APPLYING ITEM RESPONSE THEORY TO MEASURE DRIVERS' PERCEIVED COMPLEXITY OF ROADWAY ENVIRONMENTS

A Thesis
Presented to
The Academic Faculty

By

Faaiqa Atiyya Shaw

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science in Psychology

Georgia Institute of Technology
December 2018

APPLYING ITEM RESPONSE THEORY TO MEASURE DRIVERS' PERCEIVED COMPLEXITY OF ROADWAY ENVIRONMENTS

Approved by:

Dr. James Roberts, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Michael P. Hunter
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Susan Embretson
School of Psychology
Georgia Institute of Technology

Date Approved: [August 14, 2018]

بسم الله

ACKNOWLEDGEMENTS

I would like to thank my graduate school advisors, Dr. James Roberts, Dr. Michael Hunter, and Dr. Patricia Mokhtarian for supporting the undertaking of this degree in a myriad of ways, but most particularly for their patience with me as I sought to balance my endeavors between their programs of study. I would also like to thank Dr. Susan Embretson, who first taught me the principles of IRT and who supported my strange application of her beloved models from day one. I am grateful for the opportunity to work with and learn from each of these Professors.

Portions of this work (i.e. experiment development and implementation) were sponsored by the Georgia Department of Transportation (GDOT) and the National Center for Transportation Systems Productivity and Management (NCTSPM). Atiyya Shaw was funded under National Science Foundation Graduate Research Fellowship (NSFGRF), and this material is based upon work supported by the NSFGRF Program under Grants DGE-1148903 and DGE-1650044. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of GDOT, NCTSPM, or the NSF.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

2-PLM	Two Parameter Logistic Model
DETECT	Dimensionality Evaluation to Enumerate Contributing Traits
GPCM	Generalized Partial Credit Model
GRM	Graded Response Model
ICC	Item Characteristic Curve
IRB	Institutional Review Board
IRT	Item Response Theory
MAR	Missing at Random
MIRT	Multidimensional Item Response Theory
NADS	National Advanced Driving Simulator
PAF	Principal Axis Factoring

SUMMARY

Roadway environments constitute visually complex systems within which users make split-second critical decisions on a daily basis. As such, understanding transportation system user perceptions and performance across varied roadway environments is crucial for a broad array of transportation research and engineering purposes (e.g. understanding safety data trends, informing roadway design guidelines, etc.). This thesis applies item response theory (IRT) to identify and interpret the dimensions present that influence drivers' perceived complexity of roadway environments. We find that a four dimensional polytomous Graded Response Model best measures this data, and were able to ascertain that participants' perceived complexity ratings were most affected by their perception of freeway and urban environments, as well the visibility and traffic conditions of the particular roadway. This study enables not only an understanding of the factors that influence driver perception of the built environment, but demonstrates an application of multidimensional, polytomous IRT to study transportation system user perceptions; one of the first known implementations of multidimensional IRT within transportation engineering.

CHAPTER 1. INTRODUCTION

The goal of transportation systems and/or infrastructure is to connect people to places, goods, and activities in a safe, efficient, and just manner. As such, transportation systems are dynamically connected to societal wellbeing with regards to the interdependence of the infrastructure with its users. These connections have long made the study of ***people*** from a transportation systems perspective an extremely critical part of transportation research and engineering (Meyer & Miller, 2017). Meanwhile, item response theory (IRT) within the field of psychometrics represents one of the most advanced approaches for estimating traits/performance of people; however, only its simplest forms (i.e. unidimensional IRT for binary items) have been implemented in a handful of applications within transportation engineering (Federal Motor Carrier Safety Administration, 2017; Greenwood, 2015; Rowell, Gagliano, & Goodchild, 2014). As such, a methodological contribution of this work is to integrate advanced IRT models not typically used in transportation engineering to data gathered from a transportation infrastructure perspective. Specifically, this thesis will demonstrate an application of multidimensional IRT (MIRT) models to measure drivers' perceived complexity of roadway environments.

As roadway and in-vehicle environments become increasingly cluttered and complex, users of these systems (cyclists, motorists, pedestrians, etc.) are experiencing greater information processing demands, even as vehicle handling demands (i.e. automated functions/technologies on vehicles) begin to decrease. Driver perception is a critical aspect of how users' process these complex environments, with derived parameters such as perception-reaction time (i.e. time taken from detection of hazard to response that is implemented by the driver) widely used within road design and traffic engineering to calculate important inputs such as road curvature, signal timing, etc. (American Association of State Highway and Transportation Officials, 2010; Dewar & Olson, 2002; Elvik, 2006; Olson & Farber, 2003; Roess, 2011). This study obtained drivers' perceived complexities of 100 varied roadway environments, with the goal of understanding factors that influence drivers' complexity judgments. Given that perception is an important step in the driver reaction process; we posit that understanding drivers' perceived complexity may aid engineers and researchers in understanding performance differences across various environments. Researchers in the social sciences have been studying links between perception and performance for decades, but these efforts have not explicitly extended into transportation engineering. As such, this thesis applies advanced psychometric methods to study drivers' judged complexity of roadway environments, combining the aforementioned methodological contribution with an applied transportation outcome.

CHAPTER 2. LITERATURE REVIEW

The transportation engineering literature has confirmed that a wide array of roadway factors influence driver *performance*, using methods ranging from naturalistic studies, to crash data analyses, and most commonly -- driving simulator experiments (the latter of which rely on metrics like lane deviations, speed adherence, eye fixations, etc.). However, there is markedly less work that has focused on the effects of environmental factors on *perception*. Here, we discuss a few of the salient roadway environment factors that have been studied with regards to driver performance. Generally, we know that the length of time needed for visual search increases as a scene becomes more cluttered (i.e., the number of objects increases) (L. Zhang & Lin, 2013), and these findings have been extended to roadway environments with Ho et al. (2001) finding that visual clutter increases reaction time and error rates increase when participants are asked to locate specific stimuli in roadway environments. More specifically, it has also been shown that increased visual clutter in the form of roadway objects like billboards and signs (Edquist, Horberry, Hosking, & Johnston, 2011; Horberry, Anderson, Regan, Triggs, & Brown, 2006; Shaw, Park, et al., 2018; Young et al., 2009), as well as roadway configurations that are inherently more visually cluttered (such as intersections and urban environments), result in reduced performance as measured via reaction times, error rates, lateral control, speed, etc. (M. A. Abdel-Aty & Radwan, 2000; Cantin, Lavallière, Simoneau, & Teasdale, 2009; Edquist, Rudin-Brown, & Lenné, 2012; Hadi, Aruldhas, Chow, & Wattleworth, 1995; Ho et al., 2001; Kaber, Zhang, Jin, Mosaly, & Garner, 2012; Stinchcombe & Gagnon, 2010).

There are also a substantial number of studies that have examined specific roadway factors *individually*. Notably, it has been seen that increased traffic in proximity to the driver results in reduced performance on a range of metrics from increased workload, increased motor vehicle crashes, decreased speed adherence, etc. (M. Abdel-Aty, Keller, & Brady, 2005; M. A. Abdel-Aty & Radwan, 2000; Brookhuis, de Vries, & de Waard, 1991; Hadi et al., 1995; Kaber et al., 2012; Karlaftis & Golias, 2002; Milton & Mannering, 1998; Mohamedshah, Paniati, & Hobeika, 1993; Schiessl, 2008; Stinchcombe & Gagnon, 2010; Teh, Jamson, Carsten, & Jamson, 2014; Zeitlin, 1995). Work zones are commonly accepted as one of the most complex environments that road system users must navigate, and it has been found that longitudinal channelizing devices such as portable concrete barriers reduce confusion (relative to other methods such as drums) by better delineating work zones (Bryden, Andrew, & Fortuniewicz, 2000; Finley, Theiss, Trout, Miles, & Nelson, 2011; Aaron Todd Greenwood, Xu, Corso, Hunter, & Rodgers, 2016; Hunter, Rodgers, Corso, Xu, & Greenwood, 2014; Xu, Greenwood, Corso, Rodgers, & Hunter, 2015). Consistent with both the heavy traffic and work zone configuration factors discussed previously is the body of the work which suggests that increased lane maneuvering or lane configuration changes also has significant negative effects on driver performance (Schiessl, 2008; Stinchcombe & Gagnon, 2010; Teh et al., 2014).

The studies cited above have established that roadway factors influence driver performance, but there is little general understanding with regards to drivers' perception of environments. However, in a study by some members of this team, drivers' perceived complexity of dynamic (i.e. video) roadway environments were studied, and it was found that traffic was the factor most likely to influence perceived complexity, relative to work zone drums, lane merges, presence of roadway objects, and urban environments, which were the factors studied in that particular study (Shaw, Greenwood, et al., 2018). The experiment in this thesis obtained ratings of perceived task complexity (i.e., participants' rated how difficult each environment was to drive through) for 100 unique roadway environments (75 of which were on-road environments and 25 of which were simulated driving roadway environments). These ratings are used to identify the dimensions that influence drivers' perceived complexity of roadway environments, as well as to identify which of the environments studied are best at differentiating between the perceptual dimensions identified. This provides the core underlying dimensions of the roadway environments that drivers' "notice," and hence, are likely the dimensions that may influence performance differences.

CHAPTER 3. METHOD

The data used in this thesis were previously obtained as part of the author's work on a preceding project; as such, only a brief overview has been provided within this section, and additional details and background regarding data collection, the resulting dataset, and further analyses can be accessed at (Hunter et al., 2016; Shaw, Bae, Corso, Rodgers, & Hunter, 2017; Shaw et al., 2016). Institutional Review Board (IRB) approval was obtained prior to study implementation, and all associated protocols followed IRB expectations with regards to participant privacy and consent.

1.1 Participants

The data used in this study come from convenience samples of 288 participants from four populations: (1) a high school in a suburban part of Georgia; (2) a rural public university in Kentucky; (3) an urban public university in Georgia; and (4) a public festival in an urban area of Georgia. Inclusion criteria for participants at the college and public festival sites include: (1) having a valid driver's license; and (2) having at least two years of driving experience. There were no such inclusion requirements for the high school participants, although the data used in this thesis exclude participants who do not have a license or learner's permit. An overview of the participants in this experiment is included in Table 1, and additional information regarding the various data collection implementations can be seen in (Hunter et al., 2016).

Table 1. Summary of Participants

Participant Sample	Recruitment Period	Reimbursement	Male	Female	Choose Not To Answer	Total
High School Participants	Fall 2014; Fall 2015	Community service; Chicken sandwich ticket	47.7% (51)	50.5% (54)	1.9% (2)	37.2% (107)
College Participants (Urban Location)	Fall 2014	Extra credit for course	45.2% (19)	54.8% (23)	0% (0)	14.6% (42)
College Participants (Rural Location)	Fall 2014	Extra credit for course	31.6% (12)	68.4% (26)	0% (0)	13.2% (38)
Festival Participants	Fall 2015	$10 Coffee gift card	43.6% (44)	56.4% (57)	0% (0)	35.1% 101
Total	N/A	N/A	43.8% (126)	55.6% (160)	0.7% (2)	100% (288)

1.2 Procedure

Self-reported ratings of complexity and response times were obtained during multiple randomized repetitions of 100 unique roadway environments. For the purposes of this report, only data from the first repetition will be used for each participant (i.e. 100 ratings for 100 unique stimuli). For the first repetition, participants rated the images in accordance with *how difficult it would be to drive through the scene* (task complexity). Ratings were made on a five-point integer scale, ranging from one (least complex) to five (most complex). For the first repetition, non-responses comprised approximately 5.32% of the total data set.

Seventy-five of the 100 unique roadway images used in this experiment are of on-road environments (existing roadways), and twenty-five are of simulated (driving simulator) roadway environments. The on-road environments were taken on roads located in California, Georgia, Kentucky, New York, Ohio, South Carolina, and Virginia. The simulated environments were obtained using the National Advanced Driving Simulator (NADS) MiniSim® software (Figure 3). Sample on-road and simulated environments used in this experiment are shown in Figure 1 and Figure 2 respectively. As can be seen in these images, a wide range of conditions were selected to ensure that the 100 images used in this experiment captured diverse (but by no means exhaustive) combinations of roadway environment characteristics.

Figure 1. Sample On-road Environment Images

Figure 2. Sample Simulated Environment Images

Figure 3. National Advanced Driving Simulator (NADS) MiniSim™ for Simulated Environment Images

1.3 Data Analysis

Here, we provide methodological details on the analyses executed in this thesis. We begin with the preliminary analyses, followed by the application of IRT models to measure perceived complexity ratings data.

1.3.1 Preliminary Analyses

Initial exploratory analyses are conducted to obtain a preliminary understanding of the dataset and to aid in the interpretation of the IRT model estimates. In previous work on this dataset, the research team classified the roadway environments with respect to approximately 70 characteristics using a binary scale (1 = presence and 0 = absence). After removing characteristics with low occurrences (less than 5% of environments were seen to have these characteristics), 42 characteristics remained (see Table 2). Given that these classifications are binary variables, tetrachoric correlations (Holgado–Tello, Chacón–Moscoso, Barbero–García, & Vila–Abad, 2010) are used to perform iterative principal axis factoring (PAF) with oblique promax rotation to obtain a better understanding of the physical dimensions present across these environments.

Table 2. Final Roadway Characteristics used for Factor Analysis

Sub-areas[1]	**Roadway Characteristics**
Geometric Design	Freeway/Highway/Uninterrupted flow facility, Arterial/Collector facility, Rural/local roads, Vertical curves, Horizontal curves, Number of lanes, Narrow/constrained lanes, Paved shoulders
Roadway Objects or Markings	Bridge infrastructure, Overhead signs, Medians, Decorated/vegetated medians, Crosswalks/pedestrian crossing zones, Work zones, Trucks/heavy vehicles, Centerline (no passing), Centerline (passing), Barrier separated
Roadside Environment	Urban/Rural, Driveways, Roadside buildings, Parked cars, Sidewalk, Guardrail, Roadside vegetation, Noise barriers/fencing, Roadside attractions, Pedestrians, Static signage, Telephone wires/poles, Streetlights, Curb and gutter, Hydrants, Drainage channels/side slopes
Operational	Time of day: low light versus daylight, Weather: snow/rain/fog versus clear conditions, Signalized intersections, Heavy traffic, Work zone diverges/maneuvering, Pavement markings: faded/unusual, Non work zone delineation devices, Low traffic, No traffic

[1]Note that these are not factors. The factor analysis is discussed in the Results section. These were sub-areas (or domains) from which the team selected roadway characteristics to classify across the images.

Next, we model the roadway environments along an approximately interval scale of complexity using Thurstone’s Method of Successive Intervals. This characterizes each image (stimulus) with a scale value (mean discrimal process for perceived complexity) and discriminal dispersion (standard deviation; Bock & Jones, 1968; Saffir, 1937; Torgerson, 1958). This step utilizes the ratings obtained for each of the 100 environments across the 288 participants. We then estimate regression models of the complexity scale values (i.e. mean and standard deviation location for each roadway environment) using the factors extracted (from the factor analysis) previously as explanatory variables. This aids in understanding how the physical constructs present in the environments affect the rating data obtained from the participants.

To summarize, these preliminary analyses involve both the environmental characteristics present, as well as the rating data, and allow us to explore and further understand: (1) the primary environment factors present in the roadway environments being studied; and (2) how these factors influence the rating patterns observed from the participants. This step of the analysis can be considered exploratory analysis that will aid in interpreting the IRT models discussed next.

1.3.2 Item Response Theory Models

Following the preliminary analyses, we turn to the core objective of this thesis: i.e., the application of multidimensional, polytomous IRT models to the perceived complexity ratings obtained from the sample. The first step of IRT model development is to examine the number of dimensions present in one's data, particularly since the local independence assumptions of IRT model development depends on specifying either the correct number of dimensions in the data, or specifying more than the correct number of dimensions (i.e. overfitting is okay, though not desirable; underfitting is not, and will violate the local independence assumption). As such, prior to developing the IRT models, we first examine the number of dimensions present in the perceived complexity ratings data. After assessing dimensionality, we estimate a series of IRT Models to jointly represent items and persons in a latent space.

We estimate the Generalized Partial Credit Model (GPCM), as well as the Graded Response Model (GRM) and investigate the relative performance of these models. We also estimate both the GPCM and GRM models in multigroup (as well as aggregate) forms to facilitate the estimation of population trait means for each demographic group (represented by the subscripts g in the equations; as such, removing subscript g gives the aggregate or non-multigroup form(s) of the equation(s)). Multigroup model results are only reported if they increase fit relative to the decrease in parsimony.

1.3.2.1 <u>Generalized Partial Credit Model (GPCM)</u>

The GPCM (Muraki, 1992, 1993) uses a discrimination (a_i), an item location (b_i), and a set of category threshold parameters (d_{iu}) to estimate the probability of responding in the kth category for a specific item i. Following (Reckase, 2009), the unidimensional GPCM (multigroup) model is formulated as follows:

$$P\left(u_{ij} = k \middle| \ \theta_{jg}, a_i, b_i, d_{iu}\right) \ = \ \frac{e^{\left[a_i \left(k(\theta_{jg} - \ b_i\right) - \sum_{u=0}^{k} \ d_{iu})\right]}}{\sum_{h=0}^{m_i} e^{\left[a_i(h \ (\ \theta_{jg} - b_i) - \sum_{u=0}^{h} \ d_{iu})\right]}} \tag{1}$$

$$= \ \frac{e^{[\sum_{u=0}^{k} a_i(\ \theta_{jg} - \ (b_i + d_{iu}))]}}{\sum_{h=0}^{m_i} e^{[\sum_{u=0}^{h} a_i(\ \theta_{jg} - \ (b_i + d_{iu}))])}} = \frac{e^{[\sum_{u=0}^{k} a_i(\ \theta_{jg} - \ b_{iu})]}}{\sum_{h=0}^{m_i} e^{[\sum_{u=0}^{h} a_i(\ \theta_{jg} - \ b_{iu})]}} = \frac{e^{[k a_i \theta_{jg} - \ \sum_{u=0}^{k} a_i b_{iu}]}}{\sum_{h=0}^{m_i} e^{[h a_i \theta_{jg} - \ \sum_{u=0}^{h} a_i b_{iu}]}} \tag{2}$$

where, (given that the roadway environment (stimulus) index is i, and the person index is j):

u_{ij} is the observed value for the complexity rating made by the *jth* person for the

ith item,

m_i is the total number of rating categories minus 1 for each environment,

k is a possible rating value which ranges between 0 and m_i,

θ_{jg} is the trait level for person j from demographic group g = 1, 2, …, G ,

with groups defined as high school (*g*=1), urban college (*g*=2) ,
rural college (*g*=3), and festival participants (*g*=4),

b_i is the overall complexity of environment i (more commonly known as item

difficulty),

a_i is the degree that a response to environment i distinguishes trait levels (more commonly known as item discrimination, or in our case –

environment discrimination),

d_{iu} is the threshold parameter for category u, where $d_{i0} = \sum_{k=0}^{m_i} d_{ik} = 0$,

b_{iu} is the step parameter for environment i and rating category u, and is equal to $b_i + d_{iu}$

Note that when $u = 0$, then both $a_i(\theta_j - b_i)$ and $a_i(\theta_j - b_{iu})$ are defined as zero. We include three versions of the unidimensional model (see Equations 1 and 2); and Equation 1 defines the GPCM using a common mathematical form. The algebraically equivalent forms in Equation 2 serve to illustrate the concept of the step parameter (b_{iu}) which becomes more important when we discuss the multidimensional form of the model. The step parameters can be thought of as absolute locations of the threshold parameters (i.e., the threshold parameters are relative to item difficulty, while the step parameters are not).

To extend the unidimensional model to the multidimensional IRT form with D dimensions (or traits being measured), we consider that the $\boldsymbol{\theta}_j$ parameter is now a 1 by D vector of traits for each person. Additionally, item discrimination ($\boldsymbol{a}_i$) is also a 1 by D vector that indicates how well each roadway environment discriminates each trait being estimated. As such the $\boldsymbol{\theta}_j$ vector should be transposed ($\boldsymbol{\theta'}_j$) in order to be premultiplied by the $\boldsymbol{a}_i$ vector. The multidimensional GPCM model is formulated as follows, again adapted from (Reckase, 2009), with all recurring notation from Equations 1 and 2 retaining their meanings:

$$P\left(u_{ij} = k \middle| \boldsymbol{\theta}'_{jg}, \boldsymbol{a}_i, f_{iu}\right) = \frac{e^{k\boldsymbol{a}_i\boldsymbol{\theta}'_{jg} - \sum_{u=0}^{k} f_{iu}}}{\sum_{h=0}^{m_i} e^{h\boldsymbol{a}_i\boldsymbol{\theta}'_{jg} - \sum_{u=0}^{h} f_{iu}}} \tag{3}$$

where, f_{iu} is the intercept parameter for rating category u, aggregated across D dimensions as follows:

$$f_{iu} = \sum_{d=1}^{D} a_{id}(b_{id} - d_{iud}) = \sum_{d=1}^{D} a_{id}(b_{iud}) \tag{4}$$

In the multidimensional form of the model, the parameter f_{iu} corresponds to the the $a_i b_{iu}$ term in the last identity from Equation 2. However, the intercept parameter, f_{iu} cannot be estimated separately for each dimension and therefore, it is estimated as an aggregate *across* dimensions. Thus, the intercept parameter in the multidimensional case is a function of the item's discrimination and step parameters across all D dimensions.

1.3.2.2 Graded Response Model (GRM)

The GRM (Samejima, 1969) is different from the GPCM in that the GRM assumes that passing step k +1 requires more of the latent trait in question than does passing step k. In contrast, the GPCM does not rely on the assumption that an item's steps must be ordered on the latent continuum (Reckase, 2009). In the context of a graded rating scale, the GRM assumes each successive response requires more of the latent trait. Moreover, the model focuses on locating the boundaries between successive response categories on the latent continuum. The GRM is formulated by first modeling the probability of obtaining a response in or above a given response category. This probability is denoted as P^* and is defined as follows (all notation is consistent with that used in Section 1.3.2.1, unless otherwise noted):

$$P^*(u_{ij} \geq 0) = 1 , \tag{5}$$

$$P^*\left(u_{ij} \geq 1 \middle| \ \theta_{jg}, a_i, b_{i1}\right) = \frac{1}{1 + e^{[-(a_i(\theta_{jg} - b_{i1}))]}} ,$$

$$P^*\left(u_{ij} \geq 2 \middle| \ \theta_{jg}, a_i, b_{i2}\right) = \frac{1}{1 + e^{[-(a_i(\theta_{jg} - b_{i2}))]}} ,$$

$$\vdots$$

$$P^*\left(u_{ij} \geq m_i \middle| \theta_{jg}, a_i, b_{i(m_i)}\right) = \frac{1}{1 + e^{\left[-\left(a_i(\theta_{jg} - b_{i(m_i)})\right)\right]}}$$

$$P^*(u_{ij} \geq m_i + 1) = 0$$

Recall that response categories range from 0 to m_i. When the response from person j to item i (i.e., u_{ij}) is equal to 0 then the probability of obtaining that response or a higher response on the rating scale must logically be equal to 1. Likewise, the probability of obtaining a response greater than m_i (i.e., greater than or equal to $m_i + 1$) must logically be equal to 0. The values of P^* for remaining $m_i - 1$ response categories between these two extremes are derived using a special 2-parameter logistic model in which the discrimination parameter, a_i, is constrained to be constant across categories. Each of these probabilities includes an item parameter, b_{ik}, which estimates the location of the boundary between the categories $k - 1$ and k on the latent continuum. The conditional probability for each unique response on the scale is then calculated as:

$$\begin{aligned} P\left(u_{ij} = k \middle| \theta_{jg}, a_i, b_{ik}\right) \\ &= P^*\left(u_{ij} \geq k \middle| \theta_{jg}, a_i, b_{ik}\right) \\ &- P^*\left(u_{ij} \geq k+1 \middle| \theta_{jg}, a_i, b_{i(k+1)}\right) \end{aligned} \tag{6}$$

for $k = 0$ to m_i. (Note that when $k = 0$ or $k = m_i + 1$ in Equation 6, then the values of P^* are constants rather than quantities that are conditional on IRT parameters.)

The multidimensional GRM approach is discussed next, using a formulation adapted from previous literature (Bock, 1972; Chalmers, 2012; Samejima, 1969), and edited for this thesis. Again, assuming that there are m_i unique categories for roadway environment i, there would subsequently be $m_i - 1$ intercept parameters corresponding to category boundaries for a given stimulus. These parameters are denoted as d_{ik}. As was the case with step parameters in the multidimensional GPCM, the category boundaries for each dimension (i.e., b_{ikd}) cannot be estimated separately, and must be linearly combined in a composite which is weighted by discrimination parameters as follows:

$$d_{ik} = \sum_{d=1}^{D} a_{id} \ (b_{ikd}). \tag{7}$$

With these definitions in hand, the multidimensional form of the model is derived in a manner that is analogous to its unidimensional counterpart:

$$\begin{aligned}
&P^*(u_{ij} \geq 0) = 1\,, \\
&P^*\left(u_{ij} \geq 1 \middle| \, \boldsymbol{\theta}_{jg}, \boldsymbol{a}_i, d_{i1}\right) = \frac{1}{1 + e^{[-\left(\boldsymbol{a}_i \boldsymbol{\theta}'_{jg} + d_{i1}\right)]}}\,, \\
&P^*\left(u_{ij} \geq 2 \middle| \, \boldsymbol{\theta}_{jg}, \boldsymbol{a}_i, d_{i2}\right) = \frac{1}{1 + e^{[-\left(\boldsymbol{a}_i \boldsymbol{\theta}'_{jg} + d_{i2}\right)]}}\,, \\
&\qquad\qquad\vdots \\
&P^*\left(u_{ij} \geq m_i \middle| \, \boldsymbol{\theta}_{jg}, \boldsymbol{a}_i, d_{i(m_i)}\right) = \frac{1}{1 + e^{[-\left(\boldsymbol{a}_i \boldsymbol{\theta}'_{jg} + d_{i(m_i)}\right)]}}\,, \\
&P^*(u_{ij} \geq m_i + 1) = 0
\end{aligned} \tag{8}$$

$$
\begin{aligned}
P\left(u_{ij} = k \middle| \boldsymbol{\theta}_{jg}, \boldsymbol{a}_i, d_{ik}\right) \\
&= P^*\left(u_{ij} \geq k \middle| \boldsymbol{\theta}_{jg}, \boldsymbol{a}_i, d_{ik}\right) \\
&- P^*\left(u_{ij} \geq k+1 \middle| \boldsymbol{\theta}_{jg}, \boldsymbol{a}_i, d_{ik}\right)
\end{aligned} \tag{9}
$$

1.3.2.3 Model Identifiability

The multidimensional GPCM and GRM are compensatory models, and as such, they both are unidentified without further constraints. Specifically, constraints are required to obtain a unique origin, scale, angle between axes, and orientation (i.e., rotation) of the axes (Reckase, 2009). We constrain the origin and scale by setting the mean and variance for each latent trait to 0 and 1, respectively. We constrain the angle between all pairs of axes in the latent space to be 90 degrees by setting the covariance between any two latent traits to zero (note: when taken with the scale constraint, this results in a variance-covariance matrix that is an identity matrix). To implement the rotation constraint, we set the upper triangle of the item discrimination parameters to 0 when considering the first D-1 items. These constraints are widely used in applied work with multidimensional IRT models. The program used for model estimation (flexMIRT), handles the location, scale, and basis constraints by default, whereas the rotation constraint is implemented manually by the user.

CHAPTER 4. RESULTS

We first discuss preliminary analyses on the data, followed by a presentation of the application of IRT models to perceived complexity ratings data.

1.4 Preliminary Analyses

As noted previously, the preliminary analyses presented in this section aim to provide an exploratory understanding of the dataset and thereby, to aid in the interpretation of the IRT models presented in Section 1.5.

1.4.1 Classification of Roadway Environments

Here, we present the results of the PAF conducted on the binary (0/1) presence of physical characteristics in the environments. Thirty-six physical characteristics were used in the factor analysis, and these are drawn from Table 2, with some of those variables removed or condensed in cases of collinearity (for example, medians and decorated medians were condensed into one category, and “passing” and “no passing centerline” was condensed into yellow centerline, capturing a lack of barrier separation between opposing traffic). This analysis facilitates an understanding of the built environment constructs present in the roadway environments used in this experiment; and, notably, is not dependent on the ratings of perceived complexity like the rest of analyses throughout this document. Table 3 gives an overview of the factors extracted, and Appendix Table 1 provides the pattern matrix for the PAF solution. We see that four primary dimensions underlie the built environment characteristics as assessed using bootstrapped parallel analysis of eigenvalues from the tetrachoric correlation matrix. These dimensions are interpreted as urban environments, freeway environments, environmental conditions, and open/constrained conditions, and are further elaborated on in Table 3.

Table 3. Four Dimensions Extracted in Factor Analysis Solution

Retained Factors	**Interpretation**
Factor 1: *Urban Environments*	This factor had high correlations with roadway characteristics such as curb and gutter, sidewalks, crosswalks, street lights, parked cars, roadside buildings, urban environments, pedestrians, signalized intersections, etc.
Factor 2: *Freeway Environments*	This factor had high correlations with barrier-separated directions of travel and non–work zone delineation devices. It also correlated somewhat with the presence of vertical curves, bridge infrastructure, work zones, heavy traffic, overhead signs, and paved shoulders. These are all indicative of freeway environments.
Factor 3: *Environmental Conditions*	This factor had high correlations with trucks/heavy vehicles, bad weather, poorly maintained or hard-to-see pavement markings, and dimly lit conditions.
Factor 4: *Open/Constrained Conditions*	This factor correlated strongly with variables that would tend to constrain the driver's movement. These included lane width, medians, driveways, guardrail, pedestrians, and traffic.

1.4.2 Modelling Roadway Environment Complexity

In this section of the report, we model the roadway environments along a scale of complexity relative to each other; and predict their locations on the complexity scale from the roadway constructs and characteristics.

1.4.2.1 Identifying Scale of Complexity for Roadway Environments

Thurstone's Method of Successive Scaling was used to construct a complexity scale for the 100 roadway environments. This method of scaling was selected because it does not assume equal intervals. Specifically, the rating scale used ranged from 1 to 5, and in Thurstone's method, it is not assumed that the intervals between adjacent rating categories are equal. Results from this method included a scale value (mean discrimal process for perceived complexity) and discriminal dispersion (i.e., standard error) for each road environment image (Bock & Jones, 1968; Saffir, 1937; Torgerson, 1958). Figure 4 shows a conceptual representation of stimuli X and Y along an AB continuum, as first presented in (Saffir, 1937). In our case, the X and Y would be roadway images (stimuli), and the continuum represents perceived complexity (from low to high) as obtained from the self-reported complexity ratings of 288 participants across 100 images. Missing data were imputed using expectation maximization algorithm prior to scaling the environments (Little, 2002). Figure 5 illustrates that as the images become more extreme with respect to their absolute scale values for complexity, their respective standard errors increase. This is most notable for stimuli with extremely negative (i.e., least complex) scale values. The quadratic nature of this relationship suggests that drivers may agree more about what constitutes high complexity environments relative to low complexity environments.

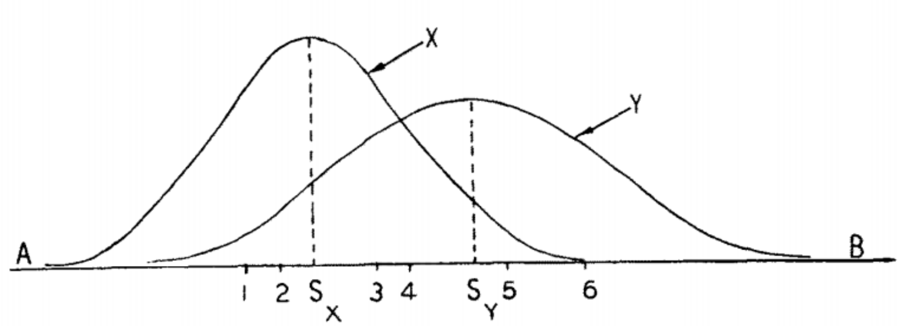


Figure 4. AB Psychological Continuum (Saffir, 1937)

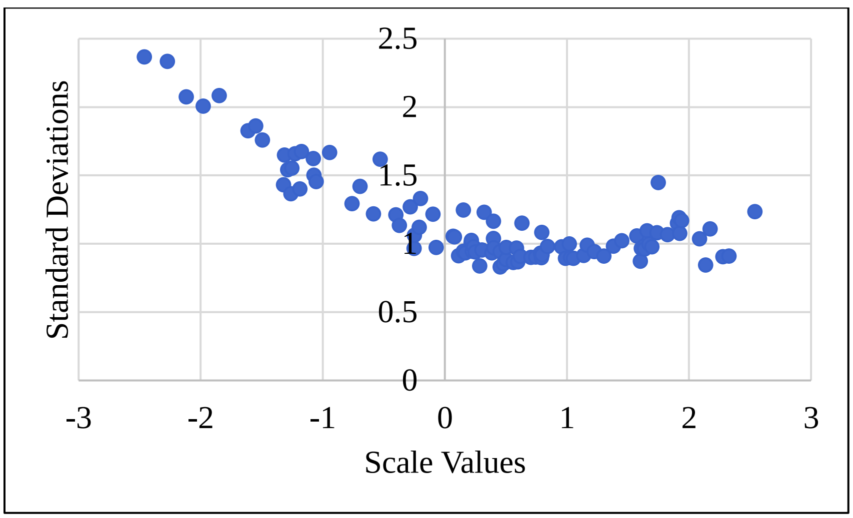


Figure 5. Standard Deviations vs. Scale Values of Perceived Complexity Ratings for each Roadway Environment

1.4.2.2 Modelling Perceived Complexity using Complexity Scale Location

Predictive models of perceived complexity with scale values (see Section 1.3.1) for each image as the dependent variable are discussed within this section. The first of these models predicted the complexity scale values for each image based on factor scores from the PAF solution described in Table 3 and Appendix Table 1. Results from this regression are shown in Table 4 and Appendix Table 2. As is seen in Table 4, three of the latent constructs (namely, urban environments, freeways, and environmental conditions) introduced as predictors were statistically significant, and together, all four constructs explain approximately 50% of the variance (i.e., this is the adjusted R square value; unadjusted R square = 0.516). Of the four predictors, environmental conditions had the greatest linear relationship with Thurstone complexity scale values conditional on the other predictors in the model, followed closely by urban environments which also contributed significantly to the likelihood of increased perceived complexity ratings.

Table 4. Scale Values as a Function of Primary Dimensions from EFA (Tetrachoric Correlations)

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig. (p-value)
	B	Std. Error	Beta		
(Constant)	-.041	.102		-.403	p = 0.688
Urban Environments	.936	.205	.335	4.564	p < 0.0001
Freeways	.572	.232	.190	2.469	p = 0.015
Environmental Conditions	1.275	.224	.450	5.689	p < 0.0001
Open/Constrained Conditions	-.367	.215	-.122	-1.705	p = 0.092

a. Dependent Variable: SCALEV
b. R-Squared: 0.5.16, Adj R-Squared: 0.495

A second model was explored in which the Thurstone complexity scale value for each image was modeled as a function of the 43 individual roadway characteristic variables (Table 2) that describe the images. The predictive value of these characteristics was assessed using a stepwise (forward selection) multiple linear regression model, with a probability of variable entry of 0.05, and a probability required to be removed from the equation of 0.10 (these parameters are commonly referred to as P-IN and P-OUT, and determine which variables enter and are retained in the equation). The results of the stepwise regression model are described in Table 5. We found that when roadway *characteristics* are used as predictors, we are able to explain more variance (84% of variance in dependent variable as measured by the adjusted R Square value) relative to when the PAF *factor scores* were used (recall that the previous model, Model 1, had an adjusted R Square of 50%). It should be noted here that stepwise prediction (as with most variable selection procedures including forward and backward regression) is known to have a series of problems such as the heavy interference of chance within the final solution (Cohen, 1983). This increased variance is attributable to the fact that some of the unique variance in each roadway characteristic was likely related to perceived complexity, and this unique variance was ignored in the PFA solution examined above.

In Table 4 we see that the environmental conditions and urban environment factors have the greatest impact on image location along the perceived complexity spectrum, while the openness of the environment has the smallest effect. These results are somewhat aligned with those of the stepwise regression model (see Table 5) which shows that variables which had high loadings on some of the constructs, are the significant predictors retained in the model. For example, all levels of traffic are significant predictors; and, correspondingly, the heavy traffic and low traffic characteristics load on two factors each in the factor analysis solution). Similarly, pavement markings and time of day also both load on two individual constructs with high loadings. Thus, we see that critical characteristics which were present in the factor analysis solution, resurfaced in the model that used individual characteristics. This is a good indicator of interpretability and robustness across solutions. We posit that the two different types of models could be useful to different practitioners, for example: Model 1 allows general conclusions to be made regarding the built environment, while Model 2 is useful for roadway engineers and designers who are interested in specific characteristics that they can control for in their design.

Table 5. Scale Values as a Function of Roadway Characteristics Retained in Stepwise Linear Regression Model

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig. (p-value)	Factors the Variables Loaded On
	B	Std. Error	Beta			
(Constant)	2.078	0.28		7.429	$p < 0.0001$	
notraffic	-2.805	0.293	-1.139	-9.57	$p < 0.0001$	Environmental
timeofday	1.249	0.162	0.333	7.725	$p < 0.0001$	Urban Environments, Environmental Conditions
drainagechannels	-0.42	0.11	-0.179	-3.81	$p < 0.0001$	Urban Environments
lowtraffic	-1.433	0.291	-0.609	-4.92	$p < 0.0001$	Environmental Conditions, Open/Constrained Conditions
parkedcars	0.459	0.196	0.106	2.341	$p = 0.021$	Urban Environments
passingCL	-0.673	0.208	-0.136	-3.23	$p < 0.0001$	Freeways, Open/Constrained Conditions
workzonediverges	0.765	0.207	0.155	3.689	$p < 0.0001$	Freeways
pavementmarkings	0.381	0.138	0.116	2.764	$p = 0.007$	Environmental Conditions, Open/Constrained Conditions
heavyTraffic	-0.803	0.308	-0.244	-2.61	$p = 0.011$	Freeways, Open/Constrained Conditions

a. Dependent Variable: SCALEV
b. R-Squared: 0.855, Adj R-Squared: 0.841

1.5 Item Response Theory Models

Within this section, we applied polytomous MIRT models (GPCM and GRM) to study the dimensions present in participants' self-reported measures of perceived complexity for 100 unique roadway environments. Approximately 5% of the data (1480 out of 28800 values) are missing, and we did not impute the missing values but rather simply retained them as missing. The missing data were treated as Missing at Random (MAR), meaning that if we condition on the parameters of the model, we can expect that the missing values do not have a pattern based on the perceived complexity of the image (for this reason MAR is often thought of as 'missing ***conditionally*** at random').

1.5.1 Dimension Identification

We performed a bootstrapped version of Horn's parallel analysis (Buja & Eyuboglu, 1992; Horn, 1965), as well as the Polytomous Dimensionality Evaluation to Enumerate Contributing Traits (also known as Poly-DETECT)(J. Zhang, 2007) to examine the number of dimensions present in the perceived complexity ratings. However, the Poly-DETECT procedure was unable to estimate the dimensionality due to the limited sample size for our data (i.e. N = 288 was too small for the 100 items); and as such, we proceeded using only bootstrapped parallel analysis. Missing values in the dataset were retained as missing for this dimensionality assessment. As shown in Figure 6 below, bootstrapped Horn's parallel analysis indicates that there are four dominant dimensions in the perceived complexity ratings, since the point at which the eigenvalues for the simulated data crosses the eigenvalues for the actual data is between four and five dimensions. After this point, we see that the eigenvalues in our real data fall below those of the simulated (random) data, indicating the variance explained by additional dimensions in the real data is now no greater than the variance explained by chance. The data in the graph are shown in Table 6. We therefore conclude that the perceived complexity ratings are 4-dimensional data.

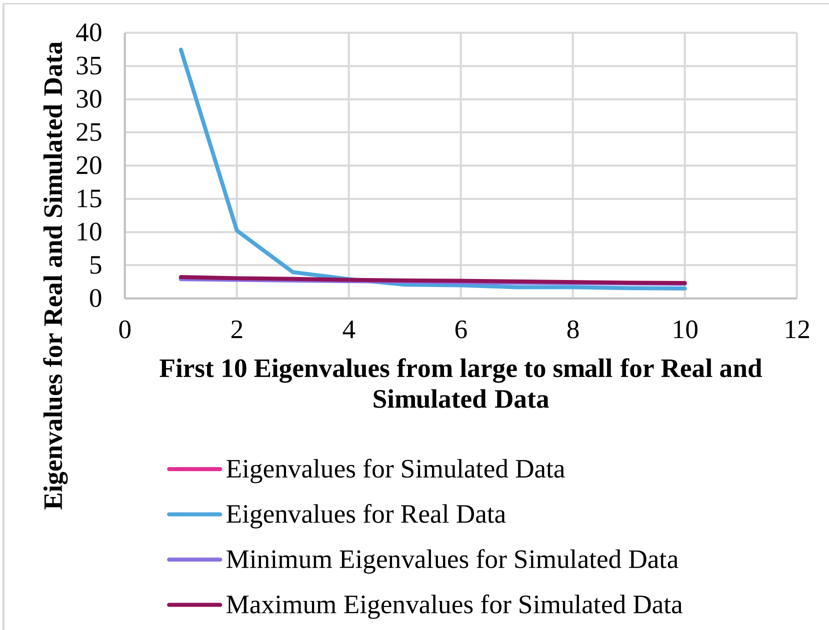


Figure 6. Horn's Parallel Analysis: Eigenvalues for Real and Simulated Data

Table 6. Horn's Parallel Analysis: Eigenvalues for Real and Simulated Data (Represented graphically in Figure 6)

	Eigenvalues for Actual Data	Eigenvalues for Simulated Data	Minimum Eigenvalues for Simulated Data	Maximum Eigenvalues for Actual Data
1	37.461171	3.058255	2.904791	3.2249875
2	10.1991582	2.9193368	2.7971754	3.0559032
3	3.970834	2.7931214	2.7227497	2.9273999
4	2.9162224	2.6983051	2.594052	2.7936186
5	2.0631407	2.6134605	2.5498726	2.688535
6	1.9761166	2.5340884	2.4274073	2.664544
7	1.6840864	2.4656546	2.3698938	2.5637439
8	1.6792898	2.3959614	2.3373967	2.4645156
9	1.5569732	2.3235062	2.2713843	2.3872561
10	1.5023627	2.2651081	2.2060926	2.3294263

As noted above, Poly-DETECT was not able to estimate the dimensionality due to the large number of total score groups (number of items + 1 = 101) relative to the number of participants (N = 288); thus, the number of participants in each group was too sparse and prevented the calculation of conditional covariance based measures of proximity for each item pair. Out of curiosity, we duplicated the data four times (i.e. we "increased" the sample size without changing the distribution of responses; N = 288*4 = 1152), and found that poly-DETECT reported four dimensions (as found in the bootstrapped parallel analysis detailed earlier). We reiterate that this execution of poly-DETECT is for exploratory purposes only, and we do not base our final model on this finding; but, rather report it here for completeness.

1.5.2 Model Development

As shown in Section 1.5, there appears to be four dominant dimensions underlying the perceived complexity ratings, and therefore, four-dimensional versions of the GPCM and GRM were selected for analysis. Given this moderately large number of dimensions, a standard marginal maximum likelihood (MML) item parameter estimation algorithm would have been computationally slow and cumbersome. This method generally relies on a fixed number of Q quadrature points for each dimension which are, in turn, crossed to form a quadrature grid. Thus, for a standard number of quadrature points per dimension (e.g., 30), a total number of 30^4 grid points would be evaluated when integrating $\boldsymbol{\theta}_j$ out of the likelihood; and this integration would be necessary on every iteration. To avoid this computational burden, we used the Metropolis-Hastings Robbins Monroe (MH-RM) technique to estimate item parameters. MH-RM uses stochastic or sampling based integration to remove $\boldsymbol{\theta}_j$ from the likelihood as opposed to the slower fixed numerical quadrature implemented in the MML method. As a result, the computational burden for MH-RM increases linearly as the dimensions increase rather than exponentially, as is the case with MML (Li Cai, 2010a, 2010b). An unfortunate consequence of using the MH-RM technique as implemented in the flexMIRT estimation program is that we do not obtain popular item fit statistics that are provided by flexMIRT when the MML procedure is used as these statistics rely on quantities that are byproducts of the MML procedure. We obtain only stochastic Yen-Bock item fit diagnostic values (instead of Orlando-Thissen-Bjourner item fit or Chen and Thissen local dependence statistics), and we also do not obtain additional Goodness of Fit (GOF) output such as the Haberman residuals table or the M_2 statistic (L. Cai, 2017). It has been documented in the literature that traditional Yen-Bock fit statistics can have problems with Type 1 error rates and power, and thus, it would be preferred to report fit statistics with more acceptable Type 1 error rates (such as that of the Orlando-Thissen-Bjorner method; (Chon, Lee, & Dunbar, 2010). The stochastic Yen-Bock statistics (developed by Li Cai) are an experimental variant of the traditional Yen-Bock statistics and as such Type 1 error rates and power characteristics are unknown, and have not yet been detailed in the literature (L. Cai, personal communication to J. Roberts, May 4, 2018).

The MH-RM procedure requires a prior distribution for person parameters. By default, the flexMIRT program used a 4-dimensional multivariate normal prior distribution along with a centroid of zeros and an identity matrix as the variance-covariance matrix. We also used a lognormal prior distribution with mean of 0, and standard deviation of 0.5 for item discrimination parameters to increase the probability that the solution would converge with the small sample that was analyzed. We did not apply a prior distribution to the intercept parameters (i.e., multidimensional step or category boundary parameters), due to a technical problem on the part of flexMIRT at the time of this writing, but would recommend doing so in future implementations. (Specifically, flexMIRT currently crashes when using prior distributions with intercepts [category boundaries] in the GRM. It will work, however, with intercepts [step parameters] in the GPCM. We ultimately chose to use identical prior distributions across models.) The flexMIRT program automatically includes all MIRT identifiability constraints discussed in Section 1.3.2.3, with the exception of the rotation constraint, which we manually implemented by setting the discrimination of dimension 2, 3, and 4 to 0 for item 1, setting the discrimination of dimension 3 and 4 to 0 for item 2, and setting the discrimination of dimension 4 to 0 for item 3 (i.e. the upper right triangle of zeroes). For the multigroup models, we constrain the model to have the items function equally across the dimensions for all four groups. Additionally, in the multigroup model, we released the constraints on the variances and centroid for $\boldsymbol{\theta}_j$ across all groups, with the exception of the reference group, in which the variance of each latent trait was constrained to be 1 and the centroid was fixed to be a null vector. For this thesis, we selected the reference group to be the festival participants, because this group has the second largest number of participants and was collected in one wave, as opposed to the largest group (high school) which was collected over a period of two years.

The total number of item parameters estimated for the four-dimensional GPCM and GRM models in this analysis was equal to 794 (each). This included a discrimination parameter on each dimension for each image excluding those that were constrained to be 0 to obtain a unique rotation (4 dimensions x 100 images – 6 rotation constraints). It also included four item intercept parameters for each item (100 items x 4 intercept parameters). The results for all models are included in Appendix Table 4 through Appendix Table 8.

1.5.2.1 Generalized Partial Credit Model

As shown in Appendix Table 5, approximately all items (98 out of 100) had admissible fit (defined as $p > 0.01$). Items 51 and 76 are the two items that were misfit by this model, as their p-values were below 0.01 and 0.0001, respectively. Four items were close to the cutoff value, meaning that they had p-values ranging from 0.02 to 0.08. When the item discrimination values are squared and summed for each dimension, we see that dimension four had the largest sum of squared discriminations with 111.37, followed by dimension two with 96.08, and dimensions one and three at 71.50 and 71.34, respectively. Overall, all of these dimensions have somewhat similar item discrimination magnitudes, confirming that all dimensions are pertinent when rating complexity of the roadway environments.

By examining the images that had the highest discriminations for each dimension, it appears that dimension 1 corresponds to the urban roadway condition, and dimension 2 corresponds to the freeway environment. Dimension 3 is associated with poor visibility and low light conditions, while dimension 4 appears to be associated with no traffic. This finding corroborates the factor analysis and regression procedures executed on the roadway constructs in Section 1.4 to some extent, and suggests that people use these four constructs when assessing the complexity of roadway environments. Appendix Table 6 contains the intercepts from the multidimensional GPCM and the corresponding multidimensional step parameters for each image. The intercepts were obtained by executing a Fourier transformation on the gamma parameters produced by flexMIRT (Houts & Cai, 2015), and the multidimensional step parameters were obtained by dividing each intercept by the maximum discrimination (MDISC) associated with item *i*:

$$MDISC_i = \sqrt{a_{i1}^2 + a_{i2}^2 + \cdots + a_{iD}^2} \tag{10}$$

We see that across 81 of the 100 stimuli, the multidimensional steps are consistently increasing in a monotonic fashion. These items require successively more of the latent trait(s) to maximize the probability of observing higher response categories. As mentioned earlier, ordinal multidimensional step parameters are not assumed in the GPCM (but an analogous assumption is required for the GRM with respect to category boundaries). For the remaining 19 items, 11 of them have disordinal steps between the final three categories whereas 2 items have decreases in steps between the first three categories.

1.5.2.2 Graded Response Model

The GRM for perceived complexity ratings performed slightly better with respect to item fit than GPCM, with 99 out of 100 items having admissible fit (defined as $p > 0.01$) and only two items (item 76 and 87) having p-values in the vicinity of the 0.01 cutoff value (Appendix Table 8). When the item discrimination values are squared and summed for each dimension, we see that dimension four (again) had the largest sum of squared item discrimination values with 154.10, followed by dimension one with 127.89, and dimensions 2 and 3 at 105.44 and 92.08, respectively. As such, these dimensions have more divergent discrimination magnitudes than the GPCM model, although none of them are negligible. This provides confidence that all four dimensions are pertinent when rating complexity of the roadway images.

As with GPCM, we examined the environments that had the highest discriminations for each dimension. The items that are most salient on each dimension remain very similar with regards to content. In the GRM, the first dimension corresponds to the freeway environment, while the second dimension is capturing poor visibility and low light conditions. The third dimension appears to be the urban environment, and the fourth appears to be capturing a lack of traffic. From a visual and numeric examination of the roadway environments that are best discriminated by these constructs, the GRM model is both more populated and more consistent. For example, for the poor visibility dimension, the GRM model had 11 items that had very strong discriminations, while the GPCM model had 7 items. Similarly for the freeway dimension, the GRM model had 39 items with strong discriminations, while the GPCM model had 20 items.

Appendix Table 9 contains the intercepts for the multidimensional GRM model along with the multidimensional category boundaries (similar to Appendix Table 6 for the GPCM model). Recall that in the GRM, each successive response requires more of the latent trait, and thus a response falling in the higher rating category means that the person has passed the preceding category boundaries. As with GPCM, to better understand these category boundaries, we standardize the intercepts across items (see Appendix Table 9) by multiplying each intercept by -1 to reverse the sign, and then dividing each intercept by the MDISC for that item. The multidimensional category boundaries always increase in a monotone fashion, which is as required by this model.

1.5.3 Model Comparisons

Table 7 summarizes overall model fit indices for the four models examined in this thesis. For all of the fit indices provided, a lower value indicates better relative fit. AIC and BIC indices are both derived from the log likelihood value, but based on differences in theoretical derivation and assumptions. The BIC index is considered to be more conservative, i.e. penalizes increased model complexity more than the AIC index (Busemeyer & Diederich, 2014). Across all four models, we see that the GRM model (highlighted) performs best across all fit indices, followed by the GPCM, and then their multiple group counterparts, respectively. This implies that at least for this sample, knowing which populations the subjects come from, does not improve the amount of information we obtain in the model relative to the cost of fitting more parameters to describe these populations. Thus, we do not report on the group means or other parameters that were estimated for these models. Based on both the relative fit indices shown here, and the patterns of absolute item fit mentioned earlier, we conclude that the four-dimensional GRM model is best of those we have investigated to explain the perceived complexity ratings data used in this experiment.

Table 7. IRT Model Comparisons (95% CI Intervals Reported for each Index)

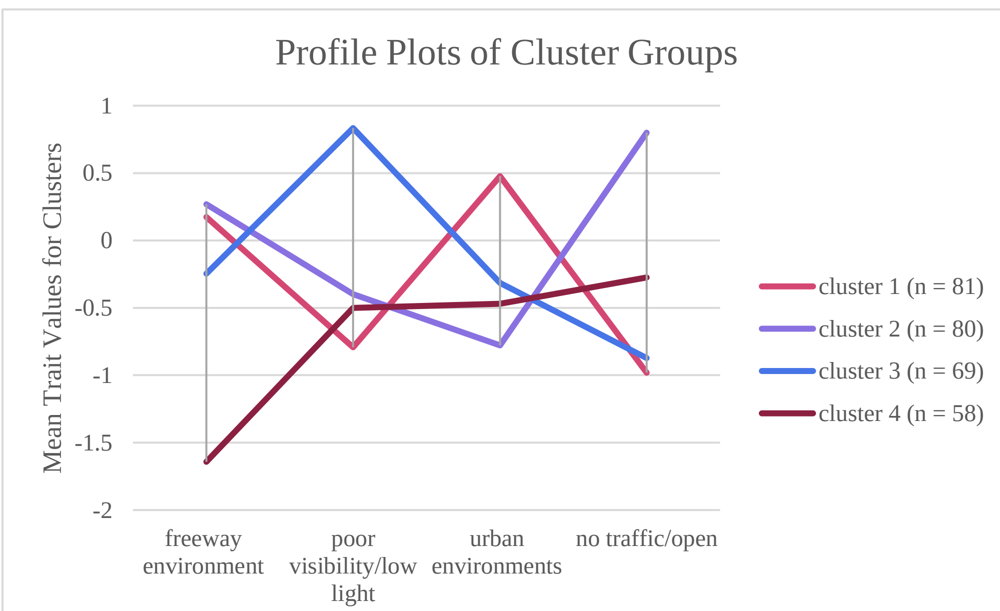
	-2loglikelihood	**AIC**	**BIC**
GPCM	54021.30, 54023.27	55609.30, 55611.27	58517.69, 58519.66
GPCM –Multigroup Model	55309.78, 55331.87	56945.78, 56967.87	59942.08, 59964.18
GRM	53714.95, 53716.64	55302.95, 55304.64	58211.34, 58213.03
GRM – Multigroup Model	54534.18, 54551.67	56170.18, 56187.67	59166.48, 59183.97

1.5.4 Trait (Theta) Estimation

Now that we have shown that the GRM model provides the best fit at both the overall level, as well as at an item level, we discuss the person (trait) estimates from that model (see Appendix Table 10). The first dimension corresponds to the freeway environment, and a high (positive) theta value would indicate that a respondent *j* perceived the stimuli from this domain as complex, whereas a low (negative) theta value on this dimension would indicate that the respondent did not perceive stimuli from this domain as being complex. Thus, relatively higher theta values on this dimension suggest that an individual will rate images of freeways as more complex than non-freeway environments. In contrast, the second dimension corresponds to poor visibility and low light conditions, and a positive theta value on this trait would indicate that person *j* rates environments with low light on the higher end of the complexity spectrum. Similarly, for dimension three, a positive theta value would result in higher perceived complexity ratings for urban environments, and for dimension four, positive theta value would result in increased complexity ratings for unconstrained environments or environments with no traffic. Correlations across the thetas indicated that the urban environment dimension had the strongest significant correlation (-0.400; $p < 0.01$) with the no traffic dimension, which is intuitive, given that one represents wide open conditions, and the other represents visually cluttered environments.

A standard two step cluster analysis (Rosenblad, 2009) on the trait (theta) values from the GRM model found four groups (see Figure 7 for profile plots for each cluster group). Notably, cluster four ($n = 58$) represented individuals with low to moderate theta values across all four dimensions, suggesting that some drivers did not perceive any of the environments as complex, regardless of which characteristics were present. Clusters one, three, and four intuitively had negative theta values on the no traffic dimension, but cluster two (n = 80) had positive values on this dimension as well as positive values on the freeway environment dimension, suggesting that this group (cluster two) finds wide open roadways to be more complex than the other clusters of respondents, and perhaps alluding to the increased perceived complexity of wide open freeways specifically. On the other hand, cluster one (n = 81) had the highest theta values on the urban environment (positive values as opposed to negative values for the other groups on this dimension), suggesting that this group perceives urban roadway environments to be more complex than other types of environments.

Figure 7. Profile Plots of Theta Values for each Cluster Group



CHAPTER 5. DISCUSSION

This thesis used exploratory factor analysis (with tetrachoric correlations), Thurstone's Method of Successive Scaling, multiple linear regression, dimensionality assessment procedures, and four polytomous multidimensional item response theory models to explore the physical characteristics present in roadway environments that influence drivers' perceived complexities of said environments. This broad swath of psychometric analyses repeatedly and reassuringly converged on slight variations of similar findings: i.e. there are four dimensions the built environment and these dimensions are similar to the four traits that affect the perceived complexity rating data. These dimensions also consistently aligned along similar themes, with freeway environments, urban environments, environmental conditions, and open/low traffic conditions capturing the latent constructs that are present.

Examining and comparing results at a more detailed level yields interesting insights into differences that these varying methods offer. For example, regression models of relative complexities of the roadway environments indicated that environmental conditions such as visibility and presence of trucks were most likely to increase perceived complexity ratings. However, the IRT models found that open/no traffic conditions followed by freeway environments, were the most discriminating traits (as indicated by their overall larger sum of squared discriminations relative to the other traits), suggesting that environments from these domains would be better at differentiating among drivers. A cluster analysis of the latent trait estimates for each person along each dimension also indicated that there are four distinct groups of participants, for whom different environments would differentiate best. Thus, we can see that IRT lends a particular insight to the data that is not possible from the typical regression modeling approach to research problems (such as the one detailed in Section 1.4.2).

The polytomous GRM model with four dimensions was found to best fit the perceived complexity rating data, relative to the GPCM model, as well as relative to the multiple group instantiations of the GRM and GPCM. The polytomous GRM model also had only one misfit item, and when the environments were separated into those that best measured each dimension, a more robust representation of those dimensions emerged. Thus, we see that the application of IRT allows us to both compare the environments relative to each other, as well as to compare the drivers in the study along the same scale of complexity used to compare the environments.

From a methodological perspective, this thesis incorporates and applies several sophisticated psychometric methods such as multidimensional IRT to the study of transportation system users. Although, as noted, there are some cases of IRT in the transportation literature, these applications are sparse. This author hopes that increasing the use of IRT in the study of transportation system users will be of significant academic benefit to the transportation research community. This is important because IRT has several key advantages over other methods that have been traditionally used in its place (e.g., Classical Test Theory, Factor Analysis, etc.). Namely, IRT facilitates use of the full information inherent in each participant's vector of responses (as opposed to factor analysis which is dependent on information between pairs of items). It allows both people and items to be placed along the same scale(s) and provides information about both. As mentioned before, IRT enables measurement specialists to create a large bank of stimuli (items) that are all calibrated to the same metric. This can be done with multiple samples of respondents so that no particular group must rate an unusually large number of stimuli. A stimulus bank such as this could then be used to implement adaptive experimentation/testing. In adaptive testing, participants are individually presented with stimuli that maximize the information about their trait levels across the dimensions. Technically speaking, many of these properties are derived from IRT's crown jewel; namely if the model truly fits the data, then IRT allows for invariant interpretations of item characteristics along with invariant interpretations of person location. Implicit in the notion of model fit is the requirement that the essential assumptions of local independence and examinee independence are valid.

The benefits of IRT are especially exciting in transportation engineering because, as noted before, transportation engineers and researchers are constantly collecting data from, and about users of our transportation systems. Such data is becoming harder and harder to obtain, as surveys and/or experiments become more time and cost intensive. Applying a property like invariance to develop a calibrated item bank of survey questions or roadway environments will facilitate not only adaptive testing and the potential for shorter questionnaires/studies, but it will also allow for comparison of data on a longitudinal basis with a consistent metric. The possibilities of IRT for transportation research and application purposes are unlimited, and it is hoped that this thesis, and its following work(s) will provide an impetus for more frequent and fruitful applications of these techniques in transportation engineering.

APPENDIX

Appendix Table 1. Pattern Matrices for Roadway Characteristics as Loaded onto Four Factors/Dimensions

Variables	Factor1	Factor2	Factor3	Factor4	Uniqueness
	Urban Environs.	Freeway Environs.	Environmental Conditions	Open / Constrained Conditions	
vertCurves	-0.2738	-0.1789	-0.3296	0.1253	0.6604
horzCurves	-0.2955	0.1031	-0.2504	0.0163	0.8313
NoLanes	0.4672	0.6245	0.2333	0.2477	0.1848
NarrowLanes	0.1409	-0.0654	0.1113	-0.3198	0.8455
PavedShoulders	-0.7354	0.2858	0.4126	-0.0167	0.2521
freeway	-0.4667	0.8651	-0.1976	0.0749	0.122
Arterial	0.8081	-0.2546	0.4195	0.1372	0.0923
rural	-0.5869	-0.6897	-0.0841	-0.1621	0.1017
bridge	-0.3852	0.7765	-0.0877	0.0295	0.3077
overhead	0.2791	0.4628	-0.0651	0.1131	0.7257
medians	0.1506	0.2127	0.0242	0.8481	0.2848
crosswalks	0.8531	-0.1245	-0.0407	-0.1711	0.2207
workzones	0.0428	0.4552	0.0861	-0.2116	0.6742
trucks	-0.145	-0.3298	0.9569	0.0547	0.2417
yellowCL	-0.1872	-0.867	0.1651	-0.3407	0.2428
barrierSep	-0.2419	0.8683	-0.1975	-0.0496	0.2686
urban	0.8354	-0.0577	-0.1367	-0.3327	0.1698
driveways	0.5265	-0.48	-0.0533	0.4186	0.3188
roadsidebuildings	0.8562	-0.1745	-0.0746	-0.0144	0.2571
Sidewalk	0.8452	-0.2468	0.0455	0.1298	0.2367
Guardrail	-0.4653	0.2373	0.1254	0.4308	0.5361
roadsideveg	-0.4122	-0.3776	-0.251	0.074	0.4765
noisebarriers	-0.1795	-0.1287	-0.0409	0.1741	0.8956
Pedestrians	0.7173	-0.1341	0.0085	-0.4382	0.2282
staticsign	0.3546	-0.092	-0.1051	-0.011	0.8642
wirespoles	-0.0069	-0.2892	-0.2599	0.2419	0.6954
streetlights	0.7075	0.2534	-0.3194	-0.0314	0.4559
curbGutter	0.9257	0.151	-0.1203	0.2888	0.1225
drainagechannels	-0.9087	-0.1693	0.0833	-0.0645	0.1727
timeofday	-0.2913	-0.332	0.7245	-0.0665	0.5164
signalizedintersections	0.7661	0.035	-0.0725	0.1146	0.431
heavyTraffic	0.1467	0.2791	0.0091	-0.5855	0.4894
pavementmarkings	0.1623	-0.0411	0.3426	-0.2253	0.7633
nonworkzonedelineation	-0.1519	0.8574	-0.2073	-0.1858	0.2733

Appendix Table 1 cont'd.

lowtraffic	-0.0356	0.0517	0.5645	0.5915	0.426
notraffic	-0.1045	-0.2455	-0.6096	-0.0894	0.4385

SPSS Model Summaries for Models 1 and 2

Appendix Table 2. Model 1 Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
				R Square Change	F Change	df1	df2	Sig. F Change
.718[a]	.516	.495	.840	.516	25.28	4	95	.000

a. Predictors: (Constant), openfactor, urbanfactor, freewayfactor, environmentalfactor

Appendix Table 3. Model 2 Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.748[a]	.559	.555	.78803055	.559	124.316	1	98	.000
2	.794[b]	.631	.623	.72513426	.071	18.738	1	97	.000
3	.835[c]	.698	.688	.65924445	.067	21.359	1	96	.000
4	.862[d]	.743	.732	.61103793	.045	16.745	1	95	.000
5	.886[e]	.785	.773	.56244896	.042	18.123	1	94	.000
6	.900[f]	.810	.797	.53146826	.025	12.278	1	93	.001
7	.906[g]	.822	.808	.51726322	.012	6.178	1	92	.015
8	.905[h]	.819	.808	.51805294	-.002	1.284	1	92	.260
9	.912[i]	.831	.818	.50322870	.012	6.560	1	92	.012
10	.916[j]	.840	.825	.49335287	.008	4.720	1	91	.032
11	.914[k]	.836	.823	.49682843	-.004	2.301	1	91	.133
12	.919[l]	.844	.831	.48581839	.009	5.217	1	91	.025
13	.925[m]	.855	.841	.47102302	.011	6.807	1	90	.011

a. Predictors: (Constant), notraffic

b. Predictors: (Constant), notraffic, Pedestrians

c. Predictors: (Constant), notraffic, Pedestrians, timeofday

d. Predictors: (Constant), notraffic, Pedestrians, timeofday, drainagechannels

e. Predictors: (Constant), notraffic, Pedestrians, timeofday, drainagechannels, lowtraffic

f. Predictors: (Constant), notraffic, Pedestrians, timeofday, drainagechannels, lowtraffic, workzones

g. Predictors: (Constant), notraffic, Pedestrians, timeofday, drainagechannels, lowtraffic, workzones, parkedcars

h. Predictors: (Constant), notraffic, timeofday, drainagechannels, lowtraffic, workzones, parkedcars

i. Predictors: (Constant), notraffic, timeofday, drainagechannels, lowtraffic, workzones, parkedcars, passingCL

Appendix Table 3 cont'd.

j. Predictors: (Constant), notraffic, timeofday, drainagechannels, lowtraffic, workzones, parkedcars, passingCL, workzonediverges
k. Predictors: (Constant), notraffic, timeofday, drainagechannels, lowtraffic, parkedcars, passingCL, workzonediverges
l. Predictors: (Constant), notraffic, timeofday, drainagechannels, lowtraffic, parkedcars, passingCL, workzonediverges, pavementmarkings
m. Predictors: (Constant), notraffic, timeofday, drainagechannels, lowtraffic, parkedcars, passingCL, workzonediverges, pavementmarkings, heavyTraffic

Appendix Table 4. Discrimination Parameters for Four Dimensional GPC Model

Item Label	P#	a1	s.e.	P#	a2	s.e.	P#	a3	s.e.	P#	a4	s.e.
Item1	1	0.26	0.07		0	----		0	----		0	----
Item2	6	0.37	0.09	7	1.05	0.14		0	----		0	----
Item3	12	0.6	0.1	13	0.65	0.1	14	0.47	0.09		0	----
Item4	19	0.68	0.1	20	0.72	0.1	21	0.47	0.09	22	0.51	0.09
Item5	27	1.4	0.16	28	1.27	0.15	29	0.72	0.12	30	0.44	0.1
Item6	35	1.16	0.14	36	0.54	0.09	37	0.57	0.1	38	0.33	0.09
Item7	43	1.25	0.15	44	1	0.13	45	0.75	0.12	46	0.29	0.1
Item8	51	0.63	0.11	52	0.84	0.12	53	0.81	0.12	54	0.94	0.13
Item9	59	0.72	0.13	60	1.55	0.18	61	0.62	0.14	62	1.04	0.15
Item10	67	1.91	0.25	68	0.47	0.1	69	0.73	0.14	70	0.48	0.12
Item11	75	1.74	0.22	76	0.51	0.1	77	0.72	0.13	78	0.39	0.11
Item12	83	1.24	0.16	84	0.56	0.09	85	0.37	0.1	86	0.43	0.1
Item13	91	1.2	0.16	92	0.5	0.09	93	0.47	0.11	94	0.41	0.1
Item14	99	1.27	0.15	100	0.79	0.11	101	0.48	0.1	102	0.54	0.1
Item15	107	1.3	0.15	108	0.72	0.1	109	0.55	0.11	110	0.59	0.1
Item16	115	1.47	0.19	116	0.54	0.1	117	0.6	0.12	118	0.4	0.11
Item17	123	0.96	0.12	124	0.86	0.11	125	0.63	0.11	126	0.24	0.09
Item18	131	1.02	0.13	132	0.7	0.1	133	0.47	0.1	134	0.51	0.1
Item19	139	1.05	0.13	140	0.81	0.11	141	0.58	0.1	142	0.37	0.09
Item20	147	0.74	0.11	148	0.76	0.11	149	0.59	0.11	150	0.99	0.13
Item21	155	0.81	0.13	156	1.19	0.15	157	0.76	0.13	158	1.11	0.15
Item22	163	0.83	0.14	164	0.87	0.14	165	0.61	0.14	166	1.34	0.19
Item23	171	0.65	0.12	172	1	0.14	173	0.59	0.12	174	1	0.15
Item24	179	0.82	0.11	180	0.74	0.1	181	0.62	0.1	182	0.59	0.1
Item25	187	0.93	0.14	188	1.1	0.15	189	0.5	0.12	190	1.34	0.17
Item26	195	0.69	0.11	196	0.82	0.12	197	0.47	0.11	198	0.83	0.12
Item27	203	0.94	0.15	204	0.7	0.12	205	0.45	0.12	206	1.42	0.2
Item28	211	0.88	0.13	212	0.39	0.08	213	0.45	0.09	214	0.72	0.11
Item29	219	0.8	0.16	220	0.81	0.16	221	1	0.18	222	1.74	0.25

Appendix Table 4 cont'd.

Item30	227	0.88	0.12	228	0.84	0.11	229	0.58	0.11	230	0.59	0.1
Item31	235	0.92	0.13	236	1.23	0.14	237	0.58	0.11	238	0.66	0.11
Item32	243	0.65	0.12	244	1.26	0.15	245	0.81	0.13	246	0.57	0.11
Item33	251	0.78	0.13	252	0.78	0.11	253	1.65	0.19	254	0.62	0.11
Item34	259	0.37	0.09	260	0.23	0.07	261	1.02	0.14	262	0.52	0.09
Item35	267	0.77	0.13	268	1.43	0.16	269	0.74	0.13	270	0.63	0.11
Item36	275	0.86	0.13	276	1.3	0.15	277	0.72	0.12	278	0.66	0.11
Item37	283	0.76	0.12	284	0.95	0.13	285	0.43	0.11	286	0.63	0.11
Item38	291	0.97	0.13	292	1.23	0.14	293	0.61	0.12	294	0.78	0.12
Item39	299	0.68	0.11	300	1.29	0.15	301	0.54	0.11	302	0.43	0.1
Item40	307	0.82	0.12	308	0.98	0.12	309	0.92	0.13	310	0.35	0.1
Item41	315	0.77	0.11	316	0.21	0.07	317	0.52	0.09	318	0.56	0.09
Item42	323	0.9	0.14	324	1.56	0.18	325	0.56	0.13	326	0.74	0.13
Item43	331	1.54	0.18	332	0.84	0.11	333	0.71	0.12	334	0.53	0.11
Item44	339	1.13	0.14	340	0.55	0.09	341	0.52	0.1	342	0.49	0.09
Item45	347	0.82	0.13	348	1.21	0.15	349	0.49	0.12	350	0.89	0.13
Item46	355	0.58	0.11	356	0.49	0.1	357	0.53	0.11	358	1.02	0.14
Item47	363	0.61	0.17	364	0.98	0.19	365	1.01	0.19	366	1.73	0.26
Item48	371	0.6	0.18	372	0.87	0.18	373	0.74	0.18	374	1.75	0.28
Item49	379	0.52	0.16	380	1.02	0.18	381	1.07	0.19	382	1.7	0.26
Item50	387	0.58	0.18	388	1.07	0.2	389	0.76	0.19	390	1.85	0.28
Item51	395	0.59	0.16	396	1.44	0.23	397	1.01	0.19	398	1.62	0.23
Item52	403	0.63	0.1	404	0.48	0.08	405	0.49	0.09	406	0.71	0.11
Item53	411	0.54	0.15	412	0.73	0.15	413	0.77	0.16	414	1.68	0.26
Item54	419	0.58	0.19	420	1.32	0.24	421	1.17	0.24	422	1.88	0.3
Item55	427	0.82	0.18	428	1.44	0.22	429	1.13	0.19	430	2	0.27
Item56	435	0.52	0.12	436	1.23	0.16	437	0.87	0.14	438	1.07	0.15
Item57	443	0.88	0.13	444	1.21	0.15	445	0.71	0.13	446	1.08	0.14
Item58	451	0.86	0.11	452	1.05	0.13	453	0.63	0.1	454	0.53	0.09
Item59	459	1.07	0.14	460	1.54	0.17	461	0.97	0.14	462	1.35	0.16
Item60	467	0.75	0.16	468	0.78	0.14	469	0.79	0.16	470	1.78	0.25
Item61	475	0.66	0.15	476	0.78	0.15	477	0.8	0.16	478	1.65	0.23
Item62	483	0.71	0.19	484	1.47	0.25	485	1.12	0.22	486	2.14	0.32
Item63	491	0.8	0.13	492	0.85	0.13	493	0.72	0.12	494	1.14	0.15
Item64	499	0.66	0.16	500	1.2	0.2	501	0.9	0.18	502	1.53	0.22
Item65	507	0.53	0.16	508	0.95	0.19	509	0.8	0.19	510	1.34	0.22
Item66	515	0.8	0.11	516	1.02	0.12	517	0.65	0.11	518	0.51	0.1
Item67	523	0.65	0.13	524	1.09	0.15	525	0.76	0.14	526	1.17	0.16
Item68	531	0.68	0.17	532	1.01	0.19	533	0.83	0.19	534	1.89	0.28
Item69	539	0.75	0.14	540	0.51	0.12	541	0.65	0.14	542	1.68	0.22

Appendix Table 4 cont'd.

Item70	547	0.7	0.13	548	1.4	0.17	549	0.81	0.14	550	0.97	0.14
Item71	555	0.68	0.12	556	1.02	0.14	557	0.94	0.14	558	1.07	0.14
Item72	563	0.85	0.13	564	0.88	0.12	565	0.83	0.13	566	1.49	0.18
Item73	571	0.54	0.12	572	1.37	0.17	573	0.7	0.13	574	0.89	0.14
Item74	579	0.54	0.1	580	0.74	0.1	581	0.54	0.1	582	0.55	0.1
Item75	587	0.5	0.11	588	0.96	0.13	589	0.78	0.12	590	0.59	0.11
Item76	595	0.47	0.13	596	0.52	0.13	597	0.83	0.15	598	1.15	0.21
Item77	603	0.43	0.17	604	1.49	0.26	605	0.75	0.19	606	1.23	0.22
Item78	611	0.66	0.15	612	0.82	0.16	613	0.52	0.15	614	1.32	0.22
Item79	619	0.68	0.14	620	0.98	0.16	621	0.58	0.14	622	1.25	0.18
Item80	627	0.81	0.15	628	1.19	0.18	629	0.86	0.16	630	1.44	0.19
Item81	635	0.66	0.12	636	0.63	0.12	637	0.54	0.12	638	0.96	0.14
Item82	643	0.54	0.12	644	0.9	0.14	645	0.83	0.14	646	0.99	0.15
Item83	651	0.61	0.13	652	1.09	0.16	653	0.83	0.15	654	1.09	0.16
Item84	659	0.61	0.16	660	0.73	0.16	661	1.05	0.19	662	1.69	0.28
Item85	667	0.78	0.13	668	1.33	0.16	669	0.58	0.12	670	0.65	0.12
Item86	675	0.4	0.14	676	0.95	0.17	677	1.08	0.18	678	1.43	0.23
Item87	683	0.76	0.2	684	1.44	0.26	685	0.75	0.21	686	1.94	0.31
Item88	691	0.92	0.14	692	1.81	0.21	693	1.34	0.17	694	0.64	0.12
Item89	699	0.42	0.1	700	0.52	0.09	701	1.1	0.14	702	0.57	0.1
Item90	707	0.67	0.11	708	0.53	0.09	709	1.06	0.13	710	0.55	0.1
Item91	715	0.62	0.13	716	1.52	0.19	717	0.79	0.14	718	0.67	0.12
Item92	723	0.59	0.13	724	0.8	0.14	725	0.76	0.14	726	0.95	0.15
Item93	731	0.66	0.1	732	0.62	0.1	733	0.75	0.11	734	0.84	0.12
Item94	739	0.78	0.14	740	0.5	0.11	741	2.03	0.27	742	0.6	0.12
Item95	747	1.12	0.16	748	0.61	0.1	749	1.96	0.24	750	0.67	0.12
Item96	755	1.06	0.16	756	0.72	0.11	757	2.08	0.26	758	0.6	0.12
Item97	763	0.86	0.15	764	0.63	0.11	765	1.87	0.23	766	0.38	0.11
Item98	771	1.05	0.14	772	0.51	0.09	773	0.67	0.12	774	0.5	0.1
Item99	779	0.93	0.13	780	1.14	0.14	781	0.7	0.12	782	0.78	0.12
Item100	787	0.41	0.1	788	0.45	0.09	789	0.95	0.13	790	0.8	0.12

Appendix Table 5. Item Fit Statistics for Four Dimensional GPC Model

Stochastic Theta Variant of Yen-Bock Item Diagnostic Tables and X2s:			
Item 1	X2(33) = 8.3, p = 1.0000	Item 51	X2(8) = 20.9, p = 0.0075
Item 2	X2(20) = 10.6, p = 0.9552	Item 52	X2(19) = 11.2, p = 0.9185
Item 3	X2(21) = 11.3, p = 0.9571	Item 53	X2(8) = 9.4, p = 0.3130
Item 4	X2(18) = 27.4, p = 0.0722	Item 54	X2(4) = 6.5, p = 0.1633
Item 5	X2(15) = 10.0, p = 0.8201	Item 55	X2(6) = 6.9, p = 0.3304
Item 6	X2(16) = 15.1, p = 0.5209	Item 56	X2(11) = 14.6, p = 0.2031
Item 7	X2(15) = 13.8, p = 0.5425	Item 57	X2(11) = 11.9, p = 0.3743

Appendix Table 5 cont’d.

Item 8	X2(15) = 9.0, p = 0.8782	Item 58	X2(17) = 15.9, p = 0.5302
Item 9	X2(11) = 13.7, p = 0.2467	Item 59	X2(11) = 11.6, p = 0.3957
Item 10	X2(9) = 11.7, p = 0.2320	Item 60	X2(9) = 18.5, p = 0.0294
Item 11	X2(11) = 4.9, p = 0.9354	Item 61	X2(8) = 7.7, p = 0.4685
Item 12	X2(16) = 10.8, p = 0.8220	Item 62	X2(5) = 3.5, p = 0.6218
Item 13	X2(15) = 14.4, p = 0.4997	Item 63	X2(11) = 14.0, p = 0.2345
Item 14	X2(15) = 13.2, p = 0.5847	Item 64	X2(7) = 7.1, p = 0.4241
Item 15	X2(14) = 8.7, p = 0.8486	Item 65	X2(6) = 8.1, p = 0.2282
Item 16	X2(13) = 15.9, p = 0.2555	Item 66	X2(16) = 8.8, p = 0.9219
Item 17	X2(17) = 14.5, p = 0.6358	Item 67	X2(11) = 9.9, p = 0.5420
Item 18	X2(16) = 21.4, p = 0.1633	Item 68	X2(6) = 11.2, p = 0.0825
Item 19	X2(16) = 11.2, p = 0.7999	Item 69	X2(11) = 11.1, p = 0.4337
Item 20	X2(15) = 13.3, p = 0.5772	Item 70	X2(11) = 9.0, p = 0.6190
Item 21	X2(11) = 8.3, p = 0.6895	Item 71	X2(12) = 12.5, p = 0.4069
Item 22	X2(9) = 13.9, p = 0.1241	Item 72	X2(14) = 5.6, p = 0.9749
Item 23	X2(13) = 8.6, p = 0.7999	Item 73	X2(12) = 7.9, p = 0.7926
Item 24	X2(17) = 24.2, p = 0.1132	Item 74	X2(17) = 13.1, p = 0.7283
Item 25	X2(11) = 8.8, p = 0.6390	Item 75	X2(15) = 12.4, p = 0.6502
Item 26	X2(14) = 12.9, p = 0.5391	Item 76	X2(10) = -1.$, p < 0.0001
Item 27	X2(10) = 9.9, p = 0.4468	Item 77	X2(7) = 6.4, p = 0.4995
Item 28	X2(18) = 9.0, p = 0.9604	Item 78	X2(9) = 7.6, p = 0.5712
Item 29	X2(6) = 6.6, p = 0.3580	Item 79	X2(9) = 8.6, p = 0.4785
Item 30	X2(17) = 5.6, p = 0.9953	Item 80	X2(7) = 6.1, p = 0.5239
Item 31	X2(14) = 10.9, p = 0.6986	Item 81	X2(12) = 13.1, p = 0.3650
Item 32	X2(12) = 9.4, p = 0.6697	Item 82	X2(12) = 10.3, p = 0.5933
Item 33	X2(14) = 10.4, p = 0.7322	Item 83	X2(10) = 9.6, p = 0.4816
Item 34	X2(18) = 5.6, p = 0.9976	Item 84	X2(5) = 7.5, p = 0.1846
Item 35	X2(13) = 17.0, p = 0.1974	Item 85	X2(13) = 20.8, p = 0.0763
Item 36	X2(15) = 8.7, p = 0.8936	Item 86	X2(9) = 11.4, p = 0.2500
Item 37	X2(15) = 12.9, p = 0.6128	Item 87	X2(4) = 7.7, p = 0.1020
Item 38	X2(14) = 10.3, p = 0.7439	Item 88	X2(10) = 7.7, p = 0.6576
Item 39	X2(15) = 11.0, p = 0.7516	Item 89	X2(19) = 16.4, p = 0.6292
Item 40	X2(16) = 8.8, p = 0.9210	Item 90	X2(17) = 14.1, p = 0.6602
Item 41	X2(17) = 15.0, p = 0.5985	Item 91	X2(12) = 5.6, p = 0.9330
Item 42	X2(10) = 8.1, p = 0.6200	Item 92	X2(9) = 5.8, p = 0.7595
Item 43	X2(11) = 9.2, p = 0.6077	Item 93	X2(15) = 21.3, p = 0.1259
Item 44	X2(16) = 12.3, p = 0.7227	Item 94	X2(12) = 5.6, p = 0.9341
Item 45	X2(12) = 13.7, p = 0.3193	Item 95	X2(12) = 11.2, p = 0.5101
Item 46	X2(15) = 14.8, p = 0.4682	Item 96	X2(12) = 6.2, p = 0.9055
Item 47	X2(7) = 17.3, p = 0.0157	Item 97	X2(10) = 10.4, p = 0.4079

Appendix Table 5 cont'd.

Item 48	X2(6) = 15.4, p = 0.0174	Item 98	X2(16) = 16.2, p = 0.4399
Item 49	X2(7) = 7.4, p = 0.3859	Item 99	X2(14) = 8.1, p = 0.8840
Item 50	X2(5) = 7.9, p = 0.1607	Item 100	0 X2(17) = 9.9, p = 0.9070

Appendix Table 6. Intercepts & Multidimensional Step Parameters for Four Dimensional GPC

Item Label	Int 1	Int 2	Int 3	Int 4	MStep 1	MStep 2	MStep 3	MStep 4
Item1	-0.42468	-0.23532	0.79532	0.90468	-1.63338	-0.90508	3.058922	3.47954
Item2	-1.08534	0.505341	1.93466	2.365341	-0.9749	0.453919	1.737796	2.124653
Item3	-1.45593	-0.86407	0.564071	1.315929	-1.45346	-0.86261	0.563114	1.313698
Item4	-1.56078	-0.07922	0.499218	1.380782	-1.29092	-0.06552	0.412901	1.14204
Item5	-2.65404	-1.86596	-0.13404	1.334041	-1.28215	-0.90143	-0.06475	0.644464
Item6	-2.20129	-1.14872	-0.69129	0.641285	-1.52963	-0.79822	-0.48036	0.445616
Item7	-2.54969	-2.14031	-0.47969	1.169691	-1.4233	-1.19477	-0.26777	0.652949
Item8	-1.1817	0.511701	1.248299	2.141701	-0.72698	0.314799	0.767955	1.317577
Item9	-0.74182	0.911823	2.508177	3.401823	-0.35418	0.435343	1.197511	1.624175
Item10	-2.76773	-3.78227	-1.77773	-0.63227	-1.28595	-1.75733	-0.82598	-0.29377
Item11	-3.59756	-2.81244	-1.92756	-0.30244	-1.80825	-1.41363	-0.96886	-0.15202
Item12	-2.51141	-1.53859	-0.58141	1.151407	-1.70368	-1.04374	-0.39441	0.781086
Item13	-2.45543	-2.01457	-1.26543	0.415427	-1.70294	-1.39719	-0.87763	0.288116
Item14	-2.68827	-1.76173	-0.45827	0.74827	-1.61844	-1.06063	-0.2759	0.450487
Item15	-2.75141	-1.81859	-0.78141	0.911407	-1.62723	-1.07554	-0.46214	0.53902
Item16	-3.04928	-2.58073	-1.35928	0.189275	-1.76862	-1.49686	-0.7884	0.109782
Item17	-2.42442	-1.14558	-0.27442	0.644422	-1.66679	-0.78759	-0.18867	0.44304
Item18	-2.35785	-1.55215	-0.84785	0.597853	-1.66252	-1.09442	-0.59782	0.421546
Item19	-2.78777	-1.37223	-0.30777	0.747767	-1.86603	-0.91852	-0.20601	0.500529
Item20	-1.37928	0.389275	0.810726	1.859275	-0.88058	0.248526	0.517594	1.187024
Item21	-1.15123	0.781234	1.338766	3.671234	-0.58431	0.396514	0.679489	1.863331
Item22	0.766203	1.153797	1.326203	3.113797	0.403069	0.606968	0.697663	1.638047
Item23	-0.35246	1.362462	1.737538	1.132462	-0.21175	0.818535	1.043872	0.680357
Item24	-1.59706	-0.72294	0.802944	1.797056	-1.14294	-0.51738	0.574632	1.286075
Item25	-0.5117	0.781701	1.998299	2.811701	-0.25208	0.385093	0.984432	1.385142
Item26	-0.51484	0.764838	1.075162	3.034838	-0.35885	0.533108	0.74941	2.115345
Item27	0.390675	1.139325	1.180675	1.649325	0.206119	0.601103	0.62292	0.870178
Item28	-1.58836	-1.09164	-0.42836	1.268356	-1.23751	-0.85051	-0.33374	0.988192
Item29	0.818837	1.931163	1.928837	2.841163	0.354888	0.836974	0.835966	1.231372
Item30	-1.83697	-0.59303	0.29303	2.13697	-1.2486	-0.40309	0.199174	1.452512
Item31	-1.73739	-0.03261	1.212614	2.717386	-0.98182	-0.01843	0.685267	1.535639
Item32	-0.7407	0.990696	1.289304	2.780696	-0.42828	0.572829	0.745487	1.607822
Item33	-2.97973	-1.30027	-0.13973	2.139727	-1.43301	-0.62533	-0.0672	1.029036

Appendix Table 6 cont'd.

Item34	-1.0977	-1.0023	-0.3777	-0.3623	-0.89609	-0.81821	-0.30833	-0.29576
Item35	-1.55505	0.035046	1.964954	2.235046	-0.8216	0.018517	1.038177	1.180879
Item36	-1.60919	-0.53081	1.570812	2.209188	-0.87482	-0.28857	0.853954	1.201001
Item37	-1.00635	0.246346	1.553654	2.246346	-0.70083	0.171558	1.081983	1.564382
Item38	-1.28182	0.251823	1.408177	2.861823	-0.69169	0.135886	0.759868	1.544271
Item39	-1.90626	-0.14374	1.563741	1.926259	-1.18153	-0.08909	0.969232	1.193927
Item40	-2.08747	-0.93253	0.592527	1.787473	-1.29417	-0.57814	0.36735	1.108181
Item41	-1.77338	-1.12662	-0.71338	0.093381	-1.60489	-1.01958	-0.6456	0.084509
Item42	-0.89839	0.688391	2.771609	3.358391	-0.44342	0.339773	1.367999	1.657621
Item43	-3.50049	-2.81951	-1.86049	-0.21951	-1.78119	-1.43469	-0.94669	-0.1117
Item44	-2.643	-1.647	-0.693	0.143	-1.82825	-1.13928	-0.47937	0.098918
Item45	-1.54023	0.770229	1.729771	3.480229	-0.86526	0.432692	0.971735	1.955092
Item46	-0.22158	0.591579	0.628421	2.281579	-0.16084	0.429425	0.456169	1.656191
Item47	0.996289	2.173711	2.186289	3.923711	0.430915	0.940173	0.945613	1.697082
Item48	1.26135	2.03865	2.76135	3.77865	0.580153	0.93767	1.270073	1.737976
Item49	1.213482	1.566518	2.873482	3.306518	0.524845	0.677537	1.242814	1.430107
Item50	0.999046	2.700955	3.079046	2.780955	0.426717	1.153643	1.315135	1.187813
Item51	0.899132	2.150868	2.449132	3.260868	0.365061	0.873284	0.994383	1.32396
Item52	-0.97602	-0.13398	0.373985	1.216016	-0.83341	-0.11441	0.319342	1.038344
Item53	1.113726	1.346274	2.773726	1.566274	0.540886	0.653824	1.347072	0.760668
Item54	1.380761	2.419239	4.100761	3.219239	0.522547	0.915558	1.551928	1.218317
Item55	0.311939	2.228061	3.151939	3.508061	0.110129	0.786614	1.112787	1.238516
Item56	-0.66149	1.041493	2.438507	2.421493	-0.34459	0.54254	1.27028	1.261417
Item57	-1.12174	0.321737	1.678263	3.601737	-0.56736	0.16273	0.848843	1.821709
Item58	-0.98057	0.130574	0.449426	1.720574	-0.61772	0.082256	0.283117	1.083881
Item59	-1.57023	0.460229	1.699771	3.450229	-0.62659	0.183653	0.678288	1.376802
Item60	0.503898	1.516102	2.423898	2.116102	0.22618	0.680517	1.08799	0.949832
Item61	0.872563	1.637437	1.782563	3.347437	0.415671	0.780041	0.849177	1.59465
Item62	1.294573	2.495427	3.164573	4.165427	0.444061	0.855975	1.085504	1.428814
Item63	-0.14229	0.95229	1.04771	2.50229	-0.07979	0.533976	0.587481	1.403104
Item64	0.84972	2.12028	2.73972	3.89028	0.379003	0.945715	1.222006	1.735194
Item65	1.319612	1.913388	32.96661	-27.0396	0.693669	1.005793	17.32926	-14.2137
Item66	-1.57291	-0.24709	0.787086	1.792914	-1.02323	-0.16074	0.512024	1.166347
Item67	-0.27907	1.339066	2.140934	2.719066	-0.14796	0.709993	1.135155	1.44169
Item68	1.229842	2.340158	2.259842	3.450158	0.51317	0.976465	0.942953	1.439629
Item69	0.046411	1.543589	1.456411	2.153589	0.023012	0.765364	0.722138	1.067822
Item70	-0.74312	0.633122	2.086878	2.863122	-0.3694	0.314718	1.037362	1.423224
Item71	-0.55602	0.866016	1.733985	1.636016	-0.29588	0.460849	0.922737	0.870603

Appendix Table 6 cont'd.

Item72	-1.15969	0.209691	1.430309	2.559691	-0.55249	0.099899	0.681416	1.219467
Item73	-0.65342	1.023417	2.016583	2.613417	-0.35176	0.550941	1.085598	1.406894
Item74	-1.35049	-0.12951	0.669512	1.930488	-1.12803	-0.10818	0.559229	1.612495
Item75	-0.66643	0.896432	1.323568	2.006432	-0.45684	0.614499	0.907299	1.375398
Item76	1.634142	1.725858	0.594142	1.605858	1.032965	1.090939	0.375566	1.015086
Item77	1.520883	2.459117	2.240883	2.539117	0.718518	1.161772	1.05867	1.199566
Item78	1.390259	1.879741	0.800259	3.129741	0.786978	1.064058	0.452999	1.77164
Item79	0.296619	1.563381	2.176619	2.163381	0.162749	0.857797	1.194269	1.187006
Item80	0.166375	1.473625	2.206375	3.673625	0.075273	0.66671	0.998227	1.662052
Item81	0.298127	0.851873	1.548127	2.461873	0.208439	0.595599	1.082394	1.721253
Item82	0.388837	1.221163	1.538837	2.411163	0.233604	0.733646	0.924498	1.44857
Item83	0.247538	1.002462	2.777538	1.732462	0.133518	0.540711	1.498158	0.934461
Item84	1.858665	1.791335	1.348665	2.721335	0.842795	0.812265	0.61154	1.233966
Item85	-1.30061	0.36061	1.89939	2.80061	-0.73442	0.203627	1.072537	1.581433
Item86	1.144817	2.085183	1.554817	2.375183	0.553769	1.008642	0.752094	1.14892
Item87	1.78537	1.81463	3.82537	5.09463	0.675903	0.68698	1.448204	1.928719
Item88	-1.91085	-0.02915	1.709147	3.830854	-0.75963	-0.01159	0.679448	1.522904
Item89	-1.92283	-0.36717	0.427172	2.022828	-1.36591	-0.26083	0.303448	1.436944
Item90	-1.88513	-0.56487	0.164868	1.325132	-1.28389	-0.38471	0.112285	0.902495
Item91	-0.90579	1.095793	1.824207	4.425793	-0.46664	0.564527	0.939789	2.280065
Item92	0.263274	1.206726	2.413274	2.596726	0.167578	0.768101	1.536089	1.65286
Item93	-0.50782	0.327818	0.932183	1.047818	-0.35142	0.226859	0.645097	0.72512
Item94	-2.85969	-2.71031	-0.76969	0.859691	-1.23759	-1.17294	-0.3331	0.372049
Item95	-3.04291	-2.51709	-1.16291	0.322914	-1.25095	-1.03478	-0.47808	0.132751
Item96	-3.36848	-2.67152	-1.14848	0.308478	-1.33902	-1.06197	-0.45654	0.122624
Item97	-3.57898	-2.99102	-1.26898	-0.00102	-1.63736	-1.36838	-0.58055	-0.00047
Item98	-2.38605	-1.63395	-0.50605	1.206051	-1.66184	-1.13801	-0.35245	0.83999
Item99	-1.86508	-0.40492	0.984918	3.325082	-1.03251	-0.22416	0.545253	1.840774
Item100	-0.28652	-0.09348	0.933482	1.806518	-0.20715	-0.06759	0.674896	1.306091

Appendix Table 7. Discrimination Parameters for Four Dimensional GRM Model

Item Label	P#	a 1	s.e.	P#	a 2	s.e.	P#	a 3	s.e.	P#	a 4	s.e.
Item1	5	0.42	0.11		0	----		0	----		0	----
Item2	10	1.45	0.16	11	0.53	0.13		0	----		0	----
Item3	16	0.96	0.13	17	0.85	0.13	18	0.75	0.13		0	----
Item4	23	0.99	0.13	24	0.67	0.14	25	0.96	0.14	26	0.67	0.13
Item5	31	1.63	0.16	32	1.13	0.15	33	1.55	0.17	34	0.42	0.13
Item6	39	0.75	0.12	40	0.91	0.14	41	1.49	0.16	42	0.35	0.12

Appendix Table 7 cont'd.

Item7	47	1.31	0.14	48	1.09	0.15	49	1.45	0.17	50	0.27	0.12
Item8	55	0.94	0.13	56	0.84	0.14	57	0.64	0.14	58	1.07	0.13
Item9	63	1.78	0.19	64	0.64	0.16	65	0.62	0.16	66	1.14	0.15
Item10	71	0.56	0.13	72	1.19	0.17	73	2.29	0.25	74	0.47	0.14
Item11	79	0.53	0.13	80	1.26	0.18	81	2.15	0.24	82	0.41	0.14
Item12	87	0.74	0.12	88	0.61	0.13	89	1.55	0.17	90	0.42	0.12
Item13	95	0.71	0.12	96	0.76	0.14	97	1.48	0.17	98	0.42	0.12
Item14	103	1.08	0.13	104	0.7	0.13	105	1.58	0.17	106	0.54	0.12
Item15	111	0.99	0.13	112	0.85	0.14	113	1.71	0.18	114	0.63	0.13
Item16	119	0.67	0.12	120	0.88	0.15	121	1.99	0.21	122	0.4	0.13
Item17	127	1.24	0.14	128	1.04	0.15	129	1.2	0.15	130	0.25	0.12
Item18	135	0.98	0.13	136	0.75	0.14	137	1.33	0.16	138	0.53	0.12
Item19	143	1.1	0.13	144	0.85	0.14	145	1.34	0.15	146	0.38	0.12
Item20	151	1.08	0.13	152	0.74	0.14	153	0.88	0.15	154	1.22	0.15
Item21	159	1.37	0.15	160	0.92	0.15	161	0.8	0.15	162	1.35	0.16
Item22	167	1.03	0.17	168	0.56	0.17	169	1.04	0.18	170	1.65	0.2
Item23	175	1.3	0.16	176	0.66	0.15	177	0.64	0.15	178	1.19	0.15
Item24	183	1.05	0.13	184	0.78	0.14	185	1.06	0.15	186	0.65	0.12
Item25	191	1.26	0.15	192	0.44	0.15	193	1	0.16	194	1.48	0.17
Item26	199	0.99	0.13	200	0.52	0.14	201	0.8	0.15	202	0.97	0.14
Item27	207	0.93	0.16	208	0.43	0.16	209	1.16	0.2	210	1.69	0.21
Item28	215	0.51	0.11	216	0.56	0.13	217	1.33	0.17	218	0.88	0.14
Item29	223	0.74	0.17	224	0.92	0.2	225	0.75	0.19	226	2.06	0.25
Item30	231	1.12	0.13	232	0.88	0.14	233	1.05	0.15	234	0.66	0.13
Item31	239	1.57	0.15	240	0.73	0.14	241	0.87	0.15	242	0.67	0.13
Item32	247	1.48	0.16	248	1.05	0.16	249	0.6	0.14	250	0.67	0.13
Item33	255	0.74	0.12	256	2.09	0.2	257	0.67	0.15	258	0.77	0.14
Item34	263	0.25	0.11	264	1.61	0.18	265	0.42	0.14	266	0.87	0.14
Item35	271	1.73	0.17	272	1.05	0.15	273	0.69	0.14	274	0.7	0.13
Item36	279	1.67	0.16	280	0.99	0.15	281	0.78	0.14	282	0.76	0.13
Item37	287	1.28	0.15	288	0.5	0.14	289	0.78	0.15	290	0.68	0.13
Item38	295	1.49	0.15	296	0.79	0.14	297	1.05	0.16	298	0.84	0.13
Item39	303	1.61	0.16	304	0.79	0.14	305	0.61	0.14	306	0.46	0.12
Item40	311	1.25	0.14	312	1.25	0.15	313	0.72	0.14	314	0.34	0.12
Item41	319	0.28	0.11	320	0.71	0.13	321	1.13	0.16	322	0.73	0.13
Item42	327	1.71	0.18	328	0.75	0.16	329	0.77	0.15	330	0.73	0.13
Item43	335	0.99	0.13	336	1.07	0.16	337	1.85	0.21	338	0.5	0.13
Item44	343	0.72	0.12	344	0.79	0.14	345	1.62	0.17	346	0.54	0.12
Item45	351	1.43	0.15	352	0.55	0.14	353	0.8	0.15	354	0.96	0.14
Item46	359	0.62	0.13	360	0.54	0.15	361	0.8	0.15	362	1.44	0.17

Appendix Table 7 cont'd.

Item47	367	1.1	0.2	368	0.88	0.2	369	0.62	0.2	370	2.17	0.27
Item48	375	0.89	0.19	376	0.64	0.2	377	0.61	0.21	378	2.04	0.28
Item49	383	0.9	0.19	384	1.03	0.22	385	0.61	0.2	386	2.34	0.29
Item50	391	1.1	0.2	392	0.66	0.2	393	0.6	0.2	394	2.18	0.27
Item51	399	1.41	0.21	400	0.84	0.2	401	0.5	0.19	402	1.91	0.23
Item52	407	0.66	0.11	408	0.59	0.13	409	0.83	0.14	410	0.97	0.13
Item53	415	0.76	0.17	416	0.65	0.19	417	0.7	0.2	418	2.15	0.26
Item54	423	1.1	0.22	424	0.87	0.23	425	0.56	0.21	426	2.15	0.29
Item55	431	1.35	0.2	432	0.91	0.19	433	0.7	0.19	434	2.16	0.25
Item56	439	1.38	0.16	440	0.96	0.16	441	0.46	0.15	442	1.29	0.16
Item57	447	1.38	0.15	448	0.8	0.15	449	0.86	0.15	450	1.22	0.15
Item58	455	1.52	0.16	456	0.83	0.15	457	0.99	0.15	458	0.57	0.12
Item59	463	1.72	0.17	464	1.1	0.16	465	0.95	0.15	466	1.44	0.16
Item60	471	0.79	0.16	472	0.71	0.18	473	0.86	0.19	474	2.06	0.23
Item61	479	0.8	0.16	480	0.66	0.18	481	0.66	0.18	482	1.98	0.23
Item62	487	1.45	0.24	488	0.88	0.23	489	0.65	0.22	490	2.43	0.31
Item63	495	0.96	0.15	496	0.85	0.15	497	0.98	0.16	498	1.46	0.17
Item64	503	1.14	0.19	504	0.75	0.19	505	0.63	0.19	506	1.67	0.21
Item65	511	0.99	0.14	512	0.99	0.15	513	0.99	0.15	514	0.99	0.14
Item66	519	1.36	0.14	520	0.95	0.14	521	0.81	0.14	522	0.57	0.13
Item67	527	1.19	0.16	528	0.83	0.16	529	0.6	0.15	530	1.29	0.16
Item68	535	0.99	0.19	536	0.69	0.2	537	0.76	0.21	538	2.14	0.27
Item69	543	0.61	0.15	544	0.5	0.16	545	0.93	0.18	546	1.86	0.22
Item70	551	1.54	0.17	552	0.9	0.16	553	0.59	0.15	554	1.11	0.15
Item71	559	1.2	0.15	560	0.92	0.16	561	0.62	0.15	562	1.33	0.16
Item72	567	0.97	0.14	568	0.81	0.15	569	0.97	0.16	570	1.85	0.19
Item73	575	1.58	0.17	576	0.87	0.16	577	0.46	0.15	578	1.03	0.15
Item74	583	1.05	0.13	584	0.77	0.14	585	0.62	0.14	586	0.76	0.13
Item75	591	1.17	0.14	592	0.89	0.15	593	0.43	0.13	594	0.77	0.13
Item76	599	0.64	0.18	600	0.78	0.2	601	0.61	0.19	602	1.54	0.22
Item77	607	1.46	0.23	608	0.56	0.21	609	0.4	0.2	610	1.52	0.22
Item78	615	1.05	0.19	616	0.51	0.19	617	0.64	0.19	618	1.55	0.21
Item79	623	1.2	0.17	624	0.49	0.16	625	0.69	0.17	626	1.41	0.18
Item80	631	1.18	0.17	632	0.75	0.17	633	0.83	0.18	634	1.64	0.19
Item81	639	0.79	0.15	640	0.47	0.15	641	0.82	0.16	642	1.2	0.16
Item82	647	1.06	0.16	648	0.76	0.16	649	0.67	0.16	650	1.18	0.16
Item83	655	1.18	0.17	656	0.84	0.18	657	0.56	0.16	658	1.38	0.17
Item84	663	0.82	0.2	664	0.86	0.22	665	0.59	0.22	666	1.98	0.28
Item85	671	1.62	0.16	672	0.81	0.16	673	0.75	0.15	674	0.72	0.14
Item86	679	1.03	0.2	680	0.99	0.21	681	0.36	0.18	682	1.78	0.23

Appendix Table 7 cont'd.

Item87	687	1.37	0.25	688	0.61	0.22	689	0.71	0.22	690	2.09	0.29
Item88	695	1.97	0.2	696	1.47	0.18	697	0.68	0.15	698	0.58	0.13
Item89	703	0.61	0.12	704	1.5	0.16	705	0.34	0.13	706	0.79	0.13
Item90	711	0.57	0.12	712	1.41	0.16	713	0.71	0.14	714	0.71	0.13
Item91	719	1.65	0.18	720	1.01	0.17	721	0.44	0.15	722	0.82	0.14
Item92	727	0.87	0.15	728	0.83	0.16	729	0.54	0.16	730	1.08	0.16
Item93	735	0.88	0.13	736	1.01	0.15	737	0.88	0.15	738	1.21	0.15
Item94	743	0.4	0.12	744	2.7	0.26	745	0.73	0.16	746	0.82	0.15
Item95	751	0.52	0.12	752	2.71	0.28	753	1.16	0.18	754	0.91	0.16
Item96	759	0.67	0.14	760	2.98	0.31	761	1	0.18	762	0.83	0.15
Item97	767	0.57	0.13	768	2.61	0.27	769	0.79	0.16	770	0.54	0.14
Item98	775	0.64	0.12	776	0.86	0.14	777	1.28	0.16	778	0.58	0.13
Item99	783	1.36	0.14	784	0.89	0.14	785	0.99	0.15	786	0.84	0.13
Item100	791	0.45	0.12	792	1.26	0.17	793	0.46	0.15	794	1.21	0.15

Appendix Table 8. Item Fit Statistics for Four Dimensional GRM Model

Stochastic Theta Variant of Yen-Bock Item Diagnostic Tables and X2s			
Item 1	X2(34) = 41.4, p = 0.1798	Item 51	X2(9) = 16.9, p = 0.0506
Item 2	X2(23) = 11.4, p = 0.9792	Item 52	X2(24) = 10.7, p = 0.9908
Item 3	X2(24) = 11.5, p = 0.9852	Item 53	X2(12) = 13.2, p = 0.3593
Item 4	X2(21) = 22.8, p = 0.3556	Item 54	X2(6) = 5.7, p = 0.4615
Item 5	X2(16) = 6.5, p = 0.9817	Item 55	X2(8) = 10.2, p = 0.2539
Item 6	X2(18) = 16.3, p = 0.5746	Item 56	X2(12) = 8.3, p = 0.7625
Item 7	X2(17) = 10.8, p = 0.8663	Item 57	X2(14) = 10.6, p = 0.7149
Item 8	X2(20) = 14.3, p = 0.8172	Item 58	X2(21) = 12.9, p = 0.9124
Item 9	X2(12) = 10.6, p = 0.5630	Item 59	X2(13) = 8.9, p = 0.7831
Item 10	X2(10) = 9.2, p = 0.5113	Item 60	X2(13) = 12.5, p = 0.4899
Item 11	X2(11) = 11.7, p = 0.3889	Item 61	X2(12) = 8.9, p = 0.7089
Item 12	X2(18) = 14.2, p = 0.7186	Item 62	X2(7) = 5.6, p = 0.5864
Item 13	X2(15) = 8.6, p = 0.8998	Item 63	X2(17) = 13.1, p = 0.7332
Item 14	X2(17) = 9.2, p = 0.9328	Item 64	X2(9) = 11.4, p = 0.2508
Item 15	X2(15) = 8.5, p = 0.9015	Item 65	X2(15) = 162.1, p < 0.0001
Item 16	X2(12) = 10.7, p = 0.5598	Item 66	X2(20) = 10.2, p = 0.9638
Item 17	X2(19) = 11.0, p = 0.9256	Item 67	X2(13) = 10.9, p = 0.6188
Item 18	X2(17) = 18.5, p = 0.3589	Item 68	X2(9) = 10.2, p = 0.3335
Item 19	X2(19) = 9.4, p = 0.9654	Item 69	X2(16) = 15.5, p = 0.4910
Item 20	X2(19) = 19.7, p = 0.4146	Item 70	X2(15) = 10.6, p = 0.7827
Item 21	X2(14) = 4.8, p = 0.9880	Item 71	X2(16) = 8.7, p = 0.9243
Item 22	X2(13) = 14.8, p = 0.3222	Item 72	X2(16) = 5.9, p = 0.9891
Item 23	X2(17) = 8.3, p = 0.9604	Item 73	X2(16) = 5.1, p = 0.9951
Item 24	X2(20) = 25.5, p = 0.1832	Item 74	X2(19) = 16.8, p = 0.6043

Appendix Table 8 cont'd.

Item 25	X2(15) = 5.6, p = 0.9858	Item 75	X2(18) = 13.7, p = 0.7505
Item 26	X2(17) = 11.6, p = 0.8215	Item 76	X2(12) = 19.8, p = 0.0700
Item 27	X2(18) = 11.8, p = 0.8596	Item 77	X2(8) = 9.4, p = 0.3136
Item 28	X2(20) = 11.7, p = 0.9270	Item 78	X2(11) = 6.6, p = 0.8302
Item 29	X2(10) = 6.7, p = 0.7565	Item 79	X2(14) = 9.8, p = 0.7767
Item 30	X2(19) = 8.9, p = 0.9752	Item 80	X2(11) = 6.4, p = 0.8494
Item 31	X2(17) = 17.4, p = 0.4271	Item 81	X2(16) = 16.8, p = 0.3990
Item 32	X2(18) = 10.3, p = 0.9204	Item 82	X2(16) = 11.5, p = 0.7794
Item 33	X2(16) = 10.6, p = 0.8345	Item 83	X2(14) = 8.2, p = 0.8799
Item 34	X2(21) = 8.2, p = 0.9940	Item 84	X2(11) = 7.2, p = 0.7798
Item 35	X2(17) = 22.6, p = 0.1610	Item 85	X2(16) = 15.2, p = 0.5125
Item 36	X2(18) = 15.6, p = 0.6248	Item 86	X2(11) = 15.7, p = 0.1534
Item 37	X2(19) = 13.6, p = 0.8095	Item 87	X2(6) = 10.9, p = 0.0898
Item 38	X2(17) = 9.4, p = 0.9285	Item 88	X2(13) = 7.8, p = 0.8579
Item 39	X2(19) = 10.3, p = 0.9444	Item 89	X2(19) = 15.2, p = 0.7127
Item 40	X2(20) = 6.8, p = 0.9974	Item 90	X2(21) = 22.8, p = 0.3572
Item 41	X2(20) = 6.8, p = 0.9972	Item 91	X2(15) = 6.5, p = 0.9706
Item 42	X2(12) = 6.8, p = 0.8738	Item 92	X2(14) = 12.5, p = 0.5666
Item 43	X2(11) = 7.5, p = 0.7621	Item 93	X2(21) = 18.0, p = 0.6524
Item 44	X2(17) = 14.3, p = 0.6458	Item 94	X2(13) = 12.4, p = 0.5004
Item 45	X2(16) = 11.7, p = 0.7681	Item 95	X2(13) = 14.2, p = 0.3644
Item 46	X2(20) = 18.6, p = 0.5464	Item 96	X2(11) = 11.0, p = 0.4484
Item 47	X2(8) = 12.8, p = 0.1202	Item 97	X2(11) = 9.1, p = 0.6123
Item 48	X2(9) = 13.7, p = 0.1350	Item 98	X2(18) = 8.7, p = 0.9655
Item 49	X2(9) = 4.5, p = 0.8739	Item 99	X2(17) = 8.2, p = 0.9621
Item 50	X2(8) = 5.2, p = 0.7318	Item 100	0 X2(22) = 14.2, p = 0.8944

Appendix Table 9. Intercepts and multidimensional category boundaries for GRM model

Item Label	Int 1	Int 2	Int 3	Int 4	MCB 1	MCB 2	MCB 3	MCB 4
Item1	1.57	0.24	-1.25	-2.63	-3.74	-0.57	2.98	6.26
Item2	1.93	-0.53	-2.68	-4.48	-1.25	0.34	1.74	2.90
Item3	3.14	1.38	-0.61	-2.45	-2.11	-0.93	0.41	1.65
Item4	2.89	0.69	-0.87	-2.7	-1.73	-0.41	0.52	1.61
Item5	4.79	2.89	0.47	-1.9	-1.88	-1.13	-0.18	0.74
Item6	4.28	2.48	0.99	-1.06	-2.22	-1.28	-0.51	0.55
Item7	4.88	3.21	0.9	-1.63	-2.17	-1.42	-0.40	0.72
Item8	1.98	-0.36	-2.13	-4.09	-1.12	0.20	1.20	2.31
Item9	1.45	-1.08	-3.51	-5.73	-0.63	0.47	1.53	2.50
Item10	6.33	5.36	2.98	0.78	-2.36	-2.00	-1.11	-0.29
Item11	6.68	4.81	2.94	0.45	-2.59	-1.86	-1.14	-0.17

Appendix Table 9 cont'd.

Item12	4.55	2.7	0.92	-1.53	-2.43	-1.44	-0.49	0.82
Item13	5.1	3.56	1.88	-0.51	-2.75	-1.92	-1.01	0.27
Item14	4.93	2.97	0.89	-1.23	-2.34	-1.41	-0.42	0.58
Item15	5.16	3.23	1.27	-1.26	-2.30	-1.44	-0.57	0.56
Item16	6.04	4.27	2.17	-0.27	-2.61	-1.85	-0.94	0.12
Item17	4.49	2.35	0.61	-1.3	-2.21	-1.16	-0.30	0.64
Item18	4.62	2.95	1.33	-0.94	-2.44	-1.56	-0.70	0.50
Item19	4.83	2.52	0.61	-1.35	-2.45	-1.28	-0.31	0.69
Item20	2.44	-0.01	-1.54	-3.51	-1.22	0.01	0.77	1.76
Item21	2.04	-0.62	-2.43	-5.53	-0.90	0.27	1.07	2.43
Item22	-0.08	-1.79	-3.12	-5.57	0.04	0.79	1.37	2.45
Item23	1.02	-1.52	-3.04	-4	-0.51	0.76	1.53	2.01
Item24	3.12	1.16	-1.05	-3.16	-1.73	-0.64	0.58	1.75
Item25	1.29	-0.93	-3.03	-5.17	-0.58	0.42	1.36	2.32
Item26	1.28	-0.77	-2.35	-5.11	-0.76	0.46	1.40	3.04
Item27	0.39	-1.32	-2.36	-3.74	-0.17	0.58	1.03	1.63
Item28	3.51	1.99	0.47	-1.85	-1.99	-1.13	-0.27	1.05
Item29	-0.36	-2.62	-3.97	-5.67	0.14	1.05	1.59	2.28
Item30	3.3	1.19	-0.6	-3.17	-1.75	-0.63	0.32	1.68
Item31	2.81	0.31	-1.8	-4.33	-1.37	-0.15	0.88	2.11
Item32	1.48	-0.96	-2.54	-4.81	-0.73	0.47	1.25	2.37
Item33	4.65	2.19	0.06	-2.88	-1.91	-0.90	-0.02	1.18
Item34	3.25	2.04	0.83	-0.37	-1.72	-1.08	-0.44	0.20
Item35	2.65	0.13	-2.52	-4.24	-1.18	-0.06	1.12	1.88
Item36	2.94	0.79	-1.86	-3.96	-1.32	-0.35	0.84	1.78
Item37	1.92	-0.21	-2.27	-4.32	-1.12	0.12	1.32	2.51
Item38	2.27	-0.06	-2.05	-4.54	-1.05	0.03	0.95	2.10
Item39	2.99	0.41	-1.86	-3.61	-1.53	-0.21	0.95	1.85
Item40	3.56	1.44	-0.74	-2.9	-1.84	-0.74	0.38	1.50
Item41	3.92	2.39	1.1	-0.51	-2.53	-1.55	-0.71	0.33
Item42	1.62	-0.82	-3.63	-5.61	-0.75	0.38	1.69	2.61
Item43	6.35	4.63	2.77	0.29	-2.64	-1.92	-1.15	-0.12
Item44	5.04	3.13	1.31	-0.56	-2.50	-1.55	-0.65	0.28
Item45	2.33	-0.57	-2.63	-5.47	-1.18	0.29	1.33	2.77
Item46	1.19	-0.57	-1.86	-4.16	-0.65	0.31	1.01	2.26
Item47	-0.6	-3.17	-4.52	-6.76	0.23	1.19	1.70	2.54
Item48	-0.91	-2.95	-4.64	-6.76	0.38	1.23	1.94	2.82
Item49	-0.73	-2.74	-4.99	-6.75	0.26	0.99	1.80	2.43

Appendix Table 9 cont'd.

Item50	-0.65	-3.41	-4.8	-5.95	0.25	1.31	1.85	2.29
Item51	-0.51	-2.83	-4.33	-6.15	0.20	1.10	1.69	2.40
Item52	2.24	0.57	-0.86	-2.59	-1.44	-0.37	0.55	1.67
Item53	-0.55	-2.26	-4.04	-4.97	0.22	0.91	1.63	2.01
Item54	-1.05	-3.5	-5.66	-6.61	0.40	1.33	2.15	2.52
Item55	0.16	-2.7	-4.81	-6.41	-0.06	0.97	1.72	2.29
Item56	1.36	-1.19	-3.72	-5.31	-0.63	0.55	1.72	2.45
Item57	2.02	-0.27	-2.54	-5.52	-0.92	0.12	1.16	2.53
Item58	2.22	0.26	-1.18	-3.14	-1.07	-0.13	0.57	1.51
Item59	2.5	-0.2	-2.49	-5.4	-0.94	0.07	0.93	2.02
Item60	0.06	-2.08	-3.55	-4.72	-0.02	0.84	1.44	1.91
Item61	-0.31	-2.33	-3.74	-6.04	0.13	1.00	1.60	2.59
Item62	-0.89	-3.24	-5.08	-7.36	0.29	1.07	1.67	2.43
Item63	0.95	-1.06	-2.39	-4.57	-0.44	0.49	1.10	2.10
Item64	-0.44	-2.85	-4.53	-6.48	0.20	1.27	2.02	2.88
Item65	0.58	0.04	-0.51	-0.53	-0.29	-0.02	0.26	0.27
Item66	2.83	0.66	-1.17	-3.16	-1.46	-0.34	0.61	1.64
Item67	0.9	-1.61	-3.62	-5.23	-0.44	0.79	1.78	2.57
Item68	-0.85	-3.09	-4.44	-6.29	0.33	1.20	1.73	2.45
Item69	0.61	-1.57	-2.72	-4.31	-0.27	0.71	1.22	1.94
Item70	1.53	-0.72	-3.05	-5.19	-0.70	0.33	1.40	2.38
Item71	1.32	-0.99	-2.77	-4.13	-0.63	0.47	1.31	1.96
Item72	2.23	-0.02	-2.13	-4.44	-0.91	0.01	0.87	1.82
Item73	1.34	-1.12	-3.14	-4.96	-0.63	0.53	1.48	2.33
Item74	2.57	0.52	-1.16	-3.44	-1.58	-0.32	0.71	2.11
Item75	1.4	-0.88	-2.57	-4.39	-0.82	0.51	1.50	2.56
Item76	-1.09	-2.62	-3.27	-4.49	0.56	1.35	1.69	2.32
Item77	-1.15	-3.33	-4.35	-5.33	0.52	1.50	1.96	2.40
Item78	-0.9	-2.59	-3.32	-5.6	0.44	1.27	1.62	2.74
Item79	0.24	-2.1	-3.7	-5.01	-0.12	1.03	1.82	2.46
Item80	0.42	-1.94	-3.82	-5.83	-0.18	0.84	1.65	2.52
Item81	0.39	-1.38	-3.04	-5	-0.23	0.80	1.77	2.91
Item82	0.21	-1.74	-3.21	-4.83	-0.11	0.92	1.71	2.57
Item83	0.43	-1.49	-3.8	-4.98	-0.21	0.72	1.83	2.40
Item84	-1.36	-2.85	-3.76	-5.53	0.57	1.20	1.58	2.32
Item85	2.22	-0.32	-2.67	-4.88	-1.06	0.15	1.28	2.34
Item86	-0.68	-2.69	-3.73	-5.24	0.29	1.16	1.61	2.27
Item87	-1.27	-2.98	-5.55	-7.98	0.48	1.12	2.08	2.99

Appendix Table 9 cont'd.

Item88	2.86	0.22	-2.31	-5.26	-1.09	-0.08	0.88	2.01
Item89	3.26	0.97	-0.79	-3.16	-1.78	-0.53	0.43	1.72
Item90	3.27	1.28	-0.31	-2.29	-1.79	-0.70	0.17	1.26
Item91	1.59	-1.07	-3.05	-6.45	-0.74	0.50	1.42	3.00
Item92	0.29	-1.75	-3.92	-5.72	-0.17	1.03	2.30	3.36
Item93	1.64	-0.18	-1.6	-2.98	-0.82	0.09	0.80	1.48
Item94	5.48	3.99	1.38	-1.34	-1.86	-1.36	-0.47	0.46
Item95	5.8	4.09	1.94	-0.59	-1.85	-1.31	-0.62	0.19
Item96	6.26	4.39	2.04	-0.61	-1.89	-1.32	-0.61	0.18
Item97	6.53	4.61	2.16	-0.2	-2.30	-1.62	-0.76	0.07
Item98	4.48	2.71	0.81	-1.65	-2.53	-1.53	-0.46	0.93
Item99	3.07	0.78	-1.43	-4.69	-1.48	-0.37	0.69	2.25
Item100	1.45	0.03	-1.61	-3.49	-0.78	-0.02	0.86	1.87

Appendix Table 10. Person scores for GRM model

Person	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	SE($\hat{\theta}_1$)	SE($\hat{\theta}_2$)	SE($\hat{\theta}_3$)	SE($\hat{\theta}_4$)
1	1.11	-0.29	-1.50	1.10	0.36	0.32	0.35	0.27
2	-1.37	-0.81	0.78	-1.76	0.49	0.35	0.44	0.62
3	-1.61	0.79	0.90	0.17	0.39	0.40	0.40	0.30
4	1.58	-0.36	0.63	-2.60	0.38	0.34	0.42	0.44
5	-0.70	-2.80	1.44	-1.04	0.46	0.45	0.44	0.49
6	-0.21	-0.35	-0.86	-0.06	0.36	0.33	0.36	0.32
7	-0.48	1.41	-0.60	-0.31	0.39	0.45	0.38	0.28
8	-2.26	1.91	-1.48	-0.67	0.40	0.47	0.40	0.45
9	0.94	-0.91	-0.19	-0.78	0.35	0.34	0.35	0.33
10	-0.76	-0.55	-0.49	0.73	0.35	0.35	0.36	0.31
11	1.66	-1.16	0.62	-1.48	0.37	0.33	0.39	0.34
12	-1.86	0.17	0.50	0.05	0.38	0.33	0.39	0.32
13	0.37	-0.58	-2.23	1.31	0.37	0.38	0.43	0.31
14	-0.93	0.12	-0.97	-0.15	0.38	0.32	0.38	0.37
15	-0.13	0.55	-1.02	0.07	0.32	0.34	0.34	0.29
16	-1.94	-0.82	-0.41	-0.79	0.46	0.38	0.39	0.53
17	0.75	-1.25	0.42	-0.98	0.36	0.33	0.37	0.36
18	-0.86	0.69	-0.87	-0.89	0.36	0.35	0.37	0.42
19	0.87	-1.45	1.14	-1.37	0.37	0.35	0.39	0.35
20	-1.18	-0.55	0.49	0.42	0.36	0.35	0.39	0.29
21	0.17	0.27	-1.55	-0.52	0.33	0.32	0.32	0.33
22	0.05	-1.30	-0.13	-0.07	0.35	0.30	0.35	0.30
23	-0.47	0.33	-0.67	-1.18	0.39	0.38	0.42	0.45
24	-0.06	-0.56	-0.72	0.78	0.33	0.33	0.32	0.25

Appendix Table 10 cont'd.

25	-0.38	0.27	-0.81	0.00	0.36	0.35	0.40	0.30
26	1.66	-0.51	-3.00	3.18	0.43	0.39	0.43	0.34
27	-0.25	-0.84	0.39	-0.21	0.32	0.31	0.35	0.29
28	-1.23	-0.78	0.00	-0.67	0.38	0.36	0.39	0.45
29	-2.38	-0.58	0.33	-0.81	0.48	0.39	0.42	0.52
30	-0.84	0.98	-1.14	0.70	0.36	0.34	0.36	0.28
31	-0.42	1.42	-1.57	0.28	0.37	0.41	0.37	0.26
32	-0.33	1.29	1.97	-0.96	0.41	0.50	0.53	0.31
33	-0.26	1.49	1.64	-2.89	0.45	0.49	0.55	0.51
34	0.56	-0.18	-0.59	0.11	0.33	0.35	0.35	0.28
35	0.59	-0.28	0.29	-0.80	0.38	0.35	0.39	0.33
36	-0.54	1.47	0.22	-2.68	0.40	0.42	0.37	0.51
37	1.45	1.10	1.14	-1.92	0.40	0.43	0.49	0.34
38	-0.92	-0.65	-0.96	-0.54	0.40	0.37	0.39	0.46
39	1.25	-0.51	-0.22	-0.69	0.46	0.38	0.50	0.36
40	-1.14	-1.71	1.49	-0.84	0.40	0.37	0.39	0.42
41	-0.33	-0.21	-0.42	-0.22	0.35	0.34	0.35	0.32
42	0.11	-1.87	-0.17	1.17	0.36	0.31	0.36	0.26
43	-0.55	-0.40	-0.40	-0.71	0.35	0.32	0.33	0.39
44	0.68	-0.94	1.91	-0.99	0.37	0.41	0.46	0.32
45	0.48	0.52	1.09	-0.60	0.39	0.37	0.48	0.29
46	-0.11	0.02	0.48	-0.89	0.37	0.34	0.43	0.34
47	-1.36	-0.40	-0.50	-1.10	0.46	0.37	0.39	0.52
48	-0.38	-2.14	0.56	-1.29	0.71	0.53	0.84	0.70
49	-2.42	-0.05	0.34	-0.62	0.47	0.36	0.41	0.47
50	-2.75	-0.74	0.92	0.08	0.52	0.33	0.42	0.47
51	-1.67	-1.66	-0.30	-1.24	0.58	0.46	0.44	0.67
52	0.34	1.00	-1.92	-0.41	0.37	0.34	0.37	0.30
53	-1.02	-0.56	1.44	-0.20	0.37	0.34	0.43	0.32
54	-0.79	0.18	-0.58	0.55	0.36	0.34	0.35	0.27
55	-1.22	0.66	0.03	-0.21	0.35	0.38	0.36	0.32
56	-0.31	-0.97	0.21	-1.55	0.36	0.39	0.38	0.48
57	0.40	0.45	-1.65	-1.09	0.36	0.37	0.35	0.42
58	-2.29	-0.86	-1.22	-0.30	0.59	0.46	0.51	0.62
59	0.43	-0.02	-1.34	-0.85	0.36	0.34	0.37	0.38
60	-0.82	0.07	2.54	-1.31	0.39	0.39	0.50	0.36
61	-1.65	-0.64	0.85	-0.55	0.39	0.36	0.38	0.40
62	0.30	-1.25	-0.09	0.62	0.34	0.39	0.34	0.26
63	-0.44	-0.90	-0.18	0.93	0.32	0.32	0.35	0.25
64	1.66	-0.94	0.74	-2.18	0.32	0.34	0.37	0.37

Appendix Table 10 cont'd.

65	-1.03	-0.75	-0.71	0.46	0.36	0.32	0.35	0.31
66	1.09	-1.36	0.15	-0.99	0.35	0.33	0.39	0.35
67	0.68	-1.10	1.26	-1.49	0.36	0.36	0.42	0.35
68	0.52	0.51	0.74	-0.31	0.34	0.38	0.45	0.28
69	-0.47	0.24	2.38	-1.29	0.36	0.36	0.53	0.33
70	-0.96	0.85	0.11	-0.52	0.37	0.38	0.37	0.32
71	0.29	-0.41	-0.85	-0.52	0.32	0.31	0.32	0.34
72	-0.93	-0.36	0.72	-0.68	0.34	0.33	0.38	0.35
73	-1.17	-2.36	-0.41	-0.31	0.49	0.50	0.41	0.51
74	0.55	0.18	0.61	0.41	0.37	0.37	0.41	0.25
75	-1.26	-0.84	-0.66	0.74	0.38	0.33	0.39	0.31
76	0.42	0.01	-1.04	-0.14	0.31	0.34	0.36	0.29
77	-0.29	-0.37	-0.92	1.30	0.36	0.33	0.36	0.25
78	0.14	-1.40	1.78	-0.66	0.35	0.38	0.43	0.30
79	-1.03	-1.56	-0.15	-0.30	0.39	0.30	0.36	0.44
80	-0.37	-0.55	1.53	-1.20	0.40	0.37	0.43	0.38
81	-1.60	-0.66	-0.97	-1.20	0.45	0.39	0.40	0.62
82	-0.09	-1.00	-0.16	-1.16	0.35	0.31	0.35	0.41
83	0.14	0.65	0.62	-1.37	0.36	0.37	0.40	0.31
84	0.05	-0.76	-0.81	-1.13	0.38	0.38	0.38	0.43
85	-0.08	0.76	0.15	0.17	0.36	0.39	0.41	0.27
86	0.26	2.06	-1.05	-1.06	0.39	0.48	0.39	0.32
87	0.55	-1.20	-0.29	-0.31	0.32	0.33	0.33	0.29
88	-0.53	0.33	-0.82	0.63	0.36	0.39	0.35	0.28
89	-2.50	-1.74	0.73	0.15	0.50	0.34	0.40	0.44
90	0.22	-1.35	0.92	-0.23	0.34	0.29	0.35	0.28
91	0.11	-1.44	-0.41	-0.80	0.37	0.36	0.40	0.41
92	-0.46	-0.05	-1.23	-0.30	0.38	0.34	0.36	0.35
93	-0.20	-0.37	-0.87	-1.49	0.36	0.33	0.38	0.49
94	1.29	-1.05	-0.71	-1.11	0.38	0.38	0.41	0.40
95	-0.22	-0.61	-1.21	0.38	0.37	0.35	0.36	0.31
96	0.53	-1.19	-0.87	0.77	0.34	0.33	0.38	0.27
97	1.32	-0.61	-0.71	-1.74	0.39	0.36	0.38	0.42
98	-2.09	-0.11	0.06	0.36	0.45	0.41	0.48	0.39
99	1.68	-1.55	0.32	-0.88	0.35	0.33	0.35	0.31
100	0.27	0.13	0.11	-1.15	0.36	0.33	0.40	0.36
101	0.14	1.00	-0.03	-0.19	0.40	0.41	0.36	0.29
102	-0.20	-0.64	0.41	-0.07	0.35	0.34	0.36	0.28
103	1.10	-0.30	-1.15	-0.70	0.35	0.35	0.34	0.33
104	0.62	1.55	-0.63	-1.82	0.41	0.41	0.38	0.38

Appendix Table 10 cont'd.

105	-2.03	-1.57	0.12	-0.06	0.48	0.35	0.41	0.48
106	-1.07	-0.98	-0.67	0.55	0.37	0.35	0.36	0.33
107	-0.98	0.63	0.00	0.61	0.33	0.36	0.40	0.25
108	-0.06	-0.59	0.15	0.65	0.33	0.31	0.36	0.25
109	-0.33	-0.91	0.36	0.95	0.36	0.38	0.38	0.27
110	-0.26	0.29	-0.82	0.94	0.36	0.34	0.34	0.25
111	-1.09	1.15	0.71	-1.10	0.39	0.43	0.41	0.35
112	-0.74	1.08	0.83	-0.39	0.38	0.44	0.43	0.33
113	1.19	-1.35	0.25	0.45	0.32	0.30	0.36	0.26
114	0.95	-0.56	0.79	-2.55	0.39	0.34	0.40	0.45
115	-1.36	-0.69	-1.63	0.15	0.43	0.35	0.40	0.41
116	-0.53	-0.16	-0.94	-1.16	0.74	0.93	0.89	0.81
117	0.24	-0.16	-0.02	0.70	0.33	0.34	0.40	0.25
118	0.24	-0.62	0.14	-1.18	0.35	0.33	0.41	0.39
119	-0.17	0.20	0.18	-0.86	0.40	0.36	0.40	0.36
120	0.36	-0.18	-0.38	-0.64	0.36	0.33	0.36	0.33
121	0.25	-1.19	-0.30	1.06	0.37	0.40	0.35	0.26
122	0.24	-0.47	-0.79	1.13	0.35	0.33	0.32	0.25
123	-0.56	0.94	-0.45	-1.43	0.40	0.36	0.36	0.40
124	-0.63	-0.53	1.75	0.10	0.39	0.34	0.45	0.30
125	-2.47	1.67	-1.36	0.49	0.46	0.46	0.51	0.38
126	1.00	0.60	-0.20	-0.41	0.37	0.39	0.39	0.26
127	0.26	-0.28	-0.64	0.71	0.32	0.31	0.35	0.25
128	0.47	-0.77	-0.14	-0.20	0.37	0.31	0.36	0.31
129	0.09	0.02	-1.22	1.63	0.33	0.35	0.34	0.25
130	-1.36	1.45	-0.35	-1.07	0.42	0.44	0.37	0.41
131	-2.62	0.18	-1.03	-1.08	0.56	0.52	0.49	0.67
132	0.65	-0.01	-2.02	1.16	0.34	0.29	0.34	0.26
133	0.10	0.36	1.19	-1.11	0.41	0.36	0.46	0.34
134	-0.02	-0.01	0.02	0.09	0.97	1.00	1.01	0.98
135	0.42	0.78	-1.03	-1.05	0.39	0.37	0.35	0.35
136	-1.55	1.40	-0.76	0.10	0.40	0.39	0.35	0.30
137	0.01	0.80	-0.55	0.84	0.36	0.39	0.39	0.28
138	0.36	-0.37	-0.98	-0.86	0.36	0.36	0.36	0.39
139	-0.84	0.71	-0.88	-0.02	0.40	0.35	0.49	0.34
140	1.30	-0.41	0.66	0.79	0.38	0.38	0.41	0.27
141	-0.57	-0.48	-1.22	-0.90	0.38	0.33	0.34	0.46
142	-2.33	0.54	0.32	-0.63	0.42	0.34	0.38	0.44
143	-0.54	0.96	-0.50	-0.87	0.36	0.36	0.38	0.35
144	-1.83	-0.28	-1.03	0.08	0.40	0.31	0.38	0.39

Appendix Table 10 cont'd.

145	-2.02	0.68	-0.63	-0.77	0.40	0.32	0.37	0.46
146	2.27	-1.32	0.68	2.26	0.39	0.46	0.48	0.28
147	-1.81	1.03	-2.80	0.84	0.38	0.35	0.38	0.35
148	0.73	-0.43	1.29	-0.76	0.37	0.36	0.42	0.30
149	0.68	-0.82	0.01	-0.66	0.34	0.33	0.36	0.32
150	-0.65	-1.55	-0.07	0.02	0.34	0.33	0.36	0.35
151	-1.26	-1.73	-1.08	-1.88	0.52	0.50	0.42	0.70
152	-1.50	0.20	-0.08	-0.08	0.38	0.34	0.37	0.36
153	-0.32	-0.38	-1.08	-0.14	0.36	0.35	0.35	0.35
154	0.03	0.39	0.15	-0.12	0.37	0.38	0.37	0.27
155	0.37	-0.89	-0.41	-1.00	0.38	0.36	0.39	0.41
156	-0.83	0.36	-1.58	1.13	0.37	0.31	0.33	0.26
157	1.43	0.30	-0.30	0.52	0.36	0.39	0.41	0.29
158	0.06	0.48	-0.69	-0.46	0.40	0.40	0.41	0.31
159	1.28	-0.16	-1.03	-1.44	0.38	0.34	0.36	0.39
160	-1.00	1.64	0.00	-0.24	0.41	0.49	0.40	0.30
161	-0.57	0.84	-0.14	0.04	0.38	0.36	0.40	0.31
162	-0.86	-0.73	1.15	-2.10	0.42	0.37	0.39	0.53
163	-0.67	-0.86	1.39	-1.77	0.40	0.34	0.42	0.43
164	-0.10	0.50	-0.24	-0.37	0.36	0.37	0.38	0.30
165	-1.72	-0.71	-0.51	-0.14	0.44	0.39	0.43	0.44
166	2.32	-0.88	-1.54	1.96	0.41	0.42	0.38	0.29
167	0.05	1.97	-0.09	-0.95	0.42	0.49	0.45	0.29
168	-1.90	-0.89	1.63	-1.11	0.47	0.36	0.44	0.44
169	-0.85	-1.62	-1.17	-0.37	0.43	0.42	0.37	0.52
170	0.97	-0.20	-3.35	3.12	0.41	0.44	0.43	0.30
171	0.06	1.97	0.57	-0.75	0.43	0.54	0.47	0.30
172	-1.51	-0.82	0.98	-0.25	0.37	0.37	0.40	0.35
173	-0.14	0.57	-1.86	0.09	0.34	0.34	0.34	0.30
174	0.11	0.59	0.03	-0.48	0.36	0.35	0.37	0.30
175	-2.24	0.30	-0.80	0.52	0.42	0.32	0.37	0.36
176	-0.33	0.86	-0.50	-1.39	0.38	0.40	0.38	0.40
177	0.80	-1.24	-0.95	0.75	0.37	0.35	0.37	0.26
178	-0.26	-0.75	1.40	-0.16	0.42	0.41	0.44	0.32
179	0.06	-1.32	-0.13	1.39	0.34	0.32	0.36	0.24
180	-0.50	-0.32	-0.08	-0.57	0.38	0.36	0.33	0.33
181	0.40	-0.62	-1.62	-0.07	0.37	0.33	0.37	0.32
182	1.65	-0.43	-0.46	1.93	0.44	0.49	0.53	0.34
183	0.83	1.32	-1.42	1.15	0.40	0.49	0.40	0.26
184	0.14	0.46	-0.83	-1.12	0.36	0.35	0.35	0.38

Appendix Table 10 cont'd.

185	-0.45	-0.24	0.15	-1.22	0.37	0.35	0.39	0.42
186	-0.21	0.34	-0.15	0.67	0.37	0.36	0.36	0.26
187	-0.01	0.15	-0.13	-1.64	0.35	0.36	0.37	0.41
188	1.96	-0.80	-1.33	1.17	0.36	0.41	0.38	0.27
189	-0.37	0.10	-1.64	0.88	0.34	0.30	0.40	0.25
190	-2.02	-0.87	-1.25	0.61	0.46	0.37	0.37	0.38
191	-1.38	0.31	-1.40	-0.51	0.41	0.37	0.40	0.45
192	0.20	0.80	-1.17	-0.73	0.36	0.35	0.37	0.32
193	0.37	0.34	-1.68	-1.26	0.36	0.35	0.33	0.42
194	0.41	-0.75	-0.23	-0.79	0.36	0.33	0.38	0.35
195	-0.52	1.62	0.24	-2.79	0.44	0.46	0.44	0.54
196	-0.13	0.95	0.75	-0.87	0.40	0.44	0.44	0.32
197	-1.00	2.34	-1.38	-0.25	0.45	0.53	0.50	0.36
198	1.08	-0.46	-1.60	2.20	0.39	0.34	0.40	0.25
199	-1.81	-0.02	-1.10	-1.29	0.46	0.36	0.39	0.56
200	0.46	0.94	-0.70	-0.36	0.35	0.38	0.37	0.26
201	-0.59	0.01	-0.80	0.41	0.36	0.31	0.38	0.30
202	1.62	-2.25	-0.86	1.82	0.34	0.30	0.35	0.26
203	0.20	-0.64	0.54	-1.47	0.36	0.35	0.38	0.40
204	-1.90	-0.06	-0.35	0.39	0.40	0.34	0.39	0.34
205	-1.32	0.95	-0.82	-0.04	0.38	0.42	0.37	0.32
206	0.53	-0.77	-0.26	0.67	0.32	0.42	0.34	0.26
207	1.16	-0.96	0.47	-1.84	0.38	0.36	0.36	0.38
208	-1.58	-0.86	-0.35	-0.62	0.45	0.35	0.37	0.47
209	2.56	-3.70	-2.41	4.60	0.41	0.44	0.36	0.34
210	1.21	0.14	-0.35	0.04	0.34	0.32	0.36	0.26
211	-0.51	-1.67	0.59	-0.92	0.44	0.37	0.42	0.47
212	-2.65	-0.79	-0.15	-0.66	0.53	0.39	0.42	0.57
213	2.05	-1.19	-1.06	2.40	0.34	0.33	0.36	0.23
214	-1.41	0.41	-0.83	0.28	0.40	0.40	0.42	0.37
215	0.78	-0.33	-0.35	1.07	0.69	0.88	0.90	0.53
216	-0.70	-0.31	-0.09	0.96	0.35	0.33	0.35	0.24
217	-0.17	0.04	-0.59	-1.34	0.39	0.34	0.39	0.47
218	-0.46	1.04	0.14	0.96	0.43	0.45	0.42	0.29
219	0.81	0.24	0.41	-2.08	0.40	0.34	0.38	0.42
220	0.00	0.05	-1.27	-1.06	0.36	0.35	0.39	0.41
221	0.07	0.85	-0.87	-1.26	0.36	0.35	0.35	0.38
222	-1.72	0.65	-0.68	-1.10	0.41	0.37	0.39	0.50
223	-0.27	-0.07	-1.30	0.09	0.35	0.33	0.39	0.30
224	0.30	-0.80	0.20	0.70	0.36	0.33	0.37	0.27

Appendix Table 10 cont'd.

225	-0.88	-2.60	-1.23	-0.02	0.49	0.52	0.43	0.50
226	0.31	-0.24	-0.09	0.11	0.36	0.33	0.40	0.28
227	-0.33	-1.36	0.46	0.41	0.33	0.32	0.35	0.27
228	0.34	0.02	1.15	-1.32	0.37	0.36	0.42	0.35
229	-0.20	-0.24	-1.38	-0.25	0.35	0.32	0.36	0.34
230	-0.40	-0.24	-0.35	-0.05	0.85	0.93	0.87	1.00
231	-0.53	-0.88	1.18	0.02	0.36	0.33	0.40	0.29
232	-0.91	-0.48	-0.16	0.88	0.33	0.34	0.38	0.27
233	-0.29	0.54	0.76	-0.90	0.37	0.36	0.41	0.30
234	-0.17	-0.99	-1.35	1.87	0.32	0.30	0.34	0.23
235	0.18	0.94	0.15	-1.80	0.36	0.37	0.36	0.40
236	-1.35	-1.18	1.04	0.59	0.36	0.35	0.35	0.27
237	-0.33	-0.28	0.33	-0.10	0.35	0.35	0.37	0.29
238	1.09	-0.81	0.09	-0.99	0.36	0.36	0.33	0.31
239	-2.15	0.27	0.84	-0.48	0.44	0.42	0.49	0.42
240	-0.51	0.49	-1.14	-0.15	0.39	0.35	0.36	0.33
241	-1.26	-0.39	-0.03	-1.58	0.42	0.38	0.37	0.54
242	-0.80	-0.68	-0.45	-0.72	0.42	0.32	0.37	0.43
243	-0.49	-0.34	1.44	-0.89	0.37	0.33	0.43	0.36
244	-0.19	-1.14	1.17	-0.89	0.37	0.36	0.42	0.37
245	-1.17	-0.21	-0.17	0.36	0.38	0.33	0.36	0.30
246	-0.29	0.46	-0.39	-0.35	0.36	0.36	0.40	0.30
247	-1.37	0.18	-0.59	0.36	0.36	0.34	0.34	0.30
248	-2.15	-0.89	0.02	0.43	0.45	0.40	0.40	0.37
249	0.50	0.55	0.66	-1.94	0.38	0.38	0.43	0.37
250	0.22	0.73	-1.58	0.32	0.35	0.34	0.33	0.27
251	0.31	0.60	0.41	-1.59	0.37	0.35	0.38	0.36
252	-0.83	-1.24	0.07	0.81	0.34	0.31	0.35	0.27
253	0.37	0.60	0.14	0.37	0.38	0.41	0.42	0.28
254	0.21	-0.55	-0.44	-0.06	0.33	0.32	0.35	0.27
255	-0.86	-1.09	-0.25	0.03	0.39	0.34	0.37	0.35
256	-0.60	-0.04	-1.12	-0.31	0.40	0.39	0.38	0.38
257	-0.90	-0.17	-0.45	0.59	0.35	0.34	0.37	0.28
258	-2.37	0.48	-1.50	0.60	0.43	0.34	0.37	0.37
259	-0.43	0.25	-0.99	0.69	0.35	0.33	0.34	0.28
260	0.93	-0.03	0.54	-2.05	0.39	0.35	0.41	0.39
261	0.15	-0.98	1.46	-0.38	0.35	0.34	0.43	0.30
262	-0.34	-0.81	-0.76	0.04	0.34	0.32	0.35	0.32
263	-0.05	-0.05	-1.35	0.04	0.34	0.35	0.34	0.30
264	0.23	-0.72	-0.61	-1.28	0.36	0.37	0.37	0.43

Appendix Table 10 cont'd.

265	0.31	-1.80	0.04	-0.71	0.38	0.35	0.38	0.37
266	-0.38	0.55	0.03	-1.89	0.38	0.35	0.38	0.47
267	-1.02	-0.90	-0.94	-1.14	0.42	0.32	0.37	0.50
268	-0.73	-1.48	0.95	0.19	0.36	0.35	0.37	0.30
269	-0.49	0.06	0.81	-0.80	0.39	0.35	0.39	0.34
270	1.56	-0.09	-0.26	-1.66	0.36	0.33	0.37	0.36
271	-0.47	-0.97	-0.56	1.48	0.36	0.30	0.37	0.26
272	0.38	-0.05	0.19	-1.00	0.36	0.34	0.39	0.34
273	-0.69	-0.62	0.98	-1.26	0.38	0.39	0.39	0.41
274	-0.36	-2.53	1.42	-0.10	0.35	0.34	0.39	0.34
275	-1.66	-1.43	-0.52	-0.48	0.47	0.32	0.38	0.49
276	1.56	-1.51	-0.58	0.45	0.32	0.32	0.33	0.24
277	0.85	0.48	-0.86	-2.21	0.36	0.32	0.34	0.45
278	-0.08	-0.51	-0.67	0.42	0.35	0.33	0.35	0.27
279	1.48	-1.60	-0.37	-1.94	0.41	0.36	0.36	0.48
280	0.17	0.16	-0.99	-0.65	0.36	0.32	0.34	0.32
281	-0.88	0.11	-0.89	-0.95	0.38	0.36	0.39	0.42
282	0.15	0.35	-0.94	-2.15	0.37	0.36	0.37	0.52
283	0.03	-1.45	-0.58	-0.24	0.37	0.34	0.36	0.39
284	-0.05	-0.27	-0.73	-0.90	0.38	0.32	0.37	0.38
285	-0.93	-0.16	-0.70	-0.05	0.35	0.37	0.38	0.32
286	-1.68	1.12	-0.06	-0.78	0.38	0.38	0.38	0.40
287	-2.34	0.60	-0.98	-1.07	0.47	0.35	0.36	0.54
288	1.62	-0.10	-0.37	-1.14	0.36	0.33	0.39	0.33

REFERENCES

Abdel-Aty, M., Keller, J., & Brady, P. (2005). Analysis of Types of Crashes at Signalized Intersections by Using Complete Crash Data and Tree-Based Regression. *Transportation Research Record: Journal of the Transportation Research Board, 1908*, 37-45. doi:10.3141/1908-05

Abdel-Aty, M. A., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention, 32*(5), 633-642. doi:http://dx.doi.org/10.1016/S0001-4575(99)00094-9

American Association of State Highway and Transportation Officials. (2010). *Highway Safety Manual* (A. A. o. S. H. a. T. Officials Ed. Vol. 2). Washington, D.C. : American Association of State Highway and Transportation Officials.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika, 37*(1), 29-51. doi:10.1007/BF02291411

Bock, R. D., & Jones, L. V. (1968). *The measurement and prediction of judgment and choice*. San Francisco: San Francisco, Holden-Day.

Brookhuis, K. A., de Vries, G., & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention, 23*(4), 309-316. doi:http://dx.doi.org/10.1016/0001-4575(91)90008-S

Bryden, J., Andrew, L., & Fortuniewicz, J. (2000). Intrusion Accidents on Highway Construction Projects. *Transportation Research Record, 1715*(1), 30. doi:10.3141/1715-05

Buja, A., & Eyuboglu, N. (1992). Remarks on Parallel Analysis. *Multivariate Behavioral Research, 27*(4), 509-540. doi:10.1207/s15327906mbr2704_2

Busemeyer, J. R., & Diederich, A. (2014). Chapter 4 - Estimation and Testing of Computational Psychological Models A2 - Glimcher, Paul W. In E. Fehr (Ed.), *Neuroeconomics (Second Edition)* (pp. 49-61). San Diego: Academic Press.

Cai, L. (2010a). High-dimensional Exploratory Item Factor Analysis by A Metropolis–Hastings Robbins–Monro Algorithm. *Psychometrika, 75*(1), 33-57. doi:10.1007/s11336-009-9136-x

Cai, L. (2010b). Metropolis-Hastings Robbins-Monro Algorithm for Confirmatory Item Factor Analysis. *Journal of Educational and Behavioral Statistics, 35*(3), 307-335. doi:10.3102/1076998609353115

Cai, L. (2017). flexMIRT (R) version 3.51: Flexible multilevel multidimensional item analysis and test scoring. Chapel Hill, NC: Vector Psychometric Group.

Cantin, V., Lavallière, M., Simoneau, M., & Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis & Prevention, 41*(4), 763-771. doi:http://dx.doi.org/10.1016/j.aap.2009.03.019

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, *48*(6). doi:10.18637/jss.v048.i06

Chon, K. H., Lee, W.-C., & Dunbar, S. B. (2010). A Comparison of Item Fit Statistics for Mixed IRT Models. *Journal of Educational Measurement, 47*(3), 318-338. doi:10.1111/j.1745-3984.2010.00116.x

Cohen, J. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.. ed.). Hillsdale, N.J.: Hillsdale, N.J. : L. Erlbaum Associates.

Dewar, R. E., & Olson, P. L. (2002). *Human Factors in Traffic Safety*. Tucson, AZ: Lawyers & Judges Publishing Company, Inc. .

Edquist, J., Horberry, T., Hosking, S., & Johnston, I. (2011). Effects of advertising billboards during simulated driving. *Applied Ergonomics*, *42*(4), 619-626. doi:http://dx.doi.org/10.1016/j.apergo.2010.08.013

Edquist, J., Rudin-Brown, C. M., & Lenné, M. G. (2012). The effects of on-street parking and road environment visual complexity on travel speed and reaction time. *Accident Analysis and Prevention, 45*, 759-765. doi:10.1016/j.aap.2011.10.001

Elvik, R. (2006). Laws of accident causation. *Accident Analysis & Prevention*, *38*(4), 742-747. doi:http://dx.doi.org/10.1016/j.aap.2006.01.005

Federal Motor Carrier Safety Administration. (2017). *Improving Motor Carrier Safety Measurement* (978-0-309-46201-3). Retrieved from Washington, DC: https://www.nap.edu/catalog/24818/improving-motor-carrier-safety-measurement

Finley, M. D., Theiss, L., Trout, N. D., Miles, J. D., & Nelson, A. A. (2011). *Studies to Determine the Effectiveness of Longitudinal Channelizing Devices in Work Zones* (FHWA/TX-11/0-6103-1). College Station, Texas.

Greenwood, A. T. (2015). *Evaluating comprehension of temporary traffic control.* (Doctor of Philosophy in Civil Engineering Dissertation), Georgia Institute of Technology, Atlanta, GA.

Greenwood, A. T., Xu, Y., Corso, G. M., Hunter, M. P., & Rodgers, M. O. (2016). *Identification of Diverges in Freeway Work Zones: The Effects of Channelizing Devices.* Paper presented at the 95th Annual Meeting of the Transportation Research Board, Washington D.C.

Hadi, M. A., Aruldhas, J., Chow, L.-F., & Wattleworth, J. A. (1995). Estimating safety effects of cross-section design for various highway types using negative binomial regression. *Transportation Research Record*(1500), 169-177.

Ho, G., Scialfa, C. T., Caird, J. K., & Graw, T. (2001). Visual search for traffic signs: The effects of clutter, luminance, and aging. *Human Factors, 43*(2), 194-207. doi:10.1518/001872001775900922

Holgado–Tello, F., Chacón–Moscoso, S., Barbero–García, I., & Vila–Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity, 44*(1), 153-166. doi:10.1007/s11135-008-9190-y

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention, 38*(1), 185-191. doi:http://dx.doi.org/10.1016/j.aap.2005.09.007

Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185. doi:10.1007/BF02289447

Houts, C. R., & Cai, L. (2015). *Flexible Mutilevel Multidimensional Item Analysis and Test Scoring*: Vector Psychometric Group, LLC.

Hunter, M. P., Rodgers, M. O., Corso, G. M., Shaw, F. A., Bae, J., & Greenwood, A. T. (2016). *Factors Influencing Visual Search in Complex Driving Environments*. Retrieved from Atlanta, GA: https://trid.trb.org/view/1474329

Hunter, M. P., Rodgers, M. O., Corso, G. M., Xu, Y., & Greenwood, A. T. (2014). *Improved Methods for Delineating Diverges in Work Zones*. (RP 10-07). Atlanta, GA.

Kaber, D., Zhang, Y., Jin, S., Mosaly, P., & Garner, M. (2012). Effects of hazard exposure and roadway complexity on young and older driver situation awareness and performance. *Transportation Research Part F: Traffic Psychology and Behaviour, 15*(5), 600-611. doi:http://dx.doi.org/10.1016/j.trf.2012.06.002

Karlaftis, M. G., & Golias, I. (2002). Effects of road geometry and traffic volumes on rural roadway accident rates. *Accident Analysis & Prevention, 34*(3), 357-365. doi:http://dx.doi.org/10.1016/S0001-4575(01)00033-1

Little, R. J. A. (2002). *Statistical analysis with missing data* (2nd ed.. ed.). Hoboken, N.J.: Hoboken, N.J. : Wiley.

Meyer, M. D., & Miller, E. J. (2017). *Urban transportation planning : a decision-oriented approach*. Not yet in print; available upon request from authors.

Milton, J., & Mannering, F. (1998). The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. *Transportation, 25*(4), 395-413. doi:10.1023/A:1005095725001

Mohamedshah, Y. M., Paniati, J. F., & Hobeika, A. G. (1993). Truck Accident Models for Interstates and Two-Lane Rural Roads. *Transportation Research Record*(1407), 35.

Muraki, E. (1992). A Generalized Partial Credit Model: Application of an EM Algorithm. *ETS Research Report Series, 1992*(1), i-30. doi:10.1002/j.2333-8504.1992.tb01436.x

Muraki, E. (1993). Information Functions of the Generalized Partial Credit Model *ETS Research Report Series, 1993*(1), i-12. doi:10.1002/j.2333-8504.1993.tb01538.x

Olson, P. L., & Farber, E. (2003). *Forensic Aspects of Driver Perception and Response* (2 ed.). Tucson, AZ: Lawyers & Judges Publishing Company, Inc. .

Reckase, M. (2009). *Multidimensional item response theory*. New York: Springer.

Roess, R. P. (2011). *Traffic engineering* (4th ed.. ed.). Upper Saddle River, NJ: Upper Saddle River, NJ : Pearson.

Rosenblad, A. (2009). Applied Multivariate Statistics for the Social Sciences, Fifth Edition by James P. Stevens. *International Statistical Review, 77*(3), 476.

Rowell, M., Gagliano, A., & Goodchild, A. (2014). Identifying truck route choice priorities: the implications for travel models. *Transportation Letters, 6*(2), 98-106. doi:10.1179/1942787514Y.0000000015

Saffir, M. (1937). A comparative study of scales constructed by three psychophysical methods. *Psychometrika, 2*(3), 179-198. doi:10.1007/BF02288395

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*. Fredericton: Fredericton : University of New Brunswick.

Schiessl, C. (2008). Subjective strain estimation depending on driving manoeuvres and traffic situation. *Intelligent Transport Systems, IET, 2*(4), 258-265. doi:10.1049/iet-its:20080024

Shaw, F. A., Bae, J., Corso, G. M., Rodgers, M. O., & Hunter, M. P. (2017). *Assessment Time for Rating Perceived Roadway Complexity.* Paper presented at the 96th Annual Meeting of the Transportation Research Board, Washington D.C.

Shaw, F. A., Greenwood, A. T., Bae, J., Corso, G. M., Rodgers, M. O., & Hunter, M. P. (2018). *Effects of Roadway Factors and Demographic Characteristics on Drivers' Perceived Complexity of Simulated Roadway Videos.* Transportation Letters: The International Journal of Transportation Research: doi:https://doi.org/10.1080/19427867.2018.1492220

Shaw, F. A., Greenwood, A. T., Bae, J., Woolery, W., Xu, Y. A., Guin, A., . . . Hunter, M. P. (2016). *Drivers' Perceived Complexity of Simulated and On-road*

Environments. Paper presented at the 95th Annual Meeting of the Transportation Research Board, Washington D.C.

Shaw, F. A., Park, S. J., Bae, J., Becerra, Z., Corso, G. M., Rodgers, M. O., & Hunter, M. P. (2018). Effects of roadside distractors on performance of drivers with and without attention deficit tendencies. *Transportation Research Part F: Traffic Psychology and Behaviour*. doi:https://doi.org/10.1016/j.trf.2018.02.013

Stinchcombe, A., & Gagnon, S. (2010). Driving in dangerous territory: Complexity and road-characteristics influence attentional demand. *Transportation Research Part F: Traffic Psychology and Behaviour, 13*(6), 388-396. doi:http://dx.doi.org/10.1016/j.trf.2010.06.008

Teh, E., Jamson, S., Carsten, O., & Jamson, H. (2014). Temporal fluctuations in driving demand: The effect of traffic complexity on subjective measures of workload and driving performance. *Transportation Research Part F: Traffic Psychology and Behaviour, 22*, 207-217. doi:http://dx.doi.org/10.1016/j.trf.2013.12.005

Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: New York, Wiley.

Xu, Y., Greenwood, A. T., Corso, G. M., Rodgers, M. O., & Hunter, M. P. (2015). *Response Time as a Surrogate Safety Measure to Evaluate Work Zone Delineation Methods.* Paper presented at the Proceedings of the 5th Road Safety and Simulation International Conference, Orlando, FL.

Young, M. S., Mahfoud, J. M., Stanton, N. A., Salmon, P. M., Jenkins, D. P., & Walker, G. H. (2009). Conflicts of interest: The implications of roadside advertising for driver attention. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*(5), 381-388. doi:http://dx.doi.org/10.1016/j.trf.2009.05.004

Zeitlin, L. R. (1995). Estimates of driver mental workload: A long-term field trial of two subsidiary tasks. *Human Factors, 37*(3), 611-621. doi:10.1518/001872095779049327

Zhang, J. (2007). Conditional Covariance Theory and Detect for Polytomous Items. *Psychometrika, 72*(1), 69-91. doi:10.1007/s11336-004-1257-7

Zhang, L., & Lin, W. (2013). *Selective Visual Attention: Computational Models and Applications*. Singapore: John Wiley & Sons Singapore Pte. Ltd. .