 1$:

$$p_{2j}(x_{2j} = 1 \mid x_{1j} = 1) = \frac{exp\ (a_{2j}(\theta_2 - b_{2j}))}{1 + exp\ (a_{2j}(\theta_2 - b_{2j}))} \tag{16}$$

where $a_{2j}$ and $b_{2j}$ are the item discrimination and item difficulty of item *j* for the second dimension. The model for dimension 3 is conditioned on $x_{1j} = 0$:

$$p_{3j}(x_{3j} = 1 \mid x_{1j} = 0) = \frac{exp\ (a_{3j}(\theta_3 - b_{3j}))}{1 + exp\ (a_{3j}(\theta_3 - b_{3j}))} \tag{17}$$

where $a_{3j}$ and $b_{3j}$ are the item discrimination and item difficulty of item *j* for the third dimension.

The generated data was fit with a three-dimensional dichotomous model as in De Boeck's research. The root mean square error (RMSE), bias and correlations were used as dependent variables to evaluate the accuracy of overall ability parameter recovery. They were computed as follows: let $\theta_i$ be the value of the generating parameter and let $\hat{\theta}_i$ be the value of the estimated parameter. The RMSE was calculated as

$$RMSE = \sqrt{\frac{1}{n} \sum (\theta_i - \hat{\theta}_i)^2} \tag{18}$$

where *n* is the number of replications. Bias was calculated as

$$Bias = \frac{1}{n} \sum (\theta_i - \hat{\theta}_i) \tag{19}$$

Furthermore I calculated the correlations between the generating and estimated scores along the first $r_{\theta_1 \hat{\theta}_1}$, second $r_{\theta_2 \hat{\theta}_2}$ and the third dimension $r_{\theta_3 \hat{\theta}_3}$. The correlations between the estimated scores $r_{\hat{\theta}_1 \hat{\theta}_2}$ on the first and second dimensions, the estimated scores

$r_{\hat{\theta}_1\hat{\theta}_3}$ on the first and third dimensions, and the estimated scores $r_{\hat{\theta}_2\hat{\theta}_3}$ on the second and third dimensions were also calculated..

*2.6.2 Results*

Table 5 shows that the estimated thetas for 3 dimensions all have a mean of 0 and variance of 1. The bias and RMSE between the true theta and estimated theta increased as the amount of missing data increased to 50%. The correlations between the generating and estimated scores along the first $r_{\theta_1\hat{\theta}_1}$, second $r_{\theta_2\hat{\theta}_2}$ and third $r_{\theta_3\hat{\theta}_3}$ dimensions decreased, which are .95, . 82 and .76. With the presence of nearly 50% missing values in dimension 2 and 3, the correlation $r_{\theta_2\hat{\theta}_2}$ and $r_{\theta_3\hat{\theta}_3}$ decreased sharply. The correlation between the estimated scores for $r_{\hat{\theta}_1\hat{\theta}_2}$, $r_{\hat{\theta}_1\hat{\theta}_3}$, and $r_{\hat{\theta}_2\hat{\theta}_3}$ are .93, .94 and .79. The correlation of the generating dimensions was .33, much lower than the recovered correlations. Results suggest that the methods employed by Partchev and De Boeck (2012) may not be able to recover distinct, moderately correlated dimensions even when they exist.

As stated in the previous section, Partchev and De Boeck's (2012) branching model leads to a large amount of missing values. The within item median split would cause at least 50% missing values on dimension 2 and dimension 3 variables. Even though the missing values in this study are MAR, the estimations of item parameters and person abilities might be affected by the sparse data condition. The large RMSE and the correlations between generated thetas and estimated thetas in the simulation results show that the large amount of missing data have great impact on the robustness of the theta estimation. In the current study, only within item median split method was conducted and correlations between different dimension are compared with Partchev and De Boeck's

(2012) study.

Table 4. Covariance between Estimated Thetas

| | Dimension 1 | Dimension 2 | Dimension 3 |
|---|---|---|---|
| Dimension 1 | 1.00 | | |
| Dimension 2 | 0.93 | 1.00 | |
| Dimension 3 | 0.94 | 0.79 | 1.00 |

Table 5. Simulation Study Measures of Person Parameter Estimation Accuracy

| | $\hat{\theta}_1$ | $\hat{\theta}_2$ | $\hat{\theta}_3$ |
|---|---|---|---|
| Mean | 0 | 0 | 0 |
| variance | 0.95 | 0.93 | 0.93 |
| Bias | 0.04 | -0.04 | 0.08 |
| RMSE | 0.33 | 0.78 | 0.81 |

The results suggest that the 50% missing values in the branching model have impact on the theta estimations. The estimated theta correlations between the three dimensions are much higher than the true theta correlations. The simulation study raises several important methodological issues in dual processing theories: treatment of missing data, coding of response variables, and the evidence of the response time and ability dimensions.

In the next chapter, this study presented a polytomous scoring method to deal with missing data in branching model. Both unidimensional and multidimensional IRT models were conducted. Raw scores were computed to provide validity evidence if there is three distinct dimensions, speed, fast accuracy and slow accuracy.

## Chapter 3: Methodology

### 3.1 Sample

This total sample is composed of 3,350 students in 3rd (1,178), 4th (1,143), and 5th (1,029) grades from 59 schools, 32 districts and14 states who participated in the *Multiple-choice Online Causal Comprehension Assessment* (MOCCA; Davison et al., 2018, and Liu et al. (2019)). Three different forms were administrated to each grade through an online platform. Table 6 and Table 7 show the sample size and demographic information for each form. The sample size of grade 3 is slightly larger than grade 4 and grade 5. The percentages of gender and students on free or reduce lunch were nearly equal across all forms. The sample from all forms were predominantly White and Hispanic. The remaining minority groups were Asian, African American, American Indian, Native Hawaiian, and two or more races.

Table 6. Sample Size by Grade and Form

| Grade | Form | Sample Size | Total |
|---|---|---|---|
| 3 | 3.1 | 403 | 1178 |
| | 3.2 | 399 | |
| | 3.3 | 376 | |
| 4 | 4.1 | 411 | 1143 |
| | 4.2 | 405 | |
| | 4.3 | 327 | |
| 5 | 5.1 | 384 | 1029 |
| | 5.2 | 327 | |
| | 5.3 | 318 | |

Table 7. Percentages of Students by Demographics

| | | Grade 3 | | | Grade 4 | | | Grade 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Form 3.1 | Form 3.2 | Form 3.3 | Form 4.1 | Form 4.2 | Form 4.3 | Form 5.1 | Form 5.2 | Form 5.3 |
| Gender | Male | 52.5% | 45.9% | 45.4% | 51.0% | 50.5% | 50.3% | 54.3% | 52.5% | 52.6% |
| | Female | 47.5% | 42.9% | 54.6% | 49.0% | 49.5% | 49.7% | 45.7% | 47.5% | 47.4% |
| Race | White | 65.2% | 65.9% | 64.7% | 66.8% | 69.9% | 63.2% | 67.0% | 68.9% | 67.1% |
| | Hispanic | 19.9% | 16.6% | 18.4% | 22.0% | 18.8% | 19.6% | 24.3% | 22.7% | 22.4% |
| | Asian | 4.4% | 2.4% | 3.8% | 1.6% | 2.4% | 3.9% | 1.9% | 1.1% | 2.0% |
| | American Indian | 1.3% | 2.1% | 2.5% | 0.7% | 0.3% | 0.0% | 0.3% | 0.8% | 1.2% |
| | African American | 6.8% | 9.5% | 10.0% | 5.6% | 5.5% | 9.6% | 3.2% | 3.8% | 5.9% |
| | Native Hawaiian | 0.4% | 0.6% | 0.0% | 0.7% | 1.0% | 1.1% | 0.3% | 0.8% | 0.4% |
| | Two or More | 0.2% | 3.0% | 3.0% | 2.6% | 2.1% | 2.5% | 2.9% | 1.9% | 1.2% |
| English Learner | Yes | 10.9% | 9.9% | 9.1% | 11.3% | 7.7% | 9.4% | 5.7% | 5.8% | 7.1% |
| | No | 89.1% | 90.1% | 90.9% | 88.7% | 92.3% | 90.6% | 94.3% | 94.2% | 92.9% |
| Free Lunch | Yes | 52.8% | 49.4% | 47.7% | 48.7% | 42.0% | 47.4% | 44.4% | 42.5% | 43.5% |
| | No | 47.2% | 50.6% | 50.3% | 51.3% | 58.0% | 52.6% | 55.6% | 57.5% | 56.5% |
| Special Education | Yes | 11.3% | 10.1% | 7.7% | 14.6% | 8.6% | 11.2% | 10.8% | 12.8% | 12.9% |
| | No | 88.7% | 89.9% | 92.3% | 85.4% | 91.4% | 88.8% | 89.2% | 87.2% | 87.1% |

## 3.2 Instrument

MOCCA is an online, multiple-choice, causal comprehension assessment for identifying comprehension processes for students in Grade 3-5. Each grade has three forms for a total of nine forms. Each form has 40 items and two different item orders: forward and backward. Participants were randomly assigned to complete one of the forms at their grade level. Each item is a seven-sentence story with the sixth sentence missing. Three answer choices are: Causal Coherent (CCI), which is the correct answer, Paraphrase (PAR) and Lateral Connect (LCN), which are two different error types. Participants are required to choose one of three choices for the missing sentence. The correct answer, causal coherent response, is the response that best completes the causal sequence of the story when inserted for the missing sentence.. The first type of incorrect response, the paraphrase, simply paraphrases prior information (i.e., generally the goal, subgoal,or a combination of the two) without advancing the story or its causal sequence. The second type of incorrect response, the lateral connection response, is an elaboration of, evaluation of, or association with the information in the story. That is, the response goes beyond the information in the story but does not complete the causal sequence. It may be an inference, and it may be accurate, but it does not fully complete the story (i.e., there is still a causal gap in the story).

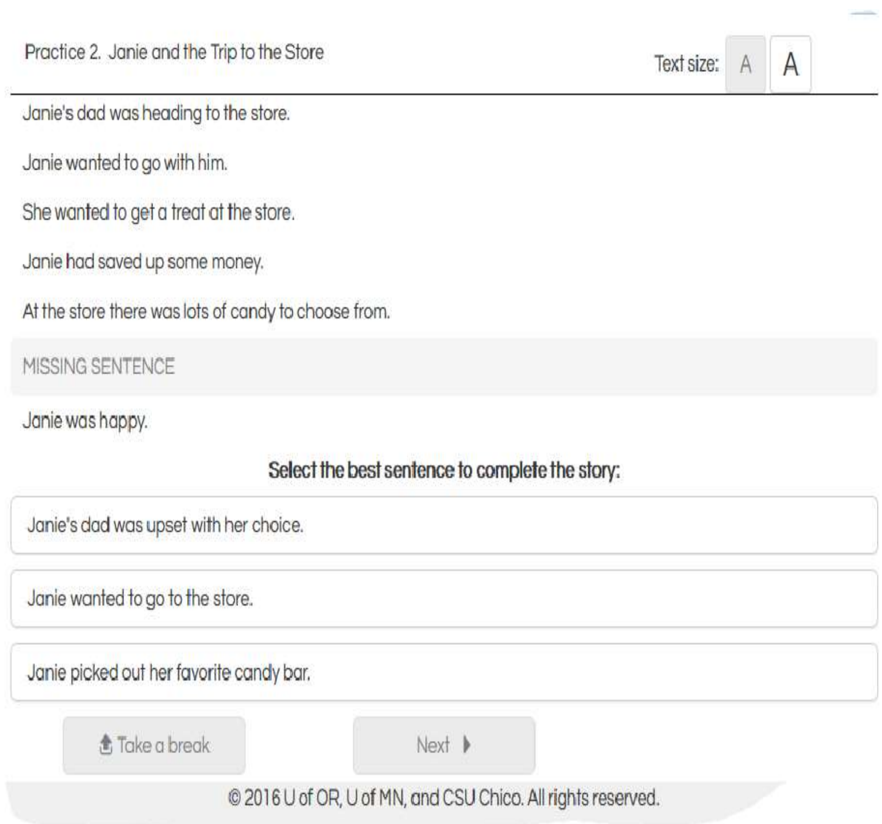



*Figure 2. Screen Shot of Practice Item with, from top to bottom, the Lateral Connect, Paraphrase, and Causal Coherent (correct) Answers Respectively.*

### 3.3 Procedures

MOCCA was administered through an online platform. Teachers were recruited to administer MOCCA through local connections and email blasts. Participants took tests in groups with computers or tablets in their classrooms or in school computer labs between March and June of 2016. Before the test, students were given instructions and an example about the MOCCA test and how to complete the test. Participants were not able to skip items. There was no time limit of finishing the test. Participants could review all the previous responses after completing all items. Participants who did not complet all 40 items were eliminated from the sample.

## 3.4 Analysis

### *3.4.1Dichotomously Scored Variables*

In this study, both dichotomous and polytomous variables were considered. Table 8 shows the dichotomous scoring for the IRT analysis. According to Partchev and De Boeck's (2012) study, the dichotomous variables were based on three categories: fast vs. slow, fast accuracy and slow accuracy. The FS variable indicates fast vs. slow, in which fast and slow are determined by median response time of each item. Participants who responded faster than the median response time are coded as 1 and those who are slower than the median are coded as 0. The MF and MS variables are for correct vs. incorrect following the fast and slow branches. For the fast correct dimension (MF), participants who were categorized as fast on the FS variable are scored as 1 for correct answers and 0 for incorrect answers. For the slow correct dimemsion (MS), those who were slow on the FS variable, are categorized into 1 for correct answers and 0 for incorrect answers. For the analyses of these three response variables, we fit the following three 2PL models:

$$p_{1j(x_{1j}=1)} = \frac{exp\ [a_j(\theta_1 - b_{1j})]}{1 + exp\ \ a_j(\theta_1 - d_{1j})} \qquad (20)$$

$$p_{2j(x_{2j}=1)} = \frac{exp\ [a_j\ (\theta_2 - b_{2j}])}{1 + exp\ [a_j(\theta_1 - b_{2j}])} \qquad (21)$$

and

$$p_{3j(x_{3j}=1)} = \frac{exp[a_j\,(\theta_3 - b_{3j})]}{1+\,exp\,[a_j(\theta_3 - b_{3j})]} \tag{22}$$

where $(a_j, b_j)$ refer to discrimination and difficulty parameters for the fast and slow correct item response variables. These are all unidimensional 2PL models that differ in the latent dimension: a fast dimension $\theta_1$ underlying the response variable $x_{1j}$, a fast accuracy dimension $\theta_2$ underlying the latent variable $x_{2j}$, and a slow accuracy dimension $\theta_3$ underlying the latent variable $x_{3j}$.

The other way to code the data avoids coding any item as missing. While not used in prior research, it was investigated as a possibility here. In this coding, participants who were categorized as fast and correct were coded as 1 in variable FCNO and all other responses were coded as 0

$$\begin{aligned} x'_{2j} &= 1 \text{ if the response to item j is fast and correct} \\ &= 0 \text{ if the response to item } j \text{ is either slow or it is incorrect.} \end{aligned} \tag{23}$$

In the coding of Equation 2 and $x_{2j}$, a response is considered as providing evidence of a student's ability along the latent fast-accuracy dimension only if it is fast. A slow response is treated as uninformative regarding the fast-correct dimension and, therefore is coded as missing. In the coding of Equation 23 and $x'_{2j}$, every response, irrespective of whether it is fast or slow, is considered informative with respect to the latent fast-accuracy dimension, even if it is slow. The corresponding coding for variables manifesting a slow-accuracy dimension (SCNO) is

$$x'_{3j} = 1 \text{ if the response to item } j \text{ is slow and correct v} \quad (24)$$

$$= 0 \text{ if the response to item } j \text{ is either fast or incorrect.}$$

Table 8. Dichotomously Scored Variables

| Response | FS | MF | MS | FCNO | SCNO |
|---|---|---|---|---|---|
| Fast and correct | 1 | 1 | Missing | 1 | 0 |
| Fast and incorrect | 1 | 0 | Missing | 0 | 0 |
| Slow and correct | 0 | Missing | 1 | 0 | 1 |
| Slow and incorrect | 0 | Missing | 0 | 0 | 0 |

The designation of "missing" in Table 8 may not entirely suggest that students did not answer a question. As the structure of branching model, it may be described as uninformative because students made responses to questions but are uninformative on sub-responses. For instance, for MF variable, the slow sub-response is not missing but rather it is uninformative as to underlying fast accuracy dimension. Similarly, for MS variable, the fast sub-response is uninformative under slow accuracy dimension. The polytomously scored variables treat responses as uninformative, rather than missing.

### *3.4.2 Polytomously Scored Variable*s

Table 9 shows the first set of polytomously scored variables. The first variable is a fast/slow variable that is the same dichotomous variable as FS in Table 9. For the purpose of measuring fast and slow intelligence, the polytomously scored variables treat one of the sub-responses as the middle category and the other sub-response are coded as 0 and 2. That is, fast and accurate responses are scored as 2. Let $y_{2j}$ be the polytomously recoded fast accuracy variable counterpart of $x_{2j}$ and let $y_{3j}$ be the polytomously re-coded slow accuracy counterpart of $x_{3j}$ .

$$y_{2j} = 2 \text{ if the response to item } j \text{ is fast and correct.} \tag{25}$$

$= 1$ if slow

$= 0$ if fast and incorrect.

$$y_{3j} = 2 \text{ if the response to item } j \text{ is slow and correct.} \tag{26}$$

$= 1$ if fast

$= 0$ if slow and incorrect.

Table 9. Polytomously Scored Variables

| Response | FS | PF | PS |
|---|---|---|---|
| Fast and correct | 1 | 2 | 1 |
| Fast and incorrect | 1 | 0 | 1 |
| Slow and correct | 0 | 1 | 2 |
| Slow and incorrect | 0 | 1 | 0 |

The polytomously scored variables analogous to a rating scale like that used in the measurement of attitudes. The rating scale for measuring attitudes with three categories of attitude is as follows: disagree, neither agree nor disagree, and agree. For example, the neutral category of the polytomous fast sub-response, which is neither fast correct nor fast incorrect, is the neutral point in the rating scale in that it does not provide information about the person's tendency to answer correctly when using a fast process. It is neutral with respect to whether the person tends to be correct when giving a fast answer. Fast correct and fast incorrect are scaled on the two ends as disagree and agree in the rating scale. Here the fast correct dimension is viewed as being bipolar with people who give all fast correct responses at one end, people who give all fast incorrect answers at the other end, and people with roughly equal numbers of fast correct and fast incorrect responses in the middle of the dimension. As the analogy of the rating scale scoring, the uninformative responses in fast- and slow- accuracy were scored as 1 in the current study.

| Disagree | Neither Agree nor Disagree | Agree |
|---|---|---|
| __________ | __________ | __________ |
| Fast incorrect | Slow incorrect<br>Slow correct | Fast correct |

This study will compare the results of unidimensional, two dimensional and three dimensional graded response models. For the polytomous variables, we fit the graded response model, although a generalized partial credit model may also be appropriate. For the three polytomous variables, the model for item *j* and response category *c* were

$$p_{2j}\left(y_{2j}=c\right)=\frac{\exp\left[a_{2j}(\theta_2-b_{2jc})\right]}{1+\exp\left[a_{2j}(\theta_2-b_{2jc})\right]}-\frac{\exp\left[a_{2j}(\theta_2-b_{2jc+1})\right]}{1+\exp\left[a_{2j}(\theta_2-b_{2jc+1})\right]} \qquad (27)$$

where $c$ = 0, 1, or 2 and $\frac{\exp\left[a_{2j}(\theta_2-b_{2jc+1})\right]}{1+\exp\left[a_{2j}(\theta_2-b_{2jc+1})\right]}\equiv 0$ when $c$ = 2 and $\frac{\exp\left[a_{2j}(\theta_2-b_{2jc})\right]}{1+\exp\left[a_{2j}(\theta_2-b_{2jc})\right]}=1$ when $c$ = 0. The graded response model for $y_{3j}$ is defined similarly:

$$p_{3j}\left(y_{3j} = c\right) = \frac{\exp\left[a_{3j}(\theta_3 - b_{3jc})\right]}{1+\exp\left[a_{3j}(\theta_3 - b_{3jc})\right]} - \frac{\exp\left[a_{3j}(\theta_3 - b_{3jc+1})\right]}{1+\exp\left[a_{3j}(\theta_3 - b_{3jc+1})\right]} \tag{28}$$

where $c = 0, 1,$ or $2$ and $\frac{\exp\left[a_{2j}(\theta_2 - b_{2jc+1})\right]}{1+\exp\left[a_{2j}(\theta_2 - b_{2jc+1})\right]} = 0$ when $c = 2$ and $\frac{\exp\left[a_{2j}(\theta_2 - b_{2j}\ )\right]}{1+\exp\left[a_{2j}(\theta_2 - b_{2jc})\right]} = 1$ when $c = 0$.

### *3.4.3 Measures*

Dichotomous and polytomous models for each form were evaluated by the following criteria. The first criteria are the global model fit statistics, which are, -2LL, information criteria AIC (Akaike, 1974), BIC (Schwartz, 1978) and RMSEA.

$$AIC = = -2\ln(L) + 2K \tag{30}$$

$$BIC = -2\ln(L) + \ln(N)\,K \tag{31}$$

where $K$ is the number of parameters and $N$ is the sample size. AIC estimates the expected relative distance between the candidate model and the true model. This study compared the model fit between the two-dimensional model and three-dimensional models to check if the results agree with Partchev and De Boeck's (2012) research.

The second criterion is the correlation of theta values and scatterplots were used to visualize the relations among dimensions. The correlations between speed and accuracy, and correlations between speed and fast correct and slow correct dimensions are compared for Partchev and De Boeck's (2012) dichotomous coding as compared to the polytomous coding.

The third criterion is the raw score measures of speed, fast accuracy and slow accuracy. This study used raw score measures of the three constructs in a dual processing theory. In this section, the number of correct items, the number of items that the examinee answers correctly with a fast response, the difference between the number of items answered correctly with a fast response and the number answered incorrectly with a fast response and the proportion of fast responses that are correct were calculated. Similarly, the number of items that the examinee answers correctly with a slow response, the difference between the number of items answered correctly with a slow response and the number answered incorrectly with a slow response and the proportion of slow responses that are correct were calculated. These raw scores will be used to aid interpretation of fast accuracy and slow accuracy dimensions in IRT models mentioned above. The correlations between those raw scores and theta scores were examined and the scatterplots will be used to demonstrate the relations between raw scores and theta values.

This study will provide two dichotomous codings and one polytomous coding. Both unidimensional and multidimensional IRT models will be conducted. The model comparison will show how well the IRT models fit the data and whether the results align with Partchev and De Boeck's (2012) results. The correlations between thetas and raw scores will provide evidence of whether the three dimensions are distinctive and whether polytomous scoring method is an alternative solution

# Chapter 4: Results

This Chapter has three sections and presents the results for unidimensional dichotomous 2PL models, unidimensional polytomous models, multidimensional dichotomous 2PL models and multidiemensional polytomous models. The first section contains descriptive statistics of response time, the means and standard deviations for the raw scores, and the percentages of missing values for dichotomously coded variables. The second section contains model comparisons. The models are compared on the basis of fit to the data and the distinctness of the three dimensions within them. This section will compare the multidimensional models mentioned above and check if the results agree with the Partchev and De Boeck's (2012) results with regard to dimensionality: two vs three dimensions. The third section is correlation between person abilities and the correlations between raw scores and person abilities. These correlations clarify the meaning of the dimensions in the alternative IRT models.

This chapter will show the tables and figures of results for Form 3.1, Form 4.1 and Form 5.1 and use these three forms as examples to illustrate the results for each grade. The tables and figures for the remaining six are shown in the Appendix.

## 4.1 Descriptive Statistics

### *4.1.1 Response time*

In this study, I used the within item median response time to split responses on the items into fast and slow. The median response time of each item for Form 3.1, Form 4.1 and Form 5.1 are listed in table 10, table 11 and table 12. The median response time ranges from 38.1 to 60.3 seconds, 37.3 to 55.1 seconds and 38.7 to 59.8 seconds for Form 3.1, Form 4.1 and From 5.1 respectively. These results show that participants across all grades took less than one minute, on average, to answer one item. The means of the 40 median response times are 40.75 for Form 3.1, 43.92 for Form 4.1 and 47.02 for Form 5.1. Participants from Grade 5 spent more time in answering one item than Grade 4 and Grade 3 students.

Table10. Form 3.1 Median Response Time in seconds of Each Item

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 40.8 | 11 | 55.9 | 21 | 45.3 | 31 | 49.2 |
| 2 | 39.2 | 12 | 38.8 | 22 | 36.3 | 32 | 44.3 |
| 3 | 34.3 | 13 | 44.9 | 23 | 46.4 | 33 | 38.3 |
| 4 | 48.2 | 14 | 44.9 | 24 | 48.2 | 34 | 44.2 |
| 5 | 45.7 | 15 | 49.2 | 25 | 37.1 | 35 | 44.2 |
| 6 | 47.6 | 16 | 44.1 | 26 | 46.2 | 36 | 42.9 |
| 7 | 44.0 | 17 | 46.1 | 27 | 39.2 | 37 | 43.8 |
| 8 | 47.4 | 18 | 39.6 | 28 | 49.0 | 38 | 42.8 |
| 9 | 44.9 | 19 | 42.4 | 29 | 39.6 | 39 | 43.0 |
| 10 | 39.4 | 20 | 47.2 | 30 | 38.5 | 40 | 47.5 |

Table 11. Form 4.1 Median Response Time in Seconds of Each Item

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 37.3 | 11 | 50.3 | 21 | 48.3 | 31 | 45.1 |
| 2 | 39.1 | 12 | 45.1 | 22 | 47.4 | 32 | 43.7 |
| 3 | 50.6 | 13 | 47.9 | 23 | 44.4 | 33 | 39.7 |
| 4 | 39.8 | 14 | 34.9 | 24 | 33.4 | 34 | 47.1 |
| 5 | 37.5 | 15 | 52.6 | 25 | 42.2 | 35 | 44.1 |
| 6 | 47.7 | 16 | 43.0 | 26 | 51.8 | 36 | 55.1 |
| 7 | 46.4 | 17 | 43.1 | 27 | 44.4 | 37 | 44.0 |
| 8 | 35.2 | 18 | 38.4 | 28 | 40.9 | 38 | 41.4 |
| 9 | 38.3 | 19 | 42.9 | 29 | 42.3 | 39 | 43.2 |
| 10 | 42.4 | 20 | 53.1 | 30 | 49.9 | 40 | 43.1 |

Table 12. Form 5.1 Median Response Time in Seconds of Each Item

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 48.3 | 11 | 46.6 | 21 | 52.3 | 31 | 54.8 |
| 2 | 44.1 | 12 | 44.4 | 22 | 58.1 | 32 | 45.0 |
| 3 | 42.8 | 13 | 46.2 | 23 | 47.4 | 33 | 46.8 |
| 4 | 55.9 | 14 | 39.9 | 24 | 42.5 | 34 | 39.3 |
| 5 | 36.8 | 15 | 54.7 | 25 | 42.7 | 35 | 46.4 |
| 6 | 55.8 | 16 | 42.6 | 26 | 54.7 | 36 | 49.9 |
| 7 | 39.8 | 17 | 53.7 | 27 | 45.2 | 37 | 40.1 |
| 8 | 38.7 | 18 | 48.8 | 28 | 61.4 | 38 | 35.8 |
| 9 | 47.4 | 19 | 50.1 | 29 | 41.4 | 39 | 45.7 |
| 10 | 49.1 | 20 | 59.8 | 30 | 42.9 | 40 | 44.0 |

### *4.1.2 Percentage of Missingness and Raw Scores*

As this study only include participants who completed all 40 items, the percentage of missingness for dichotomously coded MF and MS only depends on the FS response variable. With the within item median split, 50% responses were categorized as fast and 50% responses were slow. Therefore, the percentages of missing values for MF and MS are both 50%. As the missing values were coded in the 0 category for the dichotomous varianbles FCNO and SCNO, there is no missing data on these variables. As the missing values were scored in the neutral category in the two polytomous sub-response variables, PF and PS, there are no missing values for these response variables..

This study calculated 12 raw scores shown in Table 13. Form 4.1 (28.27) and Form 5.1 (28.21) have higher means of correct responses than Form 3.1 (24.37). As will be seen below, each of these raw scores corresponds to a latent dimension in one of our IRT solutions, and the correlation between the raw score and its latent counterpart helps in the interpretation of the latent variable. *FC* is the number of items that the examinee answers correctly with a fast response and here called the fast-correct score. Form 5.1 has higher mean than Form 4.1 and Form 3.1. As a raw score indicator of the fast- slow dimension, variable *Fast* was used here. An item response time was categorized as fast ($x_{1j} = 1$) if the response time was less than the median response time for that item. An item response was classified as slow ($x_{1j} = 0$) if the response time was greater than the median response time for that item. These three forms have similar means for *Fast*. The proportion of fast responses that are correct, *FCF*, is higher for Form 5.1 and Form 4.1 since these two forms have higher *FC*. *FC - FI* is the difference between the number of items answered correctly with a fast response and the number answered incorrectly with a fast response. FC – FI is the net fast correct, the number of fast correct responses net the number of fast incorrect responses. Form 5.1 and Form 4.1 have a much larger difference than Form 3.1. In a parallel fashion, we define three slow accuracy variables. The first is *SC*, which is slow-correct score. The second is the proportion of slow responses that are correct, *SCS* and is the proportion slow correct. The third is *SC – SI* and is the difference between the number of items answered correctly with a slow response and the number answered incorrectly with a slow response SC – SI is the net slow correct responses, the number of slow correct responses net any slow incorrect response.

Table 13. Mean and SD of Raw Scores by Forms

| | Form 3.1 | | Form 4.1 | | Form 5.1 | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| Correct | 24.37 | 10.49 | 28.27 | 9.73 | 28.21 | 9.09 |
| Incorrect | 15.63 | 10.49 | 11.73 | 9.72 | 11.79 | 9.09 |
| Fast | 20.06 | 12.23 | 20.09 | 11.37 | 20.03 | 11.18 |
| Slow | 19.94 | 12.23 | 19.91 | 11.37 | 19.97 | 11.18 |
| FC | 11.49 | 9.97 | 13.37 | 10.25 | 13.69 | 10.35 |
| SC | 12.89 | 10.12 | 14.90 | 10.34 | 14.52 | 9.71 |
| FI | 8.58 | 8.57 | 6.72 | 7.68 | 6.34 | 6.67 |
| SI | 7.05 | 6.75 | 5.01 | 5.00 | 5.45 | 5.30 |
| FC_FI | 2.91 | 14.01 | 6.64 | 14.10 | 7.36 | 13.35 |
| SC_SI | 5.84 | 12.09 | 9.89 | 11.60 | 9.07 | 10.93 |
| FCF | 0.56 | 0.32 | 0.64 | 0.31 | 0.67 | 0.29 |
| SCS | 0.61 | 0.28 | 0.71 | 0.27 | 0.71 | 0.25 |

## 4.2 Model Comparison

For each form, three two-dimensional models and three three-dimensional models were analyzed. Table 14 shows the goodness of fit results of the multidimensional 2PL and polytomous models. The three-dimensional 2PL models with response vectors ($x_{1j}$, $x_{2j}$, $x_{3j}$) have lower -2LL, AIC and BIC than the two dimensional 2PL model across all forms, which suggests that three dimensional 2PL model fits better. Partchev and De Boeck (2012)'s research has similar findings in that their three-dimensional 1PL model fits the best among all models in their study.

For the response vectors ($x_{1j}, x'_{2j}, x'_{3j}$) the AIC and BIC for the three-dimensional models have lower -2LL, AC and BIC than the two dimensional models across all forms. Finally, the three dimensional graded response models with response vector ($x_{1j}, y_{2j}, y_{3j}$) have lower -2LL, AIC and BIC than the two dimensional graded response models, which suggests that three dimensional polytomous model fit better than the two dimensional polytomous model.

Table 14. Model Comparisons of Multidimensional Models by Forms

| Form | Dimen-sions | Response Vectors | -2LL | AIC | BIC |
|---|---|---|---|---|---|
| Form 3.1 | Two Di-mensional | $x_{1j}, x_{2j}, x_{3j}$ | 33514.33 | 33996.33 | 34960.08 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 48959.32 | 49441.32 | 50405.06 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 78479.73 | 79212.73 | 80405.39 |
| | Three Di-mensional | $x_{1j}, x_{2j}, x_{3j}$ | 33481.62 | 33967.62 | 34939.36 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 47283.24 | 47769.24 | 48740.98 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 75833.42 | 76479.42 | 77771.07 |
| Form 4.1 | Two Di-mensional | $x_{1j}, x_{2j}, x_{3j}$ | 33445.26 | 33927.26 | 34895.74 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 52683.55 | 53165.55 | 54134.03 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 78935.14 | 79577.14 | 80867.11 |
| | Three Di-mensional | $x_{1j}, x_{2j}, x_{3j}$ | 33389.61 | 33875.61 | 34852.13 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 51362.70 | 51848.70 | 52825.22 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 75928.96 | 76574.96 | 77872.97 |
| Form 5.1 | Two Di-mensional | $x_{1j}, x_{2j}, x_{3j}$ | 31715.92 | 32197.92 | 33150.02 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 49415.14 | 49897.41 | 50849.51 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 73939.33 | 74581.33 | 75849.48 |
| | Three Di-mensional | $x_{1j}, x_{2j}, x_{3j}$ | 31680.21 | 32166.21 | 33126.22 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 48522.02 | 49008.02 | 49968.03 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 71562.09 | 72208.09 | 73484.15 |

## 4.3 Correlations

As the three-dimensional IRT models are the best fitting models in this study, the two-dimensional theta scores will be excluded. The raw scores, theta scores from three-dimensional and unidimensional IRT models will be analysed in the following section.

### *4.3.1 Raw Scores*

To investigate the plausibility of raw scores reflecting separate fast- and slow-accuracy dimensions, we examined the correlations of possible raw score indicators of those dimensions. Table 15-17 show the intercorrelations of raw scores variables that can be interpreted as indicators of a fast, a fast-accuracy, or a slow-accuracy dimension. There are three sets of variables that might be considered measure of speed, fast accuracy, or slow accuracy in the following analysis.

The first set of variables that can be interpreted as measuring speed, fast-accuracy and slow-accuracy would be *Fast*, *FC*, and *SC*. *Fast* is highly correlated with both *FC* and *SC*, .719 and -.834, .752 and -.898, .811 and -.811 for Form 3.1, Form 4.1 and Form 5.1 respectively. However, *FC* and *SC* are correlated only -.454, -.554 and -.591 for Form 3.1, Form 4.1 and Form 5.1, which suggest that they manifest distinct concepts. Both *FC* and *SC* are more highly correlated with *Fast* than with *Correct* suggesting that they reflect speed of response more than accuracy of response.

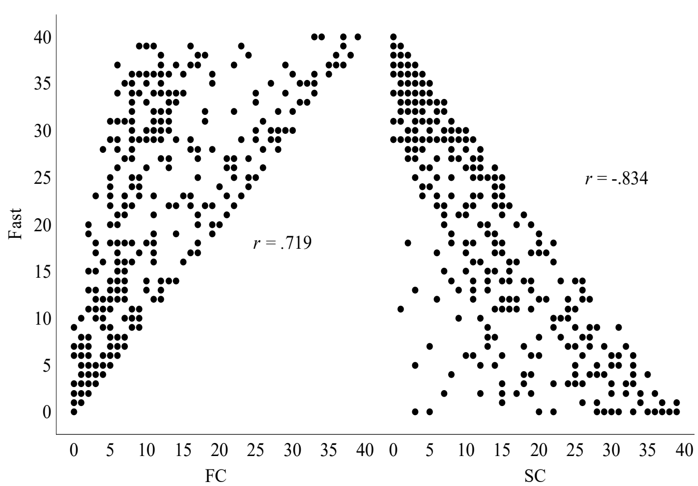
Figure 3 contains scatterplots for pairs of variables in the set of *Fast*, *FC*, *SC*. The first is a plot of *Fast* vs. *FC* that shows a distinctive triangular shape. As *Fast* increases, the theoretical upper limit on the range of *FC* increases. For a given value of *Fast*, the theoretical range of *FC* is 0-Fast: $0 < FC \leq Fast$. Increasing the upper limit of

*FC* as *Fast* increases has two effects: the expected value of *FC* given *Fast* increases, as evidenced by the correlation of .719, .752 and .811, and the range of *FC* increases. The increase in the range of *FC* as a function of *F* gives the scatterplot its distinctive triangular shape.

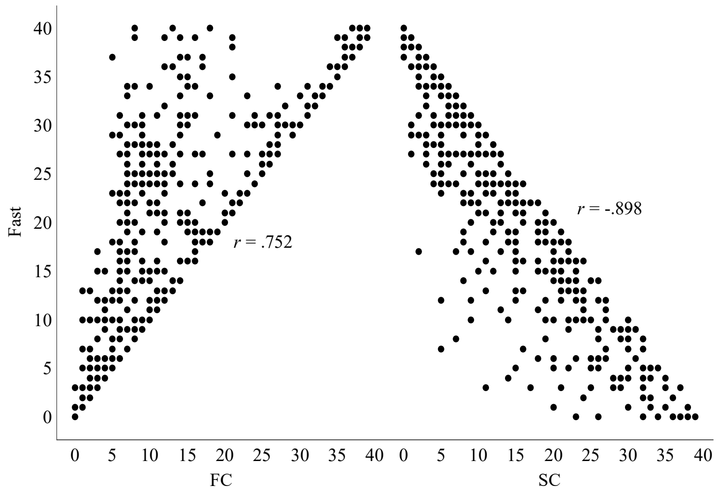
The scatterplots in Figure 3, *Fast* vs. *SC*, also show distinctive triangular shapes. As *Fast* increases, the theoretical upper limit on the range of *SC* decreases. For a given value of *Fast*, *SC* ranges from 0 – (40 – *F*) as there are 40 items: $0 < SC \leq (40 - FC)$. Decreasing the upper limit of *SC* as *Fast* increases has two effects: the expected value of *SC* given *Fast* decreases, as evidenced by the negative correlation of -.834, -.898 and -.811, and the range of SC decreases. The decrease in the range as a function of *Fast* gives the scatterplot its distinctive triangular shape.

Figure 3. Raw Score Scatterplots of *FC* and *SC* against Fast.

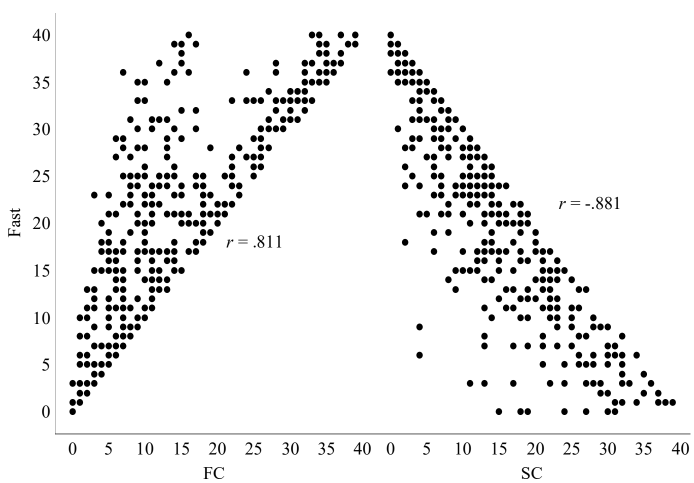
Form 3.1



Form 4.1



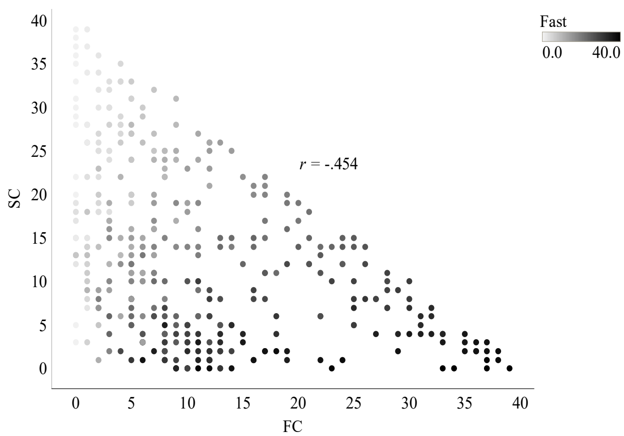

Form 5.1



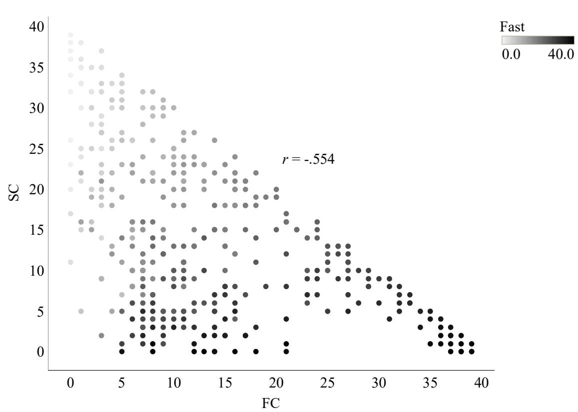

The scatterplots in Figure 4 are heat diagrams in which the shading shows the number of fast responses of the student. Darker dots indicate students with predominantly fast responses, and the shading becomes lighter as the number of fast responses by the student decreases. These scatterplots also have a distinctive triangular shape and for a similar reason. For a given value of *SC*, the theoretical range of *FC* is 0 to (40 – *SC*): $0 < FC \leq (40 - SC)$. As a result, the expectation of *FC* to decrease as *SC* increases, resulting in the correlation of -.437, and a decreasing range for *FC* as *SC* increases. The distinctive triangular shape results from the decrease in range of *FC* as *SC* increases.

Figure 4. Raw Score Scatterplots of *FC* against *SC*

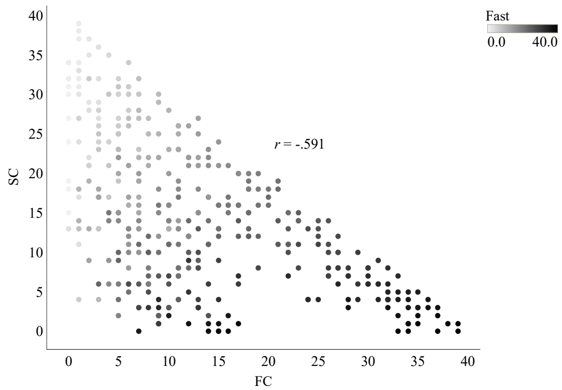
Form 3.1




Form 4.1



Form 5.1

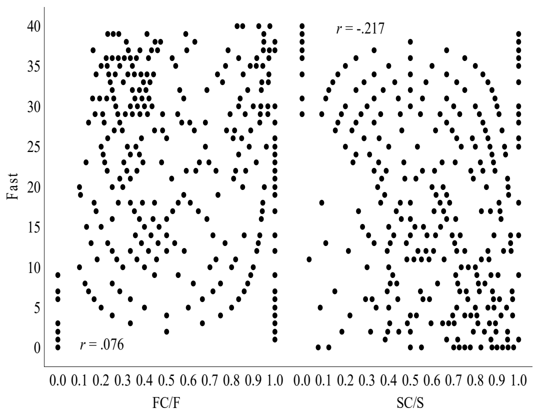



The second set of raw score variables that might be considered measures of speed, fast accuracy, or slow accuracy respectively are *Fast*, *FCF*, and *SCS*. *FCF* and *SCS* are most highly correlated with overall accuracy, *Correct.* The correlations for all forms rang from .7 to .8. Their correlations with speed *Fast* are small and not significant across all forms, which ranges for -.26 and -.017, supporting the hypothesis that there are fast- and slow- accuracy dimensions relatively independent of speed. The correlation of *FCF* and *SCS* with each other for Form 3.1, Form 4.1 and Form 5.1 are .557, .512 and .639, which gives some modest support to the hypothesis that FCF and SCS represent distinct dimensions of fast- and slow-accuracy.

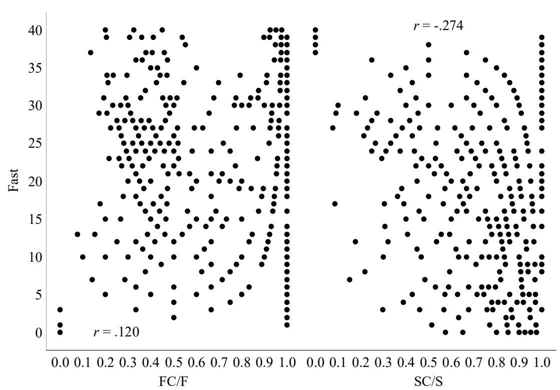
Figure 5 shows the correlation and the scatterplots for pairs of variables in the set (*Fast*, *FCF*, *SCS)*. The *Fast* variable was almost uncorrelated with either *FCF* (*r* = -.076, .120, 113) or *SCS* (*r* = -.217, -.274, -.090) for the three forms.

Figure 5. Raw Score Scatterplots of *FCF* and *SCS* against *Fast*

Form 3.1




Form 4.1




Form 5.1

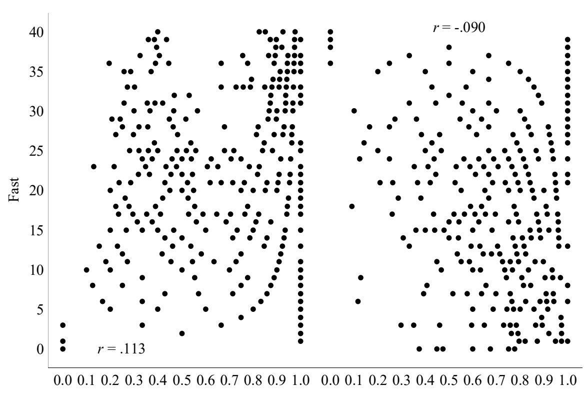
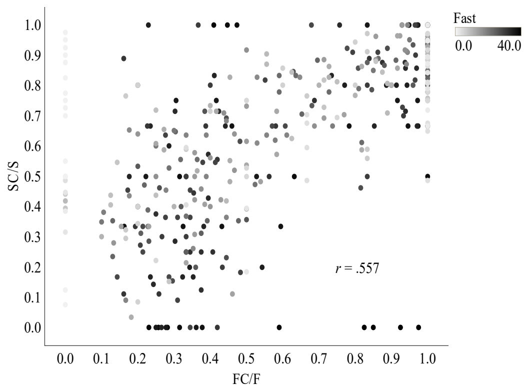
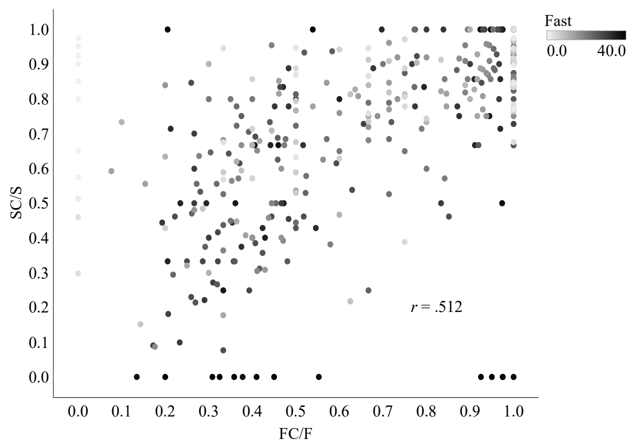


Figure 6. Raw Score Scatterplots of *FCF* against *SCS*

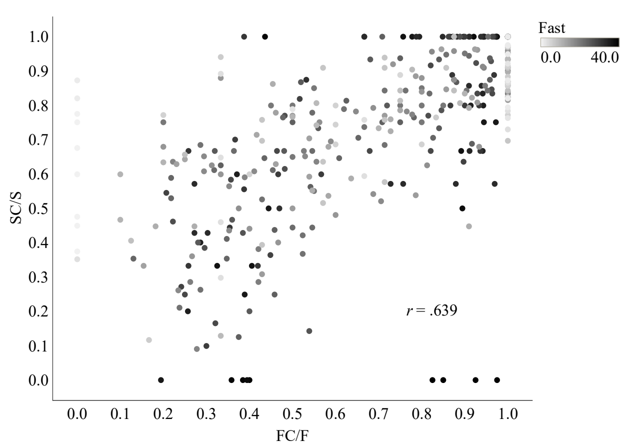
Form 3.1




Form 4.1




Form 5.1



The scatterplots in Figure 6 are heat diagrams with the shading of dots showing the number of fast responses by the student. Darker dots indicate students with predominantly fast responses, and lighter dots indicate students with predominantly slow responses. There is a substantial positive correlation between the *FCF* and *SCS* ($r$ = .557, .512 and .639)[2]. These correlations would suggest that fast accuracy and slow accuracy, as measured by *FCF* and *SCS,* are distinct from speed, but but moderately correlated with each other.

The other set of raw score variables that could manifest the dimensions fast, fast-accuracy, and slow-accuracy are *Fast*, *FC – FI*, and *SC – SI*. *Fast* has significant but rather modest relationships with both *FC – FI* and *SC – SI*: .151 and -.384 for Form 3.1, .288 and -.623 for Form 4.1 and .420 and -.541. Furthermore*, FC – FI* and *SC – SI* are significantly but only modestly correlated, .289, .137 and .113. Whereas fast-correct and slow-correct seem to reflect response speed more than response accuracy, net fast-correct and net slow-correct seem to reflect accuracy more than speed. The correlation of *Correct* with *FC – FI*, and *SC – SI*, were .834 and .769 for Form 3.1, .807 and .696 for Form 4.1, .802 and .684 for Form 5.1. Both *FC – FI* and *SC – SI* are relatively uncorrelated with each other and with *Fast,* leading to the conclusion that these three variables constitute distinct measures of speed, fast-accuracy, and slow-accuracy.

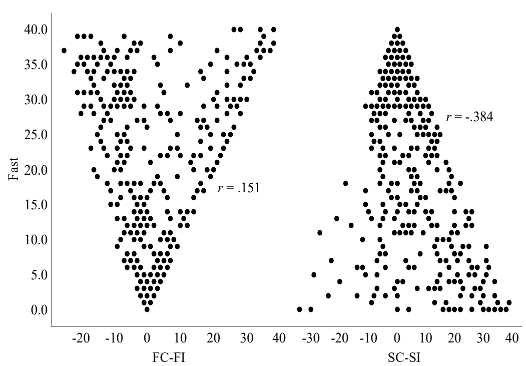
Figure 7 shows the scatterplots of *Fast* vs. *FC – FI* and *Fast* vs. *SC-SI* with a very distinctive equilateral triangular shape that looks like an inverted pyramid. For a given

[2] Given the heteroscedasticity and nonlinearity, the correlation coefficient is not a good measure of association for many of the scatterplots in Figures 2-4. We have reported the correlation coefficients, however, because they are the most commonly used measure of association in psychology, and we suspect that readers would wonder about the correlation.most commonly used measure of association in psychology, and we suspect that readers would wonder about the correlation.

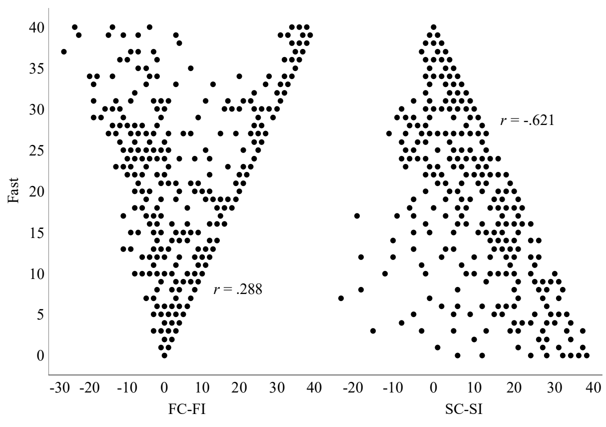
value of *Fast*, the theoretical range of *FC - FI* is -*Fast* to *Fast:* that is, -*Fast* < *FC – FI* < *Fast*. Both the theoretical upper and lower limits of *FC – FI* vary as a function of *Fast*. As *Fast* increases, there is very little change in the expected value of *FC –FI*, as evidenced by the correlation of .151, .288 and .420, but the range of *FC –FI* increases thereby giving the plot its distinctive shape. The scatterplots of *Fast* vs. *SC-SI* also have distinctive pyramidal shapes resulting from the fact that as *Fast* increases, the range of *SC – SI* decreases: (*Fast - 40*) < *SC – SI* < (40 – *Fast*). The correlations between *Fast* and *SC-SI are* negatively correlated, which are -.384, -.621 and -.541.

Figure 7. Raw Score Scatterplots of *FC-FI* and *SC-SI* against *Fast*

Form 3.1




Form 4.1



Form 5.1

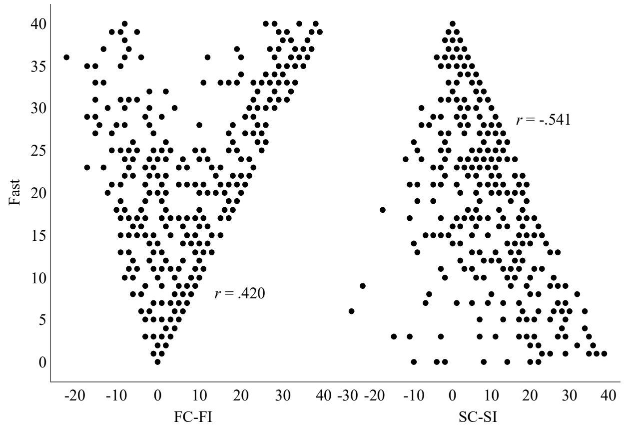
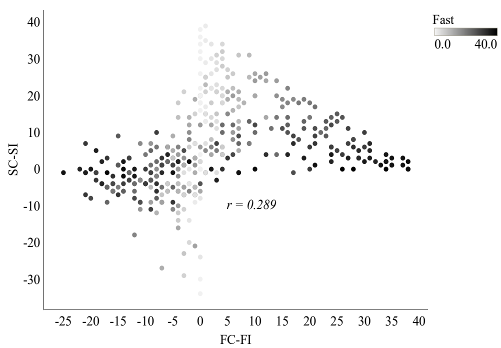



Figure 8 contains heat diagrams with the shading of points reflecting the number of fast responses given by the students for *FC –FI* vs. *SC – SI*. Students giving a large number of fast responses (dark shading) tend to be clustered at the two ends of the horizontal axis. At the right end of the axis are students who gave predominantly fast responses, most of which were correct. To the left along this axis are students who gave predominantly fast responses, most of which were incorrect. These students on the far left may be what have been variously labeled as rapid guessers or careless responders (Guo, Rios, Haberman, Liu, Wang, & Peak, 2019; Rios, Guo, Mao, & Liu, 2017; Wise & DeMars, 2006). In the middle are students with an approximately equal number of fast correct and fast incorrect responses. At the ends of this dimension, there are two somewhat fuzzy clusters, one of correct fast responders and one of incorrect fast responders.

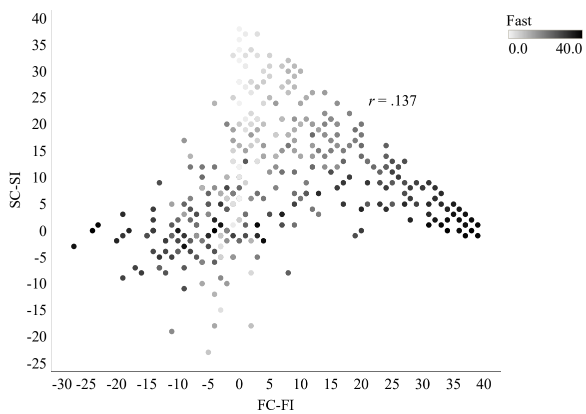
Along the vertical axis of the *FC –FI* vs. *SC – SI* graph are students whose responses were predominantly slow. Those at the top are students who have predominantly slow responses, most of which are correct. At the bottom are students whose responses are predominantly slow, but incorrect. In the middle are students whose responses display an approximately equal number of slow correct and slow incorrect responses.

Figure 8. Raw Score scatterplots of *FC-FI* against *SC-SI*

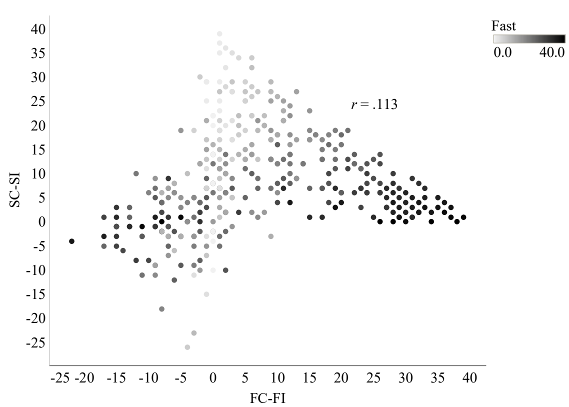
Form 3.1




Form 4.1




Form 5.1



Of particular interest are the correlations among raw scores that could represent fast and slow accuracy. Taking *FCF* and *SCS* as measures of fast and slow accuracy, the correlation between the two raw scores in the scatterplots of Figure 6 are .557, .512 and .639. Taking the raw scores *FC* and *SC* as measures of fast and slow accuracy, the correlation between the two raw scores in the scatterplots of Figure 4 are.-.454, -.554 and -.591. Taking raw scores *FC –FI* and *SC – SI* as measures of fast and slow accuracy, the correlation between the two raw scores in the scatterplots of Figure 8 are.289, .137 and .113. The raw score variables (*FC –FI*, *SC – SI*) would seem to provide stronger evidence for the existence of distinct fast and slow accuracy dimensions than would the raw scores (FC, SC) and (*FCF, SCS*).

*4.3.2 Unidimensional IRT Models*

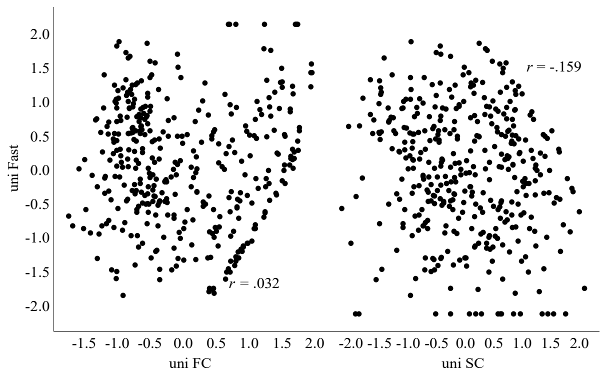
In the analyses for which results are reported in Table 15-17, we fit unidimensional models to the following seven item response variables: the dichotomous fast/slow variable $x_{1j}$, called uni Fast in the table, the dichotomous fast/correct variable $x_{2j}$, with missing data (uni FC), the dichotomous slow-correct variable $x_{3j}$ with missing data (uni SC), the dichotomous fast-correct variable $x'_{2j}$ with no missing data (uni FCNO), the dichotomous slow-correct variable $x'_{3j}$ with no missing data (uni SCNO), the polytomous fast-correct variable $y_{2j}$ (uni PFC), and the polyomous slow-correct variable $y_{3j}$ (uni PSC). We fit the 2PL model, Equations 20-22, to the dichotomous variables, and the graded response model, Equations 30-31, to the polytomous variables.

The correlations in Table 15-17 shows that each of the dimensions is almost identical with one and only one of the raw scores discussed earlier. By “almost identical,” we mean it has a correlation of .90 or greater with one and only one of the total scores. Uni Fast has a correlation of .979, .978 and .983 with total score *Fast* for Form 3.1, Form 4.1 and Form 5.1 respectively. These correlations suggest that the dimension underlying $x_{1j}$ is a speed dimension. Uni FC are most highly correlated with *FCF* ($r$ = .962, .963 and .948) and uni SC are highly correlated with *SCS* ($r$ = .947, .917 and .936) suggesting that the dimension can be interpreted as reflecting the proportion of correct fast and slow responses. Variables uni FCNO and uni SCNO are most highly correlated with *FC* and *SC*, correlations above .97, and both are more highly correlated with overall speed *Fast* than with accuracy. Finally, uni PFC and uni PSC are most highly correlated with *FC* – *FI* and *SC* – *SI* respectively, suggesting that they reflect net fast- and net slow-accuracy respectively. The dimensions underlying uni FC, uni SC, uni PFC, and uni PSC are more highly correlated with overall accuracy *Correct* than with overall speed *Fast*.

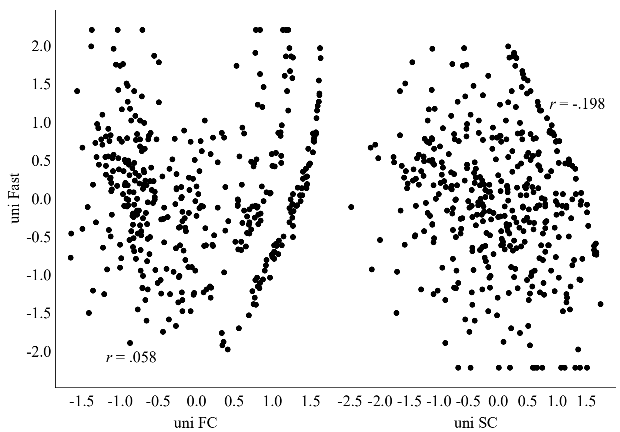
The figures below show the scatterplots for unidimensional $\theta$ values. The scatter-plots and correlations below parallel the shapes of the scatterplots and the correlations of Figure 5. In essence, the shapes of scatterplots and correlations involving the raw scores *Fast*, *FCF*, and *SCS* are similar to those for the uni Fast, uni FC and uni SC the latter two of which are the dichotomous variables with large amounts of missing data. The similar-ity of the scatterplots and correlations reflects the fact that uni Fast, uni FC and uni SC are highly correlated with the raw scores *Fast*, *FCF*, and *SCS* respectively.

Figure 9. Scatterplots of Uni FC and Uni SC against Uni Fast

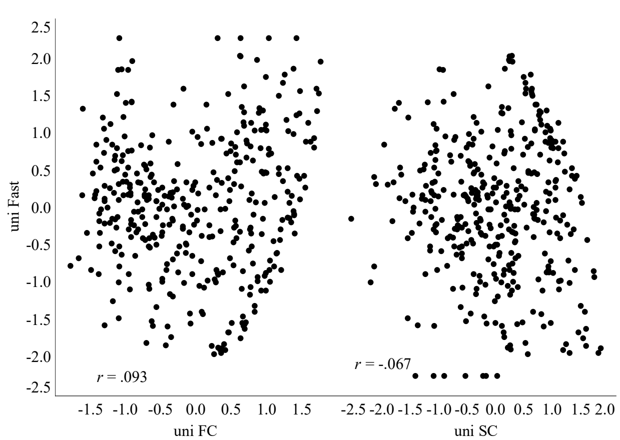
Form 3.1




Form 4.1



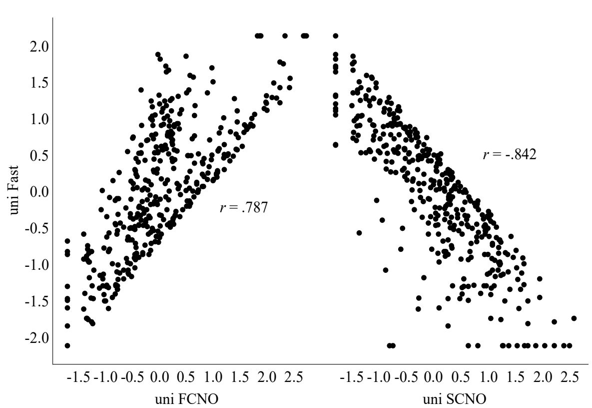

Form 5.1



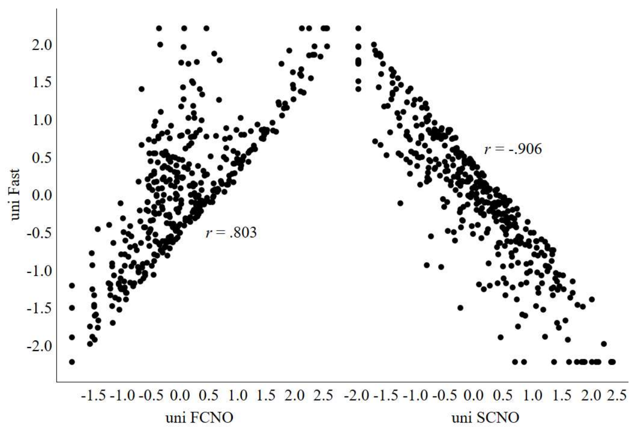
In a similar fashion, the scatterplots and correlations in Figure 10 are similar to the shapes of the scatterplots and the correlations in Figure 4. These scatterplots are imperfect triangles similar to those in Figure 5. For instance, the plots of uni FCNO against uni Fast is a triangle similar to that of *Fast* vs. *FC* in Figure 5. The hypotenuse of the right triangle is clearly visible for *Fast* vs. *FC*, although the upper left portion of the right triangle is not well filled-in with actual points. Similarly, the plot of uni Fast vs. uni SCNO in Figure 10 are also right triangles similar to *Fast* vs *SC* in Figure 5, although the lower left portion of the triangle is not well filled with actual data points. In essence, the shapes of scatterplots and correlations for the raw scores *Fast*, *FC*, and *SC* are similar to those for the dimensions underlying $x_{1j}$, $x'_{2j}$, and $x'_{3j}$. These similarities reflect the fact that the dimensions underlying $x_{1j}$, $x'_{2j}$, and $x'_{3j}$ are so highly correlated with raw scores *Fast* ($r$ = .979, .978, .983), *FC* (.969, .978, .980), and *SC* (.973, .978, .982) respectively.

Figure 10. Scatterplots of uni FCNO and uni SCNO against uni Fast

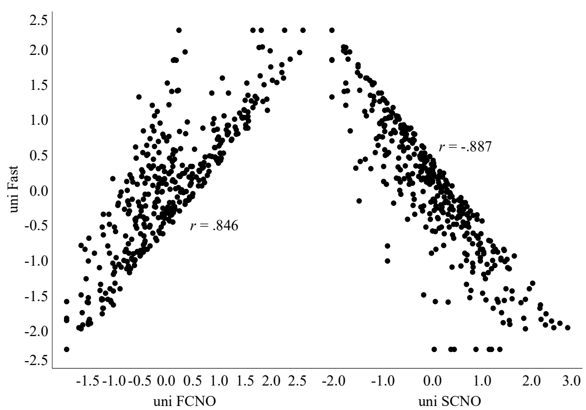
Form 3.1



Form 4.1



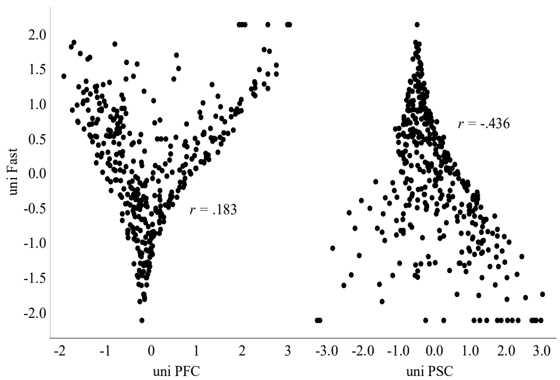

Form 5.1



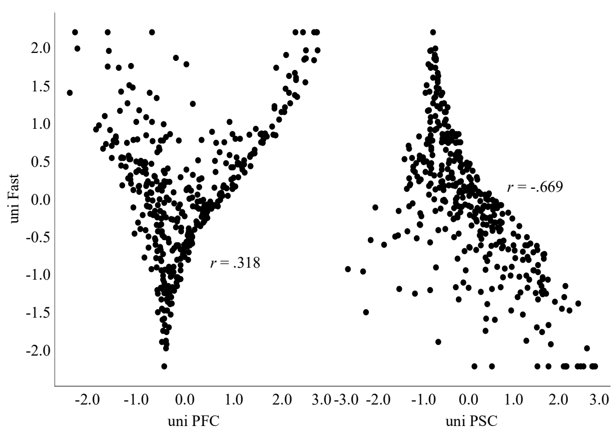

Finally, the scatter plots and correlations in Figure 7 parallel the shapes of the scatterplots and the correlations below. These scatterplots are imperfect triangles (or pyramids) similar to those in Figure 7. In essence, the shapes of scatter plots and correlations for the raw scores *Fast, FC –FI*, and *SC – SI* are similar to those for the dimensions uni Fast, uni PFC and uni PSC These similarities reflect the fact that the dimensions underlying uni Fast, uni PFC and uni PSC are so highly correlated with *Fast, FC –FI* (*r* = .991, .990 and .989), and *SC – SI* (*r* = .985, .987 and .976).

Figure 11. Scatterplots of uni PFC and uni PSC against uni Fast

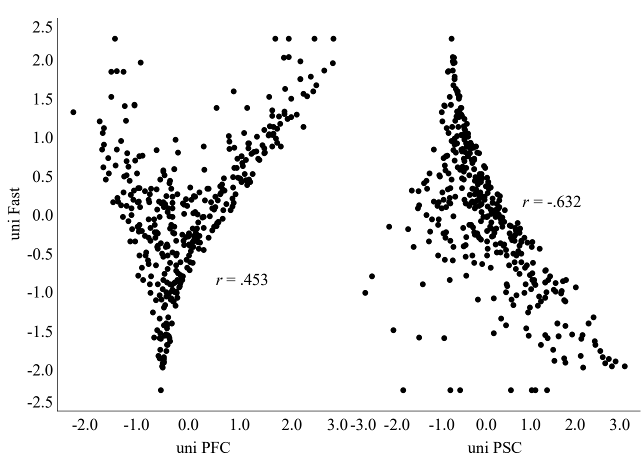
Form 3.1




Form 4.1




Form 5.1

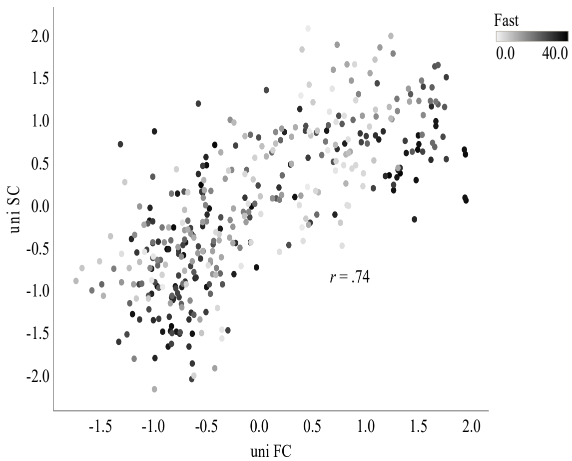


Taking the dimensions underlying uni FC and uni SC as measures of fast and slow accuracy, the correlations between the theta estimates in the scatterplots for all three forms are above .7. Taking the dimensions uni FCNO and uni SCNO as measures of fast and slow accuracy, the correlation between the theta estimates in the scatterplots range from -.491 to -.625 for the three forms Taking the dimensions uni PFC and uni PSC as measures of fast and slow accuracy, the correlation between the theta estimates in the scatterplots range from -.041 to .169. The dimensions underlying the latter set of variables, uni PFC and uni PSC, would seem to provide stronger evidence for the existence of distinct fast and slow accuracy dimensions than would the dimensions uni FC and uni SC or the dimensions uni FCNO and uni SCNO,.

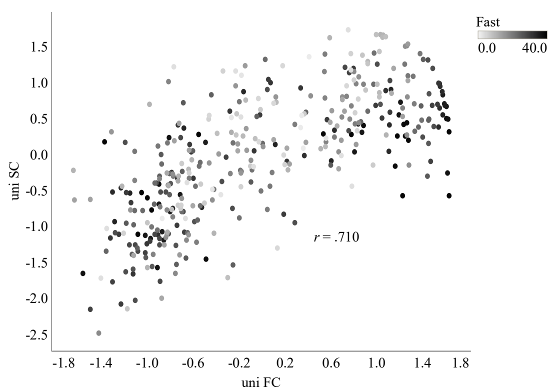
The correlations in Table 15-17 and the similar scatterplots between raw scores and unidimensional $\theta$ values suggest that the various dimensions derived from our unidimensional models all correspond, at least approximately, to one of the raw score dimensions. The intercorrelations of the unidimensional $\theta$ values parallel the correlations among their corresponding raw scores.

Figure 12. Scatterplots of uni FC against uni SC

Form 3.1




Form 4.1




Form 5.1

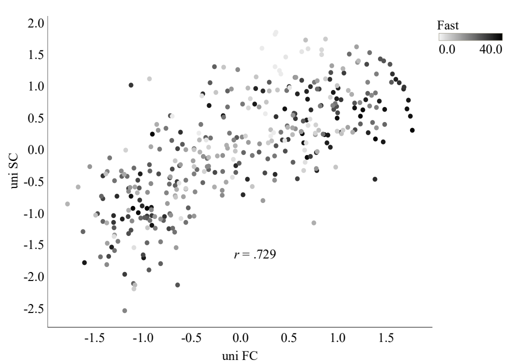
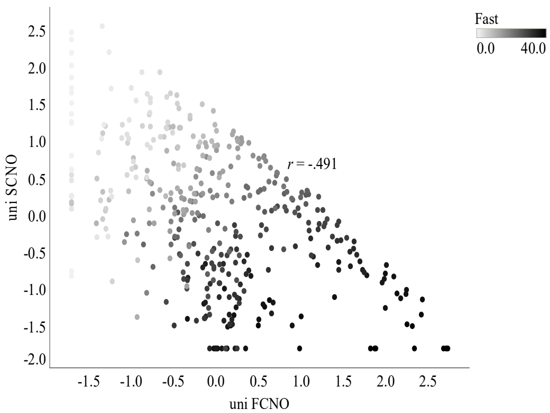
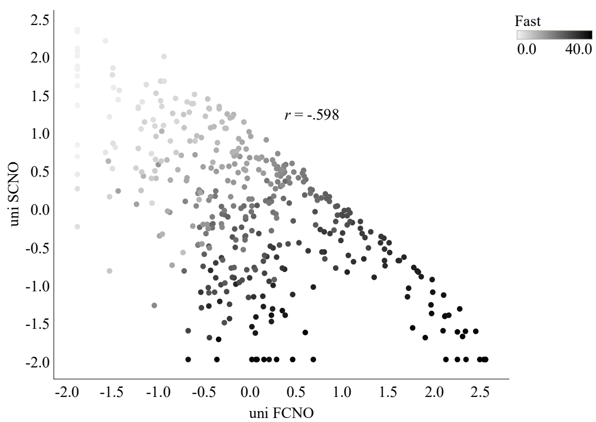


Figure 13. Scatterplots of uni FCNO against uni SCNO

Form 3.1




Form 4.1




Form 5.1

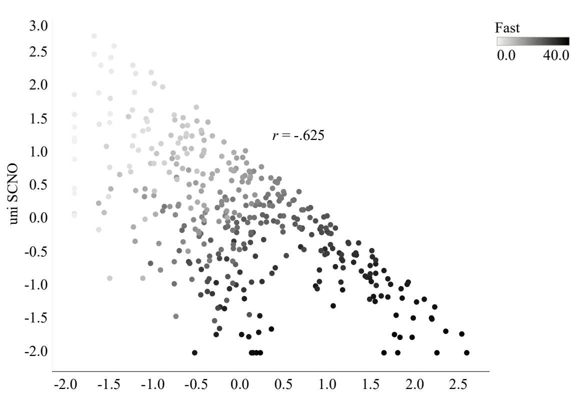
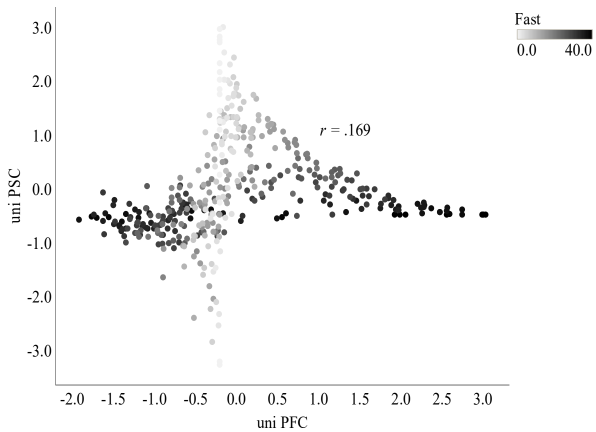
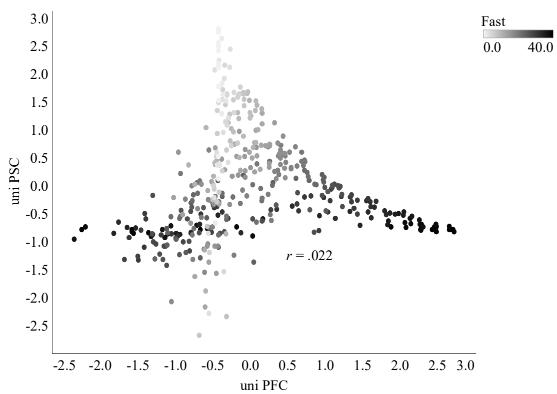


Figure 14. Scatterplots of uni PFC against uni PSC

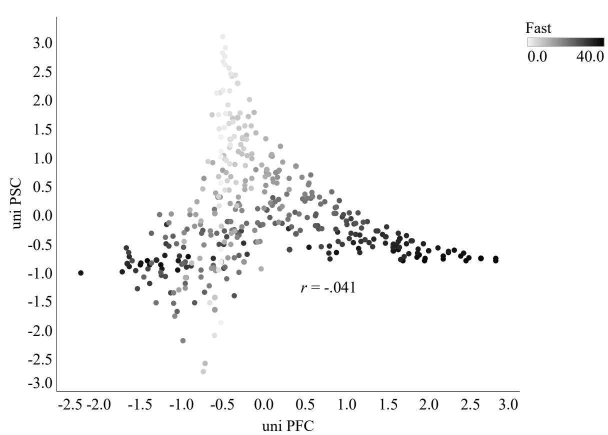
Form 3.1




Form 4.1




Form 5.1



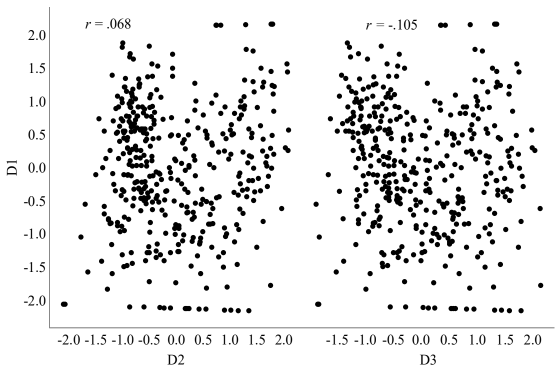
### *4.3.3 Three-Dimensional IRT Models*

The correlations of dimensions in our three-dimensional models and the raw scores are also shown in Table 15-17 for Form 3.1, Form 4.1 and Form 5.1. Correlations of the three dimensions underlying the response variables $(x_{1j}, x_{2j}, x_{3j})$ are named as D1, D2 and D3 in table 15-17. As expected, the dimension corresponding to $x_{1j}$, D1, correlated most highly with *Fast* ($r$ = .978, .978 and .983). Whereas, D2 and D3 did not correlate with *Fast*. D2, and D3 were both most highly correlated with overall accuracy *Correct,* with $r$ =.956 and .974 for Form 3.1, $r$ = .933 and .964 for Form 4.1, and $r$ = .947 and .963 for Form 5.1, not *FCF* and *SCS* as the unidimensional results had led us to expect. D1, D2 and D3 are the fast-correct and slow-correct variables of De Boeck and colleagues. The person parameters on these two dimensions are similar to those of De Boeck et al. in that the fast-accuracy and slow-accuracy dimensions in this three-dimensional model appear to be variations on a single overall accuracy dimension, rather than distinct fast- and slow-accuracy dimensions.

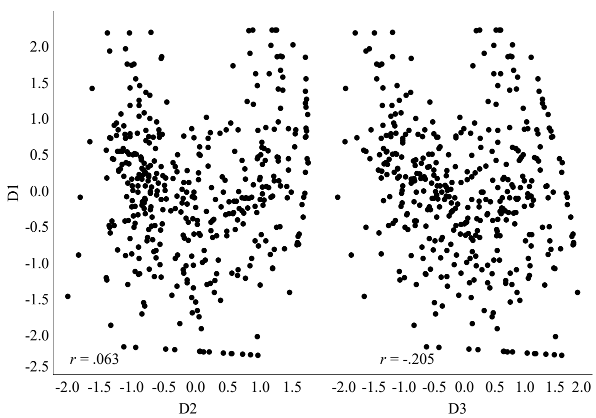
Figure 15 shows the scatterplots for multidimensional $\theta$ values. The scatterplots and correlations in Figure 15 parallel the shapes and low correlations of the scatterplots in Figure 5. The correlation of the speed dimension underlying *D1* is almost uncorrelated with the accuracy dimensions underlying *D2* and *D3*. However, Figure 16 shows fast and slow accuracy dimensions that are nearly identical with a correlation of .972, .944 and .968, which suggest that D2 and D3 seem like variations on an overall accuracy dimension rather than distinct speed accuracy dimensions.

Figure 15. Scatterplots of D2, D3 against D1

Form 3.1




Form 4.1




Form 5.1

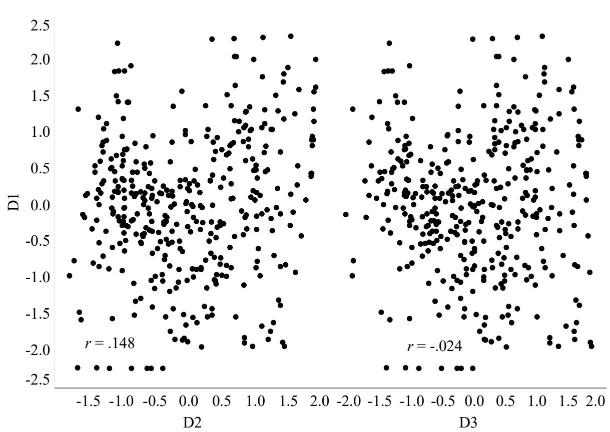
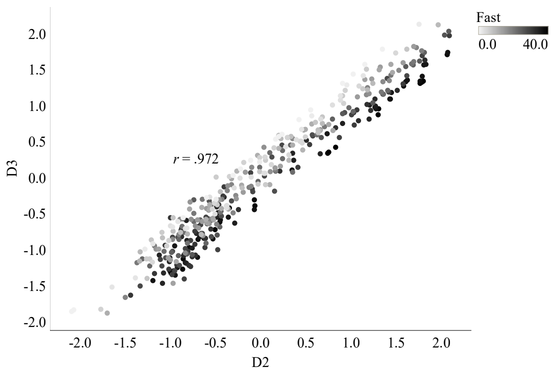
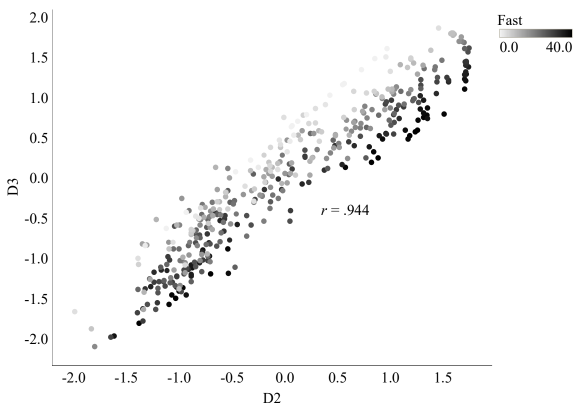


Figure 16. Scatterplots of D2 against D3

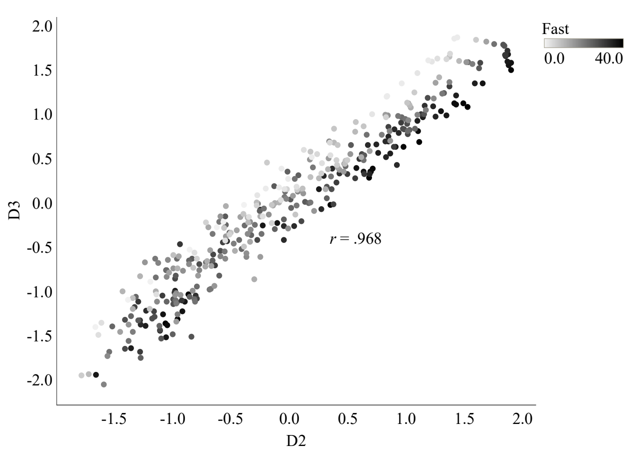
Form 3.1




Form 4.1




Form 5.1

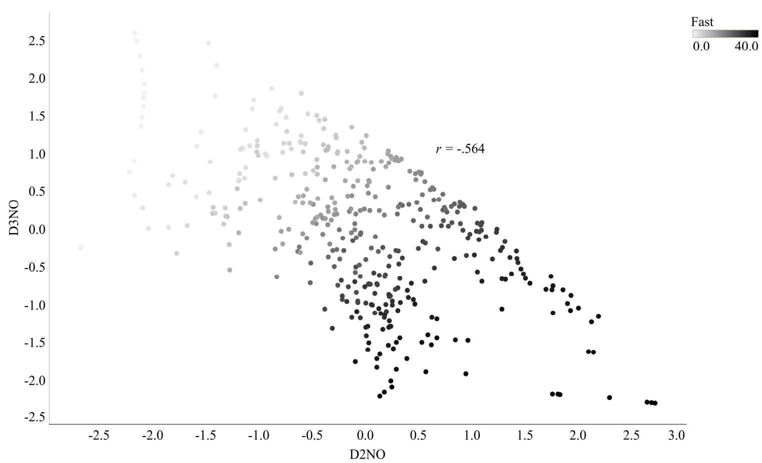


*D1NO, D2NO* and *D3NO* correspond to the response vectors $(x_{1j}, x'_{2j}, x'_{3j})$ in correlation tables 15-17. These behaved like the corresponding unidimensional dimensions. Each was most highly correlated with a single raw score; *Fast* in the case of *D1NO* with $r$ = .960, .960 and .969, *FC* in the case of *D2NO* with $r$ = .924, .950 and .957, and *SC* in the case of *D3NO* with $r$ = .952, .950 and .963.

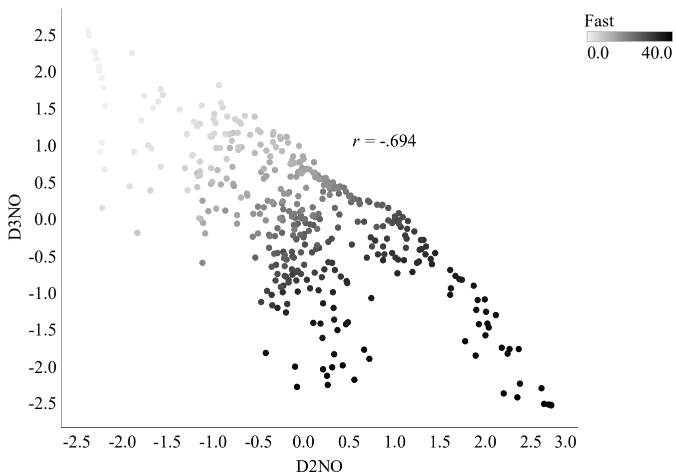
The scatter plots and correlations in Figure 17 parallel the shapes of the scatter-plots and the correlations in Figure 4. For instance, the plot of *D2NO* against *D3NO* bears some resemblance to a right triangle. The hypotenuse is clearly visible but the upper left portion is poorly filled with actual points

Figure 17. Scatterplots of D2NO against D3NO

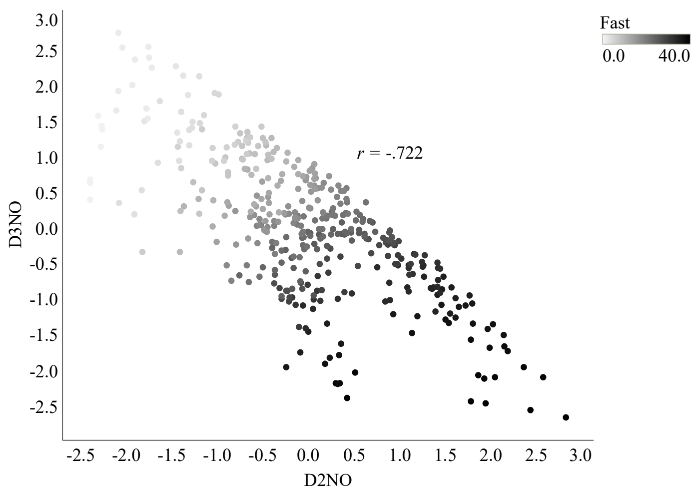
Form 3.1



Form 4.1



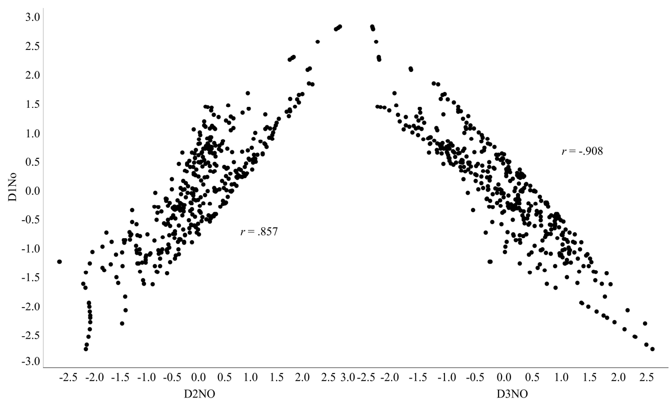

Form 5.1



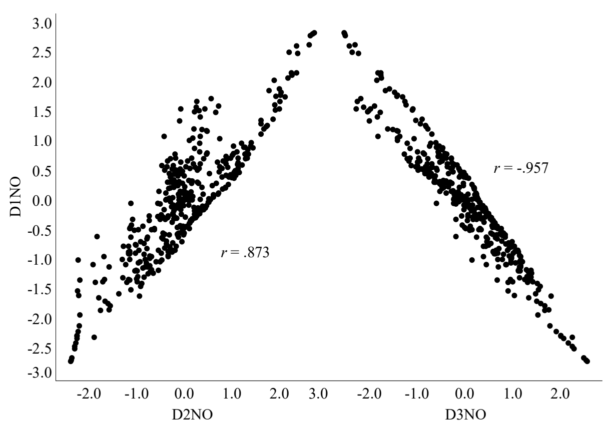

The correlations of the speed dimension, *D1NO,* are very high in absolute value with fast accuracy, *D2NO,* (.857, .873 and .905) and slow accuracy, *D3NO,* (-.908, -.949 and -.948). In essence, the shapes of scatterplots and correlations for the raw scores *Fast*, *FC*, and *SC* are crudely similar to those for the multidimensional dimensions *D1NO, D2NO* and *D3NO.* These similarities reflect the fact that *D1NO, D2NO* and *D3NO* are so highly correlated with *Fast* (.978, .960 and .969), *FC* (.924, .950 and .957), and *SC* (.952, .950 and .963) respectively.

Figure 18. Scatterplots of D2NO and D3NO against D1NO

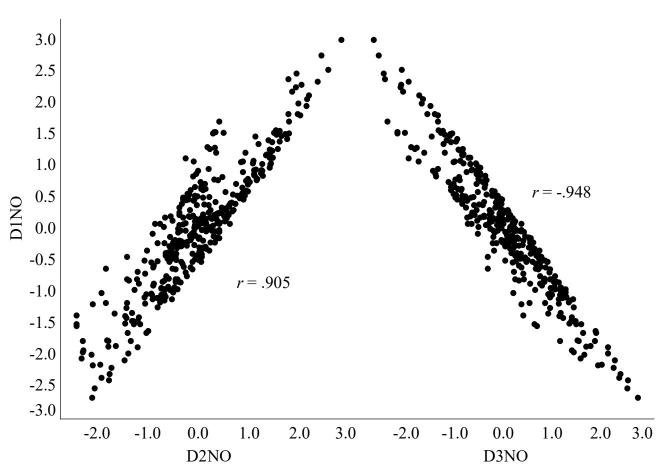
Form 3.1




Form 4.1




Form 5.1

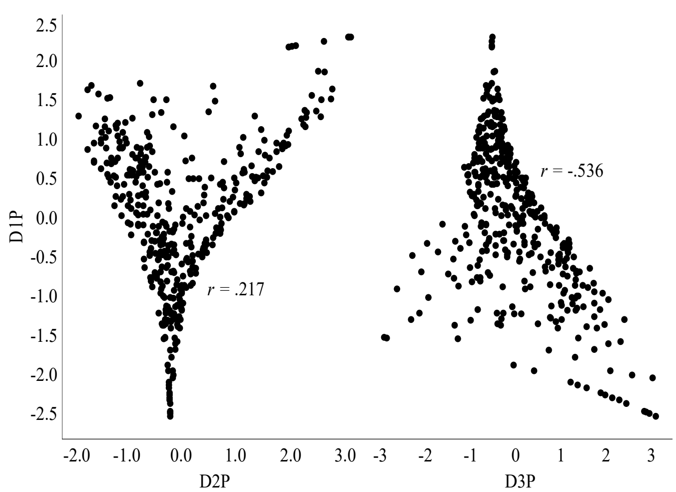


The dimensions for the last three variables ($x_{1j}, y_{2j}, y_{3j}$) are named as *D1P, D2P* and *D3P* in table 15-17. Each was most highly correlated with a single raw score; *Fast* in the case of *D1P* with *r* = .973, .971 and .977; *FC* - *FI* in the case of $y_{2j}$ with *r* = .991, .989 and .986; *SC* - *SI* in the case of $y_{3j}$ with *r* = .981, .973 and .956.

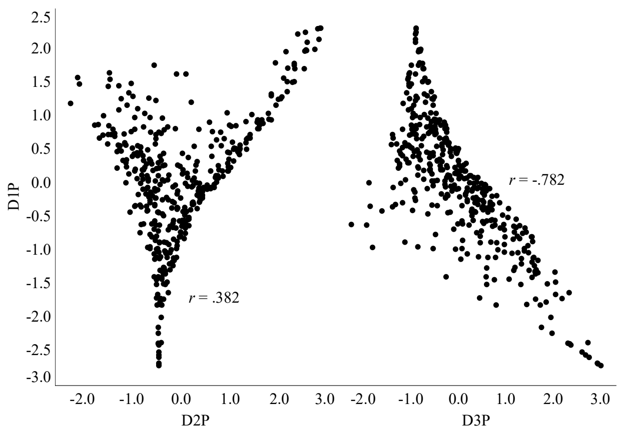
The scatterplots and correlations in Figure 7 parallel the shapes of the scatterplots and the correlations in Figure 18. These scatterplots are imperfect triangles (or pyramids) similar to those in Figure 18. In essence, the shapes of scatter plots and correlations for the raw scores *Fast, FC –FI*, and *SC – SI* are similar to those for the multidimensional dimensions *D1P, D2P* and *D3P.* These similarities reflect the fact that the dimensions *D1P, D2P* and *D3P* are so highly correlated with *Fast*, *FC –FI* , and *SC – SI* respectively.

Figure 19. Scatterplots of D2P and D3P against D1P

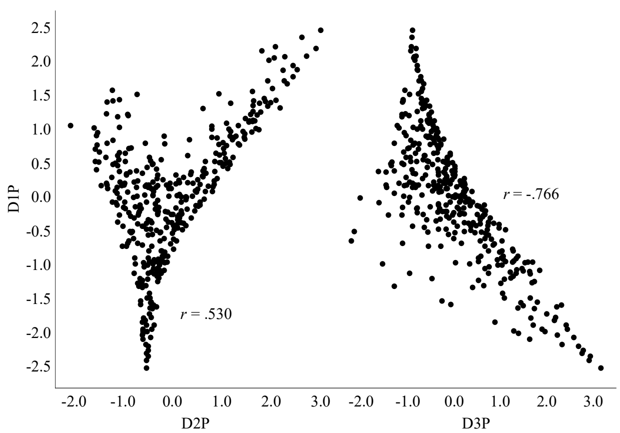
Form 3.1



Form 4.1




Form 5.1

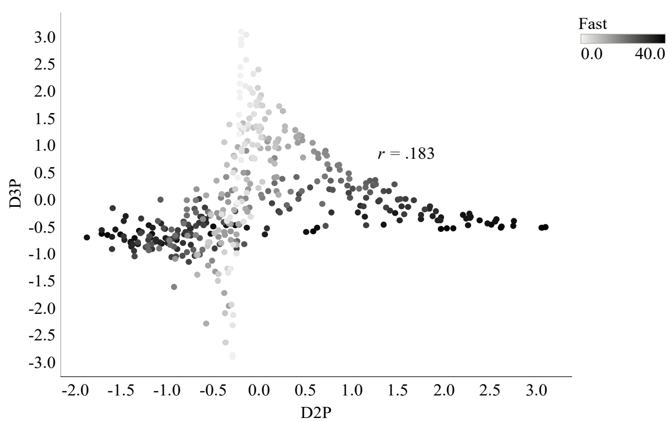



This study investigated the correlations of the dimensions that could represent fast and slow accuracy. Taking the dimensions *D2* and *D3* as measures of fast and slow accuracy, the correlations between the theta estimates in the scatterplots of Figure 13 are .972, .944 and .968 lending little if any support to the hypothesis of distinct fast and slow accuracy. Taking the dimensions *D2NO* and *D3NO* as measures of fast and slow accuracy, the correlations between the theta estimates in the scatterplots of Figure 17 are-.564, -.694 and -.772. Taking the dimensions *D2P* and *D3P* as measures of fast and slow accuracy, the correlations between the theta estimates in the scatter plots of Figure

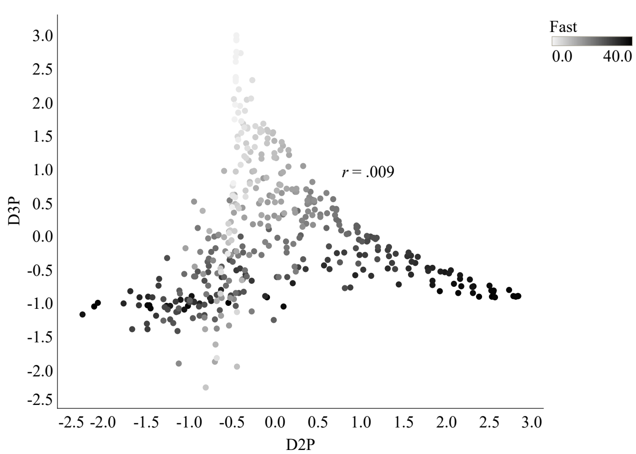
19 are .183, .009 and -.071. The dimensions of *D2NO, D3NO* and *D2P, D3P* would seem to provide stronger evidence for the existence of distinct fast and slow accuracy dimensions than would the dimensions *D2* and *D3*. However, the correlations of the fast and slow accuracy dimensions *D2NO* and *D3NO* with speed, *Fast*, were so high ($r_{D2NO}$ = .814, .829, .875 and $r_{D3NO}$ = -.879, -.924 and -.918) that they cast doubt on whether fast and slow accuracy are distinct from speed and whether these are more speed than accuracy dimensions.

Figure 20. Scatterplots of D2P against D3P

Form 3.1




Form 4.1



Form 5.1

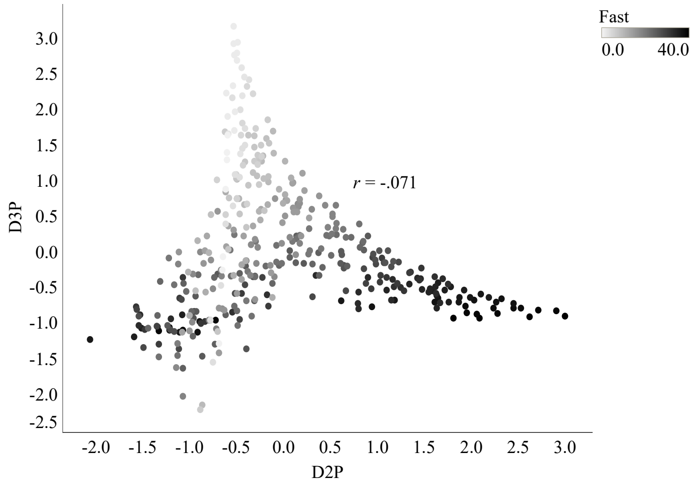

Fast
0.0
40.0
r = -.071
D3P
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
D2P

Table 15. Form 3.1 Correlations between Raw Scores and Theta Values

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.12 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.12 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.51 | 0.72 | -0.72 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.53 | -0.83 | 0.83 | -0.45 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.77 | 0.59 | -0.59 | -0.14 | -0.66 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.58 | -0.56 | 0.56 | -0.62 | 0.01 | -0.08 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.83 | 0.15 | -0.15 | 0.80 | 0.08 | -0.71 | -0.40 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.77 | -0.38 | 0.38 | -0.03 | 0.83 | -0.51 | -0.55 | 0.29 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.81 | 0.08 | -0.08 | 0.60 | 0.25 | -0.59 | -0.51 | 0.79 | 0.49 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.78 | -0.22 | 0.22 | 0.22 | 0.59 | -0.57 | -0.50 | 0.50 | 0.78 | 0.56 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.11 | 0.98 | -0.98 | 0.71 | -0.82 | 0.57 | -0.55 | 0.16 | -0.38 | 0.12 | -0.24 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.91 | 0.03 | -0.03 | 0.67 | 0.31 | -0.73 | -0.51 | 0.90 | 0.55 | 0.96 | 0.61 | 0.03 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.90 | -0.16 | 0.16 | 0.34 | 0.61 | -0.59 | -0.63 | 0.62 | 0.85 | 0.66 | 0.95 | -0.16 | 0.74 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.45 | 0.77 | -0.77 | 0.97 | -0.49 | -0.02 | -0.67 | 0.70 | -0.04 | 0.62 | 0.19 | 0.79 | 0.64 | 0.31 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.51 | -0.84 | 0.84 | -0.45 | 0.97 | -0.68 | 0.07 | 0.09 | 0.78 | 0.22 | 0.64 | -0.84 | 0.28 | 0.62 | -0.49 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.80 | 0.17 | -0.17 | 0.80 | 0.04 | -0.69 | -0.37 | 0.99 | 0.24 | 0.75 | 0.45 | 0.18 | 0.86 | 0.58 | 0.71 | 0.05 | 1.00 | | | | | | | | | | |
| uni PSC | 0.71 | -0.44 | 0.44 | -0.12 | 0.85 | -0.49 | -0.49 | 0.22 | 0.99 | 0.41 | 0.71 | -0.44 | 0.49 | 0.79 | -0.12 | 0.79 | 0.17 | 1.00 | | | | | | | | | |
| D1 | -0.11 | 0.98 | -0.98 | 0.71 | -0.82 | 0.57 | -0.54 | 0.16 | -0.39 | 0.12 | -0.24 | 1.00 | 0.04 | -0.17 | 0.79 | -0.85 | 0.19 | -0.44 | 1.00 | | | | | | | | |
| D2 | 0.96 | 0.06 | -0.06 | 0.64 | 0.36 | -0.66 | -0.65 | 0.86 | 0.66 | 0.84 | 0.70 | 0.07 | 0.96 | 0.86 | 0.60 | 0.33 | 0.82 | 0.59 | 0.07 | 1.00 | | | | | | | |
| D3 | 0.97 | -0.11 | 0.11 | 0.49 | 0.53 | -0.72 | -0.60 | 0.79 | 0.77 | 0.78 | 0.80 | -0.10 | 0.90 | 0.93 | 0.44 | 0.51 | 0.75 | 0.71 | -0.11 | 0.97 | 1.00 | | | | | | |
| D1NO | -0.10 | 0.96 | -0.96 | 0.76 | -0.85 | 0.48 | -0.46 | 0.25 | -0.45 | 0.16 | -0.29 | 0.98 | 0.11 | -0.21 | 0.82 | -0.87 | 0.28 | -0.51 | 0.98 | 0.09 | -0.10 | 1.00 | | | | | |
| D2NO | 0.39 | 0.81 | -0.81 | 0.92 | -0.51 | 0.09 | -0.71 | 0.61 | -0.03 | 0.57 | 0.16 | 0.84 | 0.57 | 0.28 | 0.98 | -0.52 | 0.61 | -0.12 | 0.84 | 0.54 | 0.38 | 0.86 | 1.00 | | | | |
| D3NO | 0.47 | -0.88 | 0.88 | -0.47 | 0.95 | -0.70 | 0.17 | 0.09 | 0.70 | 0.20 | 0.59 | -0.89 | 0.27 | 0.56 | -0.52 | 0.99 | 0.05 | 0.73 | -0.89 | 0.30 | 0.47 | -0.91 | -0.56 | 1.00 | | | |
| D1P | -0.13 | 0.97 | -0.97 | 0.72 | -0.84 | 0.55 | -0.50 | 0.18 | -0.43 | 0.12 | -0.26 | 1.00 | 0.04 | -0.19 | 0.79 | -0.86 | 0.20 | -0.49 | 1.00 | 0.05 | -0.12 | 0.99 | 0.84 | -0.91 | 1.00 | | |
| D2P | 0.80 | 0.18 | -0.18 | 0.81 | 0.03 | -0.68 | -0.39 | 0.99 | 0.24 | 0.75 | 0.46 | 0.20 | 0.86 | 0.59 | 0.72 | 0.05 | 1.00 | 0.18 | 0.20 | 0.83 | 0.76 | 0.29 | 0.62 | 0.04 | 0.22 | 1.00 | |
| D3P | 0.71 | -0.48 | 0.48 | -0.14 | 0.88 | -0.53 | -0.44 | 0.22 | 0.98 | 0.41 | 0.71 | -0.49 | 0.50 | 0.79 | -0.16 | 0.82 | 0.18 | 1.00 | -0.49 | 0.59 | 0.71 | -0.56 | -0.16 | 0.76 | -0.54 | 0.18 | 1.00 |

Table 16. Form 4.1 Correlations between Raw Scores and Theta Values

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.16 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.16 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.47 | 0.75 | -0.75 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.48 | -0.90 | 0.90 | -0.55 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.86 | 0.48 | -0.48 | -0.22 | -0.59 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.62 | -0.42 | 0.42 | -0.57 | -0.03 | 0.14 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.81 | 0.29 | -0.29 | 0.85 | -0.08 | -0.71 | -0.49 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.70 | -0.62 | 0.62 | -0.25 | 0.90 | -0.59 | -0.45 | 0.14 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.78 | 0.12 | -0.12 | 0.62 | 0.12 | -0.65 | -0.52 | 0.81 | 0.33 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.77 | -0.27 | 0.27 | 0.17 | 0.55 | -0.63 | -0.52 | 0.47 | 0.72 | 0.51 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.16 | 0.98 | -0.98 | 0.74 | -0.88 | 0.46 | -0.40 | 0.29 | -0.61 | 0.17 | -0.30 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.91 | 0.05 | -0.05 | 0.66 | 0.23 | -0.79 | -0.55 | 0.89 | 0.45 | 0.96 | 0.59 | 0.06 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.89 | -0.20 | 0.20 | 0.27 | 0.57 | -0.65 | -0.73 | 0.56 | 0.82 | 0.60 | 0.92 | -0.20 | 0.71 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.39 | 0.79 | -0.79 | 0.98 | -0.60 | -0.13 | -0.57 | 0.78 | -0.29 | 0.63 | 0.11 | 0.80 | 0.63 | 0.22 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.46 | -0.90 | 0.90 | -0.55 | 0.98 | -0.60 | 0.03 | -0.07 | 0.86 | 0.11 | 0.60 | -0.91 | 0.21 | 0.57 | -0.60 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.77 | 0.32 | -0.32 | 0.86 | -0.13 | -0.68 | -0.46 | 0.99 | 0.08 | 0.76 | 0.41 | 0.32 | 0.85 | 0.51 | 0.79 | -0.12 | 1.00 | | | | | | | | | | |
| uni PSC | 0.64 | -0.67 | 0.67 | -0.33 | 0.92 | -0.56 | -0.38 | 0.07 | 0.99 | 0.24 | 0.66 | -0.67 | 0.39 | 0.75 | -0.38 | 0.88 | 0.02 | 1.00 | | | | | | | | | |
| D1 | -0.16 | 0.98 | -0.98 | 0.74 | -0.89 | 0.46 | -0.39 | 0.29 | -0.62 | 0.17 | -0.31 | 1.00 | 0.06 | -0.21 | 0.81 | -0.91 | 0.32 | -0.68 | 1.00 | | | | | | | | |
| D2 | 0.93 | 0.06 | -0.06 | 0.63 | 0.26 | -0.75 | -0.67 | 0.86 | 0.52 | 0.84 | 0.66 | 0.06 | 0.97 | 0.81 | 0.58 | 0.24 | 0.82 | 0.45 | 0.06 | 1.00 | | | | | | | |
| D3 | 0.96 | -0.20 | 0.20 | 0.39 | 0.52 | -0.81 | -0.63 | 0.72 | 0.74 | 0.73 | 0.80 | -0.20 | 0.89 | 0.93 | 0.32 | 0.51 | 0.68 | 0.68 | -0.21 | 0.94 | 1.00 | | | | | | |
| D1NO | -0.13 | 0.96 | -0.96 | 0.78 | -0.90 | 0.38 | -0.32 | 0.36 | -0.66 | 0.21 | -0.34 | 0.98 | 0.12 | -0.25 | 0.84 | -0.93 | 0.40 | -0.71 | 0.99 | 0.09 | -0.19 | 1.00 | | | | | |
| D2NO | 0.33 | 0.83 | -0.83 | 0.95 | -0.64 | -0.04 | -0.57 | 0.71 | -0.32 | 0.59 | 0.07 | 0.85 | 0.57 | 0.18 | 0.99 | -0.64 | 0.73 | -0.41 | 0.85 | 0.51 | 0.26 | 0.87 | 1.00 | | | | |
| D3NO | 0.39 | -0.92 | 0.92 | -0.59 | 0.95 | -0.58 | 0.14 | -0.11 | 0.79 | 0.05 | 0.54 | -0.95 | 0.16 | 0.48 | -0.65 | 0.99 | -0.16 | 0.81 | -0.95 | 0.17 | 0.44 | -0.96 | -0.69 | 1.00 | | | |
| D1P | -0.16 | 0.97 | -0.97 | 0.76 | -0.91 | 0.42 | -0.34 | 0.32 | -0.66 | 0.19 | -0.31 | 0.99 | 0.08 | -0.24 | 0.82 | -0.92 | 0.36 | -0.72 | 0.99 | 0.06 | -0.21 | 0.99 | 0.87 | -0.95 | 1.00 | | |
| D2P | 0.76 | 0.34 | -0.34 | 0.87 | -0.14 | -0.66 | -0.48 | 0.99 | 0.08 | 0.76 | 0.41 | 0.34 | 0.84 | 0.51 | 0.81 | -0.14 | 1.00 | 0.02 | 0.35 | 0.82 | 0.67 | 0.43 | 0.74 | -0.19 | 0.38 | 1.00 | |
| D3P | 0.62 | -0.73 | 0.73 | -0.37 | 0.95 | -0.60 | -0.29 | 0.06 | 0.97 | 0.21 | 0.64 | -0.74 | 0.39 | 0.72 | -0.43 | 0.91 | 0.02 | 0.99 | -0.74 | 0.42 | 0.66 | -0.77 | -0.47 | 0.86 | -0.78 | 0.01 | 1.00 |

Table 17. Form 5.1 Correlations between Raw Scores and Theta Values

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.02 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.02 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.51 | 0.81 | -0.81 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.40 | -0.88 | 0.88 | -0.59 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.82 | 0.42 | -0.42 | -0.19 | -0.56 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.69 | -0.50 | 0.50 | -0.63 | 0.03 | 0.14 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.80 | 0.42 | -0.42 | 0.87 | -0.18 | -0.65 | -0.56 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.68 | -0.54 | 0.54 | -0.22 | 0.87 | -0.57 | -0.46 | 0.11 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.87 | 0.11 | -0.11 | 0.56 | 0.21 | -0.68 | -0.63 | 0.78 | 0.49 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.82 | -0.09 | 0.09 | 0.31 | 0.44 | -0.63 | -0.62 | 0.55 | 0.69 | 0.64 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.02 | 0.98 | -0.98 | 0.80 | -0.87 | 0.41 | -0.49 | 0.41 | -0.53 | 0.13 | -0.11 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.91 | 0.09 | -0.09 | 0.62 | 0.20 | -0.79 | -0.59 | 0.87 | 0.45 | 0.95 | 0.65 | 0.09 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.90 | -0.06 | 0.06 | 0.35 | 0.48 | -0.64 | -0.75 | 0.60 | 0.78 | 0.73 | 0.94 | -0.07 | 0.73 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.46 | 0.84 | -0.84 | 0.98 | -0.62 | -0.11 | -0.64 | 0.82 | -0.24 | 0.56 | 0.27 | 0.85 | 0.58 | 0.31 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.37 | -0.88 | 0.88 | -0.60 | 0.98 | -0.56 | 0.07 | -0.18 | 0.84 | 0.20 | 0.46 | -0.89 | 0.17 | 0.47 | -0.63 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.76 | 0.45 | -0.45 | 0.88 | -0.23 | -0.61 | -0.54 | 0.99 | 0.06 | 0.73 | 0.52 | 0.45 | 0.82 | 0.56 | 0.83 | -0.24 | 1.00 | | | | | | | | | | |
| uni PSC | 0.59 | -0.64 | 0.64 | -0.34 | 0.92 | -0.54 | -0.33 | 0.01 | 0.98 | 0.40 | 0.60 | -0.63 | 0.36 | 0.69 | -0.37 | 0.88 | -0.04 | 1.00 | | | | | | | | | |
| D1 | -0.01 | 0.98 | -0.98 | 0.80 | -0.87 | 0.40 | -0.48 | 0.42 | -0.54 | 0.14 | -0.11 | 1.00 | 0.10 | -0.07 | 0.85 | -0.89 | 0.46 | -0.63 | 1.00 | | | | | | | | |
| D2 | 0.95 | 0.14 | -0.14 | 0.62 | 0.22 | -0.73 | -0.71 | 0.85 | 0.54 | 0.90 | 0.74 | 0.14 | 0.96 | 0.85 | 0.58 | 0.20 | 0.81 | 0.44 | 0.15 | 1.00 | | | | | | | |
| D3 | 0.96 | -0.02 | 0.02 | 0.47 | 0.40 | -0.76 | -0.69 | 0.74 | 0.69 | 0.84 | 0.83 | -0.03 | 0.89 | 0.94 | 0.42 | 0.39 | 0.70 | 0.60 | -0.02 | 0.97 | 1.00 | | | | | | |
| D1NO | -0.01 | 0.97 | -0.97 | 0.83 | -0.90 | 0.33 | -0.41 | 0.48 | -0.60 | 0.14 | -0.14 | 0.98 | 0.15 | -0.11 | 0.87 | -0.92 | 0.52 | -0.69 | 0.99 | 0.16 | -0.03 | 1.00 | | | | | |
| D2NO | 0.39 | 0.88 | -0.88 | 0.96 | -0.66 | -0.02 | -0.65 | 0.75 | -0.27 | 0.50 | 0.22 | 0.89 | 0.51 | 0.27 | 0.99 | -0.67 | 0.77 | -0.40 | 0.89 | 0.52 | 0.36 | 0.90 | 1.00 | | | | |
| D3NO | 0.30 | -0.92 | 0.92 | -0.64 | 0.96 | -0.55 | 0.17 | -0.22 | 0.77 | 0.14 | 0.39 | -0.93 | 0.13 | 0.38 | -0.67 | 0.99 | -0.28 | 0.83 | -0.93 | 0.13 | 0.31 | -0.95 | -0.72 | 1.00 | | | |
| D1P | -0.03 | 0.98 | -0.98 | 0.82 | -0.90 | 0.36 | -0.41 | 0.46 | -0.60 | 0.11 | -0.13 | 0.99 | 0.11 | -0.11 | 0.86 | -0.91 | 0.50 | -0.70 | 0.99 | 0.13 | -0.04 | 1.00 | 0.90 | -0.95 | 1.00 | | |
| D2P | 0.75 | 0.49 | -0.49 | 0.90 | -0.25 | -0.58 | -0.57 | 0.99 | 0.05 | 0.72 | 0.51 | 0.49 | 0.81 | 0.56 | 0.85 | -0.26 | 1.00 | -0.05 | 0.50 | 0.80 | 0.70 | 0.56 | 0.79 | -0.31 | 0.53 | 1.00 | |
| D3P | 0.57 | -0.71 | 0.71 | -0.39 | 0.95 | -0.59 | -0.23 | -0.01 | 0.96 | 0.37 | 0.57 | -0.71 | 0.35 | 0.65 | -0.43 | 0.92 | -0.06 | 0.99 | -0.71 | 0.41 | 0.57 | -0.76 | -0.47 | 0.88 | -0.77 | -0.07 | 1.00 |

## Chapter 5: Conclusion and Discussion

### 5.1 Limitations and Future Research

This section discusses the limitations of the current study and future research investigating multidimensional branching models.

As IRTPRO only reported -2LL, AIC and BIC for goodness of fit, this study only used these three statistics to compare different models. Moreover, models based on different response variables cannot be compared by using the AIC and BIC. Hence, the multidimensional 2PL models cannot be compared with the multidimensional polytomous models. Even though the simple polytomous scoring method provides different results from Partchev and De Boeck's (2012) study, we still cannot say that results based on the polytomous coding is better.

The large amount of missing values in Partchev and De Boeck (2012) approach are caused by the branching model, not by nonresponses. This study only included students who completed all 40 items. The missing values in dimension 2 and dimension 3 depend on the speed dimension, not because of omit responses. The polytomous scoring method used in this study may not apply to other types of missing values. In the future work, it is necessary to conduct simulation studies in evaluating the polytomous scoring method for its effectiveness in dealing with missing data.

Another limitation is the method of dichotomizing response times into fast and slow. Dichotomizing at the item median, presumes that for every item, half the people use a fast process and half use a slow process. It seems highly unlikely that the proportion using the fast and slow processes would be the same on any single item, let alone every item. At this point, however, there is no clearly better alternative method of classifying into fast and slow of which we are aware. Alternatives need to be explored in future research. The classification of response time is based solely on response speed. In most dual processing theories, the "fast" process is defined by characteristics in addition to speed. For instance, the "fast" process may consume less working memory. It may consume less conscious attention. It may take place in a different location of the brain. It may be more difficult to unlearn. In future, we think better ways of classifying responses will be based on more than just speed of response. MRI and ERP technology may provide a way to better classify responses by the process used to arrive at a solution. Our method of classifying fast and slow should be viewed as a "baby step" to move us a short distance toward dual processing theories of individual differences.

A fourth limitation, again involving our classification into fast and slow, is that it requires dichotomizing a continuous response time variable with an attendant loss of information. However, the loss of information from dichotomizing single item response times might be partially offset by combining the dichotomized response time information across items. We hypothesize that as the number of items increases, the estimate of a person's location on a speed dimension based on dichotomized response times becomes more similar to their location based on continuous response times. This is another question for future research.

Prior research has produced substantial evidence for the predictive validity of raw and IRT scores of intelligence, aptitude, and achievement assessments. Indirectly, this research supports the external validity of scores in this research.The scores *FC* and *SC* sum to the total accuracy score *Correct* when there is no missing data. Therefore, when there is no missing data, the equally weighted sum of *FC* and *SC* will have as much predictive validity as *Correct*, and an optimally weighted sum may have higher validity. Likewise, given no missing data, (*FC* – *FI*) + (*SC* – *SI*) = *C* – *I* and *C* – *I* will be perfectly correlated with the total score *Correct*. Therefore, the equally weighted sum of (*FC* – *FI*) + (*SC* – *SI*) will have predictive validity equal to that of *C* and an optimally weighted sum may have a higher validity. Similarly, *F***FCF* – *S***SCS* = *C*, and an optimally weighted sum of these two variables may have a higher validity than *Correct*. Taken together, fast correct and slow correct should be able to predict external variables as well as *Correct.* The question for future research is not whether fast- and slow accuracy possess validity, but rather whether an optimally weighted combination can out predict overall accuracy. In recent years, a number of methods have been developed for addressing the question of whether an optimal linear combination of variables can out predict a total score. These include multiple regression (Bulut & Desjardins, 2013; Davison & Davenport, 2002; Davison, Davenport, Chang, Vue & Su, 2015, Desjardins, 2012), structural equation modeling (Davison, Chang, & Davenport, 2014), and meta-analytic techniques (Wiernik, Wilmot, Davison, & Ones, in press).

Graded response model was used, rather than nominal response model. The fast accuracy and slow accuracy were coded in order. The polytomous coding $y_{2j}$ and, $y_{3j}$ were scored as 0 for incorrect responses and 2 for correct responses. Missing values in the dichotomous variables $x_{2j}$, and, $x_{3j}$ were scored as 1 in the polytomous model. At first, the nominal response model may seem appropriate for this data, but the dual-process theory posits that the different categories are functions of different underlying dimensions. The nominal model can incorporate multidimensionality between items, but not between response categories.

Lastly, in this study, students could take as much or as little time on an item as desired. This means that some students' responses were almost all fast, some were almost all slow, and most were somewhere in between. This means that, for the fast-accuracy dimension, the precision of scores varied substantially across students depending, in part, on the number of fast responses given by the student. Likewise, the precision of the slow-accuracy dimension varied substantially across students depending on the number of slow responses. This raises the question of whether it is reasonable and fair to compare the accuracy scores of students who vary substantially in their distribution of slow and fast responses.

Goldhammer (2015) has considered the possibility of placing time limits on items. In this scheme, each item would have its own time-limit, and students would receive credit only if they answered within that time-limit. Semmes (Davison, Semmes, Huang, & Close, 2012; Semmes, Davison, & Close, 2011) compared performance of students with and without item time-limits. If tight time limits are applied, then every student would be forced to give a fast response on every item. If the goal is to measure fast-accuracy performance, then this would ensure that every student had enough "fast" responses so as to yield a reliable fast-accuracy score for every student. In the testing scenario of this research, only some students had enough fast responses to yield a fast-accuracy score with a small standard error. Likewise, only some students had enough slow responses to yield a slow-accuracy score.

To measure fast-accuracy, it may be necessary to modify the standard testing scenario. One possible modification would be to apply item time limits. In high stakes situation, fairness would dictate that students be told that their performance would be judged based on a combination of their response accuracy and response time. If the dimension measured is a net fast-accuracy dimension, they would also need to be told that incorrect guesses would be counted against them.

### 5.2 Conclusion

The current study investigated multidimensional IRT analyses similar to those of Partchev and De Boeck's (2012) for measuring individual differences in dual processing. The goal was to determine if there was a way to derive separate dimensions of speed, fast accuracy and slow accuracy that are interpretable and distinct from each other. Raw score, separate unidimensional, and multidimensional approaches were compared. Reaching the goal required, among other things, addressing the large amount of missing data in the approach of Partchev and De Boeck (2012)

This study developed a polytomous scoring method for dealing with missing data in the branching model. In the prior literature on response time and response accuracy, researchers have proposed unidimensional models where scores along the dimension depend on both response time and accuracy. More commonly, researchers have considered two dimensional models with one dimension reflecting accuracy and a second dimension reflecting response time (Davison, Huang, Semmes, & Close, 2012; Semmes, Davison, & Close, 2011; van der Linden, 2009). Molenaar, Tuerlinckz, & van der Maas (2015) contains an excellent review of these models. This study departs from this prior literature in three major respects. First response times are dichotomized. Second, the research posits that differences in response time are, in part, a result of qualitative individual difference in cognitive processes used to solve problems. Third, our models contain three dimensions, not two or one. The dimensional dichotomy between fast- and slow-accuracy follows from the assumption of two underlying cognitive processes.

In order to compare the results of the impact of f missing values, multidimensional polytomous models were also fitted. Chapter 1 stated the research questions to be addressed in this study: how well will model fitting techniques work with large amounts of missing values? How well will this new approach perform in estimating distinct person parameters on speed and accuracy dimensions? This chapter summarizes the results presented in Chapter 4 and discusses the limitations of the current study.

Based on the results presented in Chapter 4, the following conclusions can be drawn. First, the current study agrees with Partchev and De Boeck's (2012) study on two- versus three-dimensional model comparison. For both the 2PL and polytomous models, the lower AIC and BIC indicates that three dimensional IRT models fit better than two dimensional IRT models. Second, the large amount of missing values in the branching model may have an impact on theta estimations. There is a strong positive correlation between the fast accuracy dimension and the slow accuracy dimension in Partchev and De Boeck's (2012) study. However, the theta correlations of the three dimensions from three dimensional polytomous models are different from their study. In the current study, the theta correlations between dimension 2 and dimension 3 are less than .2 in absolute values for the three dimensional polytomous model. This study presented correlations between raw scores and theta values to provide validity evidence useful in interpreting the dimensions in our polytomous model. The first dimension is a speed dimension highly correlated with the number of fast responses by a student. The fast accuracy and slow accuracy dimensions are highly correlated with students' net number of fast correct and net number of slow correct responses. These seem to be fast- and slow-accuracy dimensions that place an emphasis on a clean performance, one with few errors.

Regarding the responses with dichotomous coding and with some responses as missing, $x_{2j}$ and, $x_{3j}$, the 2PL models of our analysis recovered unidimensional person parameters along fast accuracy and slow accuracy dimensions that were most highly correlated with percent fast correct, *FCF,* and percent slow correct, *SCS*. The two dimensions were substantially correlated, as were *FCF* and *SCS*, which provided some evidence for the distinctness of the two dimensions. In the three-dimensional model for $x_{1j}$, $x_{2j}$, and, $x_{3j}$, the fast-accuracy and slow-accuracy dimensions (person parameters) were so highly correlated that they did not support an hypothesis of two distinct accuracy dimensions. In the size of correlations between the fast- and slow-dimensions, these three-dimensional results were very similar to those of De Boeck. This indicates that $x_{2j}$ and, $x_{3j}$ or percent fast-correct and percent slow-correct provide little evidence to support an hypothesis of distinct dimensions.

The second way of dichotomously coding dimensions without missing values are $x'_{2j}$ and $x'_{3j}$ which leads to dimensions corresponding to number of fast correct *FC* and number of slow correct *SC* responses. When we fitted a 2PL model to these response codings, the unidimensional person parameters of the fast-accuracy and slow-accuracy dimension had correlations over .95 with the *FC* and *SC* raw scores. *FC* and *SC* had relatively small correlations as did the dimensions underlying $x'_{2j}$ and, $x'_{3j}$. These results support the hypothesis of two distinct fast- and slow-accuracy dimensions. Results of fitting the three dimensional model based on $x_{1j}$, $x'_{2j}$, and, $x'_{3j}$ also supported an hypothesis of distinct fast- and slow-correct dimensions. However, the raw scores *FC* and *SC* as well as both the uni- and multidimensional dimensions were more highly correlated with the overall fast variable *Fast* than the overall accuracy raw score *Correct*, suggesting that these dimensions may reflect speed of response more than accuracy of response. We conclude that $x'_{2j}$ and $x'_{3j}$ support the distinctness of the corresponding raw scores and fast- and slow-accuracy dimensions, but the interpretation of these dimensions as speed, rather than accuracy dimensions, is in some doubt.

Lastly, the polytomous coding $y_{2j}$ and, $y_{3j}$ were fitted with a graded response model. The unidimensional person parameters of the fast-accuracy and slow-accuracy dimension had correlations over .95 with the raw scores *FC - FI* and *SC – SI*. Raw scores *FC - FI* and *SC - SI* had relatively small correlations with each other as did the dimensions underlying $y_{2j}$ and $y_{3j}$. These results support the hypothesis of two distinct fast- and slow-accuracy dimensions. Results of fitting our three dimensional model based on $x_{1j}$, $y_{2j}$, and, $y_{3j}$ also supported an hypothesis of distinct fast- and slow-correct dimensions. The raw scores *FC - FI* and *SC – SI* as well as the dimensions underlying $y_{2j}$ and, $y_{3j}$ were more highly correlated with overall accuracy *C* than with overall speed *Fast*, supporting the interpretation of these dimensions as fundamentally accuracy concepts. We conclude that the polytomous coding $y_{2j}$ and, $y_{3j}$ yielded dimensions providing the strongest evidence for distinct fast- and slow-accuracy dimensions with the clearest accuracy interpretation. However, these dimensions represent a very specific concept of accuracy such that high scores require a strong performance, meaning a large number of correct responses, as well as a clean performance, meaning very few incorrect responses.

## Reference

Akaike, H. (1974). A new look at the statistic model identification. *IEEE Transactions on Automatic Control, 19*, 716-723.

Ake, C. F. (2005, April). Rounding after multiple imputation with non-binary categorical covariates. In *annual meeting of the SAS Users Group International, Philadelphia, PA*.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, *2*(4), 89-195.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological bulletin*, *130*(4), 553.

Bernaards, C. A. & Sijtsma, K. (2000). Influence of imputation and EM methods on factor analysis when item nonresponse in questionnaire data is nonignorable. *Multivariate Behavioral Research*, *35*, 321-364.

Bernheim, B. D., & Rangel, A. (2004). Addiction and cue-triggered decision processes. *American economic review*, *94*(5), 1558-1590.

Betsch, T. (2008). The nature of intuition and its neglect in research on judgment and decision making. *Intuition in judgment and decision making*, 3-22.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, *39*(5), 752.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. Guilford Press.

Davison, M. L., Biancarosa, G., Carlson, S. E., Seipel, B., & Liu, B. (2018). Preliminary Findings on the Computer-Administered Multiple-Choice Online Causal Comprehension Assessment, a Diagnostic Reading Comprehension Test. *Assessment for Effective Intervention*, *43*(3), 169-181.

Davison, M. L., Semmes, R., Huang, L., & Close, C. N. (2012). On the reliability and validity of a numerical reasoning speed dimension derived from response times collected in computerized testing. *Educational and Psychological Measurement*, *72*(2), 245-263.

Davison, M. L., Semmes, R., Huang, L., & Close, C. N. (2012). On the reliability and validity of a numerical reasoning speed dimension derived from response times collected in computerized testing. *Educational and Psychological Measurement*, *72*(2), 245-263.

De Ayala, R. J., Plake, B. S., & Impara, J. C. (2001). The impact of omitted responses on the accuracy of ability estimation in item response theory. *Journal of educational measurement*, *38*(3), 213-234.

De Boeck, P., Chen, H., & Davison, M. (2017). Spontaneous and imposed speed of cognitive test responses. *British Journal of Mathematical and Statistical Psychology*, *70*(2), 225-237.

De Boeck, P. and Partchev, I. (2012). IRT Trees: Tree-Based response modeling of the GLMM family. *Journla of Statistical Software, 48,* 1-28.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and counterexample retrieval for causal conditionals. *Thinking & Reasoning*, *11*(2), 123-150.

Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*(1), 222.

Entink, R. K., Fox, J. P., & van der Linden, W. J. (2009). A multivariate multilevel approach to the modeling of accuracy and speed of test takers. *Psychometrika*, *74*(1), 21.

Entink, R. K., Fox, J. P., & van der Linden, W. J. (2009). A multivariate multilevel approach to the modeling of accuracy and speed of test takers. *Psychometrika*, *74*(1), 21.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, *59*, 255-278.

Evans, J. S. B., & Over, D. E. (1996). Rationality in the selection task: Epistemic utility versus uncertainty reduction.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223-241.

Evans, J. S. B., & Wason, P. C. (1976). Rationalization in a reasoning task. *British Journal of Psychology*, *67*(4), 479-486.

Finch, H. (2008). Estimation of item response theory parameters in the presence of missing data. *Journal of Educational Measurement*, *45*(3), 225-245.

Fodor, J. A. (1983). *The modularity of mind*. MIT press.

Furneaux, W.D. (1961). Intellectual abilities and problem solving behavior. In H.J. Eysenck (Ed.), The handbook of abnormal psychology. London: Pitman.

Glasser, M. (1964). Linear regression analysis with missing observations among the independent variables. *Journal of the American Statistical Association*, *59*(307), 834-

844.

Harmes, J. C., Parshall, C. G., & Kromrey, J. D. (2003). Recalibration of IRT item parameters in a CAT: sparse data matrices and missing data treatments. In *annual meeting of the National Council on Measurement in Education, Chicago: IL*.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, *49*, 81.

Leite, W., & Beretvas, S. N. (2010). The performance of multiple imputation for Likert-type items with missing data. *Journal of Modern Applied Statistical Methods*, *9*(1), 8.

Liu, B., Kennedy, P. C., Seipel, B., Carlson, S. E., Biancarosa, G., & Davison, M. L. (2019). Can we learn from student mistakes in a formative, reading comprehension assessment?. *Journal of Educational Measurement*, *56*(4), 815-835.

Liu, Y., Li, Z., Liu, H., & Luo, F. (2019). Modeling Test-Taking Non-effort in MIRT Models. *Frontiers in psychology*, *10*, 145.

Lord, F. M. (1974). Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika*, *39*(2), 247-264.

Lord, F. M. (1980). *Applications of item response theory to pr actical testing problems*. Routledge.

Ludlow, L. H., & O'leary, M. (1999). Scoring omitted and not-reached items: Practical data analysis implications. *Educational and Psychological Measurement*, *59*(4), 615-630.

Maris, G., & Van der Maas, H. (2012). Speed-accuracy response models: Scoring rules based on response time and accuracy. *Psychometrika*, *77*(4), 615-633.

Markovits, H., Fleury, M. L., Quinn, S., & Venet, M. (1998). The development of conditional reasoning and the structure of semantic memory. *Child development*, *69*(3), 742-755.

Mislevy, R. J., & Wu, P. K. (1988). Inferring examinee ability when some item responses are missing. *ETS Research Report Series*, *1988*(2), i-75.

Mislevy, R. J., & Wu, P. K. (1996). Missing responses and IRT ability estimation: Omits, choice, time limits, and adaptive testing. *ETS Research Report Series*, *1996*(2), i-36.

Mislevy, R. J. (2013). Missing responses in item response theory. In van der Linden, W. J., & Hambleton, R. K. (Eds.), *Handbook of modern item response theory* (pp.171-193). Springer Science & Business Media.

Molenaar, D., Tuerlinckx, F., & van der Maas, H. L. (2015). A bivariate generalized linear item response theory modeling framework to the analysis of responses and response times. *Multivariate Behavioral Research*, *50*(1), 56-74.

Moustaki, I., & O'Muircheartaigh, C. (2000). A one dimension latent trait model to infer attitude from nonresponse for nominal data. *Statistica*, *60*(2), 259-276.

Partchev, I., & De Boeck, P. (2012). Can fast and slow intelligence be differentiated? *Intelligence, 40*(1), 23-32.

Ranger, J., & Ortner, T. (2012). The case of dependency of responses and response times: A modeling approach based on standard latent trait models. *Psychological Test and Assessment Modeling*, *54*(2), 128.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of experimental psychology: General*, *118*(3), 219.

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current directions in psychological science*, *13*(2), 60-66.

Rios, J. A., Guo, H., Mao, L., & Liu, O. L. (2017). Evaluating the impact of careless responding on aggregated-scores: To filter unmotivated examinees or not?. *International Journal of Testing*, *17*(1), 74-104.

Rose, N., Davier, M., & Xu, X. (2010). Modeling nonignorable missing data with item response theory (IRT). *ETS Research Report Series*, *2010*(1), i-53.

Schafer, J. L., Khare, M., & Ezzati-Rice, T. M. (1993). Multiple imputation of missing data in NHANES III. In *Proceedings of the Annual Research Conference* (pp. 459-487).

Schwarz, G. E. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*, 461-464.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127.

Sijtsma, K., & Van der Ark, L. A. (2003). Investigation and treatment of missing item scores in test and questionnaire data. *Multivariate Behavioral Research*, *38*(4), 505-528.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological bulletin*, *119*(1), 3.

Smith, E. R., & Collins, E. C. (2009). Dual-process models: A social psychological perspective. *J. St. BT Evans & K. Frankish (Eds.), In two minds: Dual processes and beyond*, 197-216.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and social psychology review*, *4*(2), 108-131.

Smits, N., Mellenbergh, G. J., & Vorst, H. C. M. (2002). Alternative missing data techniques to grade point average: Imputing unavailable grades. *Journal of Educational Measurement*, *39*, 187–206.

Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.

Stanovich, K. E. (2004). Metarepresentation and the great cognitive divide: A commentary on Henriques'"Psychology Defined". *Journal of clinical psychology*, *60*(12), 1263-1266.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and social psychology review*, *8*(3), 220-247.

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological review*, *112*(1), 159.

Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2001). Using multivariate statistics.

Thissen, D. (1983). Timed testing: An approach using item response theory. In *New horizons in testing* (pp. 179-203). Academic Press.

van der Linden, W. J. (2007). A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika*, *72*(3), 287.

van der Linden, W. J. (2010). *Elements of adaptive testing* (pp. 31-56). C. A. Glas (Ed.). New York, NY: Springer.

van der Linden, W. J., & Glas, C. A. (2010). Statistical tests of conditional independence between responses and/or response times on test items. *Psychometrika*, *75*(1), 120-139.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005). A dual-process specification of causal conditional reasoning.*Thinking & Reasoning*,*11*(3), 239-278.

Wang, C., & Xu, G. (2015). A mixture hierarchical model for response times and response accuracy. *British Journal of Mathematical and Statistical Psychology,* 68(3), 456-477.

Wason, P. C., & Evans, J. S. B. (1974). Dual processes in reasoning *Cognition*,*3*(2), 141-154.

Wilson, T. D. (2004). *Strangers to ourselves*. Harvard University Press.

Wise, S. L., & DeMars, C. E. (2006). An application of item response time: The effort-moderated IRT model. *Journal of Educational Measurement*, *43*(1), 19-38.

Zhang, B., & Walker, C. M. (2008). Impact of missing data on person—model fit and person trait estimation. *Applied Psychological Measurement*, *32*(6), 466-479.

.

## Appendix

Table 18. Median response time for Form 3.2

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 35.6 | 11 | 50.5 | 21 | 42.8 | 31 | 35.0 |
| 2 | 37.1 | 12 | 40.8 | 22 | 45.5 | 32 | 36.2 |
| 3 | 38.0 | 13 | 39.0 | 23 | 38.8 | 33 | 38.4 |
| 4 | 44.3 | 14 | 36.3 | 24 | 34.9 | 34 | 41.1 |
| 5 | 42.6 | 15 | 44.8 | 25 | 42.2 | 35 | 39.2 |
| 6 | 39.9 | 16 | 41.2 | 26 | 43.1 | 36 | 32.0 |
| 7 | 54.9 | 17 | 36.5 | 27 | 42.1 | 37 | 42.5 |
| 8 | 36.6 | 18 | 42.1 | 28 | 40.9 | 38 | 40.1 |
| 9 | 42.1 | 19 | 42.6 | 29 | 36.3 | 39 | 41.1 |
| 10 | 39.5 | 20 | 49.4 | 30 | 39.1 | 40 | 45.2 |

Table 19. Median response time for Form 3.3

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 49.0 | 11 | 47.7 | 21 | 51.9 | 31 | 53.6 |
| 2 | 35.4 | 12 | 40.2 | 22 | 46.9 | 32 | 34.7 |
| 3 | 37.5 | 13 | 46.6 | 23 | 47.9 | 33 | 47.2 |
| 4 | 48.7 | 14 | 48.8 | 24 | 38.8 | 34 | 45.1 |
| 5 | 43.1 | 15 | 53.6 | 25 | 34.7 | 35 | 48.6 |
| 6 | 37.5 | 16 | 46.4 | 26 | 51.2 | 36 | 43.2 |
| 7 | 35.1 | 17 | 40.5 | 27 | 43.7 | 37 | 45.4 |
| 8 | 52.5 | 18 | 36.2 | 28 | 46.2 | 38 | 41.6 |
| 9 | 44.8 | 19 | 39.1 | 29 | 40.0 | 39 | 54.7 |
| 10 | 44.2 | 20 | 50.0 | 30 | 46.6 | 40 | 38.4 |

Table 20. Median response time for Form 4.2

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 41.5 | 11 | 42.8 | 21 | 49.9 | 31 | 43.4 |
| 2 | 42.1 | 12 | 44.4 | 22 | 45.7 | 32 | 47.5 |
| 3 | 36.2 | 13 | 34.4 | 23 | 46.3 | 33 | 41.1 |
| 4 | 44.5 | 14 | 42.9 | 24 | 53.1 | 34 | 49.9 |
| 5 | 49.1 | 15 | 50.5 | 25 | 39.5 | 35 | 44.5 |
| 6 | 44.6 | 16 | 45.9 | 26 | 40.8 | 36 | 46.0 |
| 7 | 45.9 | 17 | 39.5 | 27 | 42.4 | 37 | 44.4 |
| 8 | 38.8 | 18 | 45.1 | 28 | 53.8 | 38 | 35.3 |
| 9 | 38.9 | 19 | 52.9 | 29 | 37.7 | 39 | 44.3 |
| 10 | 42.9 | 20 | 51.2 | 30 | 49.6 | 40 | 42.7 |

Table 21. Median response time for Form 4.3

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 38.7 | 11 | 43.4 | 21 | 41.7 | 31 | 52.3 |
| 2 | 40.8 | 12 | 40.8 | 22 | 45.0 | 32 | 41.8 |
| 3 | 44.4 | 13 | 46.8 | 23 | 45.0 | 33 | 35.9 |
| 4 | 46.3 | 14 | 43.0 | 24 | 42.3 | 34 | 49.3 |
| 5 | 62.9 | 15 | 44.2 | 25 | 41.6 | 35 | 50.4 |
| 6 | 42.1 | 16 | 41.0 | 26 | 47.0 | 36 | 49.2 |
| 7 | 46.9 | 17 | 44.0 | 27 | 40.2 | 37 | 41.1 |
| 8 | 50.2 | 18 | 45.8 | 28 | 42.5 | 38 | 35.7 |
| 9 | 44.8 | 19 | 41.3 | 29 | 37.3 | 39 | 36.5 |
| 10 | 46.6 | 20 | 46.1 | 30 | 48.2 | 40 | 40.0 |

Table 22. Median response time for Form 5.2

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 38.7 | 11 | 47.0 | 21 | 54.1 | 31 | 43.9 |
| 2 | 43.7 | 12 | 47.0 | 22 | 55.1 | 32 | 50.7 |
| 3 | 49.3 | 13 | 45.0 | 23 | 48.9 | 33 | 49.1 |
| 4 | 43.6 | 14 | 51.7 | 24 | 38.6 | 34 | 51.7 |
| 5 | 43.4 | 15 | 45.7 | 25 | 59.6 | 35 | 43.3 |
| 6 | 49.1 | 16 | 43.8 | 26 | 44.9 | 36 | 39.5 |
| 7 | 38.3 | 17 | 45.7 | 27 | 44.9 | 37 | 48.4 |
| 8 | 40.7 | 18 | 51.4 | 28 | 54.5 | 38 | 42.4 |
| 9 | 42.7 | 19 | 40.2 | 29 | 43.1 | 39 | 44.6 |
| 10 | 48.2 | 20 | 45.7 | 30 | 43.1 | 40 | 42.9 |

Table 23. Median response time for Form 5.3

| Item | Response Time | Item | Response Time | Item | Response Time | Item | Response Time |
|---|---|---|---|---|---|---|---|
| 1 | 50.65 | 11 | 45.55 | 21 | 53.8 | 31 | 44.05 |
| 2 | 31.6 | 12 | 46 | 22 | 45.4 | 32 | 52.6 |
| 3 | 46.45 | 13 | 42.05 | 23 | 48.1 | 33 | 48.15 |
| 4 | 40.05 | 14 | 50.55 | 24 | 37.5 | 34 | 47.95 |
| 5 | 49.85 | 15 | 43.75 | 25 | 40.65 | 35 | 41.95 |
| 6 | 40.8495 | 16 | 44.95 | 26 | 56.8 | 36 | 44.35 |
| 7 | 35.7495 | 17 | 52.4 | 27 | 50.7 | 37 | 51.45 |
| 8 | 42.55 | 18 | 43.25 | 28 | 41.7 | 38 | 38.9 |
| 9 | 40.15 | 19 | 54.3 | 29 | 43.05 | 39 | 40.3995 |
| 10 | 45.8 | 20 | 52.5 | 30 | 36.5 | 40 | 41.35 |

Table 24. Mean and SD of Raw Scores for Grade 3

| | Form 3.2 | | Form 3.3 | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| Correct | 23.92 | 9.82 | 26.05 | 9.91 |
| Incorrect | 16.08 | 9.82 | 13.95 | 9.91 |
| Fast | 20.07 | 11.78 | 20.02 | 11.99 |
| Slow | 19.93 | 11.78 | 19.98 | 11.99 |
| FC | 10.91 | 9.29 | 11.88 | 9.68 |
| SC | 13.01 | 10.24 | 14.17 | 10.70 |
| FI | 9.17 | 8.15 | 8.14 | 8.31 |
| SI | 6.91 | 5.72 | 5.81 | 5.29 |
| FC_FI | 1.74 | 12.91 | 3.74 | 13.48 |
| SC_SI | 6.10 | 11.66 | 8.36 | 11.88 |
| FC/F | 0.53 | 0.30 | 0.59 | 0.32 |
| SC/S | 0.62 | 0.27 | 0.66 | 0.27 |

Table 25. Mean and SD of Raw Scores for Grade 4

| | Form 4.2 | | Form 4.3 | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| Correct | 28.25 | 9.50 | 27.80 | 9.64 |
| Incorrect | 11.75 | 9.50 | 12.20 | 9.64 |
| Fast | 20.07 | 11.27 | 20.02 | 11.34 |
| Slow | 19.93 | 11.27 | 19.98 | 11.34 |
| FC | 13.44 | 10.39 | 12.93 | 9.55 |
| SC | 14.81 | 9.98 | 14.88 | 10.26 |
| FI | 6.63 | 7.37 | 7.09 | 7.78 |
| SI | 5.12 | 5.22 | 5.10 | 5.10 |
| FC_FI | 6.81 | 14.05 | 5.83 | 13.23 |
| SC_SI | 9.69 | 11.25 | 9.77 | 11.58 |
| FC/F | 0.66 | 0.30 | 0.65 | 0.31 |
| SC/S | 0.73 | 0.23 | 0.71 | 0.25 |

Table 26. Mean and SD of Raw Scores for Grade 5

| | Form 5.2 | | Form 5.3 | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| Correct | 29.27 | 9.02 | 28.67 | 9.76 |
| Incorrect | 10.73 | 9.02 | 11.33 | 9.76 |
| Fast | 20.09 | 11.48 | 20.01 | 11.68 |
| Slow | 19.91 | 11.48 | 19.99 | 11.68 |
| FC | 14.39 | 10.80 | 13.74 | 10.82 |
| SC | 14.88 | 9.93 | 14.94 | 10.53 |
| FI | 5.70 | 6.56 | 6.27 | 7.29 |
| SI | 5.03 | 5.46 | 5.05 | 5.28 |
| FC_FI | 8.69 | 13.70 | 7.46 | 14.29 |
| SC_SI | 9.84 | 11.18 | 9.88 | 11.88 |
| FC/F | 0.67 | 0.31 | 0.67 | 0.31 |
| SC/S | 0.73 | 0.25 | 0.73 | 0.25 |

Table 27. Model Comparisons of Multidimensional Models for Grade 3

| Form | Dimensions | Models | -2LL | AIC | BIC |
|---|---|---|---|---|---|
| Form 3.2 | Two Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 34441.71 | 34923.71 | 35885.05 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 48330.55 | 48812.55 | 49773.89 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 78934.59 | 79576.59 | 80857.05 |
| | Three Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 34401.14 | 34888.14 | 35857.46 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 47151.40 | 47637.40 | 48606.71 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 76205.47 | 76851.47 | 78139.90 |
| Form 3.3 | Two Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 314046.34 | 31528.34 | 32475.37 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 46196.58 | 46678.85 | 47625.88 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 72976.84 | 73618.84 | 74880.23 |
| | Three Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 30989.62 | 31475.62 | 32430.51 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 44931.12 | 45417.12 | 46372.01 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 70112.36 | 70758.36 | 72027.62 |

Table 28. Model Comparisons of Multidimensional Models for Grade 4

| Form | Dimensions | Models | -2LL | AIC | BIC |
|---|---|---|---|---|---|
| | Two Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 33004.37 | 33486.37 | 34451.31 |

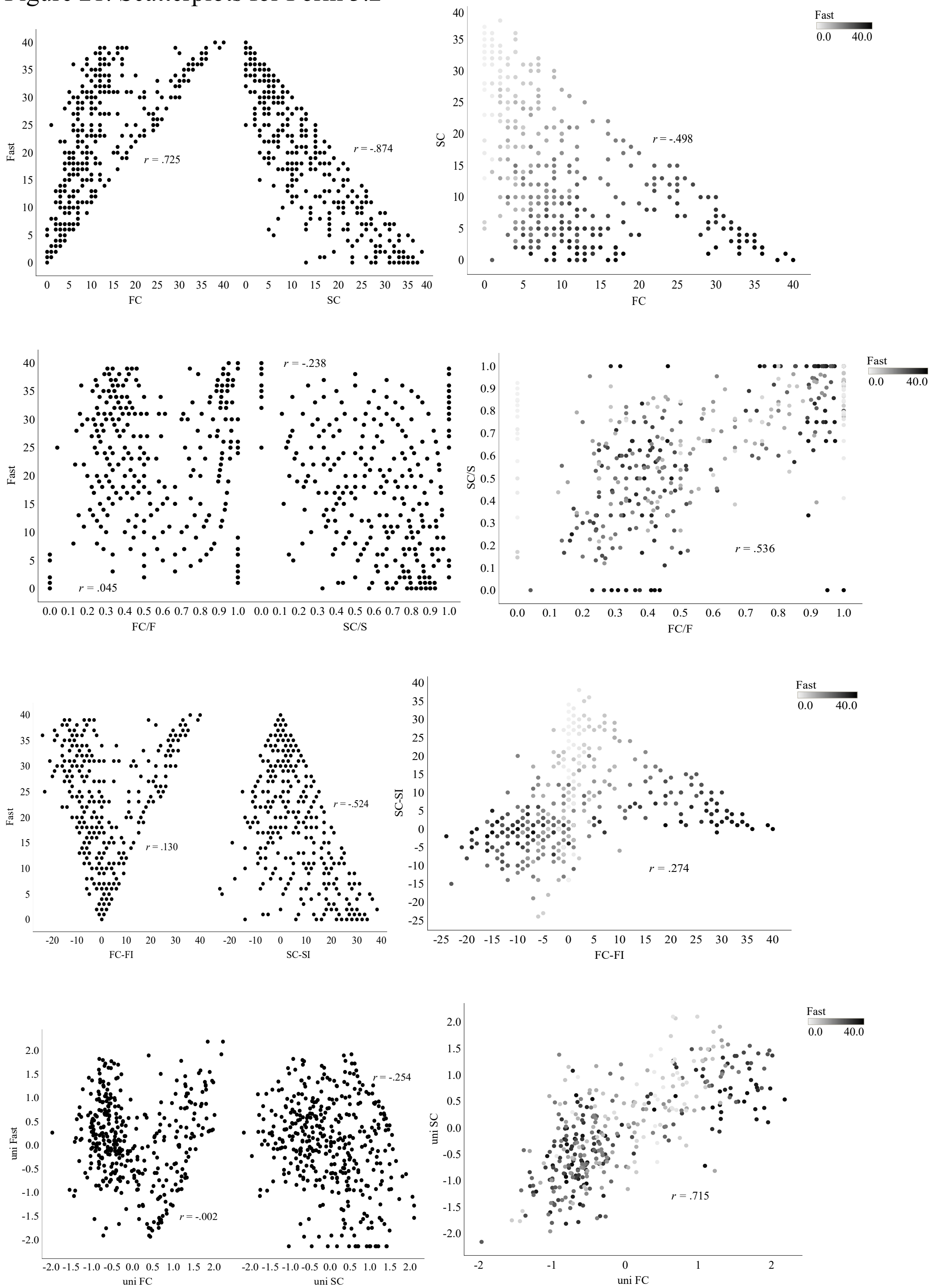
| | | | | | |
|---|---|---|---|---|---|
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 52043.39 | 52525.39 | 53490.32 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 77696.92 | 78282.92 | 79614.16 |
| Form 4.2 | Three Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 32960.73 | 33446.73 | 34419.67 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 50913.40 | 51399.40 | 52372.34 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 74972.37 | 75618.37 | 76911.63 |
| Form 4.3 | Two Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 46386.87 | 46872.87 | 47822.54 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 47532.48 | 48014.48 | 48956.33 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 71423.96 | 72065.96 | 73320.46 |
| | Three Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 30212.87 | 30694.87 | 31636.72 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 30162.04 | 30648.04 | 31597.71 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 68675.90 | 69321.90 | 70584.21 |

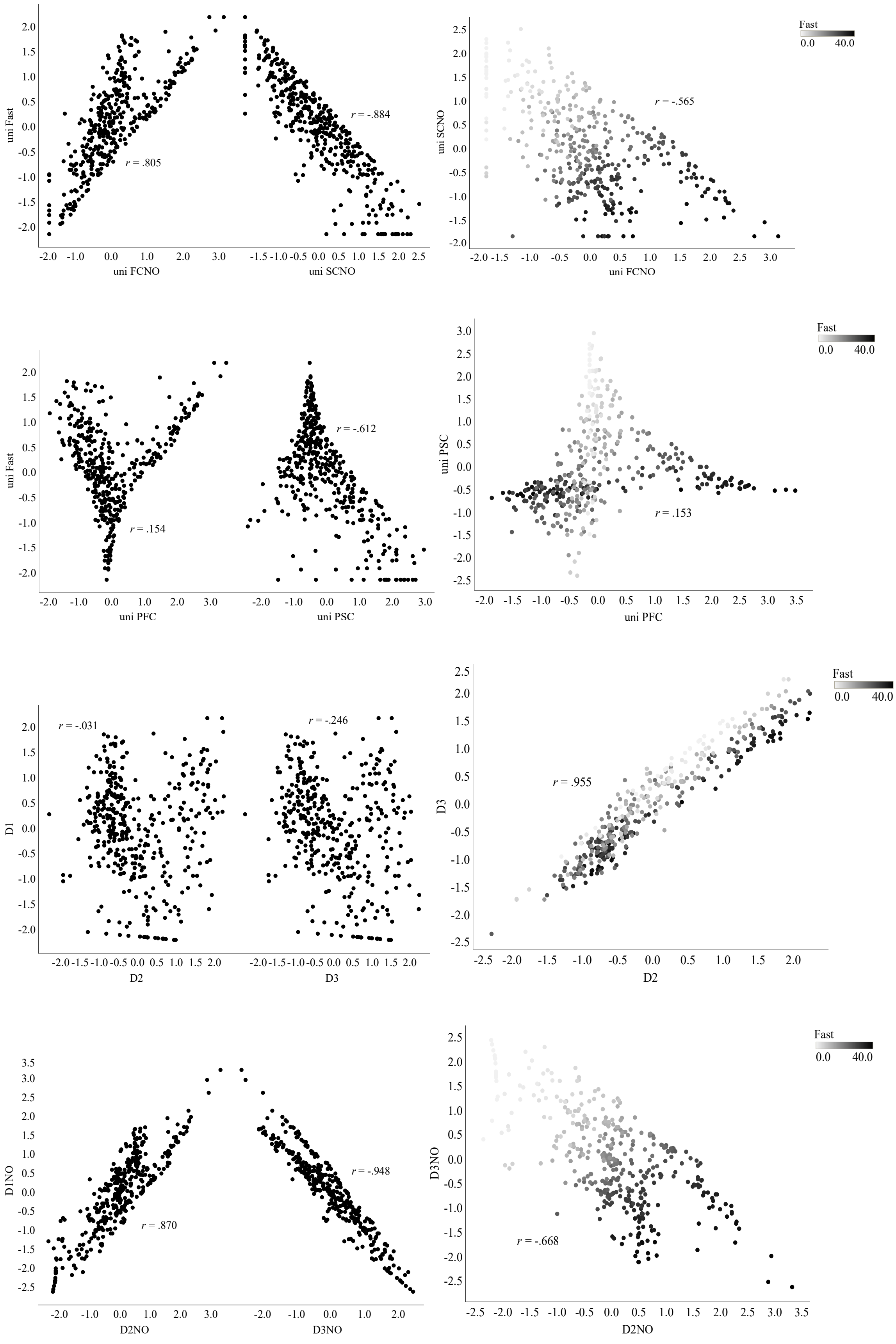
Table 29. Model Comparisons of Multidimensional Models for Grade 5

| Form | Dimensions | Models | -2LL | AIC | BIC |
|---|---|---|---|---|---|
| Form 5.2 | Two Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 25861.52 | 26343.52 | 27256.90 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 41620.33 | 42102.33 | 43015.71 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 61688.26 | 62330.26 | 63546.84 |

| | | | | | |
|---|---|---|---|---|---|
| | Three Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 25824.34 | 26310.34 | 27231.30 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 41648.91 | 42134.91 | 43055.87 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 59470.43 | 60116.43 | 61340.59 |
| Form 5.3 | Two Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 25030.30 | 25512.30 | 26418.95 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 39848.95 | 40330.95 | 41237.61 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 60021.40 | 60663.40 | 61871.02 |
| | Three Dimensional | $x_{1j}, x_{2j}, x_{3j}$ | 24972.34 | 25458.34 | 26372.34 |
| | | $x_{1j}, x'_{2j}, x'_{3j}$ | 38909.93 | 39395.93 | 40310.11 |
| | | $x_{1j}, y_{2j}, y_{3j}$ | 57408.98 | 58054.98 | 59270.12 |

Figure 21. Scatterplots for Form 3.2




uni Fast
uni FCNO
uni SCNO
r = .805
r = -.884
Fast
0.0
40.0
uni SCNO
uni FCNO
r = -.565
uni Fast
uni PFC
uni PSC
r = .154
r = -.612
Fast
0.0
40.0
uni PSC
uni PFC
r = .153
D1
D2
D3
r = -.031
r = -.246
Fast
0.0
40.0
D3
D2
r = .955
D1NO
D2NO
D3NO
r = .870
r = -.948
Fast
0.0
40.0
D3NO
D2NO
r = -.668

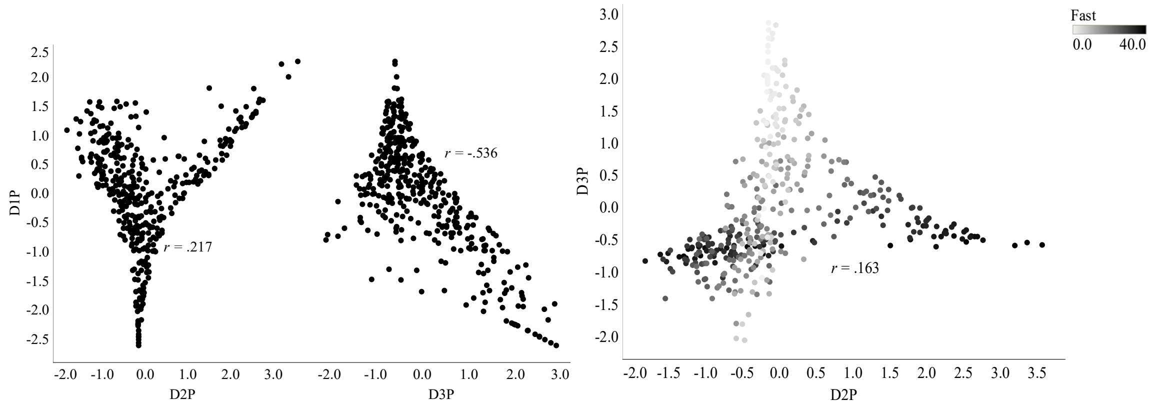
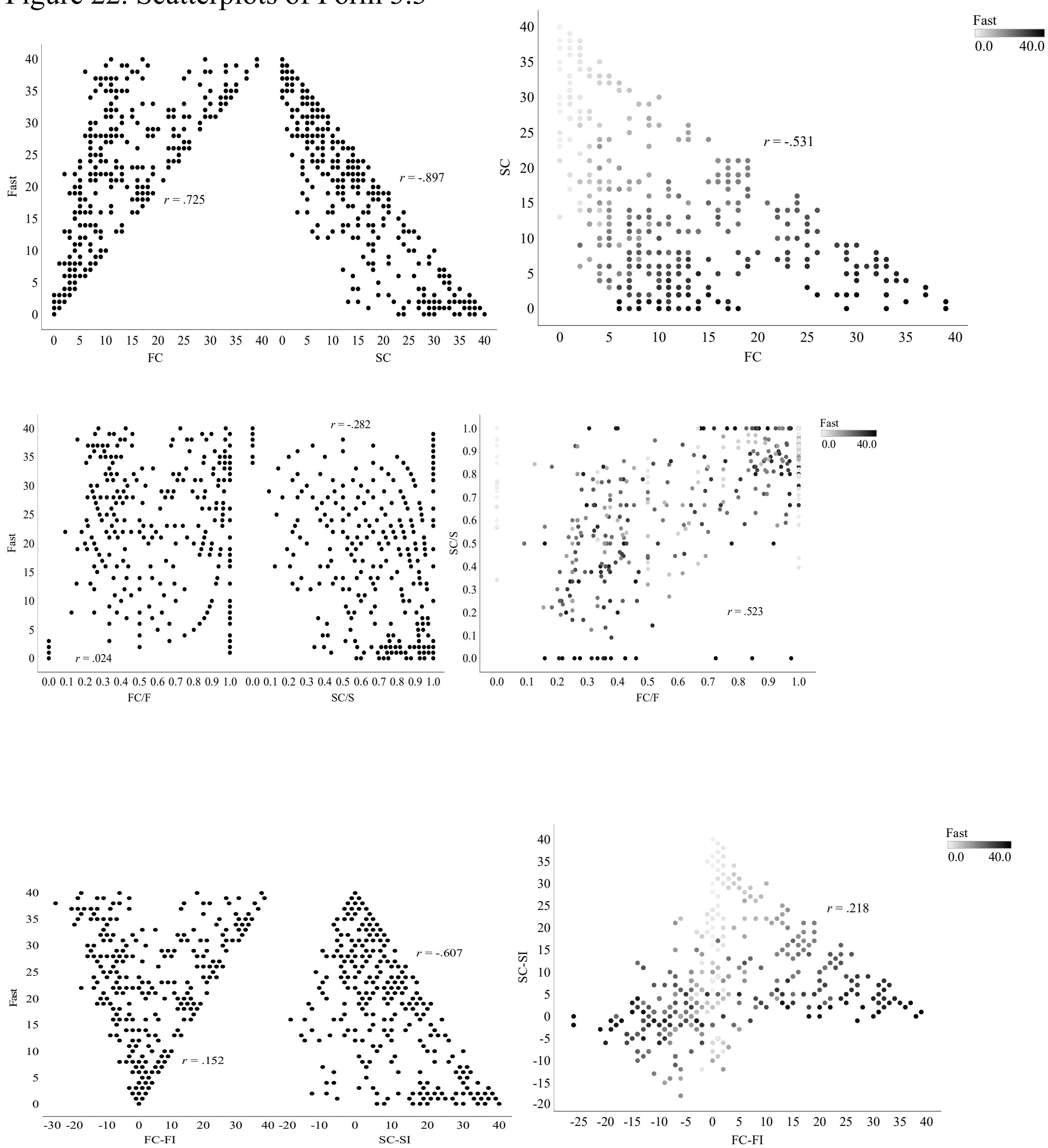
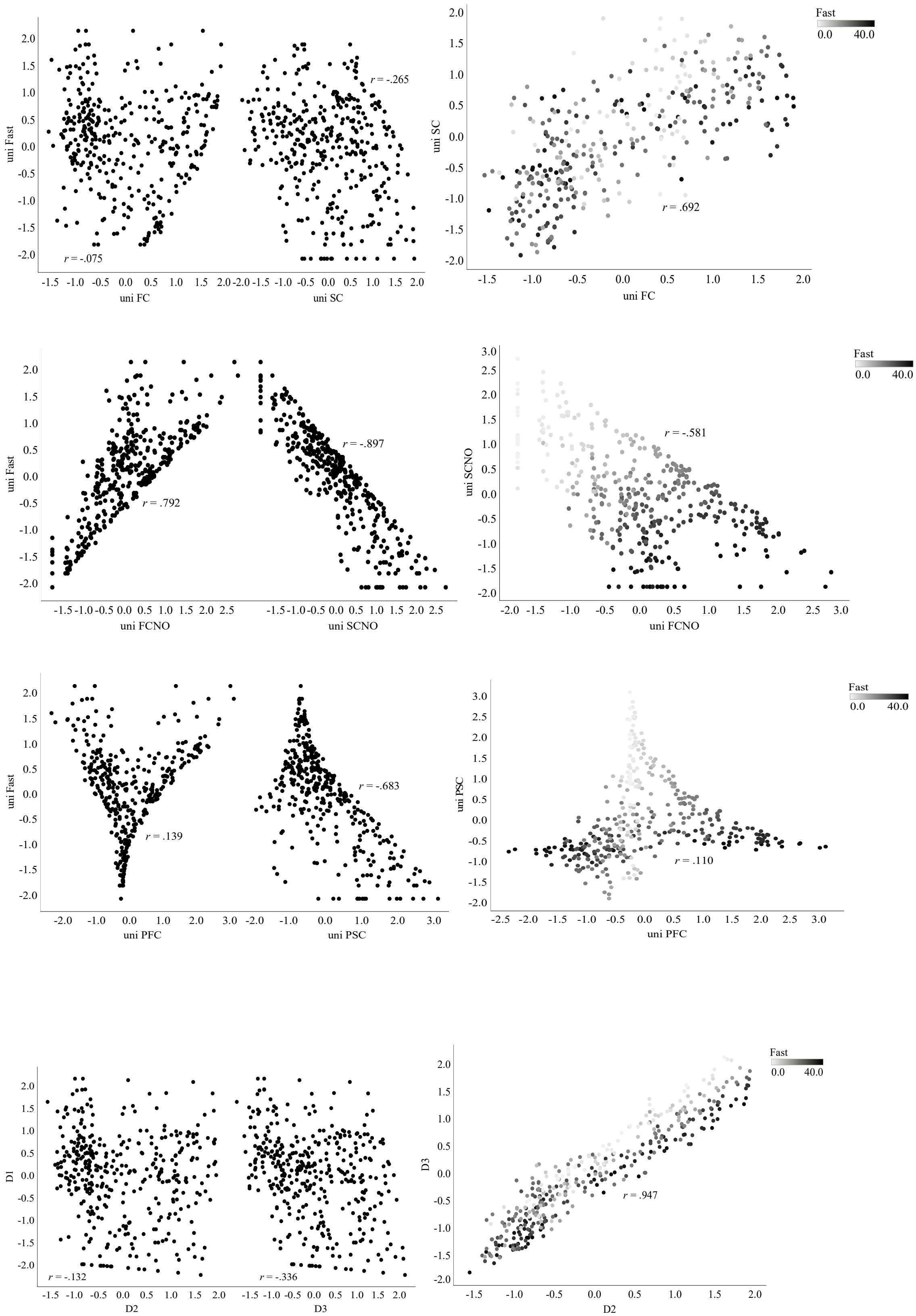



Figure 22. Scatterplots of Form 3.3





uni Fast
uni FC
uni SC
r = -.075
r = -.265
uni SC
uni FC
Fast
0.0 40.0
r = .692
uni Fast
uni FCNO
uni SCNO
r = .792
r = -.897
uni SCNO
uni FCNO
Fast
0.0 40.0
r = -.581
uni Fast
uni PFC
uni PSC
r = .139
r = -.683
uni PSC
uni PFC
Fast
0.0 40.0
r = .110
D1
D2
D3
r = -.132
r = -.336
D3
D2
Fast
0.0 40.0
r = .947

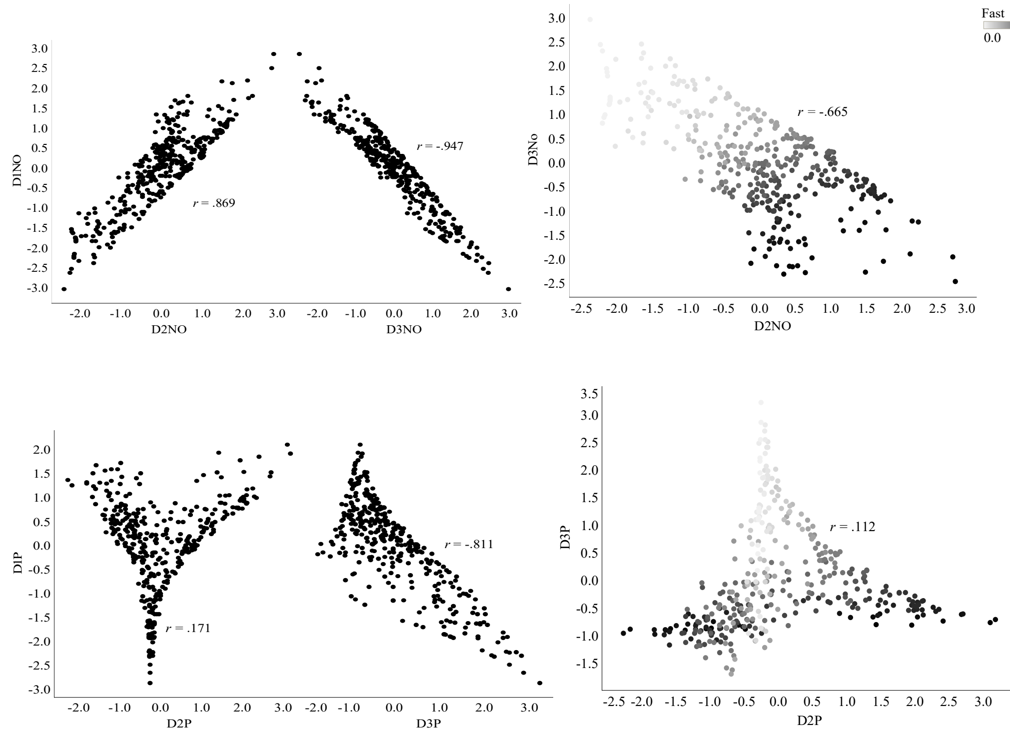
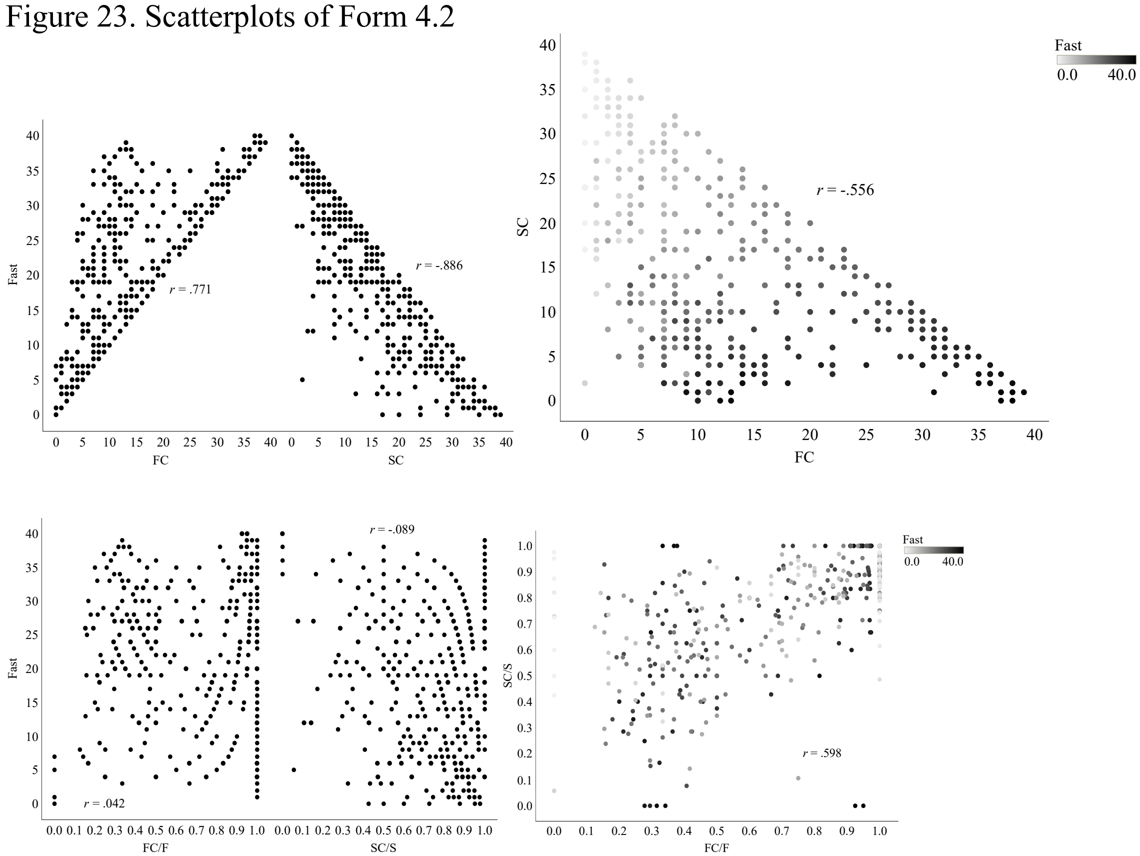
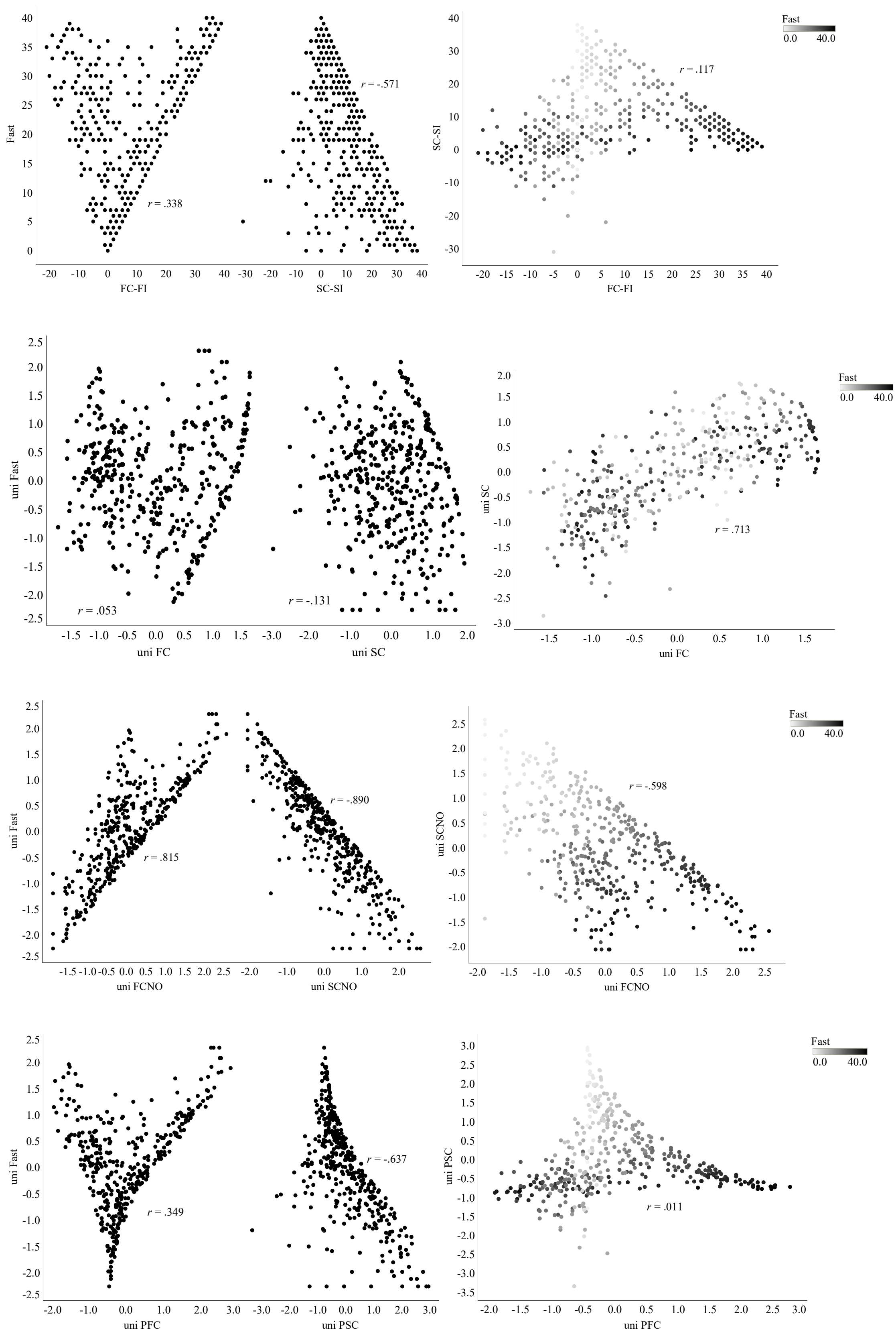



Figure 23. Scatterplots of Form 4.2





Fast
FC-FI
SC-SI
r = .338
r = -.571
r = .117
uni Fast
uni FC
uni SC
r = .053
r = -.131
r = .713
uni FCNO
uni SCNO
r = .815
r = -.890
r = -.598
uni PFC
uni PSC
r = .349
r = -.637
r = .011

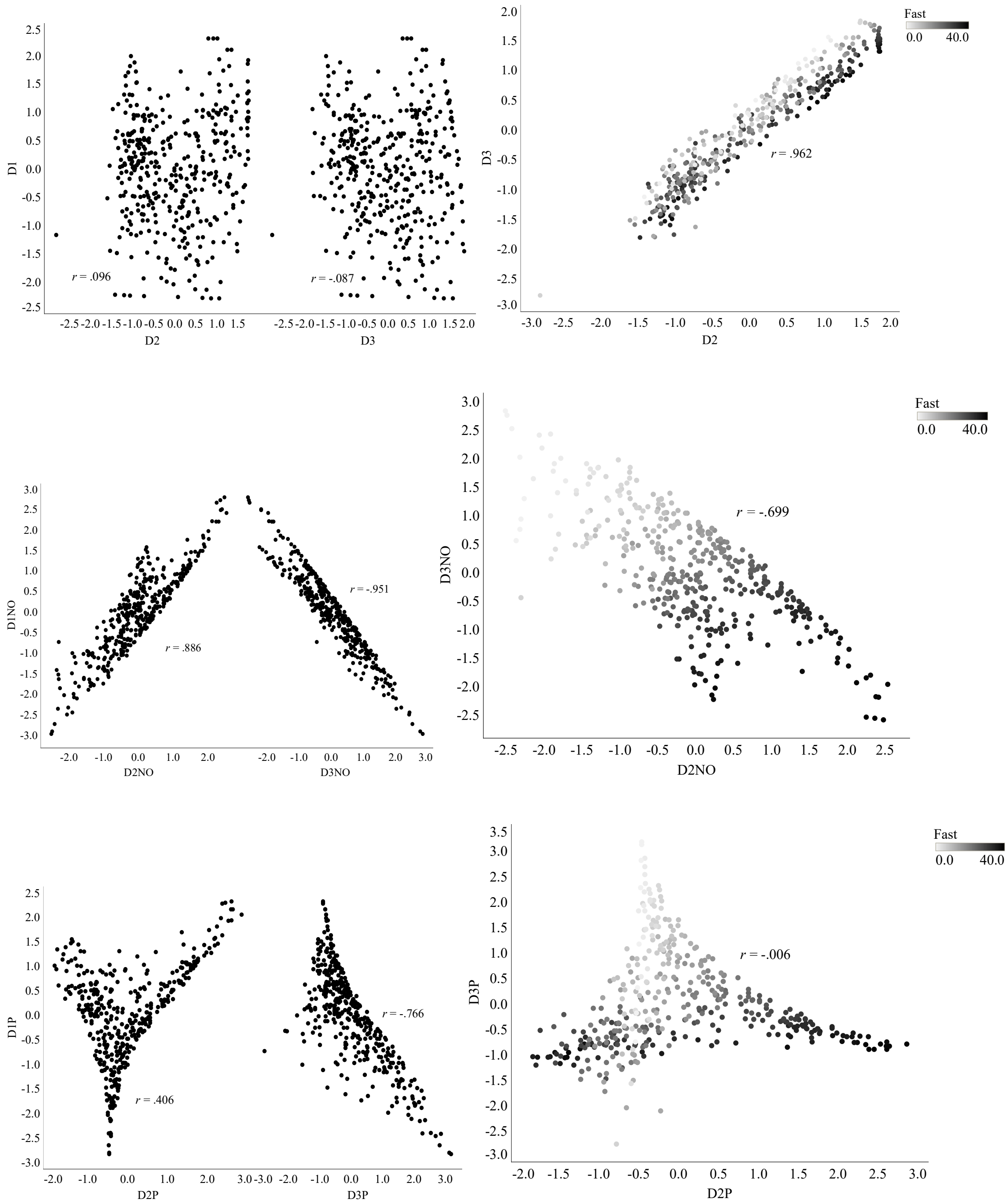
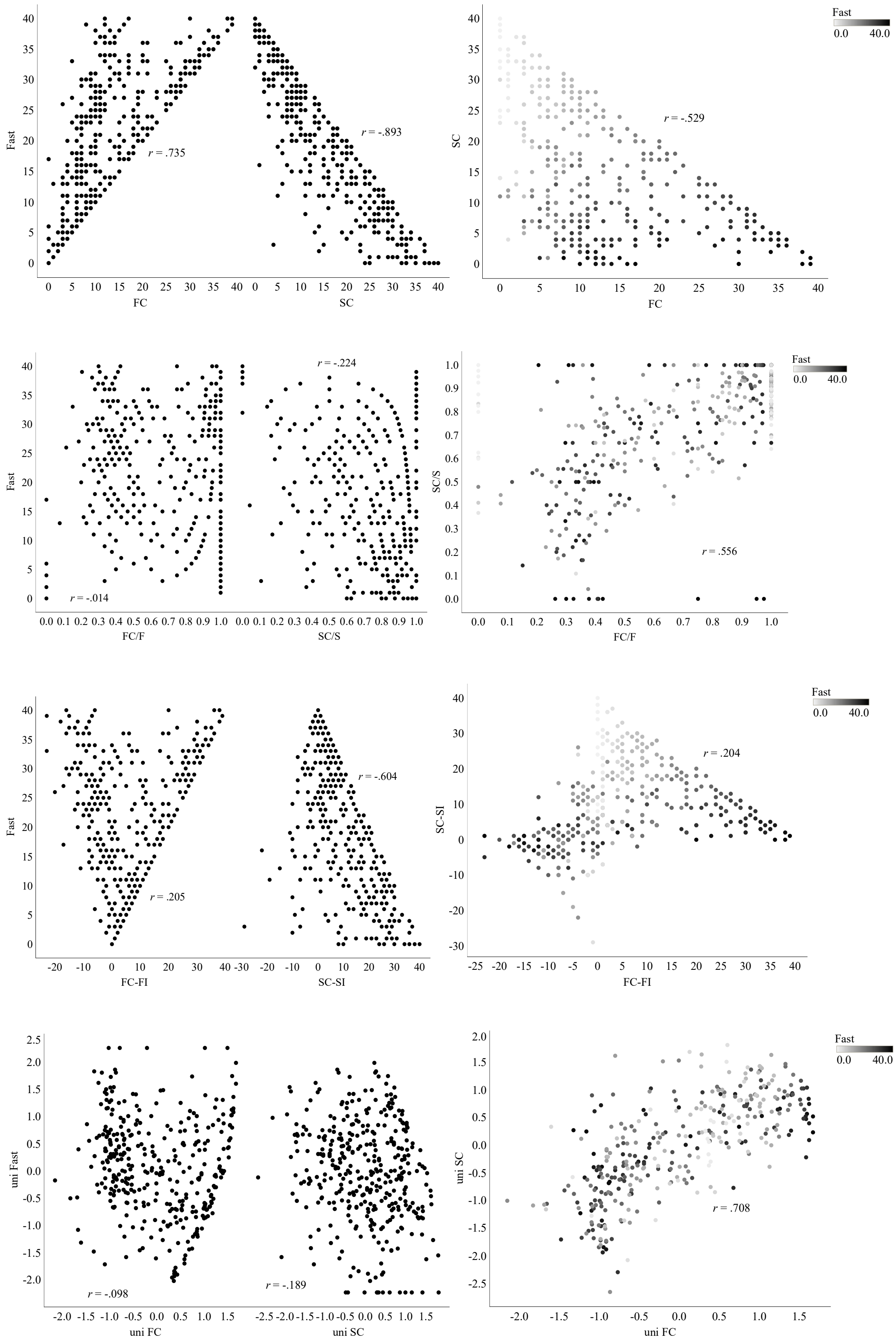
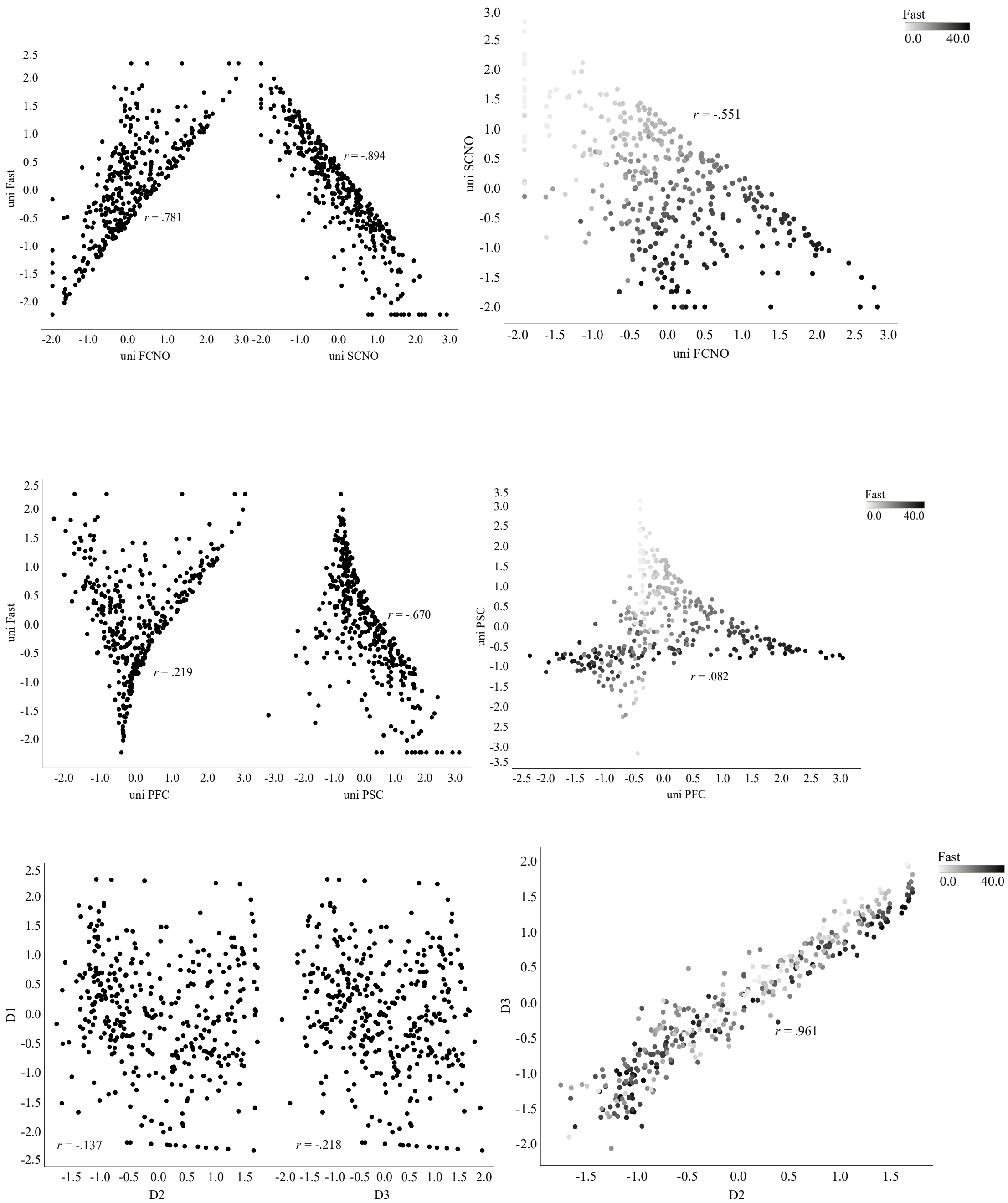



Figure 24. Scatterplots of Form 4.3


Fast
FC
SC
r = .735
r = -.893
Fast
0.0 40.0
SC
FC
r = -.529
Fast
FC/F
SC/S
r = -.014
r = -.224
Fast
0.0 40.0
SC/S
FC/F
r = .556
Fast
FC-FI
SC-SI
r = .205
r = -.604
Fast
0.0 40.0
SC-SI
FC-FI
r = .204
uni Fast
uni FC
uni SC
r = -.098
r = -.189
Fast
0.0 40.0
uni SC
uni FC
r = .708


uni Fast
uni FCNO
uni SCNO
r = .781
r = -.894
Fast
0.0
40.0
uni SCNO
uni FCNO
r = -.551
uni Fast
uni PFC
uni PSC
r = .219
r = -.670
Fast
0.0
40.0
uni PSC
uni PFC
r = .082
D1
D2
D3
r = -.137
r = -.218
Fast
0.0
40.0
D3
D2
r = .961

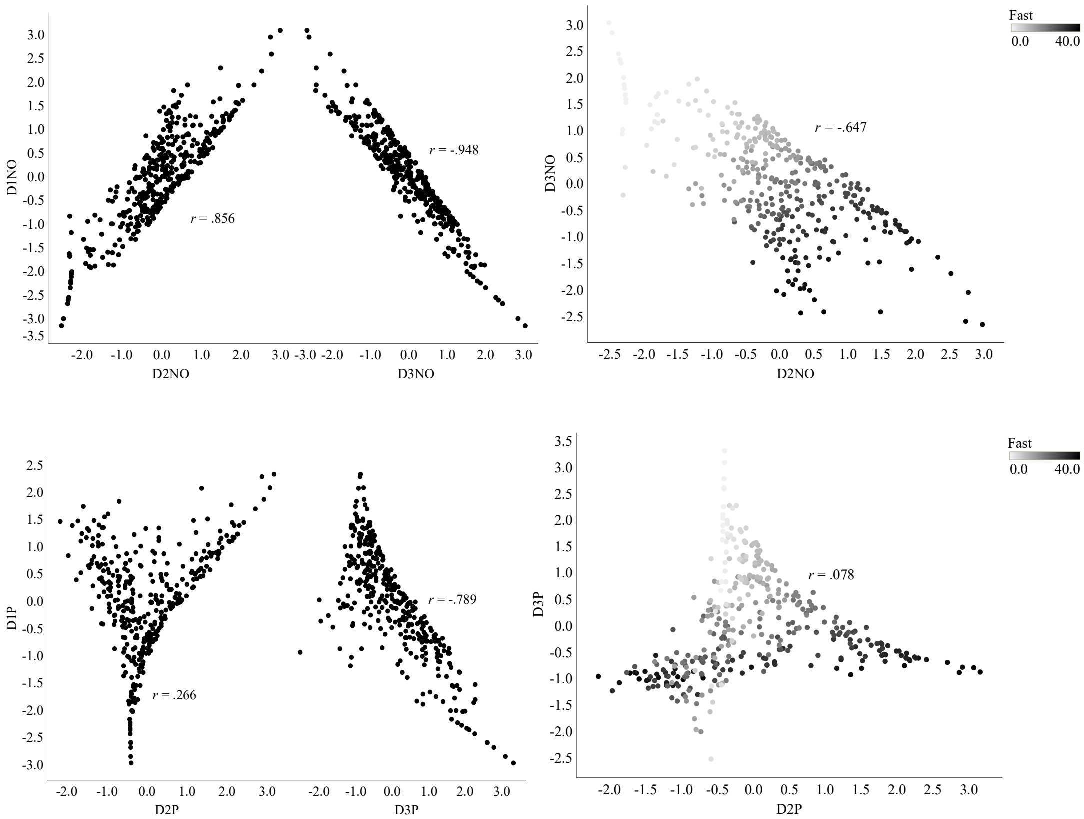
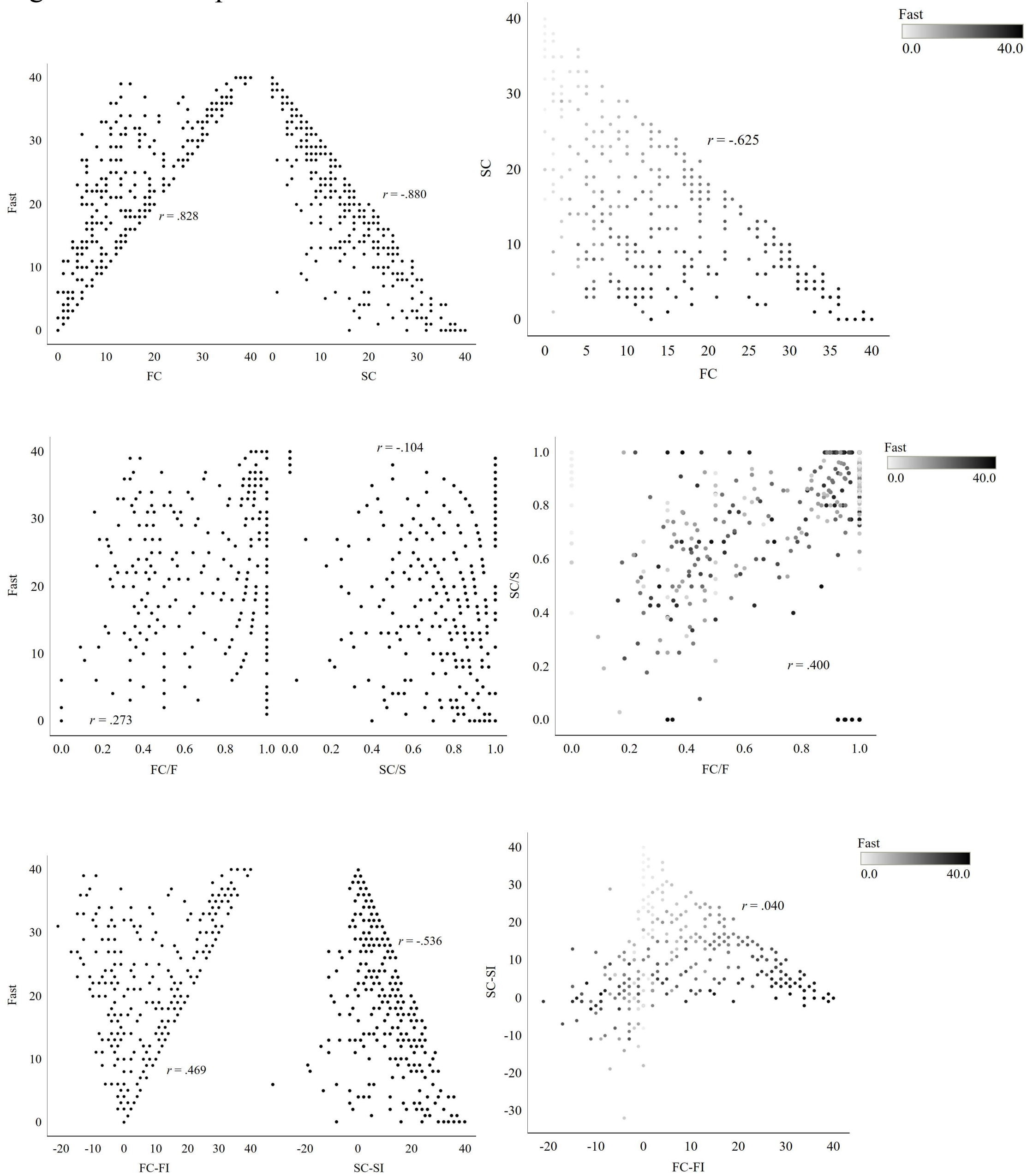
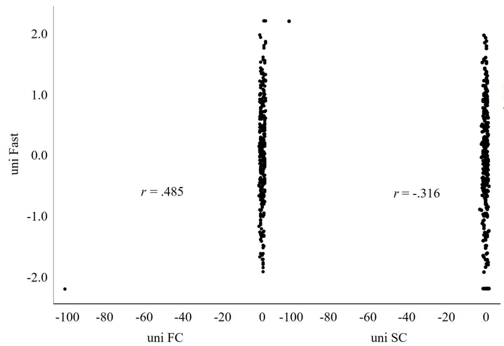
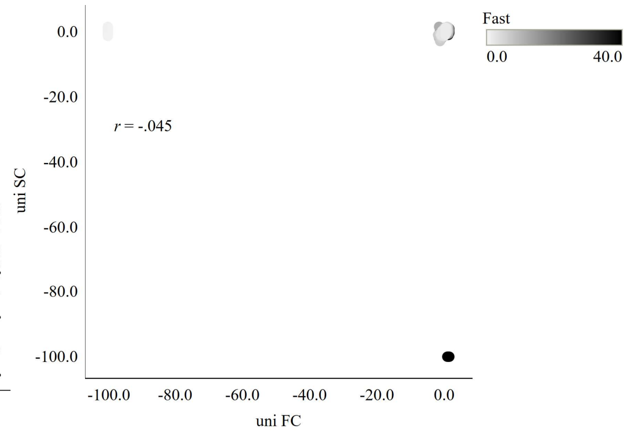
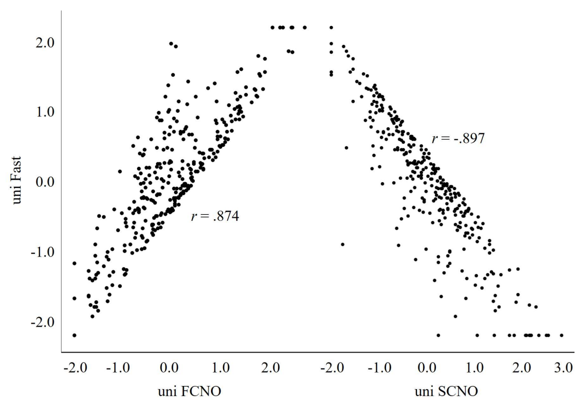

D1NO
D2NO
D3NO
r = .856
r = -.948
D3NO
D2NO
r = -.647
Fast
0.0
40.0
D1P
D2P
D3P
r = .266
r = -.789
D3P
D2P
r = .078
Fast
0.0
40.0

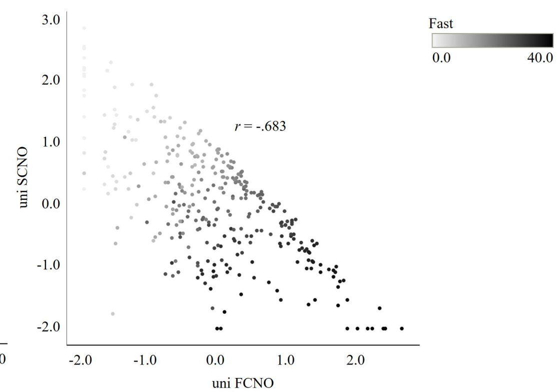
Figure 25. Scatterplots of Form 5.2

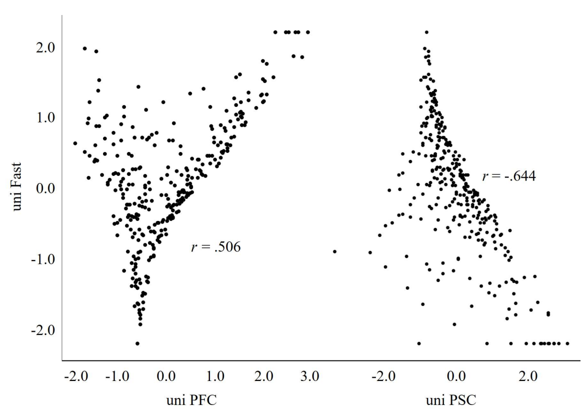


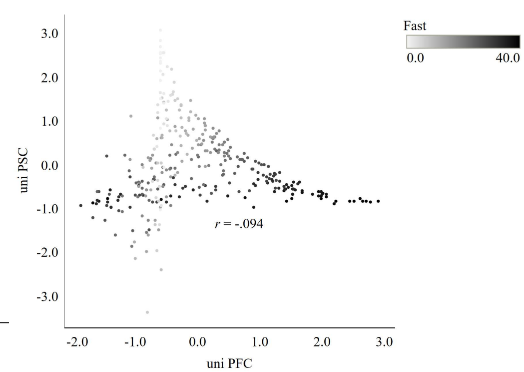

uni Fast
r = .485
r = -.316
uni FC
uni SC


Fast
0.0
40.0
r = -.045
uni SC
uni FC


uni Fast
r = -.897
r = .874
uni FCNO
uni SCNO


Fast
0.0
40.0
r = -.683
uni SCNO
uni FCNO


uni Fast
r = -.644
r = .506
uni PFC
uni PSC


Fast
0.0
40.0
r = -.094
uni PSC
uni PFC

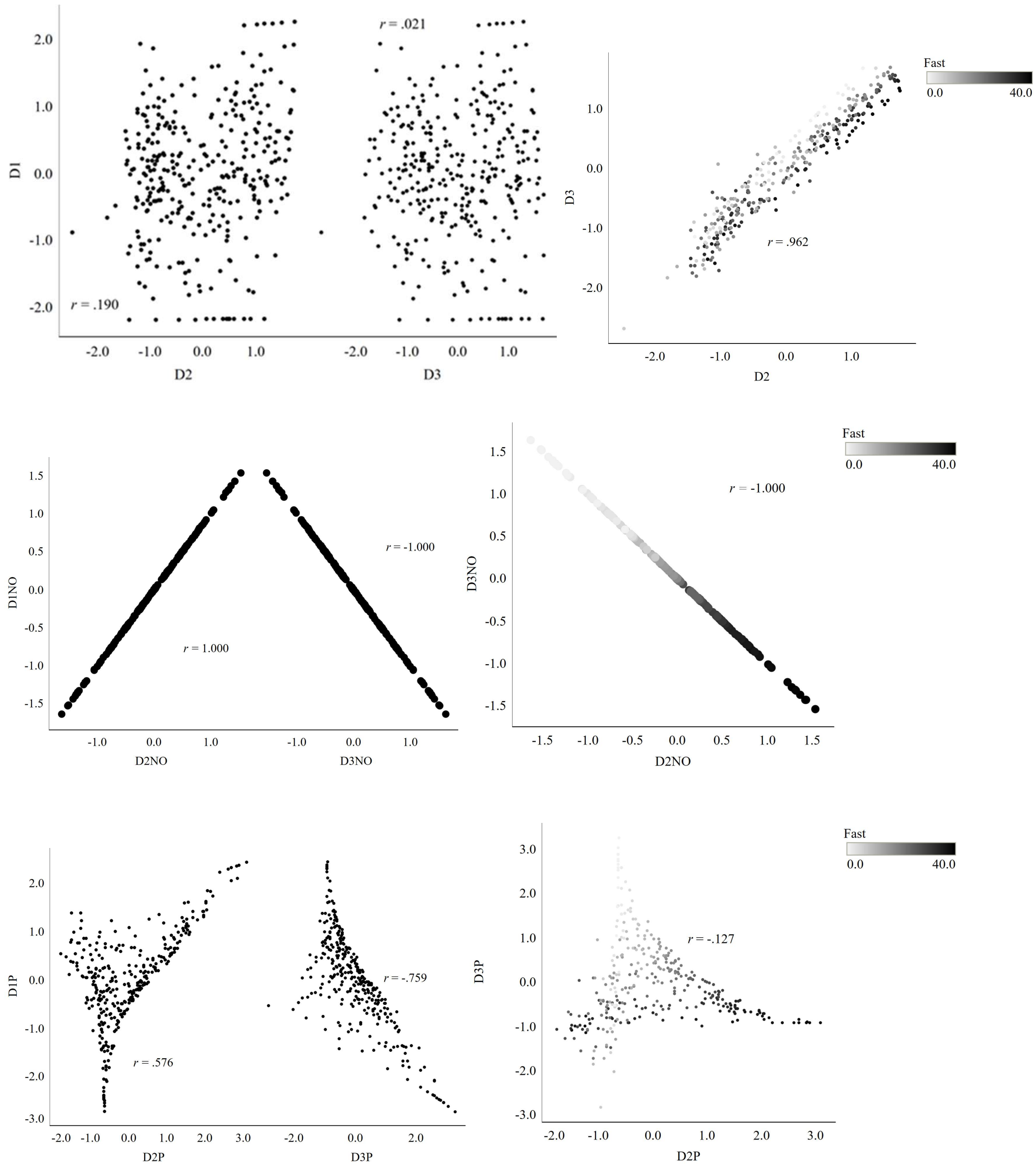
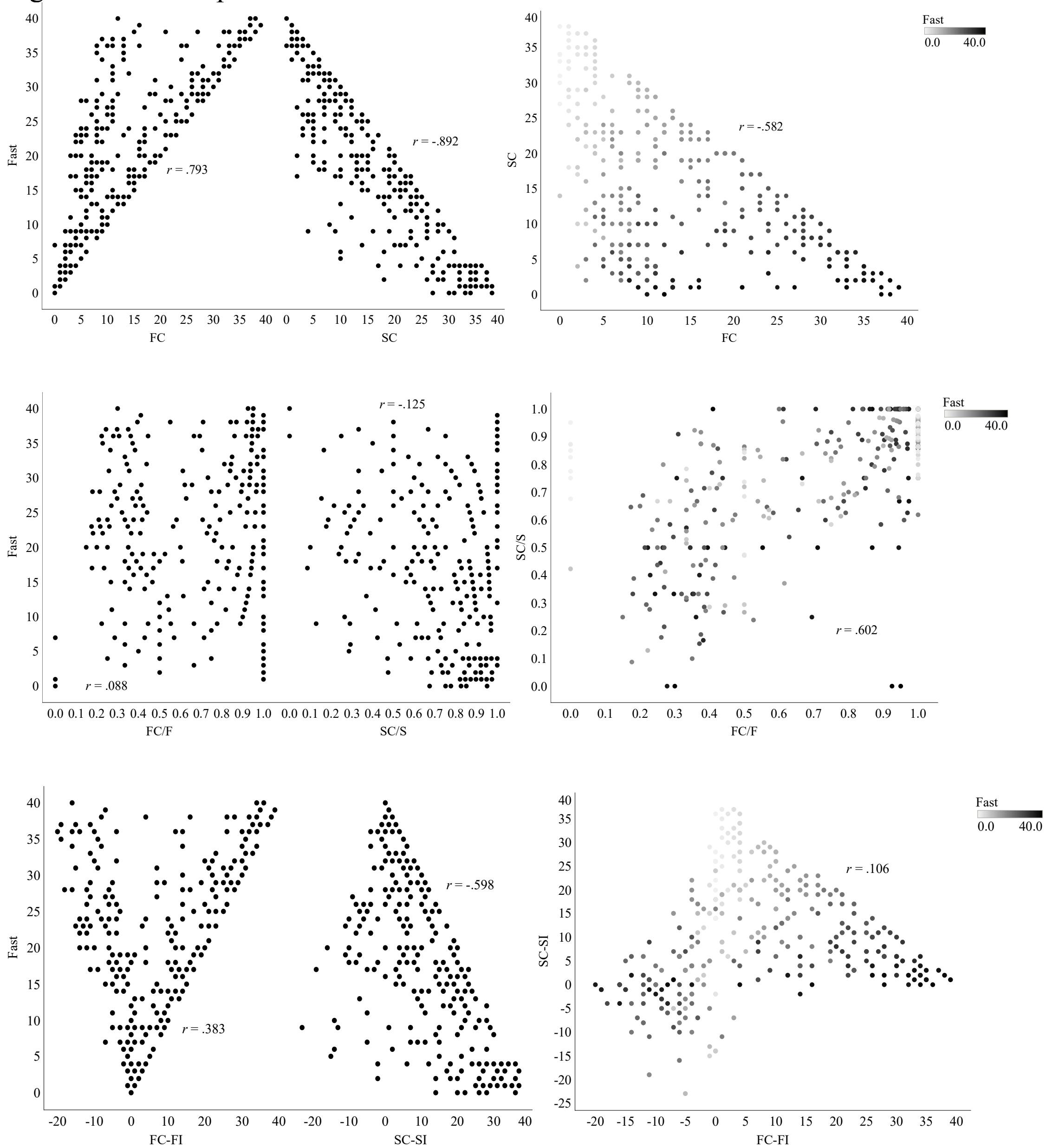
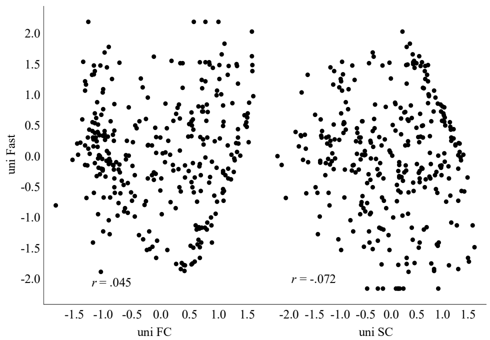
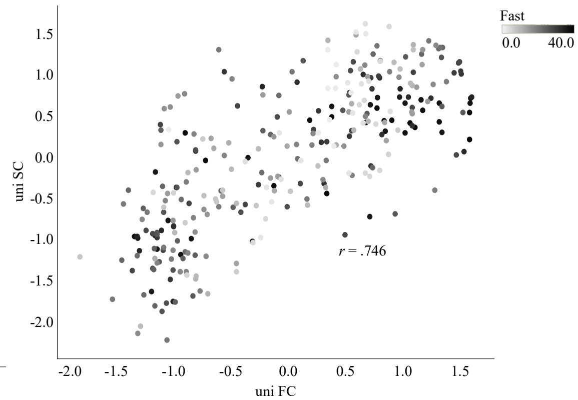
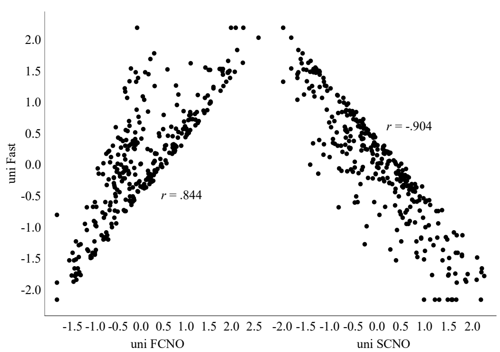

D1
D2
D3
r = .190
r = .021
Fast
0.0
40.0
r = .962
D1NO
D2NO
D3NO
r = 1.000
r = -1.000
r = -1.000
D1P
D2P
D3P
r = .576
r = -.759
r = -.127

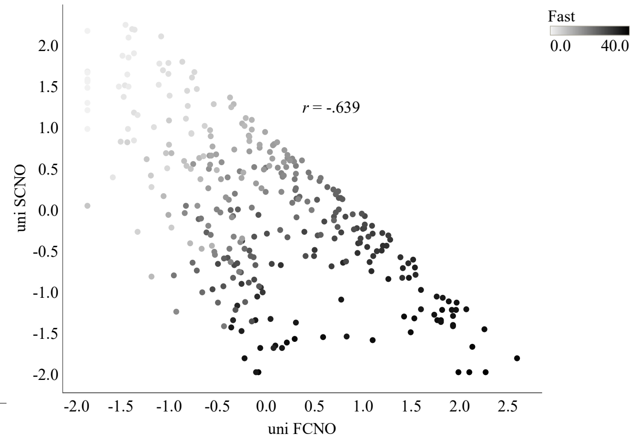
Figure 26. Scatterplots of Form 5.3

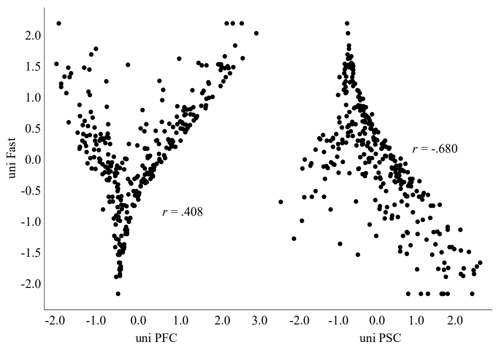


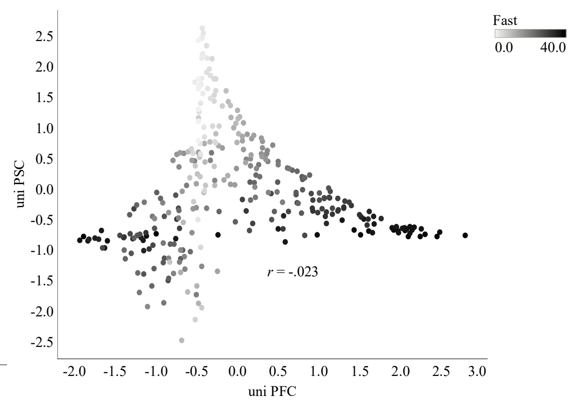

uni Fast
r = .045
r = -.072
uni FC
uni SC


Fast
0.0
40.0
uni SC
r = .746
uni FC


uni Fast
r = .844
r = -.904
uni FCNO
uni SCNO


Fast
0.0
40.0
uni SCNO
r = -.639
uni FCNO


uni Fast
r = .408
r = -.680
uni PFC
uni PSC


Fast
0.0
40.0
uni PSC
r = -.023
uni PFC

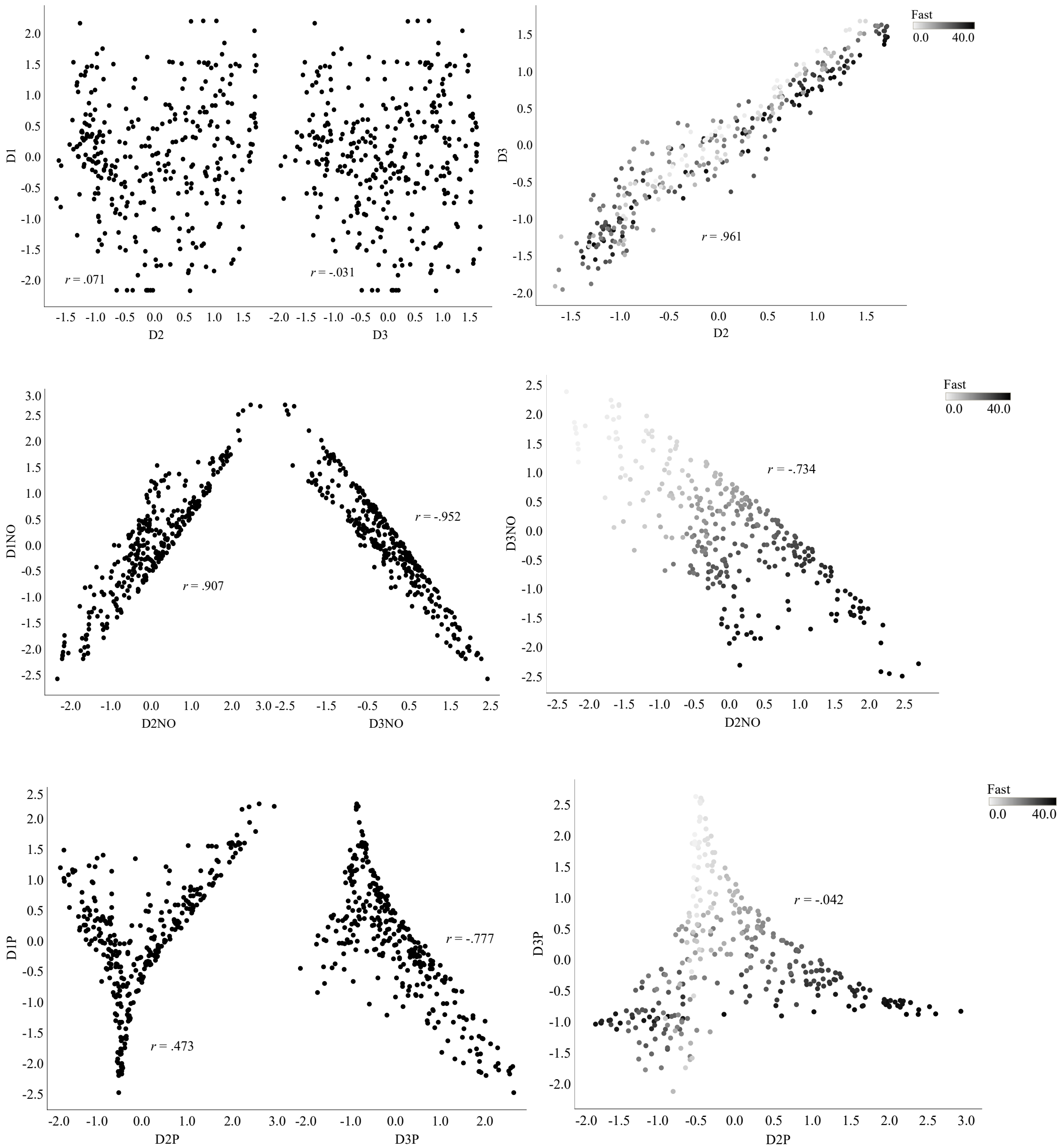

D1
D2
D3
r = .071
r = -.031
r = .961
Fast
0.0
40.0
D1NO
D2NO
D3NO
r = .907
r = -.952
r = -.734
D1P
D2P
D3P
r = .473
r = -.777
r = -.042

Table 30. Correlations between Raw Scores and Theta Values for Form 3.2

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.23 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.23 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.43 | 0.72 | -0.72 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.57 | -0.87 | 0.87 | -0.50 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.81 | 0.62 | -0.62 | -0.09 | -0.70 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.56 | -0.50 | 0.50 | -0.60 | 0.01 | -0.03 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.82 | 0.13 | -0.13 | 0.78 | 0.08 | -0.70 | -0.41 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.78 | -0.52 | 0.52 | -0.14 | 0.87 | -0.60 | -0.48 | 0.27 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.73 | 0.05 | -0.05 | 0.56 | 0.19 | -0.57 | -0.44 | 0.77 | 0.38 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.84 | -0.24 | 0.24 | 0.25 | 0.58 | -0.63 | -0.54 | 0.58 | 0.77 | 0.54 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.24 | 0.98 | -0.98 | 0.71 | -0.87 | 0.61 | -0.46 | 0.13 | -0.54 | 0.09 | -0.26 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.91 | 0.00 | 0.00 | 0.67 | 0.29 | -0.75 | -0.49 | 0.92 | 0.51 | 0.94 | 0.65 | 0.00 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.90 | -0.24 | 0.24 | 0.26 | 0.63 | -0.63 | -0.64 | 0.59 | 0.86 | 0.58 | 0.94 | -0.25 | 0.71 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.31 | 0.79 | -0.79 | 0.97 | -0.58 | 0.05 | -0.60 | 0.67 | -0.21 | 0.55 | 0.16 | 0.81 | 0.61 | 0.17 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.56 | -0.88 | 0.88 | -0.49 | 0.98 | -0.72 | 0.07 | 0.10 | 0.82 | 0.19 | 0.61 | -0.88 | 0.28 | 0.63 | -0.57 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.78 | 0.16 | -0.16 | 0.79 | 0.04 | -0.67 | -0.39 | 0.99 | 0.22 | 0.73 | 0.54 | 0.15 | 0.89 | 0.55 | 0.68 | 0.06 | 1.00 | | | | | | | | | | |
| uni PSC | 0.71 | -0.59 | 0.59 | -0.23 | 0.90 | -0.58 | -0.40 | 0.20 | 0.98 | 0.31 | 0.70 | -0.61 | 0.44 | 0.80 | -0.31 | 0.85 | 0.15 | 1.00 | | | | | | | | | |
| D1 | -0.25 | 0.98 | -0.98 | 0.71 | -0.88 | 0.61 | -0.44 | 0.12 | -0.55 | 0.09 | -0.27 | 1.00 | 0.00 | -0.27 | 0.80 | -0.89 | 0.15 | -0.62 | 1.00 | | | | | | | | |
| D2 | 0.95 | -0.02 | 0.02 | 0.60 | 0.36 | -0.71 | -0.62 | 0.88 | 0.62 | 0.81 | 0.75 | -0.02 | 0.96 | 0.84 | 0.51 | 0.36 | 0.85 | 0.55 | -0.03 | 1.00 | | | | | | | |
| D3 | 0.98 | -0.22 | 0.22 | 0.40 | 0.58 | -0.77 | -0.58 | 0.77 | 0.79 | 0.70 | 0.86 | -0.24 | 0.88 | 0.94 | 0.29 | 0.58 | 0.74 | 0.73 | -0.25 | 0.96 | 1.00 | | | | | | |
| D1NO | -0.23 | 0.96 | -0.96 | 0.75 | -0.90 | 0.54 | -0.38 | 0.19 | -0.60 | 0.12 | -0.30 | 0.98 | 0.05 | -0.31 | 0.83 | -0.91 | 0.23 | -0.67 | 0.99 | -0.01 | -0.25 | 1.00 | | | | | |
| D2NO | 0.23 | 0.83 | -0.83 | 0.92 | -0.62 | 0.15 | -0.61 | 0.57 | -0.25 | 0.49 | 0.12 | 0.86 | 0.52 | 0.12 | 0.99 | -0.61 | 0.58 | -0.35 | 0.86 | 0.43 | 0.21 | 0.87 | 1.00 | | | | |
| D3NO | 0.50 | -0.92 | 0.92 | -0.53 | 0.96 | -0.72 | 0.18 | 0.08 | 0.75 | 0.15 | 0.53 | -0.93 | 0.25 | 0.54 | -0.62 | 0.99 | 0.03 | 0.79 | -0.94 | 0.30 | 0.51 | -0.95 | -0.67 | 1.00 | | | |
| D1P | -0.26 | 0.97 | -0.97 | 0.71 | -0.90 | 0.59 | -0.39 | 0.14 | -0.59 | 0.09 | -0.28 | 0.99 | 0.00 | -0.29 | 0.81 | -0.90 | 0.17 | -0.67 | 1.00 | -0.05 | -0.26 | 0.99 | 0.86 | -0.94 | 1.00 | | |
| D2P | 0.79 | 0.17 | -0.17 | 0.80 | 0.03 | -0.66 | -0.41 | 0.99 | 0.23 | 0.73 | 0.55 | 0.17 | 0.89 | 0.57 | 0.69 | 0.05 | 1.00 | 0.16 | 0.17 | 0.85 | 0.74 | 0.24 | 0.59 | 0.02 | 0.18 | 1.00 | |
| D3P | 0.71 | -0.65 | 0.65 | -0.27 | 0.93 | -0.63 | -0.33 | 0.20 | 0.97 | 0.29 | 0.69 | -0.68 | 0.45 | 0.78 | -0.36 | 0.88 | 0.16 | 1.00 | -0.68 | 0.54 | 0.72 | -0.73 | -0.41 | 0.83 | -0.72 | 0.16 | 1.00 |

Table 31. Correlations between Raw Scores and Theta Values for Form 3.3

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.26 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.26 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.40 | 0.73 | -0.73 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.56 | -0.90 | 0.90 | -0.53 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.85 | 0.60 | -0.60 | -0.12 | -0.68 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.54 | -0.45 | 0.45 | -0.57 | 0.01 | 0.01 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.81 | 0.15 | -0.15 | 0.79 | 0.04 | -0.70 | -0.42 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.75 | -0.61 | 0.61 | -0.22 | 0.90 | -0.61 | -0.44 | 0.22 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.74 | 0.02 | -0.02 | 0.55 | 0.19 | -0.61 | -0.43 | 0.77 | 0.36 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.82 | -0.28 | 0.28 | 0.21 | 0.57 | -0.65 | -0.51 | 0.56 | 0.74 | 0.52 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.28 | 0.98 | -0.98 | 0.69 | -0.89 | 0.60 | -0.42 | 0.13 | -0.61 | 0.05 | -0.31 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.89 | -0.05 | 0.05 | 0.62 | 0.29 | -0.78 | -0.46 | 0.90 | 0.47 | 0.94 | 0.63 | -0.08 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.88 | -0.25 | 0.25 | 0.23 | 0.61 | -0.64 | -0.66 | 0.56 | 0.84 | 0.58 | 0.92 | -0.26 | 0.69 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.28 | 0.80 | -0.80 | 0.97 | -0.62 | 0.02 | -0.56 | 0.68 | -0.31 | 0.53 | 0.12 | 0.79 | 0.55 | 0.14 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.57 | -0.89 | 0.89 | -0.50 | 0.98 | -0.71 | 0.05 | 0.08 | 0.86 | 0.20 | 0.62 | -0.90 | 0.31 | 0.62 | -0.58 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.78 | 0.16 | -0.16 | 0.79 | 0.01 | -0.69 | -0.39 | 0.99 | 0.18 | 0.73 | 0.53 | 0.14 | 0.87 | 0.53 | 0.68 | 0.06 | 1.00 | | | | | | | | | | |
| uni PSC | 0.69 | -0.67 | 0.67 | -0.31 | 0.92 | -0.60 | -0.35 | 0.14 | 0.98 | 0.28 | 0.67 | -0.68 | 0.41 | 0.77 | -0.41 | 0.88 | 0.11 | 1.00 | | | | | | | | | |
| D1 | -0.30 | 0.98 | -0.98 | 0.69 | -0.90 | 0.61 | -0.40 | 0.12 | -0.63 | 0.04 | -0.33 | 1.00 | -0.08 | -0.28 | 0.79 | -0.91 | 0.13 | -0.70 | 1.00 | | | | | | | | |
| D2 | 0.94 | -0.10 | 0.10 | 0.52 | 0.40 | -0.75 | -0.59 | 0.84 | 0.62 | 0.81 | 0.73 | -0.12 | 0.96 | 0.83 | 0.43 | 0.41 | 0.80 | 0.56 | -0.13 | 1.00 | | | | | | | |
| D3 | 0.96 | -0.30 | 0.30 | 0.31 | 0.61 | -0.79 | -0.56 | 0.71 | 0.80 | 0.69 | 0.85 | -0.32 | 0.85 | 0.94 | 0.20 | 0.63 | 0.67 | 0.75 | -0.34 | 0.95 | 1.00 | | | | | | |
| D1NO | -0.27 | 0.97 | -0.97 | 0.74 | -0.92 | 0.54 | -0.33 | 0.20 | -0.68 | 0.08 | -0.34 | 0.98 | -0.01 | -0.32 | 0.82 | -0.93 | 0.22 | -0.75 | 0.99 | -0.11 | -0.34 | 1.00 | | | | | |
| D2NO | 0.19 | 0.84 | -0.84 | 0.93 | -0.66 | 0.13 | -0.56 | 0.58 | -0.35 | 0.47 | 0.07 | 0.85 | 0.46 | 0.09 | 0.99 | -0.63 | 0.59 | -0.45 | 0.85 | 0.34 | 0.12 | 0.87 | 1.00 | | | | |
| D3NO | 0.54 | -0.91 | 0.91 | -0.51 | 0.96 | -0.72 | 0.13 | 0.08 | 0.81 | 0.18 | 0.57 | -0.93 | 0.30 | 0.55 | -0.60 | 0.99 | 0.05 | 0.84 | -0.94 | 0.39 | 0.59 | -0.95 | -0.66 | 1.00 | | | |
| D1P | -0.31 | 0.97 | -0.97 | 0.70 | -0.92 | 0.58 | -0.34 | 0.14 | -0.68 | 0.04 | -0.33 | 0.99 | -0.07 | -0.32 | 0.80 | -0.92 | 0.16 | -0.75 | 0.99 | -0.15 | -0.36 | 0.99 | 0.86 | -0.95 | 1.00 | | |
| D2P | 0.79 | 0.18 | -0.18 | 0.80 | 0.00 | -0.68 | -0.41 | 0.99 | 0.18 | 0.73 | 0.54 | 0.15 | 0.87 | 0.54 | 0.69 | 0.05 | 1.00 | 0.11 | 0.15 | 0.81 | 0.68 | 0.23 | 0.60 | 0.04 | 0.17 | 1.00 | |
| D3P | 0.68 | -0.73 | 0.73 | -0.36 | 0.95 | -0.64 | -0.26 | 0.14 | 0.97 | 0.26 | 0.65 | -0.75 | 0.41 | 0.74 | -0.46 | 0.92 | 0.11 | 0.99 | -0.76 | 0.54 | 0.73 | -0.80 | -0.52 | 0.88 | -0.81 | 0.11 | 1.00 |

Table 32. Correlations between Raw Scores and Theta Values for Form 4.2

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.09 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.09 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.50 | 0.77 | -0.77 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.43 | -0.89 | 0.89 | -0.57 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.84 | 0.44 | -0.44 | -0.23 | -0.56 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.64 | -0.47 | 0.47 | -0.58 | 0.00 | 0.11 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.81 | 0.34 | -0.34 | 0.86 | -0.13 | -0.69 | -0.49 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.68 | -0.57 | 0.57 | -0.23 | 0.89 | -0.55 | -0.46 | 0.12 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.85 | 0.04 | -0.04 | 0.56 | 0.23 | -0.72 | -0.52 | 0.79 | 0.44 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.81 | -0.09 | 0.09 | 0.31 | 0.45 | -0.57 | -0.67 | 0.53 | 0.71 | 0.60 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.10 | 0.98 | -0.98 | 0.75 | -0.88 | 0.44 | -0.44 | 0.32 | -0.57 | 0.05 | -0.11 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.90 | 0.07 | -0.07 | 0.64 | 0.21 | -0.79 | -0.54 | 0.88 | 0.43 | 0.95 | 0.64 | 0.05 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.88 | -0.11 | 0.11 | 0.32 | 0.52 | -0.62 | -0.74 | 0.57 | 0.80 | 0.68 | 0.94 | -0.13 | 0.71 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.43 | 0.81 | -0.81 | 0.98 | -0.61 | -0.13 | -0.59 | 0.79 | -0.26 | 0.54 | 0.26 | 0.82 | 0.60 | 0.28 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.43 | -0.88 | 0.88 | -0.55 | 0.98 | -0.57 | 0.03 | -0.11 | 0.86 | 0.22 | 0.47 | -0.89 | 0.20 | 0.53 | -0.60 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.78 | 0.36 | -0.36 | 0.87 | -0.16 | -0.67 | -0.46 | 0.99 | 0.07 | 0.74 | 0.49 | 0.35 | 0.84 | 0.52 | 0.80 | -0.15 | 1.00 | | | | | | | | | | |
| uni PSC | 0.62 | -0.63 | 0.63 | -0.31 | 0.91 | -0.53 | -0.38 | 0.05 | 0.99 | 0.39 | 0.63 | -0.64 | 0.38 | 0.72 | -0.35 | 0.88 | 0.01 | 1.00 | | | | | | | | | |
| D1 | -0.10 | 0.98 | -0.98 | 0.76 | -0.88 | 0.44 | -0.44 | 0.33 | -0.58 | 0.06 | -0.12 | 1.00 | 0.06 | -0.14 | 0.82 | -0.89 | 0.35 | -0.64 | 1.00 | | | | | | | | |
| D2 | 0.94 | 0.10 | -0.10 | 0.63 | 0.23 | -0.73 | -0.67 | 0.85 | 0.52 | 0.87 | 0.72 | 0.09 | 0.96 | 0.83 | 0.59 | 0.23 | 0.82 | 0.45 | 0.10 | 1.00 | | | | | | | |
| D3 | 0.96 | -0.07 | 0.07 | 0.47 | 0.43 | -0.77 | -0.67 | 0.75 | 0.69 | 0.80 | 0.83 | -0.09 | 0.89 | 0.93 | 0.41 | 0.43 | 0.71 | 0.63 | -0.09 | 0.96 | 1.00 | | | | | | |
| D1NO | -0.08 | 0.97 | -0.97 | 0.79 | -0.91 | 0.36 | -0.36 | 0.40 | -0.64 | 0.09 | -0.15 | 0.98 | 0.11 | -0.18 | 0.85 | -0.92 | 0.43 | -0.70 | 0.99 | 0.11 | -0.09 | 1.00 | | | | | |
| D2NO | 0.36 | 0.85 | -0.85 | 0.95 | -0.65 | -0.03 | -0.60 | 0.72 | -0.30 | 0.48 | 0.22 | 0.87 | 0.52 | 0.24 | 0.99 | -0.64 | 0.73 | -0.38 | 0.87 | 0.52 | 0.34 | 0.89 | 1.00 | | | | |
| D3NO | 0.37 | -0.92 | 0.92 | -0.59 | 0.96 | -0.57 | 0.14 | -0.14 | 0.79 | 0.18 | 0.39 | -0.94 | 0.16 | 0.43 | -0.65 | 0.99 | -0.18 | 0.82 | -0.94 | 0.17 | 0.36 | -0.95 | -0.70 | 1.00 | | | |
| D1P | -0.11 | 0.97 | -0.97 | 0.77 | -0.91 | 0.41 | -0.37 | 0.36 | -0.63 | 0.05 | -0.14 | 0.99 | 0.07 | -0.17 | 0.83 | -0.91 | 0.38 | -0.70 | 0.99 | 0.08 | -0.11 | 1.00 | 0.88 | -0.95 | 1.00 | | |
| D2P | 0.77 | 0.38 | -0.38 | 0.88 | -0.18 | -0.65 | -0.49 | 0.99 | 0.07 | 0.74 | 0.49 | 0.38 | 0.84 | 0.53 | 0.81 | -0.17 | 1.00 | 0.01 | 0.38 | 0.82 | 0.71 | 0.45 | 0.75 | -0.20 | 0.41 | 1.00 | |
| D3P | 0.60 | -0.70 | 0.70 | -0.35 | 0.94 | -0.57 | -0.29 | 0.04 | 0.97 | 0.36 | 0.60 | -0.71 | 0.37 | 0.69 | -0.41 | 0.92 | 0.00 | 0.99 | -0.72 | 0.42 | 0.61 | -0.76 | -0.45 | 0.87 | -0.77 | -0.01 | 1.00 |

Table 33. Correlations between Raw Scores and Theta Values for Form 4.3

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.22 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.22 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.43 | 0.74 | -0.74 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.54 | -0.89 | 0.89 | -0.53 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.85 | 0.55 | -0.55 | -0.16 | -0.65 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.59 | -0.43 | 0.43 | -0.57 | -0.03 | 0.08 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.81 | 0.20 | -0.20 | 0.81 | 0.00 | -0.70 | -0.46 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.74 | -0.60 | 0.60 | -0.22 | 0.90 | -0.61 | -0.46 | 0.20 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.78 | -0.01 | 0.01 | 0.53 | 0.24 | -0.67 | -0.45 | 0.78 | 0.41 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.82 | -0.22 | 0.22 | 0.24 | 0.54 | -0.63 | -0.59 | 0.54 | 0.74 | 0.56 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.23 | 0.98 | -0.98 | 0.71 | -0.88 | 0.55 | -0.41 | 0.19 | -0.60 | 0.03 | -0.25 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.91 | -0.10 | 0.10 | 0.57 | 0.36 | -0.81 | -0.48 | 0.87 | 0.53 | 0.96 | 0.64 | -0.10 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.88 | -0.18 | 0.18 | 0.30 | 0.56 | -0.62 | -0.72 | 0.58 | 0.80 | 0.60 | 0.94 | -0.19 | 0.71 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.36 | 0.78 | -0.78 | 0.97 | -0.57 | -0.07 | -0.58 | 0.74 | -0.25 | 0.54 | 0.19 | 0.78 | 0.54 | 0.25 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.55 | -0.89 | 0.89 | -0.50 | 0.98 | -0.68 | 0.00 | 0.04 | 0.87 | 0.23 | 0.58 | -0.89 | 0.35 | 0.57 | -0.55 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.77 | 0.23 | -0.23 | 0.82 | -0.04 | -0.67 | -0.43 | 0.99 | 0.15 | 0.73 | 0.50 | 0.22 | 0.82 | 0.54 | 0.75 | -0.01 | 1.00 | | | | | | | | | | |
| uni PSC | 0.68 | -0.66 | 0.66 | -0.30 | 0.92 | -0.59 | -0.38 | 0.13 | 0.98 | 0.33 | 0.67 | -0.67 | 0.48 | 0.73 | -0.35 | 0.89 | 0.08 | 1.00 | | | | | | | | | |
| D1 | -0.24 | 0.98 | -0.98 | 0.71 | -0.89 | 0.56 | -0.40 | 0.18 | -0.61 | 0.02 | -0.26 | 1.00 | -0.11 | -0.20 | 0.78 | -0.90 | 0.21 | -0.68 | 1.00 | | | | | | | | |
| D2 | 0.95 | -0.12 | 0.12 | 0.50 | 0.42 | -0.79 | -0.59 | 0.83 | 0.63 | 0.81 | 0.73 | -0.13 | 0.97 | 0.83 | 0.44 | 0.42 | 0.79 | 0.57 | -0.14 | 1.00 | | | | | | | |
| D3 | 0.96 | -0.20 | 0.20 | 0.40 | 0.53 | -0.77 | -0.64 | 0.74 | 0.75 | 0.72 | 0.84 | -0.21 | 0.88 | 0.94 | 0.33 | 0.54 | 0.70 | 0.69 | -0.22 | 0.96 | 1.00 | | | | | | |
| D1NO | -0.22 | 0.96 | -0.96 | 0.75 | -0.91 | 0.48 | -0.32 | 0.26 | -0.66 | 0.08 | -0.29 | 0.98 | -0.04 | -0.24 | 0.81 | -0.92 | 0.30 | -0.73 | 0.99 | -0.11 | -0.21 | 1.00 | | | | | |
| D2NO | 0.28 | 0.82 | -0.82 | 0.94 | -0.61 | 0.04 | -0.59 | 0.66 | -0.28 | 0.49 | 0.15 | 0.84 | 0.46 | 0.21 | 0.99 | -0.60 | 0.67 | -0.38 | 0.83 | 0.36 | 0.27 | 0.86 | 1.00 | | | | |
| D3NO | 0.49 | -0.92 | 0.92 | -0.53 | 0.96 | -0.68 | 0.11 | 0.02 | 0.80 | 0.19 | 0.52 | -0.93 | 0.33 | 0.49 | -0.59 | 0.99 | -0.03 | 0.84 | -0.94 | 0.38 | 0.48 | -0.95 | -0.65 | 1.00 | | | |
| D1P | -0.25 | 0.97 | -0.97 | 0.73 | -0.91 | 0.52 | -0.34 | 0.22 | -0.66 | 0.04 | -0.27 | 0.99 | -0.09 | -0.23 | 0.79 | -0.91 | 0.25 | -0.73 | 0.99 | -0.14 | -0.23 | 1.00 | 0.85 | -0.95 | 1.00 | | |
| D2P | 0.77 | 0.25 | -0.25 | 0.84 | -0.05 | -0.66 | -0.46 | 0.99 | 0.15 | 0.73 | 0.50 | 0.24 | 0.82 | 0.55 | 0.76 | -0.02 | 1.00 | 0.08 | 0.23 | 0.79 | 0.71 | 0.31 | 0.69 | -0.04 | 0.27 | 1.00 | |
| D3P | 0.67 | -0.73 | 0.73 | -0.34 | 0.95 | -0.63 | -0.29 | 0.12 | 0.97 | 0.30 | 0.65 | -0.74 | 0.49 | 0.70 | -0.40 | 0.92 | 0.08 | 0.99 | -0.75 | 0.56 | 0.67 | -0.78 | -0.45 | 0.88 | -0.79 | 0.08 | 1.00 |

Table 34. Correlations between Raw Scores and Theta Values for Form 5.2

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | 0.02 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | -0.02 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.51 | 0.83 | -0.83 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.35 | -0.88 | 0.88 | -0.62 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.80 | 0.39 | -0.39 | -0.20 | -0.51 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.69 | -0.50 | 0.50 | -0.61 | 0.03 | 0.12 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.78 | 0.47 | -0.47 | 0.88 | -0.25 | -0.63 | -0.53 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.65 | -0.54 | 0.54 | -0.26 | 0.87 | -0.51 | -0.46 | 0.04 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.73 | 0.27 | -0.27 | 0.64 | -0.04 | -0.58 | -0.51 | 0.78 | 0.22 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.68 | -0.10 | 0.10 | 0.16 | 0.45 | -0.44 | -0.59 | 0.33 | 0.69 | 0.40 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | 0.01 | 0.98 | -0.98 | 0.81 | -0.88 | 0.37 | -0.46 | 0.46 | -0.56 | 0.31 | -0.17 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | -0.09 | 0.38 | -0.38 | 0.31 | -0.42 | 0.15 | -0.03 | 0.17 | -0.36 | 0.50 | -0.10 | 0.49 | 1.00 | | | | | | | | | | | | | | |
| uni SC | -0.09 | -0.24 | 0.24 | -0.28 | 0.23 | 0.05 | 0.08 | -0.25 | 0.17 | -0.10 | 0.46 | -0.32 | -0.05 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.43 | 0.86 | -0.86 | 0.97 | -0.67 | -0.10 | -0.59 | 0.82 | -0.30 | 0.66 | 0.09 | 0.87 | 0.43 | -0.31 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.32 | -0.88 | 0.88 | -0.63 | 0.98 | -0.49 | 0.06 | -0.27 | 0.85 | -0.07 | 0.49 | -0.90 | -0.43 | 0.32 | -0.68 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.74 | 0.51 | -0.51 | 0.89 | -0.30 | -0.59 | -0.51 | 0.99 | -0.02 | 0.73 | 0.25 | 0.51 | 0.16 | -0.33 | 0.83 | -0.33 | 1.00 | | | | | | | | | | |
| uni PSC | 0.57 | -0.61 | 0.61 | -0.35 | 0.90 | -0.49 | -0.36 | -0.04 | 0.98 | 0.11 | 0.60 | -0.64 | -0.45 | 0.16 | -0.41 | 0.88 | -0.09 | 1.00 | | | | | | | | | |
| D1 | 0.01 | 0.98 | -0.98 | 0.82 | -0.88 | 0.36 | -0.46 | 0.47 | -0.56 | 0.32 | -0.17 | 1.00 | 0.48 | -0.32 | 0.88 | -0.90 | 0.52 | -0.64 | 1.00 | | | | | | | | |
| D2 | 0.93 | 0.20 | -0.20 | 0.64 | 0.15 | -0.71 | -0.69 | 0.84 | 0.47 | 0.78 | 0.56 | 0.18 | 0.00 | -0.13 | 0.58 | 0.13 | 0.80 | 0.38 | 0.19 | 1.00 | | | | | | | |
| D3 | 0.96 | 0.04 | -0.04 | 0.48 | 0.34 | -0.72 | -0.71 | 0.73 | 0.65 | 0.68 | 0.69 | 0.02 | -0.10 | -0.07 | 0.41 | 0.32 | 0.67 | 0.57 | 0.02 | 0.96 | 1.00 | | | | | | |
| D1NO | 0.07 | 0.94 | -0.94 | 0.87 | -0.88 | 0.21 | -0.37 | 0.58 | -0.60 | 0.40 | -0.19 | 0.97 | 0.51 | -0.37 | 0.92 | -0.90 | 0.63 | -0.69 | 0.97 | 0.25 | 0.06 | 1.00 | 1.00 | | | | |
| D2NO | 0.07 | 0.94 | -0.94 | 0.87 | -0.88 | 0.21 | -0.37 | 0.58 | -0.60 | 0.40 | -0.19 | 0.97 | 0.51 | -0.37 | 0.92 | -0.90 | 0.63 | -0.69 | 0.97 | 0.25 | 0.06 | 1.00 | 1.00 | | | | |
| D3NO | -0.07 | -0.94 | 0.94 | -0.87 | 0.88 | -0.21 | 0.37 | -0.58 | 0.60 | -0.40 | 0.19 | -0.97 | -0.51 | 0.37 | -0.92 | 0.90 | -0.63 | 0.69 | -0.97 | -0.25 | -0.06 | -1.00 | -1.00 | 1.00 | | | |
| D1P | 0.00 | 0.97 | -0.97 | 0.83 | -0.90 | 0.33 | -0.39 | 0.50 | -0.61 | 0.34 | -0.20 | 0.99 | 0.51 | -0.33 | 0.89 | -0.92 | 0.54 | -0.70 | 0.99 | 0.18 | 0.00 | 0.99 | 0.99 | -0.99 | 1.00 | | |
| D2P | 0.73 | 0.54 | -0.54 | 0.91 | -0.33 | -0.55 | -0.53 | 0.98 | -0.03 | 0.72 | 0.24 | 0.54 | 0.17 | -0.36 | 0.85 | -0.36 | 1.00 | -0.10 | 0.55 | 0.79 | 0.67 | 0.66 | 0.66 | -0.66 | 0.58 | 1.00 | |
| D3P | 0.55 | -0.68 | 0.68 | -0.40 | 0.93 | -0.53 | -0.27 | -0.06 | 0.96 | 0.07 | 0.57 | -0.71 | -0.49 | 0.17 | -0.47 | 0.91 | -0.11 | 0.99 | -0.71 | 0.36 | 0.54 | -0.74 | -0.74 | 0.74 | -0.76 | -0.13 | 1.00 |

Table 35. Correlations between Raw Scores and Theta Values for Form 5.3

| | Correct | Fast | Slow | FC | SC | FI | SI | FC-FI | SC-SI | FCF | SCS | uni Fast | uni FC | uni SC | uni FCNO | uni SCNO | uni PFC | uni PSC | D1 | D2 | D3 | D1NO | D2NO | D3NO | D1P | D2P | D3P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correct | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fast | -0.08 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| Slow | 0.08 | -1.00 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| FC | 0.48 | 0.79 | -0.79 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| SC | 0.43 | -0.89 | 0.89 | -0.58 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| FI | -0.85 | 0.43 | -0.43 | -0.21 | -0.56 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| SI | -0.68 | -0.43 | 0.43 | -0.59 | -0.02 | 0.18 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| FC-FI | 0.80 | 0.38 | -0.38 | 0.87 | -0.15 | -0.67 | -0.54 | 1.00 | | | | | | | | | | | | | | | | | | | |
| SC-SI | 0.69 | -0.60 | 0.60 | -0.25 | 0.90 | -0.58 | -0.46 | 0.11 | 1.00 | | | | | | | | | | | | | | | | | | |
| FCF | 0.82 | 0.09 | -0.09 | 0.56 | 0.18 | -0.69 | -0.56 | 0.78 | 0.41 | 1.00 | | | | | | | | | | | | | | | | | |
| SCS | 0.83 | -0.13 | 0.13 | 0.29 | 0.47 | -0.63 | -0.67 | 0.54 | 0.71 | 0.60 | 1.00 | | | | | | | | | | | | | | | | |
| uni Fast | -0.09 | 0.98 | -0.98 | 0.78 | -0.88 | 0.42 | -0.41 | 0.38 | -0.60 | 0.12 | -0.15 | 1.00 | | | | | | | | | | | | | | | |
| uni FC | 0.91 | 0.04 | -0.04 | 0.60 | 0.25 | -0.81 | -0.57 | 0.85 | 0.48 | 0.96 | 0.67 | 0.05 | 1.00 | | | | | | | | | | | | | | |
| uni SC | 0.90 | -0.06 | 0.06 | 0.37 | 0.46 | -0.64 | -0.78 | 0.61 | 0.75 | 0.70 | 0.95 | -0.07 | 0.75 | 1.00 | | | | | | | | | | | | | |
| uni FCNO | 0.40 | 0.83 | -0.83 | 0.98 | -0.64 | -0.11 | -0.57 | 0.80 | -0.31 | 0.55 | 0.23 | 0.84 | 0.55 | 0.30 | 1.00 | | | | | | | | | | | | |
| uni SCNO | 0.42 | -0.90 | 0.90 | -0.58 | 0.98 | -0.57 | 0.02 | -0.15 | 0.86 | 0.18 | 0.49 | -0.90 | 0.23 | 0.46 | -0.64 | 1.00 | | | | | | | | | | | |
| uni PFC | 0.75 | 0.41 | -0.41 | 0.87 | -0.20 | -0.64 | -0.51 | 0.99 | 0.05 | 0.73 | 0.49 | 0.41 | 0.81 | 0.55 | 0.81 | -0.20 | 1.00 | | | | | | | | | | |
| uni PSC | 0.62 | -0.67 | 0.67 | -0.34 | 0.92 | -0.56 | -0.36 | 0.03 | 0.98 | 0.33 | 0.63 | -0.68 | 0.41 | 0.66 | -0.41 | 0.89 | -0.02 | 1.00 | | | | | | | | | |
| D1 | -0.09 | 0.98 | -0.98 | 0.78 | -0.88 | 0.41 | -0.41 | 0.38 | -0.60 | 0.12 | -0.15 | 1.00 | 0.05 | -0.07 | 0.85 | -0.90 | 0.41 | -0.68 | 1.00 | | | | | | | | |
| D2 | 0.93 | 0.07 | -0.07 | 0.59 | 0.26 | -0.77 | -0.67 | 0.84 | 0.52 | 0.87 | 0.73 | 0.07 | 0.97 | 0.84 | 0.53 | 0.24 | 0.79 | 0.45 | 0.07 | 1.00 | | | | | | | |
| D3 | 0.95 | -0.03 | 0.03 | 0.48 | 0.39 | -0.75 | -0.72 | 0.75 | 0.67 | 0.80 | 0.84 | -0.03 | 0.89 | 0.95 | 0.41 | 0.38 | 0.70 | 0.58 | -0.03 | 0.96 | 1.00 | | | | | | |
| D1NO | -0.07 | 0.97 | -0.97 | 0.82 | -0.90 | 0.34 | -0.34 | 0.44 | -0.65 | 0.15 | -0.17 | 0.99 | 0.09 | -0.11 | 0.87 | -0.92 | 0.48 | -0.72 | 0.99 | 0.09 | -0.03 | 1.00 | | | | | |
| D2NO | 0.33 | 0.87 | -0.87 | 0.95 | -0.68 | -0.02 | -0.57 | 0.73 | -0.35 | 0.49 | 0.18 | 0.89 | 0.48 | 0.26 | 0.99 | -0.68 | 0.75 | -0.45 | 0.89 | 0.46 | 0.35 | 0.91 | 1.00 | | | | |
| D3NO | 0.35 | -0.93 | 0.93 | -0.62 | 0.96 | -0.56 | 0.13 | -0.18 | 0.80 | 0.12 | 0.41 | -0.94 | 0.19 | 0.37 | -0.69 | 0.99 | -0.24 | 0.84 | -0.94 | 0.18 | 0.30 | -0.95 | -0.73 | 1.00 | | | |
| D1P | -0.09 | 0.98 | -0.98 | 0.80 | -0.91 | 0.38 | -0.35 | 0.41 | -0.65 | 0.12 | -0.16 | 0.99 | 0.06 | -0.10 | 0.86 | -0.92 | 0.45 | -0.73 | 0.99 | 0.07 | -0.04 | 1.00 | 0.90 | -0.95 | 1.00 | | |
| D2P | 0.75 | 0.44 | -0.44 | 0.89 | -0.22 | -0.62 | -0.53 | 0.99 | 0.04 | 0.72 | 0.49 | 0.44 | 0.80 | 0.56 | 0.83 | -0.22 | 1.00 | -0.03 | 0.44 | 0.79 | 0.70 | 0.51 | 0.77 | -0.26 | 0.47 | 1.00 | |
| D3P | 0.60 | -0.73 | 0.73 | -0.38 | 0.95 | -0.60 | -0.29 | 0.02 | 0.97 | 0.31 | 0.61 | -0.74 | 0.41 | 0.62 | -0.46 | 0.92 | -0.03 | 0.99 | -0.74 | 0.42 | 0.56 | -0.77 | -0.51 | 0.88 | -0.78 | -0.04 | 1.00 |