8-12-2014

Examination of Parameter Estimation Using Recursive Bayesian Analysis in Simulated Item Response Theory Applications

Robert Hendrick

Follow this and additional works at: http://scholarworks.gsu.edu/eps_diss

ACCEPTANCE

This dissertation, EXAMINATION OF PARAMETER ESTIMATION USING RECURSIVE BAYESIAN ANALYSIS IN SIMULATED ITEM RESPONSE THEORY APPLICATIONS, by ROBERT CLAYTON HENDRICK, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty. The Dean of the College of Education concurs.

_____   _____

William L. Curlette, Ph.D.   T. Chris Oshima, Ph.D.
Committee Chair   Committee Member

_____   _____

 Hongli Li, Ph.D.   Gwendolyn Benson, Ph.D.
 Committee Member   Committee Member

_____

Date

_____

William L. Curlette, Ph.D.
Chairperson, Department of Educational Policy Studies

_____

Paul A. Alberto, Ph.D.
Dean, College of Education

VITA

Robert Clayton Hendrick

ADDRESS: 6059 Driftwood Trail
                            Mableton, Georgia 30126

EDUCATION:
          Ph.D.   2014   Georgia State University
                          Educational Policy Studies
                          Research, Measurement, and Statistics
          M.S.   1991   University of West Georgia
                          Educational Leadership
          M.S.   1986   University of West Georgia
                          Middle Grades Mathematics
          B.S.   1984   University of West Georgia
                          Middle Grades Education

PROFESSIONAL EXPERIENCE:
          2010-Present Research Associate II
                        Georgia State University, Atlanta, GA
          2008-2010   Mathematics Teacher
                        Cobb County School System, Marietta, GA
          2002-2008   Director of Correlation Services
                        PLATO Learning, Inc., Atlanta, GA
          1998-2002   Executive Director of Implementation Services
                        NETSchools Corporation, Atlanta, GA
          1992-1998   Assistant Superintendent and Director of Technology
                        Carrollton City Schools, Carrollton, GA
          1989-1992   Assistant Principal Carrollton Junior High School
                        Carrollton City Schools, Carrollton, GA
          1984-1992   Mathematics Teacher
                        Carrollton City Schools, Carrollton, GA

PRESENTATION AND PUBLICATIONS:

Hendrick, R. C., Curlette, W. L., Ogletree, S. L., Benson, G. T. (2013, April). *Meta-analysis: New findings connect clinical teaching with evaluating student achievement in high-needs, urban, partnership schools*, paper presented at the 2013 AERA Annual Meeting, San Francisco, CA.

Curlette, W. L., Hendrick, R. C., & Granville, H. G. (2013, April). *Anchor action research as an aspect of clinical teaching*, poster session presented at the 2013 AERA Annual Meeting, San Francisco, CA.

Hendrick, R. C., Curlette, W. L., Ogletree, S. L., Benson, G. T., Martin, C., Grier, C., & Veazie, W. (2013, February). *Clinical teaching and evaluating student achievement in high-needs, professional development classrooms*, Paper

presented at the 2013 Professional Development Schools National Conference, New Orleans, LA.

Curlette, W. L., Hendrick, R. C., Ogletree, S. L., Benson, G. T., & Granville, H. G. (2012, April). *Empirical support for fidelity of implementation based on NCATE PDS standards: A new online instrument*, Roundtable session presented at the 2012 AERA Annual Meeting, Vancouver, B.C.

Granville, H. G., Hendrick, R. C., Curlette, W. L. (2012, April). *CCLCs: A case study of a CFG model with promising school-wide NCATE PDS standards outcomes.* Paper presented at the 2012 AERA Annual Meeting, Vancouver, B.C.

Curlette, W. L., Hendrick, R. C., Ogletree, S. L., Benson, G. T., Martin, C., & Granville, H. G. (2012, March). *Using the NAPDS nine essentials and the NCATE PDS standards to support partnerships: A new online instrument*, Paper presented at the 2012 Professional Development Schools National Conference, Las Vegas, NV.

Curlette, W. L., Granville, H. G., Feinberg, J., Clarke, P. J., Benson, G. T., Ogletree, S. L., Hendrick, R. C., & Veazie, W. (2012, March). *Teacher candidates in a PDS year long residency program: Implementation and evaluation,* Paper presented at the 2012 Professional Development Schools National Conference, Las Vegas, NV.

Curlette, W. L., Ogletree, S. L., Benson, G. T., Granville, H. G., & Hendrick, R. C., (2011, April). *Fidelity of implementation instrument based on NCATE PDS standards: A 4-year study and future directions,* Roundtable session presented at the 2011 AERA Annual Meeting, New Orleans, LA.

Curlette, W. L., Benson, G. T., Ogletree, S. L., & Hendrick, R. C. (2011, March). *Benefits of engaging classrooms in anchor action research and relationships to a fidelity survey*, Paper presented at the 2011 Professional Development Schools National Conference, New Orleans, LA.

PROFESSIONAL SOCIETIES AND ORGANIZATIONS:

       2010-Present   Nation Association of Professional Development Schools

       2010-Present   American Educational Research Association

ABSTRACT

EXAMINATION OF PARAMETER ESTIMATION USING RECURSIVE BAYESIAN
ANALYSIS IN SIMULATED ITEM RESPONSE THEORY APPLICATIONS

by

Robert Hendrick

For the past several years, high-stakes testing has been the predominant indicator used to assess students' academic ability. School systems, teachers, parents, and students are dependent upon the accuracy of academic ability estimates designated, *θs,* by item response theory (IRT) computer programs. In this study, the accuracy of 3 parameter logistic (3PL) IRT estimates of academic ability were obtained from the BILOG-MG and WinBUGS computer programs which were employed to compare the use of non-informative and informative priors in θ estimation. The rationale for comparing the output of these two computer programs is that the underlying statistical theory employed in these two computer programs is different, and there may be a notable difference in the accuracy of θ estimation when an informative prior is used by WinBUGS in analyzing skewed populations. In particular, the θ parameter estimates of BILOG-MG using traditional IRT analysis with non-informative priors in each situation and the θ parameter estimates of WinBUGS using Recursive Bayesian Analysis (RBA) with informative priors are compared to the true simulated θ value using Root Mean Square Errors (RMSEs). To make this comparison, Monte Carlo computer simulation is used across three occasions within three conditions giving nine comparison situations. For the priors and data generated, results show similar θ estimation accuracy for a normally distributed latent trait (RMSE = 0.35), a more accurate θ estimation process using RBA compared to traditional analysis (RMSEs of 0.36 compared to 0.76) when using latent trait distributions skewed in a similar direction, and less accurate θ estimation using RBA

compared to traditional analysis (RMSEs of 1.48 compared to 0.80) when using

extremely skewed negative then positive distributions in a longitudinal setting.

Implications for further research include extensions to other IRT models, developing

prior elicitation equations, and applying Bayesian informative prior elicitations in

BILOG-MG.

EXAMINATION OF PARAMETER ESTIMATION USING RECURSIVE BAYESIAN
ANALYSIS IN SIMULATED ITEM RESPONSE THEORY APPLICATIONS

by

Robert Hendrick

A Dissertation

Presented in Partial Fulfillment of Requirements for the
Degree of
Doctor of Philosophy
in
Research, Measurement, and Statistics
in
the Department of Educational Policy Studies
in
the College of Education
Georgia State University

Atlanta, GA
2014

ACKNOWLEDGMENTS

I thank my major professor, Dr. William L Curlette, for his hours of discussion and support. Also the contributions of the committee members, Dr. Chris Oshima, Dr. Hongli Li, and Dr. Gwendolyn Benson have improved the structure and development of the simulation. My wife, Dr. Susan Ogletree, has encouraged me throughout the process of completing this dissertation and continues to inspire academic development in the area of educational research. Many other people have made positive contributions to conceptualizing this simulation especially Vincent, Harley, Don, and others that have allowed me to express ideas and have helped bring clarity to the analytical process.

I dedicate this endeavor to my late mother. She always believed in the good in everyone, and me in particular. Love you mom.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

CCRPI   college and career ready performance index

DIC   deviance information criterion

EAP   expected a posteriori

EM   expectation maximization

ICC   item characteristic curve

IRT   Item Response Theory

JML   joint maximum likelihood

LPE   logistic positive exponent

MCMC   Monte Carlo Markov Chain

ML   maximum likelihood

MML   marginal maximum likelihood

PL   parameter logistic

RBA   recursive Bayesian analysis

RLPE   reflection of logistic positive exponent

RMSE   root mean squared error

CHAPTER 1

THE PROBLEM

Researchers, both Bayesian and traditional, have voiced concern over poorly

specified informative priors in item response theory (IRT) analysis (Fox, 2010; Mislevy,

1986). The observation is that a poorly specified distribution as the prior may adversely

affect the accuracy of item and/or person parameter estimation. Priors are part of a

Bayesian process in which the belief about a latent trait prior to seeing the observed data

related to the latent trait is a part of the outcome that estimates the latent trait. Traditional

analysis holds that incorporating a prior is too subjective; therefore, only the data should

be analyzed. In the event of needing a prior distribution, the traditional analysis uses a

non-informative prior that provides little information to the analysis about the outcome.

Therefore in many traditional analyses that use Bayesian techniques, priors used are non-

informative so the outcomes are based primarily on the data. Conversely, Bayesian

analysis uses the belief about the state of nature and changing beliefs about the state as

subsequent Bayesian analyses are performed. The prior may start out as non-informative

but is combined with the likelihood to produce the posterior or the calculated probability

of the outcome. This intermediate posterior governs our belief in a subsequent Bayesian

analysis in which the posterior from the first Bayesian analysis informs the prior in a

subsequent Bayesian analysis. Recursive Bayesian analysis (RBA) occurs when the

informed prior is based upon the previous posterior and the analysis is iterative. Among

the current applications using RBA are RADAR tracking, robotics, navigation, and

profile building software for commercial sites. This dissertation examines the accuracy of

the latent trait estimation when using RBA compared to the accuracy of latent trait

estimation using traditional analysis using a combination of normal and skewed latent trait distributions in a longitudinal setting.

Overview of Previous Research

Many comparisons of the Monte Carlo Markov Chain (MCMC) using Gibbs sampling to marginal maximum likelihood/expected a posteriori (MML/EAP) methods have indicated "little difference in item parameter recovery with samples of 300 or more" (Kieftenbeld & Natesan, 2012, p. 415). Therefore, one can assume that BILOG-MG and WinBUGS will produce similar item and person parameters using test lengths of 40 and a sample size of 500. Both programs allow the user to adjust the specified prior, but in this study, I use WinBUGS with RBA and BILOG-MG with traditional analysis in each situation across three occasions in three conditions. The RBA uses the posteriors from previous situations to inform the prior in subsequent situations simulating a longitudinal application represented by three occasions. There is a need for further examination of the creation of informative priors in IRT analysis. Proctor, Teo, Hou, and Hsieh (2005) indicated that exact or informative priors increased the accuracy of estimation of item parameters in a 2 parameter logistic (PL) IRT model using the WinBUGS software (Lunn, Thomas, Best, & Spiegelhalter, 2000). Studies by Patz and Junker (1999), Proctor et al. (2005), Liu and Gambrell (2009), and Kieftenbeld and Natesan (2012) have compared BILOG-MG (Mislevy & Bock, 1990) and the MCMC with Gibbs sampling used in WinBUGS. For samples of 300 or more simulated users with non-informative priors used in each program, BILOG-MG and WinBUGS have been found to produce relatively comparable item and person parameter estimates using actual data.

Overview of Simulated Situations

In this study, I employ a use of RBA compared with traditional analysis employing a 3PL IRT parameter estimation process in each simulated situation. Three conditions are examined with this simulation: Baseline, Favorable to RBA, and Unfavorable to RBA. All three conditions are analyzed across three occasions in a simulated longitudinal manner. Using a simulated sample size of 500 in each situation and actual 3PL IRT item parameters, the BILOG-MG estimation uses traditional analysis and non-informative priors and WinBUGS uses RBA with informative priors across the three occasions in each condition. In each case the c item parameter, or asymptomatic parameter is fixed at .17 for all situations. I will describe the simulated situations prior to stating the research questions in a technical sense; however, the following questions are addressed by the Baseline, Favorable to RBA, and Unfavorable to RBA conditions respectively:

1. Does the use of Recursive Bayesian Analysis (RBA) improve the accuracy in estimating the latent trait compared to traditional analysis in a longitudinal setting when the simulated sample distributions are normal in each of three situations?

2. Does the use of Recursive Bayesian Analysis (RBA) improve the accuracy in estimating the latent trait compared to traditional analysis in a longitudinal setting when the simulated sample person ability distributions progress from slightly to moderately to highly negative-skewed in each of three situations?

3.   Does the use of Recursive Bayesian Analysis (RBA) improve the accuracy
in estimating the latent trait compared to traditional analysis in a
longitudinal setting when the simulated sample person ability distributions
are <u>normal</u>, <u>highly negative-skewed normal</u>, and <u>highly positive-skewed</u>
<u>normal</u>, respectively, in each of three situations?

The accuracy of the BILOG-MG and WinBUGS parameter estimation for items
and persons are compared to the known true item parameters and the true simulated
sample parameters in all conditions and occasions for each situation. BILOG-MG and
WinBUGS were found by researchers to produce similar accurate estimates of actual item
and person parameters using actual data in sample sizes of over 300 (Kieftenbeld &
Natesan, 2012). Little evidence of IRT simulations applied in longitudinal settings has
been found, but cross-sectional simulated data has been identified by researchers
comparing BILOG-MG and WinBUGS (Proctor et al., 2005; Liu & Gambrell, 2009;
Kieftenbeld & Natesan, 2012).

Conditions: Baseline, Favorable to RBA, and Unfavorable to RBA

The first condition is Baseline 3PL IRT parameter estimation in a longitudinal
setting which in situations 1, 2, and 3 the sample of 500 simulated users is drawn from a
normal distribution. I hypothesize that the informative prior will lead to more accurate
estimation of item and/or person parameters. Even though the simulated sample is
different in each situation, it is drawn from a normal distribution. According to findings
of Proctor et al. (2005), WinBUGS using exact (informative) priors was more accurate
than the estimation of WinBUGS or BILOG-MG using non-informative priors in a cross-

sectional comparison (Proctor et al., 2005). As stated the first condition serves as a baseline comparison with a normal distribution used in each situation.

In the condition Favorable to RBA, the sample of 500 simulated users is drawn from a slightly negative-skewed normal distribution in situation 4, a moderately negative-skewed normal distribution in situation 5, and a moderate to highly negative-skewed distribution in situation 6. In the Favorable to RBA and Unfavorable to RBA conditions the adjusted Fisher-Pearson standardized moment coefficient is used to indicate skew. These informative priors may increase or decrease accuracy of estimation for item and/or person parameters across the three situations in the condition Favorable to RBA because the informative priors are based on the θ posterior distribution from the previous situation. This study will produce evidence to help determine the influence of informative priors on RBA of 3PL IRT applications in a simulated longitudinal manner. The skew begins as slight (- 0.40); progresses to moderate (-0.80); then ends with a moderate to high negative skew (-1.20). The informative prior of the latent trait may produce more accurate estimates of item and person parameters or conversely, the informative priors may restrict the estimation process toward the center of the prior distribution and produce a less accurate estimate (Fox, 2010; Mislevy, 1986).

In the condition Unfavorable to RBA, the sample of 500 simulated users is drawn from a normal distribution in situation 7, a moderate to high negative-skewed (-1.20) normal distribution in situation 8, and a moderate to high positive-skewed (1.20) normal distribution in situation 9. The informative priors are based on the θ posterior distribution from the previous situation, these informative priors may increase or decrease accuracy of estimation for item and/or person parameters across the three situations in the

condition Unfavorable to RBA and indicate if there is a trend (if any) between situations. This study will provide evidence of the effect of RBA with informative priors on typical test development and deployment using 3PL IRT in a longitudinal manner.

Occasion 1 represents the initial situation in each condition (Situations 1, 4, and 7). In the initial occasion, RBA uses a non-informative prior distribution for θ because there is no previous posterior distribution to inform the initial prior. For those situations using RBA in occasions 2 and 3 (situations 2, 3, 5, 6, 8, and 9) the posterior distribution from the previous situation informs the subsequent prior. These simulated situations across occasions simulate a 3PL IRT application in a longitudinal setting. In occasion 1 the test items administered to the sample total 40 and include 4 items that have extreme discrimination or difficulty values. Typically these 4 items with "poor" discrimination or difficulty parameters would not be detected by IRT tools until after occasion 1. Therefore, in occasion 1 in all three conditions the same 40 test items are administered. To simulate the replacement of test items, which occurs in typical longitudinal IRT applications, a portion of the items with "poor" discrimination or difficulty parameters are removed on each of the following occasions (2 and 3) in all conditions and replaced with items from the actual data from the New Mexico Proficiency Test (*Fundamentals of Item Response Theory*, Appendix A) This action of item removal and replacement is typical of actual IRT application across time. The following are research questions phrased in a more technical sense that guided my research:

1. To what extent, if any, does RBA with WinBUGS using informative priors increase the accuracy of estimation for person parameters compared to traditional analysis with BILOG-MG using non-informative priors

employing a 3PL IRT model in a longitudinal setting when the simulated

sample distributions are normal on each of three situations?

2. In what way does RBA with WinBUGS using informative priors affect the

accuracy of estimation for person parameters compared to traditional

analysis with BILOG-MG using non-informative priors employing a 3PL

IRT model in a longitudinal setting when the simulated sample person

ability distributions progress from slightly to moderately to highly

negative-skewed on each of three situations?

3. In what way does RBA with WinBUGS using informative priors affect the

accuracy of estimation for person parameters compared to traditional

analysis with BILOG-MG using non-informative priors employing a 3PL

IRT model in a longitudinal setting when the simulated sample person

ability distributions are normal, highly negative-skewed normal, and

highly positive-skewed normal, respectively, on each of three situations?

There are three conditions (Baseline, Favorable to RBA, and Unfavorable to

RBA) and three occasions (representing administrations of the test) yielding nine

situations for each comparison of software platforms (BILOG-MG and WinBUGS). In

each situation within all three conditions, item parameter estimates (*a* and *b*) and person

parameter estimates (θ) are generated by BILOG-MG with traditional analysis and

WinBUGS with RBA. The researcher uses root mean square error (RMSE) to compare

the item parameter estimates and person parameter estimates generated by BILOG-MG

and WinBUGS with known true item parameters and true simulated person parameters in

each situation. The RMSE comparison of actual item parameters with estimated item

parameters and simulated true person parameters with estimated person parameters indicates whether more accurate estimates are produced by BILOG-MG using a traditional analysis or WinBUGS using RBA within the specific situation.

CHAPTER 2

LITERATURE REVIEW

Classical Test Theory was developed more than 100 years ago, and Item

Response Theory grew from improvements in measurement used for Classical Test

Theory (Baker, 2001; Hambleton, Swaminathan, & Rogers, 1991; Lord, 1952). Basically,

Item Response Theory allows one to assess the item characteristics within a multiple item

test and estimate the examinee's ability given the item parameters and the response

pattern to the test by that examinee. Classical Test Theory examined the true score of the

test; therefore, the observed score is equal to the true score plus error. The true score is

not observable and is similar to the examinee's latent ability found in Item Response

Theory (Baker, 2001; Hambleton et al., 1991). To estimate the latent ability of an

examinee, the items of the test produce a probability or likelihood that the item will be

answered correctly by the examinee. The formal mathematics for i items and j examinees,

where X is the item response, x is a correct response, θ is latent ability, and P is

probability, is as follows:

1

For a difficult item, the probability of an examinee of average ability answering the item

correctly would be lower than an easier item presented to the same examinee. Therefore,

the Item Characteristic Curve (ICC) is a nondecreasing function in θ (Rupp, 2003). The

graph of the examinee's latent trait on the x-axis and the probability on the y-axis

produces a cumulative density function (curve) similar to the one shown in Figure 1.
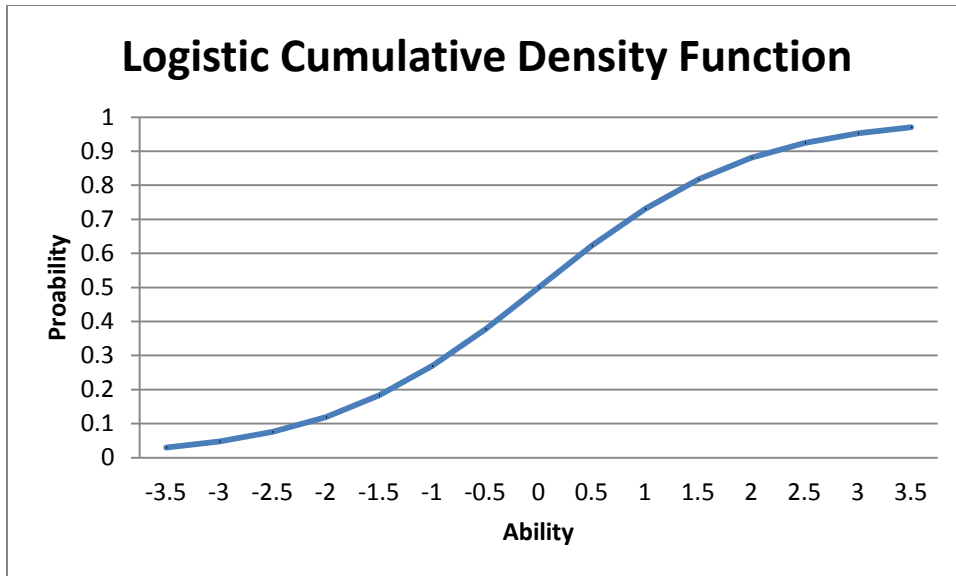
Figure 1. Logistic cumulative density function.

In Figure 1, the curve is a logistic ogive that indicates a latent ability of 0 (average

ability) and a probability of answering correctly for the item is .5; in the curve in Figure

1, the item parameters are $a = 1$ (discrimination) and $b = 0$ (difficulty). The

discrimination or a parameter is the slope of the curve at the point of inflection. The

resulting equation for plotting this curve is as follows for the 2 parameter logistic model:

$$\overline{\rule{3cm}{0pt}}$$ 2

In Equation 2, $D = 1.7$ which is a scaling coefficient for the logistic curve; $a = 1$,

which is the discrimination parameter, and $b = 0$, which is the difficulty parameter. Theta

or θ is the latent ability of the examinee and e is a constant that is approximately 2.718. In

Figure 1, $P(\theta)$, the probability of theta having a correct answer, is shown on the y-axis

and the latent ability (theta) is shown on the x-axis. Therefore, as theta increases, the

probability of a correct answer also increases. This premise forms the basis for Item

Response Theory (IRT) (Baker, 2001; de Ayala, 2009; Hambleton et al., 1991).

There are at least three basic models that are typically used in IRT analysis. The 1PL model makes the assumption that all items have the same discrimination; therefore, in all 1PL models the *a* is assumed to equal 1 and the parameter does not influence the estimated latent ability. The equation for the 1PL model can be written as follows:

$$\overline{\hspace{3em}} \tag{3}$$

The 1PL model is similar to the Rasch Model in that the item difficulty is used to estimate the latent ability of the examinee (Hambleton et al., 1991).

Though an examinee's latent ability may be below average, that examinee may be able to answer more difficult items correctly by context clues or by guessing. The 3PL model allows for the probability of examinees within the lower latent ability range to get an answer correct by chance. The *c* parameter or pseudo-chance-level parameter within the 3PL model tries to account for just such an occurrence (Hambleton et al., 1991). The *c* parameter elevates the lower asymptote of the curve and influences both the *a* and *b* parameters as follows:

$$\overline{\hspace{3em}} \tag{4}$$

Therefore, a test with a number *p* of items and *n* examinees using the 3PL model in which i = an item, j is an examinee, $D = 1.7$ scaling coefficient, θ is the latent ability, *a* is the discrimination parameter, *b* is the difficulty parameter, and c is the pseudo-chance parameter, the probability that the *j*th examinee answers the *i*th item correctly is written as follows:

$$\overline{\hspace{3em}} \tag{5}$$

This equation can be expressed in a simpler form:

in which $logit\ (x) = log$ —— (Curtis, 2010).

For this study, an IRT analysis assumes that the test is unidimensional, that the test items are independent and that local independence is maintained. Typically, the unidimensionality of a test is determined using factor analysis. Most tests have multiple factors that could influence the determination of a latent ability. For example, an individual who does not read English very well would have a more difficult task completing a word problem test in mathematics if all of the items included English passages. The specific mathematics problem-solving ability cannot be adequately measured for examinees who have difficulty reading English passages. Also, the structures of many Criterion Reference Tests (CRTs) are typically unidimensional for specified groups within a population. In factor analysis, a scree plot is sometimes used to determine unidimensionality. When the first factor of a test loads at a high eigenvalue and is separated from the other factors by a wide margin, the test typically meets the assumption of unidimensionality (Hambleton et al., 1991). The idea of unidimensionality also pertains to local independence of test answers. Hambleton et al. (1991) claimed, "when the assumption of unidimensionality is true, local independence is obtained: in this sense the two concepts are equivalent" (p. 10).

When an IRT analysis fits the data, the examinees' ability estimates are not test dependent and the item parameters are not group dependent (de Ayala, 2009). Data analyzed are dichotomous within this study; however, using polytomous data and applying multidimensional IRT algorithms have been developed by other researchers and

continued refinement of those concepts are underway (de Ayala, 2009; Fox, 2010; van

der Linden & Hambleton, 2010).

Some recent studies have examined a skewed distribution of the latent ability

(Bazan, Branco, & Bolfarine, 2006; Bolfarine & Bazan, 2008; Broccoli & Cavrini, 2010).

Using a skew latent distribution relaxes an IRT assumption that the latent ability should

be transformed to a normal distribution with mean equal of 0 and a standard deviation of

1 (Hambleton et al., 1991). These skew distribution studies propose an asymmetrical

distribution of the latent variable rather than a symmetrical normal distribution. The

studies develop the ideas that Samejima (1997, 2000) expressed regarding modeling

using asymmetric item characteristic curves, and Bazan et al. (2006) introduced a skew

variable λ as an added parameter. Bazan et al. (2006) showed that if $\lambda = 0$, the normal

probability density function is created. Normal distributions of latent ability may not

occur in many educational testing situations and assuming an asymmetrical or skewed

distribution may hold practical value (Samejima, 2000).

In another study, Broccoli and Cavrini (2010) determined that using a "skew

normal distribution" for the latent variable appeared more effective when model fit was

improved. In their study, they administered a five-item health questionnaire to a large

sample of school children with the hypothesis that because the questionnaire measures

self-perception of health, the supposed latent ability distribution of children completing

the questionnaire would be skewed to the right. Three IRT models were used to analyze

the data: a partial credit model; IRT mixed model; and IRT mixed model with a skewed

latent variable. The program WinBUGS was used to estimate the parameters and the fit

was indicated by the Deviance Information Criterion (DIC). Broccoli and Cavrini (2010)

found that by using a prior skewed latent distribution the DIC index was improved slightly from the second model (model 3 DIC was 129.8, model 2 DIC was 133.3). Also, it is observed that convergence occurred in less time using the skew latent distribution, which could be another indication of better model fit (Broccoli & Cavrini, 2010).

In another recent study, Bolfarine and Bazan (2008) used simulated data from known parameters of 18 items and 974 students of a fourth grade mathematical test. The distribution of the observed scores indicated that the latent variable was skewed in the population, so a skewed logistic model was applied to the simulated data. In this analysis, WinBUGS was used to estimate the parameters and DIC was also used to indicate model fit. Bolfarine and Bazan (2008) found that the skewed latent distribution model fit better than the IRT models without skewed latent distribution. The researchers proposed a new ICC to IRT named the reflection of the logistic positive exponent (LPE) model or RLPE. These two models represent a portion of the skewed logistic models proposed by Samejima (1997, 2000) and support the use of asymmetrical latent distributions and corresponding ICCs.

There have been many applications of IRT using one, two, and three parameter models during the past 40 years (e.g., Baker, 2001; de Ayala, 2009; Fox, 2010; van der Linden & Hambleton, 2010). As desktop computers became more powerful in the decades following the 80s, more researchers started applying IRT to test analysis and construction because of the advantage over Classical Test Theory in which the examinee scores are dependent. Baker (1985) released a small book, *The Basics of Item Response Theory*, and included software written for the current state-of-art computers (Apple II and IBM PC). This publication inspired other researchers to begin IRT work with the

available computer systems and currently, there are many computer applications

available for ability and parameter estimation (Hambleton et al., 1991). One of the more

prominent computer applications of IRT is BILOG-MG (Mislevy & Bock, 1984; Rupp,

2003).

BILOG-MG is commercially available software for IRT applications from

Mislevy and Bock (1984) and is regularly updated. The current version is 3.0 and can

handle 1PL, 2PL, and 3PL IRT models. Though one can choose Joint Maximum

Likelihood (JML) and Marginal Maximum Likelihood procedures, the default procedure

is Expected A Posteriori using a default value of 20 Expectation Maximization (EM)

cycles followed by a maximum of 2 Newton cycles if needed (de Ayala, 2009). The

default procedure typically performs faster and yields more accurate parameter estimates

of the procedures for dichotomous data (Foley, 2010).

There have been numerous studies comparing BILOG-MG to many other

programs that generate parameters and latent ability levels from response data (Foley,

2010). One of the earliest comparisons of software was conducted by Yen in 1987 (cited

in Rupp, 2003) in which LOGIST and BILOG-MG were compared using a simulated test

of 10 to 40 items and 1000 simulated examinees. In Yen's study (1987) BILOG-MG was

identified as estimating parameters more accurately than LOGIST for short tests. In a

similar simulation, Mislevy and Stocking (1989) also found that BILOG-MG performed

better than LOGIST on a 15-item test with 1500 simulated examinees. In 1993, Yoes

(cited in Foley, 2010) conducted a large simulation study using LOGIST, ASCAL, and

BILOG-MG. Within Yoes' simulation several sample sizes (250, 500, 1000, and 2000),

differing test lengths (15, 20, 50, 75, and 100), and differing latent ability distributions

(normal, uniform, and negatively skewed) were used. BILOG-MG was identified in the majority of the studies as estimating the item parameters accurately with sample sizes of 500 to 1000 examinees and latent ability levels more accurately using shorter test lengths (15 to 20 items) (Foley, 2010). An interesting study by Abdel-fattah in 1994 comparing BILOG-MG and LOGIST used non-informative default priors for LOGIST and BILOG-MG, but also used informative priors with BILOG-MG only. Comparison of the three estimation methods indicated that BILOG-MG using informative priors produced more accurate estimates than either LOGIST or BILOG-MG using default priors.

Patsula and Gessaroli presented a comparison of BILOG-MG and TESTGRAF at the 1995 National Council for Measurement in Education Annual Meeting which indicated that though close in estimation, TESTGRAF had more accurate estimations of latent ability given shorter tests. However, as test length increased the estimates of latent ability were similar and as sample size increased the item parameter estimates were similar between BILOG-MG and TESTGRAF. Yoes, who had completed a comparison in 1993, expanded his study to include ASCAL, BILOG-MG, LOGIST, and XCALIBRE in 1995. The 1995 Yoes' study used two of the same data sets used in the 1993 study, test lengths and sample sizes were the same as 1993, but the varied distributions for the latent ability were not used in the 1995 study. The evaluation was consistent with Yoes' 1993 results. BILOG-MG and XCALIBRE produced more accurate estimates of item parameters for small sample sizes and all programs performed similarly regarding larger sample sizes and longer test lengths. In numerous studies, BILOG-MG emerged as the program that produces the most accurate estimates of item and ability parameters for

sample sizes from 300 to 2000 and test lengths from 20 to 60 items. These studies are

indexed in table 1.

Table 1

Research Studies Comparing BILOG-MG with Other Programs

Researcher(s)	Year	Model	Programs Compared	Test Length	Sample Size	Ability Distribution
Yen	1987	3PL	BILOG-MG LOGIST	10, 40	1000	Normal Right skew Left skew
Mislevy Stocking	1989	3PL	BILOG-MG LOGIST	15, 45	1500	Normal
Yoes	1993	3PL	ASCAL BILOG-MG LOGIST	15, 20, 50, 75, 100	200, 500, 1000, 2000	Normal, Uniform, Negative Skew
Abdel-fattah	1994	3PL	BILOG-MG LOGIST	20, 60	200, 500, 1000, 2000	Normal (informative item parameters)
Patsula Gessaroli	1995	3PL	BILOG-MG LOGIST	20, 40	100, 250, 500, 1000	Normal
Yoes	1995	3PL	ASCAL BILOG-MG LOGIST XCALIBRE	15, 20, 50, 75, 100	200, 500, 1000, 2000	Normal

In most of the comparison studies, the point estimates of parameters have been

compared to true parameters using the Root Mean Squared Error (RMSE). The RMSE of

the parameter estimations shows the difference between the predicted parameters and the

true parameters for BILOG-MG and compared estimation programs. The RMSE is

calculated using the following formula:

8

where the estimated parameter for the *ith* item is   ,   is the true parameter for the *i*th

item and n represents the sample size.

Central to the computing ability of BILOG-MG is the Expectation Maximization

(EM) algorithm (Hsu, Ackerman, & Fan, 1999; Mislevy & Bock, 1984). Though Bock

has characterized the EM cycles within BILOG-MG as being related to the EM

algorithm, Hsu et al. (1999) provided a mathematical proof that the EM cycle used in

BILOG-MG is a special case of the EM algorithm when θ is assumed to be continuous or

discrete. The modified EM algorithm used in BILOG-MG is a rather complicated series

of equations that estimate a parameter through selecting expected values of the

parameters given established quadrature intervals and response patterns of examinees.

The values of the parameters are then maximized to produce the maximum likelihood

estimate of the parameters. Using a Bayesian prior, these parameter estimates are within a

typical distribution, thus eliminating inappropriate estimations. The kernel of the EM

implementation with respect to the posterior probability for θ is as follows from the Rupp

(2003) article:

$$\overline{\qquad\qquad}$$ 9

In the above equation, $g(\theta)$ represents prior beliefs and $d(\theta)$ represents observed data

regarding the population distribution of θ. Beginning with this premise, the EM algorithm

executes three steps iteratively until convergence of the item parameter estimates is

reached.

First, the posterior probability of θ is calculated for each examinee *i* at each

quadrature point *k* using the following equation:

Second, the posterior probabilities incorporate artificial data for each item j at each

quadrature point k in the following:

Third, the derivatives with respect to item parameters of the following equation

are set to 0 to determine the maxima of the equation. This maximizes the estimated item

parameters and is rewritten using artificial data in the former equation. The iterative

process continues until the program reaches the specified convergence value (typically

0.001). The Bayesian prior distributions for each item parameter can be set by the

BILOG-MG user or default program settings can be used to estimate item parameters (de

Ayala, 2009; Mislevy & Bock, 1984; Rupp, 2003).

BILOG-MG has been compared to LOGIST, ASCAL, TESTGRAF, XCALIBRE,

ICL, and other item parameter estimation programs (Foley, 2010; Mead, Morris, & Blitz,

2007). In the majority of comparisons, BILOG-MG has provided the best estimates of

item parameters using smaller sample sizes and has been as successful as other programs

in estimating item parameters using large datasets (Foley, 2010). BILOG-MG also makes

use of Bayesian informative priors for the latent variable as well as the estimated

parameters. When using proper informative priors, BILOG-MG has performed the

estimation of parameters more accurately than using non-informative priors (Foley,

2010).

Mislevy (1986) applied a Bayesian framework to the estimation process of BILOG-MG to improve the performance of the software. Taking a Bayesian approach and applying it to the entire IRT estimation process, Jean-Paul Fox (2010) released a book, *Bayesian Item Response Modeling*, in which the program WinBUGS is used to estimate item parameters. Fox and Glas (2001) use the Bayesian approach to IRT using a Monte Carlo Markov Chain (MCMC) with Gibbs sampling instead of the EM algorithm used by other researchers. A handful of young researchers have continued to use the WinBUGS program to extend the IRT modeling development to multidimensional IRT, skewed distributions, and bi-factor multidimensional IRT (Suh, 2010; Sheng, 2005; Fukuhara, 2009; Bolfarine & Bazan, 2010).

WinBUGS is a Bayesian modeling software that runs within the Microsoft Windows environment. Called BUGS (Bayesian inference Using Gibbs Sampling), it is open source and can run within R (Lunn, Thomas, Best, & Spiegelhalter, 2000). The MCMC simulation with Gibbs sampling uses the Metropolis-Hastings algorithm to estimate unknown parameters (Congdon, 2010; Fox, 2010). The item parameter estimation process within WinBUGS is different from that of the EM algorithm used in the BILOG-MG program. The Gibbs sampling technique does not rely on integration but on the Markov Chain properties. There are instances in which the solving of the integral is very difficult, so instead of solving or approximating the parameter directly, a random member within the parameter's known distribution is selected as an initial value for the parameter within a transition process. The initial value of the parameter is exchanged in the Markov procedure for a random value within the known distribution of the parameter and is conditional on the other unknown parameters and the known observed responses.

In estimating an unknown parameter, θ, Gibbs sampling manages a transition process and the form of the conditional densities. The transition process begins by making draws at iteration $m+1$ from the conditional probability density function as follows:

The transition from   is completed if

$$\rule{2cm}{0.4pt}$$

Where   is the mth observation from random variable U. In this manner, the Gibbs sampling continues until the process reaches convergence. The initial replications of the process are deleted and called a "burn-in" period in which the values are more extreme or volatile (Fox, 2010).

There are some issues with MCMC calculation in that the extreme estimates are brought closer to the mean because of informative priors (Mislevy, 1986; Fox, 2010). However, without the influence of the priors, the maximization of parameters can lead to inconsistent and inflated parameter estimates (Mislevy, 1986). There have been comparisons of BILOG-MG and MCMC algorithms used to estimate item parameters. Patz and Junker (1999) examined MCMC methods of estimation compared to the EM algorithm of BILOG-MG. The study used similar normal priors for the latent ability and the b parameter, while a log-normal prior was used for the a parameter. One Markov chain was used to provide the parameter estimates with a 400 iteration burn-in and 7000 iterations after the initial burn-in period. The results were practically identical for both

programs. Other studies that compared the MCMC process to BILOG-MG, and

PARDUX were conducted into the early 2000s (Proctor, Teo, Hou, & Hsieh, 2005) with

similar results using actual test data; however, when using simulated data with known

parameters the MCMC method appeared to estimate the parameters better than BILOG-

MG (Proctor et al. 2005).

The use of the MCMC method for estimating IRT parameters was difficult to

access for many IRT users. The MCMC method required the user to code the algorithm

in S-Plus or other software (Proctor et al., 2005). However, subsequent to the release of

WinBUGS, the IRT models could be programmed directly into the software which uses

the MCMC method to estimate parameters. Proctor et al. (2005) decided to compare the

parameter estimation of WinBUGS and BILOG-MG within their study. Proctor et al.

(2005) used SAS^{TM} to simulate data for a 30 item dichotomous test and 1000 examinees.

Fifteen datasets were generated and the dataset with the fewest responses that were all

correct or incorrect was used for the simulation. The reason for the selection is that the

Maximum Likelihood (ML) algorithm used in BILOG-MG does not process the extreme

item responses of all correct or all incorrect.

All of the simulated dataset was processed through BILOG-MG and WinBUGS to

produce estimates of the item parameters. Default settings were used for the priors in

BILOG-MG (latent ability was normal [N(0, 1)], difficulty parameter was the default

[N(0, 2)], and the discrimination parameter was the default [LN(1, 1)]. Within

WinBUGS, there were "exact" or informative priors set and non-informative priors. For

the "exact" or informative priors, latent ability and difficulty were set at N(0, 1) and the

discrimination parameter was set at LN(0,.5). For their non-informative priors, Proctor et

al. (2005) set difficulty at N(0, 32), discrimination at LN(0, 32), and latent ability at N(0, 1). Three Markov chains were completed for each of the exact prior settings and non-informative prior settings. The iterations for burn-in were 1500 for the exact priors and 2500 for the non-informative priors. The iterations used to estimate the parameters were 2500 for the exact priors and 1500 for the non-informative priors (Proctor et al., 2005). The graphs produced for autocorrelation indicate that the chains converged on the stationary distribution.

The researchers used a 95% confidence interval around the true parameters to indicate accuracy rather than RMSE results. Their findings indicate that WinBUGS using exact priors was better in estimating item parameters. WinBUGS using exact priors was marginally better than WinBUGS using non-informative priors, and both WinBUGS scenarios performed better than BILOG-MG at estimating the item parameter values (Proctor et al., 2005). The results of this study also demonstrate that setting the exact priors in WinBUGS gives the program an advantage in estimating item parameters. The researchers called for more research in this area that involves actual data to measure the accuracy of item parameter estimates, varied test length, and varied numbers of examinees.

Liu and Gambrell (2009) referenced Proctor et al. (2005) and expanded upon their findings. Liu and Gambrell (2009) analyzed simulated response data using a 36 item, multiple choice test with published item parameters (a, b, and c) treated as true parameters. Item responses were simulated using the 3PL IRT model and published item parameters for 200, 500, 1000, and 2000 users from a normal [N(0, 1)] distribution. The item parameter estimates were then conducted using BILOG-MG based upon the

simulated item response data. The item parameter estimates were also conducted in
WinBUGS using the same simulated item response data. There were four item parameter
estimations made with each of the programs, BILOG-MG and WinBUGS, to indicate the
item parameter estimation accuracy compared to the true item parameters with 200, 500,
1000, and 2000 simulated users. The results of the item parameter estimation were
mixed: WinBUGS was superior to BILOG-MG in estimating the a parameter when
simulated users were 500 or less, BILOG-MG was superior to WinBUGS in estimating
the b parameter when simulated users were 500 or less, and WinBUGS was superior to
BILOG-MG in estimating the c parameter at all sample sizes. For both programs the item
parameters a and b were better estimated for items without extreme a and b true
parameters, and the c parameter was better estimated for more difficult items by both
programs. Similar uninformative priors were used for both WinBUGs and BILOG-MG
(Liu & Gambrell, 2009). The study did not report the accuracy of estimation regarding
the person parameter, θ, or use informative priors for either of the programs.

Recently, Kieftenbeld and Natesan (2012) compared the parameter estimation of
Marginal Maximum Likelihood with MCMC using Samejima's grade response model.
During this extensive study, test length, sample size, latent trait distribution, and
estimation procedures were investigated yielding 120 fully crossed conditions. Some of
the results indicated that test length influenced the estimation of person parameters but
did not have much effect on item parameter estimation. Conversely, sample size had a
measureable influence on estimation of item parameters using both estimation
techniques, but sample size had little influence on estimation of person parameters. The
independent variables of test length were 5, 10, 15 and 20 items. The estimation of person

parameter was found to improve from a test having a length of 10 items to 15 items, but the improvement between 15 and 20 items was negligible (Kieftenbeld & Natesan, 2012). Five sample sizes were examined: 75, 150, 300, 500, and 1000 simulated users. Results indicate that item parameter estimation improves as sample size increases using both MML and MCMC; however, "there is little difference in item parameter recovery with samples of 300 people or more" (Kieftenbeld & Natesan, 2012, p. 415). This finding regarding required sample size is typical within studies using dichotomous IRT models. In contrast to Bazan et al. (2006) and Bolfarine and Bazan (2008), Kieftenbeld and Natesan (2012) used a normal prior latent trait distribution [N(0, 1)] to estimate person parameters given a normal, uniform, and skewed user distribution. In this study, Kieftenbeld and Natesan (2012) found that "the unit normal person parameter worked well even when it did not match the uniform or skewed latent trait distribution in the sample" (p. 416), indicating that the normal distribution prior for the latent trait did not appreciably improve the θ estimate when differing latent distributions are simulated in the sample. A common suggestion for further research concerns establishing more informative priors used in the Bayesian analysis (Kieftenbeld & Natesan, 2012; Proctor et al., 2005).

Kieftenbeld and Natesan (2012), Liu and Gambrell (2009), and Proctor et al. (2005), have all added to the published literature pertaining to the comparison of BILOG-MG and WinBUGs by previous researchers. There is agreement that in the simulated situations examined that both programs function well with samples of 500 or more, that a test about 15 items in length is the minimum for estimating a person parameter that has a high degree of accuracy, and the use of informative priors may improve the Bayesian

estimation of item and person parameters. However, there are different opinions

regarding the estimation of latent parameters using a normal distribution when the

distribution of the latent parameters is known to be skewed (Samejima, 1997, 2000;

Bolfarine & Bazan, 2008; Kieftenbeld & Natesan, 2012). A typical use of IRT in a test

making context across a number of occasions may allow use of informative priors in a

recursive Bayesian format. The use of informative priors has been reported to increase

the accuracy of estimation by Proctor et al. (2005); however, the use of informative priors

may not be beneficial when the informative prior is based on past data and the current

data has a different distribution, creating a nonconjugate prior (Mislevy, 1986). RBA uses

the Bayes theorem, in which the posterior [probability of the trait given the data, p(A|B)]

is proportional to the prior [probability of A, p(A)] combined with the likelihood

[probability of the data given the trait, p(B|A)]. The Bayes theorem for two states of

nature is as follows:

$$\rule{3in}{0.5pt} \qquad 15$$

In Equation 15, A forms an observed dichotomous variable (= 1, = 0). The left side of

the equation represents the posterior probability, p(A|B). In applying RBA, the posterior

distribution found in occasion 1 of data analysis is retained and informs the prior

distribution for occasion 2 analysis. The posterior distribution of occasion 2 is retained

and informs the prior distribution for occasion 3, and so on in a longitudinal process. In

RBA, the information retained in the subsequent posterior is used to inform the following

prior, thus possibly increasing the degree of accuracy of the estimation. This is a

possibility at this point because the posterior/prior is reflective of the sample tested,

which may change slightly from year to year in practical educational applications of IRT

(Hambleton, et al. 1991). The informative prior could possibly be too restrictive for different samples and could result in a worse estimate than non-informative priors for the person parameters (Fox, 2010; Mislevy, 1986). Thus, the study simulates three years of item responses so that I can analyze the longitudinal process with BILOG-MG using non-informative conjugate priors compared with WinBUGS and RBA with informative priors.

With this research, I seek to expand upon the previous work comparing BILOG-MG and WinBUGS with the following variations. For test length, I use one initial set of item length of 40 items and replace "poor" items in subsequent situations so the test length of 40 items is maintained. In this manner, I simulate the act of removing "poor" test items from the initial set and replacing those items with published items. This action is typically done when creating a high-stakes test (Sinharay, 2006). In addition, the number of examinees is fixed for each test length comparison using 500 simulated examinees. The sample size of 500 was identified by researchers as being effective for item parameter estimation by both BILOG-MG and WinBUGS (Ayala, 2009; Kieftenbeld & Natesan, 2012; Proctor et al., 2005).

In the Baseline condition, 500 simulated users are drawn from a normal distribution in each simulated situation, 1 through 3. In the condition Favorable to RBA, 500 simulated users will be drawn from a slightly negative-skewed normal distribution (-.40) for situation 4, a moderately negative-skewed normal distribution (-.80) in simulated situation 5, and a moderate to highly negative-skewed normal distribution (-1.20) in simulated situation 6. In the condition Unfavorable to RBA, 500 simulated users are drawn from a normal distribution for situation 7, a moderate to highly negative-

skewed normal distribution (-1.20) in simulated situation 8, and a moderate to highly

positive-skewed normal distribution (1.20) in simulated situation 9. A different sample of

500 simulated users is drawn from the aforementioned distributions for each respective

condition and occasion. Thus, the estimation accuracy of item and person parameters by

BILOG-MG and WinBUGS is compared in each situation across three occasions within

each of the three conditions. Both BILOG-MG and WinBUGS uses the same distribution

in each of the simulated situations to estimate item and person parameters.

There are four items in which the discrimination or difficulty parameter are

artificially assigned extreme values. These are the four items that are replaced within the

test, three just prior to occasion 2 in all conditions and the remaining item just prior to

occasion 3 in all conditions. By the slight difference in true test item parameters, varying

distributions of examinees (in the condition Favorable to RBA and Unfavorable to RBA),

utilization of informative priors and posteriors, and repeating the process across three

occasions, the researcher attempts to simulate a typical and practical use of the 3PL IRT

model in a longitudinal setting.

Root mean square error (RMSE) is determined for the item parameter estimations

and person parameter estimations for each of the situations with the item parameters and

person parameters being compared with the true parameters in each case. This effort

substantially expands the research of Kieftenbeld and Natesan (2012), Liu and Gambrell

(2009), and Proctor et al. (2005) and draws upon the established comparison methods

used by Patz and Junker (1999) and others to determine accurate parameter estimates.

CHAPTER 3

METHOD

I use SAS^{TM} Software (Fan, Xitao, Felsovalyi, Silvo, & Keenan, 2001) to

simulate test responses with known item parameters for the 3PL model and use the

simulated response data to estimate item parameters within BILOG-MG and WinBUGS

using the 3PL model. The degree of estimation accuracy of parameters of WinBUGS

using informative priors compared with the degree of estimation accuracy of BILOG-MG

using non-information priors is examined for three conditions. Each condition is

simulated across three occasions of incorporating new data. The 3PL IRT simulation

model for the longitudinal situations within the Baseline condition is shown in Figure 2.

For the Baseline condition in situation 1, the sample distribution of the latent trait or

person parameter is normal [N(0, 1)] for both programs. Non-informative priors are used

in both programs, those are as follows:

For the Baseline condition in situation 1, there is a simulated sample size of 500 users

drawn at random from a normal distribution with a mean of 0 and a standard deviation of

1. The minimum number of recommended examinees for the 2PL and 3PL model is 1000

(Hambleton et al., 1991); however, Kieftenbeld and Natesan (2012) indicated that there is

little improvement in parameter estimation between 500 and 1000 users. Therefore, the

researcher simulates the number of examinees at 500 to provide a comparison using a

Baseline Condition - 3PL model, 40 items

	Occasion 1 Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$ Non-informative Person Parameter Prior Distribution Theta $= N(0,1)$		Occasion 2 Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$		Occasion 3 Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$
Simulation Sample: 500 Simulated Users From $N(0,1)$ (Fixed c Parameter at 0.17)	Situation: Baseline.1.BILOG-MG Traditional Analysis Item Paramter Estimation Person Parameter Estimation	Simulation Sample: 500 Simulated Users From $N(0,1)$ (Fixed c Parameter at 0.17)	Situation: Baseline.2.BILOG-MG Traditional Analysis Non-informative Person Parameter Distribution Theta $= N(0,1)$ Item and Person Parameter Estimation	Simulation Sample: 500 Simulated Users From $N(0,1)$ (Fixed c Parameter at 0.17)	Situation: Baseline.3.BILOG-MG Traditional Analysis Non-informative Person Parameter Distribution Theta $= N(0,1)$ Item and Person Parameter Estimation
	Situation: Baseline.1.WinBUGS14 Recursive Bayesian Analysis (Initial Phase) Item Paramter Estimation Person Parameter Estimation		Situation: Baseline.2.WinBUGS14 Recursive Bayesian Analysis Informative Person Parameter Prior Distribution (Informed by previous Theta Posterior Distribution) Item and Person Parameter Estimation		Situation: Baseline.3.WinBUGS14 Recursive Bayesian Analysis Informative Person Parameter Prior Distribution (Informed by previous Theta Posterior Distribution) Item and Person Parameter Estimation

Figure 2. Baseline Condition for 3PL IRT Simulation

relatively small group within the 3PL platform. To make the comparison of item parameter estimation more accurate, the c parameter is fixed at 0.17 within the 3PL model. The number of test items is set at 40 and maintained at that length across situations 1, 2, and 3. The previous studies recommended using 30 to 60 items, so the consistency is maintained with item count and number of examinees. These 500 simulated users have their latent ability and the relative difficulty of the item form a probability of a correct answer. That probability is then compared with a random number drawn from a uniform distribution bounded by 0 and 1. A response of 1 occurs if the probability of a correct answer is greater than the random draw from a uniform distribution; the response is 0 otherwise. The dichotomous response is recorded for all items and persons creating a response matrix consisting of either 1 or 0 that is 500 X 40 for all situations. The generation of the response matrix process is repeated 50 times in an iterative manner on all situations, yielding a total number of simulated users of 25,000 per situation (500x50=25,000). Those dichotomous response matrices are analyzed by BILOG-MG using the aforementioned non-informative priors in all situations. This is a typical use of BILOG-MG when the true parameters of the sample are unknown and there are changes in the item parameter distributions because of item replacement (Sinharay, 2006). The same dichotomous response matrices are analyzed by WinBUGS using the same non-informative priors in situation 1 and using informative priors (informed by the posterior distribution) for situations 2 and 3. Analysis of parameter estimation compared to the true parameter is conducted using RMSE compared to the true parameters for both programs in all situations.

The process for situation 2 in the Baseline condition is similar to the situation 1 process. Differences are that three items are removed and replaced with three published items maintaining the length of the test, yielding 40 test items for situation 2. This action simulates typical use of IRT diagnostic information and is indicative of practical scenarios. Three of the "poor" items are replaced within the test. Also, there is a new generation of the sample of 500 simulated users and the analysis is completed for 50 replications. To simulate typical use, the BILOG-MG analysis uses non-informative priors because the typical sample latent trait presents unknown parameters. The choice of non-informative priors for the BILOG-MG analysis is typical for situation 2. Conversely, using RBA allows the researcher to incorporate prior information that can help in estimating person parameters. The posterior distribution from the person parameter estimation process in situation 1 informs the prior distribution for the person parameter estimation process in situation 2. In this manner, data are not duplicated in the analysis, but the information from the previous analysis is retained within the RBA. However, with the actions of replacing items and creating a different simulated sample in situations 2 and 3, the prior information may narrow the analysis and actually reduce the degree of accuracy of the estimation if the prior distribution is a poor predictor of the changing sample (Mislevy, 1986). By simulating typical use of IRT in test development and deployment, the practical use of RBA can be investigated through the RMSE analysis of parameters in each situation. With the aforementioned exceptions, all other processes in situation 2 are the same as situation 1 in the Baseline condition.

The process for situation 3 in the Baseline condition has some differences from situation 2. The test length is maintained at 40 through the simulated replacement of

another "poor" item and the sample of 500 simulated users is drawn from a normal distribution. RBA is again used by WinBUGS by incorporating the situation 2 person parameter posterior distribution to inform the situation 3 person parameter prior distribution and the BILOG-MG priors remain non-informative. All other processes in the Baseline condition in situation 3 proceed as described for situation 1. The Baseline condition simulates a typical application of the IRT process across three occasions in which the items in the test are refined and the actual samples change across the three occasions. However, in each situation samples are drawn from the normal distribution.

For the Favorable to RBA condition in situation 4, the sample distribution of the latent trait or person parameter is a slightly negative-skew normal distribution for both programs shown in Figure 3. The same non-informative item and person priors are used in both programs as in the previous condition. The Favorable to RBA condition in situation 4 uses a simulated sample size of 500 users drawn at random from a slightly negative-skew normal distribution (-.40) with a mean of 0.5 and standard deviation of 1. To make the comparison of item parameter estimation more accurate, the c parameter is fixed at 0.17 within the 3PL model. The number of test items is set at 40 and maintained at that length across situations 4, 5, and 6. The creation and analysis of the dichotomous response pattern for the favorable to RBA condition is similar to the Baseline condition and is repeated 50 times in an iterative manner in all situations.

The process for situation 5 in the favorable to RBA condition is similar to the situation 4 process. Differences are that three items are removed and replaced with three published items maintaining the length of the test, yielding 40 test items for

Favorable to Recursive Bayesian Analysis Condition- 3PL model, 40 items

	Occasion 1		Occasion 2		Occasion 3
	Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$ Non-informative Person Parameter Prior Distribution Theta $= N(0,1)$		Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$		Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$
Simulation Sample: 500 Simulated Users From Normal with Slight Neg. Skew (-.40) (Fixed c Parameter at 0.17)	Situation: Favorable.4.BILOG-MG Traditional Analysis Item Paramter Estimation Person Parameter Estimation	Simulation Sample: 500 Simulated Users From Normal with Moderate Neg. Skew (-.80) (Fixed c Parameter at 0.17)	Situation: Favorable.5.BILOG-MG Traditional Analysis Non-informative Person Parameter Distribution Theta $= N(0,1)$ Item and Person Parameter Estimation	Simulation Sample: 500 Simulated Users From Normal with Moderate to High Neg. Skew(-1.20) (Fixed c Parameter at 0.17)	Situation: Favorable.6.BILOG-MG Traditional Analysis Non-informative Person Parameter Distribution Theta $= N(0,1)$ Item and Person Parameter Estimation
	Situation: Favorable.4.WinBUGS Recursive Bayesian Analysis (Initial Phase) Item Paramter Estimation Person Parameter Estimation		Situation: Favorable.5.WinBUGS Recursive Bayesian Analysis Informative Person Parameter Prior Distribution Item and Person Parameter Estimation		Situation: Favorable.6.WinBUGS Recursive Bayesian Analysis Informative Person Parameter Prior Distribution Item and Person Parameter Estimation

Figure 3. Favorable to RBA Condition for 3PL IRT Simulation

situation 5. This action simulates typical use of IRT diagnostic information and is indicative of practical scenarios. Three of the "poor" items are replaced within the test. Also, there is a new generation of the sample of 500 simulated users using a moderately negative-skew distribution (-.80) with a mean of 0.75 and standard deviation of 1. To simulate typical use, the BILOG-MG analysis uses non-informative priors because the sample distribution presents unknown parameters for this condition.

The choice of non-informative priors for the BILOG-MG analysis is typical for situation 5. Conversely, using RBA allows the researcher to incorporate prior information that can help in estimating person parameters. The posterior distribution from the person parameter estimation process in situation 4 informs the prior distribution for the person parameter estimation process in situation 5. In this manner, data are not duplicated in the analysis, but the information from the previous analysis is retained within the RBA. However, with the actions of replacing items and creating a different simulated sample from a slightly different distribution in situations 5 and 6, the prior information may narrow the analysis and actually reduce the degree of accuracy of the estimation if the prior distribution is a poor predictor of the changing sample (Mislevy, 1986). By simulating typical use of IRT in test development and deployment, the practical use of RBA can be investigated through the RMSE analysis of parameters in each situation. With the aforementioned exceptions, all other processes in situation 5 are the same as situation 4 in the Favorable to RBA condition.

The process for situation 6 in the Favorable to RBA condition has some differences from situation 5. The test length is maintained at 40 through the simulated replacement of another "poor" item and the sample of 500 simulated users is drawn from

a moderate to highly negative-skew normal distribution (-1.20) with a mean of 1.0 and a standard deviation of 1. RBA is again used by WinBUGS by incorporating the situation 5 person parameter posterior distribution to inform the situation 6 person parameter prior distribution and the BILOG-MG priors remain non-informative. All other processes in the Favorable to RBA condition in situation 6 proceed as described in situation 4. The Favorable to RBA condition simulates a typical application of the IRT process across three occasions in which the items in the test are refined and the actual samples change across situations 4, 5, and 6. Each of the three samples is drawn from a negatively skewed distribution that is progressively more skewed in each situation.

For the Unfavorable to RBA condition, shown in Figure 4, in situation 7, the sample distribution of the latent trait or person parameter is a normal distribution for both programs. The same non-informative priors will also be used in both programs as in the previous initial occasion. The Unfavorable to RBA condition in situation 7 uses a simulated sample size of 500 users drawn at random from a normal distribution with a mean of 0 and a standard deviation of 1. To make the comparison of item parameter estimation more accurate, the c parameter is fixed at 0.17 within the 3PL model. The number of test items is set at 40 and maintained at that length across situations 7, 8, and 9. The creation and analysis of the dichotomous response pattern for the Unfavorable to RBA condition is similar to the previous conditions and is repeated 50 times in an iterative manner in all situations.

The process for situation 8 in the Unfavorable to RBA condition is similar to the situation 7 process. Differences are that three items are removed and replaced with

Unfavorable to Recursive Bayesian Analysis Condition - 3PL model, 40 items

	Occasion 1 Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$ Non-informative Person Parameter Prior Distribution Theta = $N(0,1)$	Occasion 2 Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$	Occasion 3 Common to all situations: Non-informative Item Parameter Prior Distributions $a = LN(0,0.5)$ $b = N(0,2)$
Simulation Sample: 500 Simulated Users From Normal Distribution (Fixed c Parameter at 0.17)	Situation: Unfavorable.7.BILOG-MG Traditional Analysis Item Parameter Estimation Person Parameter Estimation Situation: Unfavorable.7.WinBUGS Recursive Bayesian Analysis (Initial Phase) Item Parameter Estimation Person Parameter Estimation	Simulation Sample: 500 Simulated Users From Normal with Moderate to High Neg. Skew (-1.20) (Fixed c Parameter at 0.17) — Situation: Unfavorable.8.BILOG-MG Traditional Analysis Non-informative Person Parameter Distribution Theta = $N(0,1)$ Item and Person Parameter Estimation Situation: Unfavorable.8.WinBUGS Recursive Bayesian Analysis Informative Person Parameter Prior Distribution Item and Person Parameter Estimation	Simulation Sample: 500 Simulated Users From Normal with Moderate to High Pos Skew (1.20) (Fixed c Parameter at 0.17) — Situation: Unfavorable.9.BILOG-MG Traditional Analysis Non-informative Person Parameter Distribution Theta = $N(0,1)$ Item and Person Parameter Estimation Situation: Unfavorable.9.WinBUGS Recursive Bayesian Analysis Informative Person Parameter Prior Distribution Item and Person Parameter Estimation

Figure 4. Unfavorable to RBA Condition for 3PL IRT Simulation

three published items maintaining the length of the test, yielding 40 test items for situation 8. This action simulates typical use of IRT diagnostic information and is indicative of practical scenarios. Three of the "poor" items are replaced within the test. Also, there is a new generation of the sample of 500 simulated users using a moderate to highly negative-skew distribution (-1.20) with a mean of 1.0 and a standard deviation of 1. To simulate typical use, the BILOG-MG analysis uses non-informative priors because the sample distribution presents unknown parameters for this situation. The choice of non-informative priors for the BILOG-MG analysis is typical for situation 8. Conversely, using RBA allows the researcher to incorporate prior information that can help in estimating person parameters. The posterior distribution from the person parameter estimation process in situation 7 informs the prior distribution for the person parameter estimation process in situation 8. In this manner, data are not duplicated in the analysis, but the information from the previous analysis is retained within the RBA. However, with the actions of replacing items and creating a different simulated sample from a highly different distribution in situations 8 and 9, the prior information may narrow the analysis and actually reduce the degree of accuracy of the estimation if the prior distribution is a poor predictor of the changing sample (Mislevy, 1986). By simulating typical use of IRT in test development and deployment, the practical use of RBA is investigated through the RMSE analysis of parameters in each situation. With the aforementioned exceptions, all other processes in situation 8 are the same as situation 7 in the Unfavorable to RBA condition.

The process for situation 9 in the Unfavorable to RBA condition has some differences from situation 8. The test length is maintained at 40 through the simulated

replacement of another "poor" item and the sample of 500 simulated users is drawn from a moderate to highly positive-skew normal distribution (1.20) with a mean of -1.0 and a standard deviation of 1. RBA is again used by WinBUGS by incorporating the situation 8 person parameter posterior distribution to inform the situation 9 person parameter prior distribution and the BILOG-MG priors remain non-informative. This condition is termed "Unfavorable to RBA" because the sample distribution undergoes an extreme change from negative skew -1.20 in situation 8 to positive skew 1.20 in situation 9. The person parameter prior distribution is unlike the given sample distribution to test the accuracy of using recursive Bayesian estimation within extreme situations. All other processes in the Unfavorable to RBA condition in situation 9 proceed as described in situation 7.

An analysis using RMSE is completed for the estimated parameters for each condition in each situation compared with the true parameters used for each condition in each situation. In that manner an investigation of using RBA with informative priors is assessed in a typical IRT application when the sample distribution remains normal (Baseline condition), experiences gradual change in a consistent direction (Favorable to Recursive Bayesian Analysis condition), and experiences extreme change in different directions (Unfavorable to Recursive Bayesian Analysis condition).

A SAS[TM] software program is used to simulate the dichotomous response data for the test items. For the 40 item test, items 10, 20, 30, and 40 comprise the "poor" items that are removed from the test across occasions 2 and 3 in all conditions. These artificially assigned parameters help to identify issues in the estimation of extreme item parameter values by the WinBUGS and BILOG-MG programs. The parameters are shown in Table 2 and reflect extreme low and high discrimination and difficulty within

those items. Items 10 20 and 30 are removed prior to situation 2 and item 40 is removed

prior to occasion 3. After these items are removed, those items will be replaced with

published items from the New Mexico Proficiency Test (*Fundamentals of Item Response

Theory*, Appendix A). The removal and replacement of items is intended to simulate a

typical IRT diagnostic process of identifying "poor" test items regarding difficulty and

discrimination parameters and making adjustments to the test in the test development

process (Sinharay, 2006).

Table 2

Parameters of "Poor" Items

Item	Item Parameters		
	a	b	c
10	0.02	0.0	0.17
20	1.0	-3.0	0.17
30	2.0	0.0	0.17
40	1.0	3.0	0.17

The 40 items in the test help show the accuracy of person parameter estimation of

WinBUGS and BILOG-MG when using authentic data from actual tests in the 3PL IRT

model. The initial test items are taken from actual data from the New Mexico Proficiency

Test (*Fundamentals of Item Response Theory*, Appendix A) showing actual item

parameters for the 3PL model in Hambleton et al. (1991). There are a total of 75 items

that have 3PL true parameters listed. The initial test (first occasion for all conditions)

includes the first 36 items and four "poor items"; the second test (second occasion for all

conditions) includes the first 39 items and one "poor item"; and the third test (third occasion for all conditions) includes the first 40 items.

The simulation is actually a nested model with two programs (BILOG-MG parameter estimation using traditional analysis and WinBUGS parameter estimation using RBA) in each situation across three occasions within each of three conditions (3 × 3 = 9 situations). For both WinBUGS and BILOG-MG parameter estimation in each condition and occasion described below, the SAS^{TM} software 3PL simulation script uses predefined item parameters and randomly generated person parameters (abilities) (θ). Given the item parameters and the simulated abilities, response probabilities for each item within each condition and occasion are calculated and compared with random numbers drawn from a uniform distribution (0,1) to generate a discrete variable. The SAS^{TM} software program will read the true parameter file given *a, b*, and *c* parameters from a data file for the 40 items for simulation in each condition and occasion. The text of these files is shown in Appendix C.

The number of simulated examinees is 500 and is a random draw from a simulated population of examinees with normal distribution, mean = 0 and standard deviation = 1 in the Baseline condition in all situations 1 through 3. In the Baseline condition, a separate random draw of 500 simulated users from a normal distribution is made for each occasion and within each situation there are two programs compared to the true parameters. The parameter estimates are calculated using WinBUGS with RBA, and parameter estimates are calculated using BILOG-MG. In each estimation process within a condition and situation both programs will use the same sample and response pattern in the analysis and estimation of the person parameter. This Baseline measurement is to

determine how RBA may perform in a longitudinal study when there is no difference in the sample's normal distribution between occasions. The normal distribution of the latent variable may be found in tests that are norm-referenced or tests that measure a latent trait or ability that is normally distributed in the population.

In the Favorable to RBA condition, the simulated sample distributions across situation 4 through 6 become progressively negative skewed. In the Favorable to RBA condition in situation 4, the simulated sample distribution is negative skewed by -.40; in situation 5, the simulated sample is negative skewed by -.80; and in situation 6, the simulated sample is negative skewed by -1.20. The distribution in all three situations is negative skewed in increasing amounts. This simulates a test in which the students are given the criteria and pertinent curriculum that the test is based upon. On subsequent occasions the tested population gradually improves their test scores through more direct and effective teaching of the concepts tested. This may be indicative of Criterion-Referenced-Tests (CRT) where the outcome may be negative skewed.

In the Unfavorable to RBA condition the simulated sample distributions across situation 7 through 9 show extreme change. In the Unfavorable to RBA condition in situation 7, the simulated sample distribution is normal; in situation 8, the simulated sample is negative skewed by -1.20; in situation 9, the simulated sample is positive skewed by 1.20. The Unfavorable to RBA condition simulates changing test administration or population treatments that result in extreme shifts in the scoring outcomes across occasions.

In the Favorable and Unfavorable to RBA conditions the variations for occasions 1, 2, and 3 involve a change from the normal distribution. To simulate the skewed normal

distributions, Fleishman's (1978) power transformation was used as a part of the SAS^{TM} code (Appendix A). To further simulate the skewed normal distributions the mean of the distribution was also slightly shifted in each of the transformations as follows: slight negative skewed mean = 0.5, moderate negative skewed mean = 0.75, highly negative skewed mean = 1.0, and highly positive skewed mean = −1.0. These changes simulate typical population skewed distributions found in many educational settings in which CRT testing is utilized and are consistent with skewed distributions simulated by Bazan et al. (2006), Kieftenbeld and Natesan (2012), and Samejima (1997). For the Favorable and Unfavorable to RBA conditions the detailed description of the sample distributions are shown in Table 3.

Table 3

Sample Distributions in the Favorable and Unfavorable to RBA Conditions

	Favorable			Unfavorable	
Situation	4	5	6	8	9
Mean	0.5	0.75	1.0	1.0	-1.0
SD	1	1	1	1	1
Skew	-0.40	-0.80	-1.20	-1.20	1.20
Excess kurtosis	0.4	0.7	1.4	1.4	1.4

Transformed from the 0 to 1 range of the typical beta distribution to a -3 to 3 range, the progressively negative skewed normal distributions generated in the Favorable to RBA condition can be modeled as β(7.4,5), β(8.7,4), and β(16.1,3) respectively. The extreme change in the Unfavorable to RBA condition goes from a normal distribution to

$\beta(16.1,3)$ to $\beta(3,16.1)$ in each of three situations showing extreme changes in the shape of the distribution. Each of these skewed distributions is slightly, moderately, or highly skewed from normal and may be representative of a typical population in which a CRT may be administered. In each of three conditions there are three occasions that are examined within this study with parameters *a, b*, and θ being estimated by WinBUGS and BILOG-MG for a total of 9 situations (three conditions x three occasions). In each situation the estimates from WinBUGS and BILOG-MG are compared with the true item parameters used to generate the dichotomous responses and the true θ values selected by the SAS^{TM} simulation. The conditions represent a consistent normal distribution in the Baseline condition, a progressively negative skewed distribution in the Favorable to RBA condition, and an extreme shift in the distributions in the Unfavorable to RBA condition. The situations simulate outcome data across three occasions within each condition and simulate outcome distributions from year to year in typical IRT practical applications. The SAS ^{TM} code for each condition and occasion is shown in Appendix B.

The root mean square error of the parameter estimations is calculated for the estimated item parameters compared with the true item parameters and placed in tables by the designated situation. The RMSE shows the difference between the estimated parameters and the true parameters for BILOG-MG and WinBUGS within two sections of the tables for each situation. Sample items will be provided in a comprehensive table to examine trends, if any, among the situations within the simulated data. The RMSE is calculated using the following formula:

$$\frac{\overline{}}{\overline{}}$$

16

where   is the estimated parameter for the *i*th item,   is the true parameter for the *i*th

item, and n = 25,000 (500X50). In addition to calculating RMSE for the item parameters,

RMSE is calculated for the person parameters estimated in each situation. The true

person parameters are indications of the latent ability or trait being examined. Within

each situation, 500 simulated users are draw at random from differing distributions. The

true person parameters drawn by SAS[TM] are compared to the person parameters estimated

by BILOG-MG and WinBUGS for each situation. The intent of this study is to

investigate the accuracy of item and person parameter estimation using RBA in a

simulated longitudinal process of test development using 3PL IRT.

Numerous studies have shown that BILOG-MG and WinBUGS are similar at

estimating item and person parameters with sample sizes of 500 and more (Kieftenbeld &

Natesan, 2012; Liu & Gambrell, 2009; and Proctor et al., 2005). More research has called

for investigating the use of informative priors in Bayesian IRT analysis (Kieftenbeld &

Natesan, 2012; and Proctor et al., 2005). The RBA allows the researcher to retain the

information from prior and posterior information and apply that information in a

subsequent analysis. Classical analysis typically starts with a "clean state" and does not

consider prior information in a current analysis. This study seeks to simulate a typical use

of a 3PL IRT process within conditions and across occasions to provide evidence of the

comparable accuracy of item and person parameter estimation in a simulated longitudinal

process.

The process used in forming the prior for the following situation consists of

evaluating the posterior distribution estimated by WinBUGS in the former situation. In

all three conditions: Baseline, Favorable to RBA, and Unfavorable to RBA, the first

occasion uses a non-informative prior in the Bayesian analysis. In each condition, occasions 2 and 3 use the former posterior distribution to inform the subsequent prior. The posterior is analyzed using the four central moments of the posterior distribution: mean, variance, skewness, and kurtosis. These four central moments describe the location, dispersion, and shape of the posterior distribution. The informative prior for the following situation is a transformed beta distribution with similar moments to the posterior distribution found in the previous situation. The beta distribution shape parameters are adjusted to a similar location, dispersion, and shape as measured by the first four central moments of the beta distribution. Finally Pearson's χ^2 statistic is used to examine the "goodness of fit" of the current prior in transformed beta distribution format and former posterior distributions.

Some IRT Monte Carlo simulations include large samples and few replications, others select an arbitrary number of replications that may produce confusing outcomes; however, this Monte Carlo simulation sets the replications at 50 and the sample size at 500 for a total of 25,000 users in each situation (50 x 500 = 25,000). When estimating the *a, b*, and θ parameters, the total sample size number times the number of replications provides a large simulated population of 25,000 users. A recommended minimum number of replications is 25 (Harwell, Stone, Hsu, & Kirisci, 1996), but because the person parameter, θ, is the focus of the study, 50 replications and 500 sample size are combined to yield the desired precision.

CHAPTER 4

RESULTS AND DISCUSSION

The purpose of this study is to examine the increase, if any, in the accuracy of person parameter estimates using RBA compared to traditional analysis with 3 PL IRT applications in a longitudinal setting. The simulation procedures described in chapter 3 were completed and the results for each of the conditions: Baseline, Favorable to RBA, and Unfavorable to RBA are reported.

Restating from chapter 3, there are three occasions within each condition, the Baseline condition, Favorable to RBA condition, and Unfavorable to RBA condition. Each situation in the occasion applies the same item data and simulated user data to estimate the item and person parameters a, b, and θ using BILOG-MG and WinBUGS. Microsoft Excel is used to calculate and analyze the RMSE of the estimates compared to the true parameters. In each situation within all conditions the sample size was 500 with 50 replications of the two programs (BILOG-MG and WinBUGS) which are automated within a SAS ^{TM} program (Appendix B). BILOG-MG results are recorded as BILOG-MG output files (.PH1, .PH2, and .PH3) in a designated directory. Likewise, WinBUGS output files are placed in another designated directory and the simulated true person parameters are recorded by SAS ^{TM} in yet another directory. Estimates of θ and the item parameters, a and b, from all of the output files are cataloged and RMSE calculations are made for all 50 replications per situation. Therefore, the analysis of RMSE is provided for each situation within all conditions.

Baseline Condition Results

Each program, BILOG-MG and WinBUGS, used non-informative priors in the initial occasion of the Baseline condition. The RMSE when estimating the latent ability or trait ranges from 0.34 to 0.38 for BILOG-MG and from 0.32 to 0.37 for WinBUGS across the 50 replications. The means of the RMSE for θ estimation are also similar with 0.36 for BILOG-MG and 0.35 for WinBUGS, which indicates slightly more precision for WinBUGS; however, the accuracy of the θ estimation is virtually the same.

In the Baseline condition the posterior of situation 1 informs the prior distribution of situation 2. The total sample size for the situation 1 posterior distribution is 25,000 (500×50) and the posterior distribution is calculated using the total sample size. Because the posterior distribution is normal, the informed prior distribution for situation 2 will only use the first two central moments of the former posterior distribution (mean and variance). The informed prior used in situation 2 was N(0.00898, 0.93012) which is very similar to the default prior and the latent trait distribution of N(0, 1) in the Baseline condition. The RMSE of the θ estimates compared to the simulated true θ values is 0.37 for BILOG-MG and 0.36 for WinBUGS for situation 2. Similarly, the informed prior distribution for situation 3 in the Baseline condition is N(0.0162, 0.9642) and the RMSE of the θ estimates compared to the simulated true values are 0.36 for BILOG-MG and 0.35 for WinBUGS using RBA in situation 3. Because the informative prior distribution used in the RBA is very similar to the non-informative priors and the simulated latent ability distributions in situations 2 and 3 of the Baseline condition, the RMSE of the traditional analysis using BILOG-MG and the RMSE of the RBA are very similar for a normally distributed latent variable. This finding indicates that the use of RBA with

WinBUGS compared to traditional analysis with BILOG-MG does not increase the

accuracy of the θ estimation using a sample size of 500 or more that is distributed

normally, N(0, 1). The finding is consistent with the published literature which indicates

that BILOG-MG and WinBUGS have similar accuracy when estimating the θ parameter

Table 4

Root Mean Square Error (RMSE) Calculations for Baseline Condition

Estimated Parameter	BILOG-MG	WinBUGS
Situation 1 (500 × 50)		
θ	0.36	0.35
a	0.14	0.19
b	0.19	0.28
Situation 2 (500 × 50)		
θ	0.37	0.36
a	0.12	0.16
b	0.16	0.28
Situation 3 (500 × 50)		
θ	0.36	0.35
a	0.13	0.14
b	0.15	0.24

using a sample size of 500 or more that have a normal N(0, 1) distribution. This finding

also suggests a baseline level of precision of approximately 0.35 for the estimation of the

θ parameter for both BILOG-MG with traditional analysis and WinBUGS using RBA in

this simulation.

Though the accuracy of estimating the item parameters, *a* and *b*, are not the focus of this study, item parameter estimates by BILOG-MG and WinBUGS are compared with the true item parameters (Appendix C) using RMSE. The findings of this study show that in the Baseline condition the *a*, or discrimination parameter estimates are similar for BILOG-MG and WinBUGS as shown in Table 4. The *b*, or difficulty parameter estimates in the Baseline condition from BILOG-MG are more precise with a RMSE of 0.19, 0.16, and 0.15 in situations 1, 2, and 3 respectively. The difficulty item parameter estimates for WinBUGS have a RMSE that is 0.28, 0.28, and 0.24 for the same three situations using a normal distribution of the latent variable. The default prior of the difficulty parameter for BILOG-MG and the prior for the difficulty parameter set in WinBUGS was N(0, 0.5) for all situations within the Baseline condition.

Favorable to RBA Condition Results

Each program, BILOG-MG and WinBUGS, use non-informative priors in the initial occasion of the Favorable to RBA condition. The RMSE when estimating the latent ability or trait ranges from 0.51 to 0.67 for BILOG-MG and from 0.47 to 0.61 for WinBUGS across the 50 replications. The means of the RMSE for θ estimation are also similar with 0.61 for BILOG-MG and 0.55 for WinBUGS (see Table 5), which indicates slightly more precision for WinBUGS; however, the accuracy of the θ estimation is virtually the same with the difference indicated approximately 0.06. Both programs generate less accurate θ estimates than the RMSE calculations in situation 4 compared to the RMSE value previously found in the Baseline condition. The sample is drawn from a negative skewed distribution with mean = 0.5, standard deviation = 1, skew = -0.40, and excess kurtosis = 0.4 as shown in chapter 3, table 2. Because the dichotomous response

Table 5

Root Mean Square Error (RMSE) Calculations for Favorable to RBA
Condition

Estimated Parameter	BILOG-MG	WinBUGS
Situation 4 (500 × 50)		
θ	0.61	0.55
a	0.16	0.20
b	0.62	0.42
Situation 5 (500 × 50)		
θ	0.76	0.36
a	0.19	0.21
b	0.87	0.30
Situation 6 (500 × 50)		
θ	0.79	0.37
a	0.34	0.28
b	1.16	0.25

matrix is also generated from the comparison of the probability of a correct answer with a

random uniform variable, each simulated true θ value for each sample of 500 users has a

mean range of 0.36 to 0.57 and the mean is 0.49; however in a skewed distribution the

median is a better indicator of central tendency, therefore the median range is 0.36 to 0.63

and the median is .53. The standard deviation of the simulated sample is 0.92 and the

skew has a wide range of -0.70 to 0.01 among the 50 samples, with the total skew equal

to -0.31. The following chart in figure 5 shows the prior, latent trait, and posterior distributions for the initial occasion in the Favorable to RBA condition.

Though the latent trait distribution in typical IRT applications is unknown, it is known in these simulated data and the latent trait distribution yields data which are negative skewed which is dissimilar to the non-informative prior in simulation 4, N(0, 1). This difference in the prior produces parameter estimates that have less accuracy in predicting person parameters than instances when the latent trait is distributed normally around a mean of 0, as demonstrated by elevated RMSE values in the Favorable to RBA condition, situation 4 than those RMSE values previously shown in the Baseline condition.
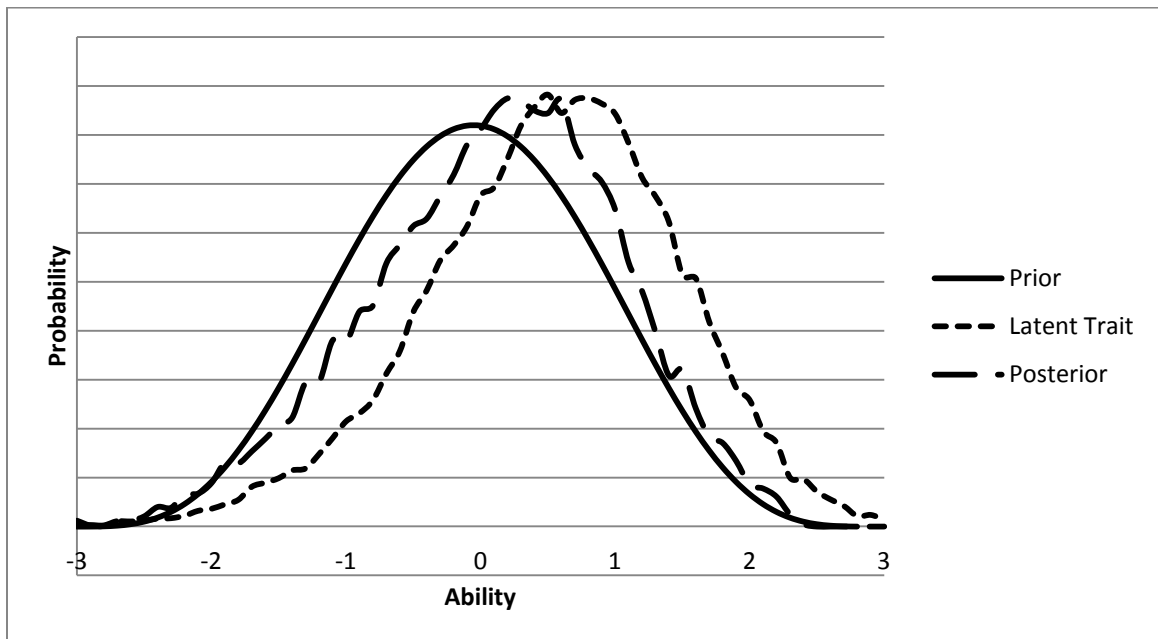

Figure 5. Bayesian analysis for Favorable condition, situation 4 – using non-informative prior.

In the Favorable to RBA condition the posterior of situation 4 informs the prior distribution of situation 5. The total sample size for the situation 4 posterior distribution

is 25,000 (500 x 50) and the posterior distribution is calculated using the total sample size. The informed prior distribution for situation 5 uses the first four central moments of the former posterior distribution (mean, variance, skew, and kurtosis), because the posterior distribution shown in Figure 5 is negative skewed. The informed prior used in situation 5 is represented as a transformed beta distribution which uses shape parameters α and β. The beta distribution can approximate a uniform, normal, negative skew, and positive skew distribution. Also α and β values are easily interpreted into mean, variance, skew, and excess kurtosis using the following equations:

$$\underline{\hspace{3cm}} \hspace{4cm} 17$$

$$\underline{\hspace{4cm}} \hspace{3cm} 18$$

$$\underline{\hspace{2cm}} \quad \underline{\hspace{2cm}} \hspace{4.5cm} 19$$

$$\underline{\hspace{6cm}} \hspace{2cm} 20$$

In equation 17, which indicates the first central moment, μ equals the mean, *min* is the minimum of the range, *max* is the maximum of the range, and α and β are shape parameters. The mean is the first central moment and is considered the strongest indication of central tendency in normal distributions. However, because the posterior distribution from the Favorable to RBA condition, situation 4 has a negative skew, the stronger indication of the center of the distribution may be the median or mode. In equations 18, 19 and 20 which indicate the second, third, and fourth central moments, equals the variance, γ equals the skewness, and k equals the excess kurtosis. Typically, the mean and variance are used to describe a distribution because those are the measures with the greatest precision when there are homogenous data. In using the skewness and

kurtosis to inform the prior probability distribution, the precision is far less than that of the mean and variance with the uncertainty of the skewness statistic about 2.45 times the uncertainty of the central tendency (Wheeler, 2011). The kurtosis statistic has more variation at about 4.9 times the uncertainty of the central tendency of the distribution. Therefore, caution is recommended when using either the skewness or kurtosis statistic of the posterior distribution to inform the subsequent prior distribution. A large sample (25,000) and homogeneous data are used to inform the subsequent prior because the uncertainty is greater using the skewness and kurtosis.

The first four central moments typically provide the characteristics of a distribution and by carefully matching the first four central moments of the posterior with a beta distribution transformed to the -3 to 3 range the next prior distribution can be informed. The previous posterior distribution central moments are used to inform the subsequent prior distribution in a transformed beta, given a large sample and homogeneous data. It is interesting that even distributions with exactly the same mean, variance, skew, and kurtosis may not be exactly the same; however, those distributions will be similar (Wheeler, 2011). In analyzing the situation 4 posterior distribution the characteristics are shown in the Table 6.

Table 6 shows the first four central moment values of the Favorable to RBA condition, situation 4 posterior distribution of the total sample size across all replications $(500 \times 50 = 25,000)$. The transformed beta distribution α shape parameter is assigned the value of 5.4 (because an α value of 5.4 creates a beta distribution with similar shape to

Table 6

Characteristics of Posterior Distribution Favorable Condition, Situation 4

Mean	0.059703259
Standard Error	0.005819662
Median	0.1223
Mode	1.064
Standard Deviation	0.920169374
Sample Variance	0.846711676
Excess Kurtosis	-0.232830971
Skewness	-0.277216436
Range	5.48
Minimum	-3.149
Maximum	2.331
Sum	1492.581473
Count	25000

the posterior distribution) and the β parameter is calculated using the formula shown in Equation 21.

$$\rule{3cm}{0.4pt} \qquad\qquad\qquad\qquad 21$$

where α is a beta shape parameter of 5.4, *min* is the minimum value in the range of the distribution, *max* is the maximum value, and μ is the measure of central tendency. Typically in a normal distribution, μ is recognized as the mean; however, because the posterior distribution has a skew statistic of -0.28, the posterior distribution has a slight negative skew which indicates that a better measure of central tendency may be the

median or mode. Substituting the values for the min, max, α, and setting μ equal to the

median (0.12), the β value is 5.06 producing a transformed distribution beta (5.4, 5.06).

That beta (5.4, 5.06) distribution has a mean of 0.15, variance of 0.89, skew of -0.04, and

excess kurtosis of -0.44. Though the mean and variance are similar to the central

tendency value (median) and variance of the posterior distribution the skew statistic is not

similar with the posterior indicating greater negative skew than the beta; thus, additional

fitting is needed.

Another process of fitting uses the Pearson χ^2 statistic that compares the posterior

distribution with the transformed beta prior distribution. Using 13 equivalent intervals

within the range of both distributions the χ^2 statistic was calculated using the following:

$$\underline{\hspace{3cm}} \hspace{6cm} 22$$

where i is the number of a specific interval, n is the total number of intervals,   is the

observed count for interval i, and   is the expected count for interval i. In this

simulation the posterior distribution contains the observed value and the subsequent prior

distribution contains the expected value. The β parameter is adjusted to minimize the χ^2

statistic in an effort to fit the transformed beta prior distribution to the observed posterior

distribution. The adjustment results in a reduction in the β parameter from 5.06 to 4.8; at

which the χ^2 statistic is minimized.

The posterior distribution is shown in graphical format in Figure 6, with the

transformed beta (5.4, 4.8) distribution (in process of fitting) superimposed. The

transformed beta (5.4, 4.8) distribution has characteristics: μ = 0.32,   = 0.82, γ = -0.10,

and k = -0.45. The mean of the transformed beta distribution is between the median and
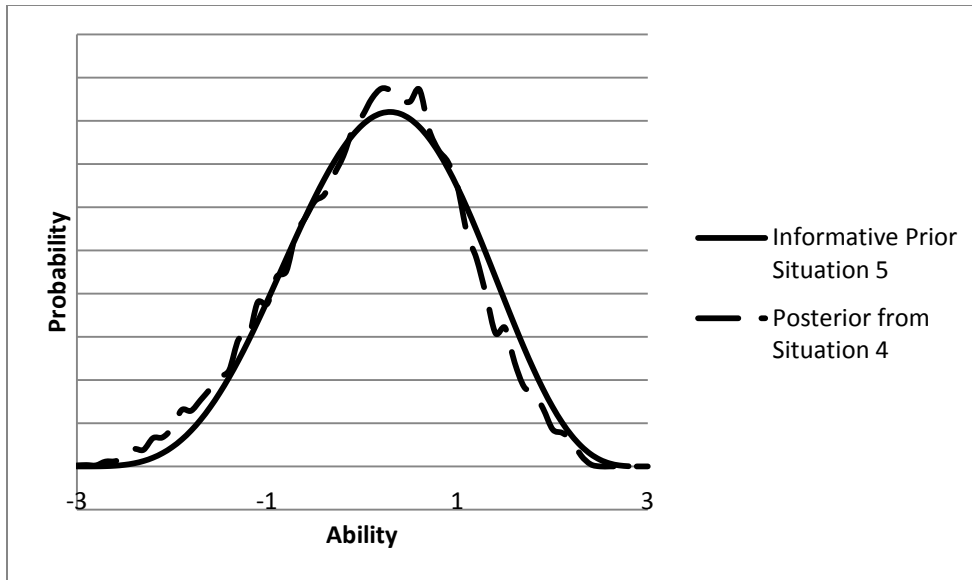
Figure 6. Bayesian analysis for Favorable condition, situation 5 – prior distribution beta (5.4, 4.8) informed by posterior distribution situation 4.

mode of the situation 1 posterior distribution which is consistent with determining the central tendency of a skewed sample. The variance is similar (0.82 and 0.85); however, the skew values are not similar (-0.10 and -0.28) and the excess kurtosis values are not similar (-0.45 and -0.23). Though these distributions indicate similar central moment values for variance, the values for skewness are not similar. Because the central tendency indicator of a skewed distribution is better indicated by the median or mode, the mean for the transformed beta distribution is compared to the median or mode of the posterior. Also the skew value has a slightly negative skew in the posterior and the skew of the transformed beta distribution is currently much less skewed. The skewness value of the beta (5.4, 4.8) distribution is -0.10 which is not similar to the skewness value of the posterior distribution of -0.28

The application of the Pearson   fit statistic is somewhat problematic. The χ^2 statistic is minimized at about beta (5.4, 4.8) and fits the transformed beta prior to the posterior distribution when the bins to describe the distributions are limited to the -1.5 to

1.5 range, or focused on the center of the distributions. The statistic for that central

portion of the distributions is,   which is less than   .

However, when the tails are included in the bins, the   statistic is 398.87, which greatly

exceeds the critical value. This inconsistency between the central part of the distribution

and the tails indicates that the transformed beta distribution tails may not conform to the

posterior tails so that the fit statistic becomes an insignificant indicator of fit. Due to the

fact that the transformed beta distribution needs to be similar to the posterior distribution,

the application of the   fit statistic may lead to over fitting the prior distribution to the

former posterior distribution because the central portion of the posterior distribution was

influenced by the non-informative situation 4 prior and biased toward the mean of that

non-informative prior.

　　　　The lower accuracy in θ estimation is evidenced by the elevated RSME (0.55 and

0.61) for θ estimation by WinBUGS and BILOG-MG, shown in table 4 for the Favorable

to RBA condition, situation 4. The lower accuracy may be attributed to the slightly

skewed likelihood and resulting posterior distribution. To make estimates with higher

accuracy, the researcher proposes that the skew values may be made more similar

between the former posterior distribution and subsequent prior distribution by further

adjustment of the β parameter. It is the belief of the researcher that the second and third

central moments (variance and skew) of the posterior distribution should be primarily

used to inform the prior distribution because of the inconsistency in skew values between

the posterior and prior and decreased dependence of the mean as an indicator of central

tendency in a skewed distribution. Since the posterior distribution is influenced by the

non-informative prior in situation 4, the posterior distribution's skew value may indicate

that the transformed beta distribution used for the prior in situation 5 needs to increase

the negative skew. Typically the negative skew is increased by reducing the β shape

parameter in a beta distribution; therefore, the informative prior is adjusted to beta (5.4,

3.8), thus reducing the β shape parameter from the previous beta distribution. The

transformed beta (5.4, 3.8) distribution has characteristics: $\mu = 0.57$, $= 0.86$, $\gamma = -.20$,

and k = -0.44. With this new distribution the mean is still between the median and mode

(0.122 and 1.064) of the posterior distribution, the variance is consistent, the skew values

are similar (-0.20 and -0.28), and the kurtosis is relatively unchanged. Since a slightly

more platykurtic prior distribution is desirable for a prior distribution, the excess kurtosis

of -0.44 is acceptable. The transformed beta (5.4, 3.8) distribution superimposed upon the

posterior distribution from situation 4 is shown in graphic form in Figure 7.
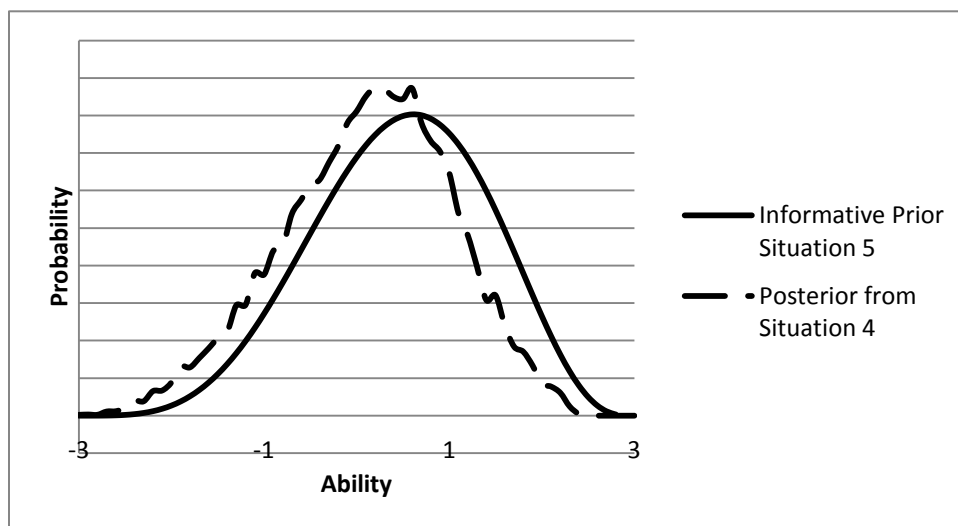
*Figure 7.* Bayesian analysis for Favorable condition, situation 5 – prior distribution beta
(5.4, 3.8) informed by posterior distribution situation 4.

Therefore, the Favorable to RBA condition, situation 5 prior distribution is beta

(5.4, 3.8) for WinBUGS and N (0,1) (the default) for BILOG-MG. The RMSE when

estimating the latent ability or trait ranges from 0.69 to 0.84 for BILOG-MG and from

0.32 to 0.39 for WinBUGS across the 50 replications. The means of the RMSE for θ

estimation are not similar with 0.76 for BILOG-MG and 0.36 for WinBUGS, which

indicates more precision for WinBUGS; in this case the accuracy of the θ estimation

indicates a 95% credible interval for WinBUGS at ± 0.71 and the BILOG-MG 95%

credible interval at ±1.53. In this simulation, the accuracy in estimating θ is improved by

using an informative prior in the Favorable to RBA condition, situation 5.

   The simulated sample for situation 5 is drawn from a negative skewed distribution

with mean = 0.75, standard deviation = 1, skew = -0.80, and excess kurtosis = 0.7. The

dichotomous response matrix is also generated from the comparison of the probability of

a correct answer with a random uniform variable and the total simulated sample of true θ

values for 25,000 users (500 users and 50 replications) has a total mean of 0.66; however,

in a skewed distribution the median (0.82) is a better indicator of central tendency. The

standard deviation of the simulated sample is 0.78 and the total skew equal to -1.12,

which is highly skewed. Figure 8 shows the prior, latent trait, and posterior distributions

for situation 5 in the Favorable to RBA condition.

   Figure 8 shows that the posterior distribution is more consistent with the latent

trait distribution in the Favorable to RBA condition, situation 5 using the informative

transformed beta prior. The RMSE is similar to the estimation error previously displayed

by both programs in the Baseline condition, thus the Favorable to RBA θ parameter
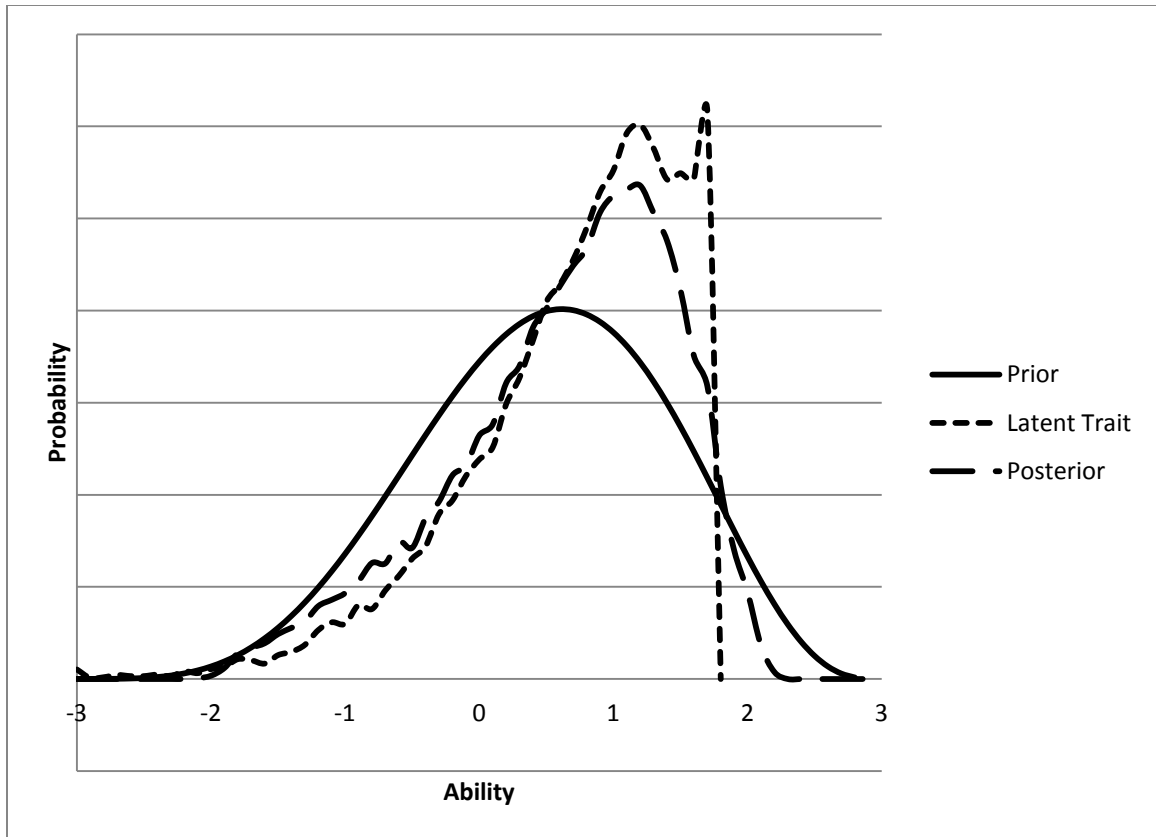
Figure 8. Bayesian analysis for Favorable condition, situation 5 – using informative prior beta (5.4, 3.8).

estimation may have the most precision achieved in this simulation with a negatively skewed sample and the RMSE is much more accurate than using non-informative N(0,1) priors when the sample is skewed. In analyzing the situation 5 posterior distribution to inform the situation 6 prior, the characteristics of the situation 5 posterior distribution are shown in Table 7.

Table 7 shows the first four central moment values of the Favorable to RBA condition, situation 5 posterior distribution of the total sample size across all replications (500x50=25,000). A process using the estimated shape and calculation of the β parameter that is very similar to the fitting process for the situation 5 prior is completed. During that process, a transformed beta prior distribution for situation 6 using shape parameters of α

equal to 7.5 and β equal to 4.73 is fitted to the situation 5 posterior distribution. The

posterior distribution is shown in graphical format in Figure 9, with the transformed beta

(7.5, 4.73) distribution superimposed.

Table 7

Characteristics of Posterior Distribution Favorable Condition, Situation 5

Mean	0.582571581
Standard Error	0.005250524
Median	0.7268
Mode	1.209
Standard Deviation	0.830180725
Sample Variance	0.689200036
Excess kurtosis	-0.033306977
Skewness	-0.69732275
Range	4.279
Minimum	-2.151
Maximum	2.128
Sum	14564.28952
Count	25000

In comparing the situation 5 posterior distribution with the transformed beta (7.5,

4.73) prior distribution, the mean of the prior, 0.729 is similar to the posterior median

0.727 and falls between the median and mode. Also the variance is similar at 0.645 for

the prior distribution and 0.689 for the posterior. The skew value for the prior is -0.24 and

the skew value for the posterior shows more negative skew at -0.69, while the adjusted

negative skew from normal aided the accuracy of θ parameter estimation in situation 5

the researcher believes that using a moderate to extreme skew value in the prior might

produce a more leptokurtic distribution (as shown by the posterior).A more leptokurtic
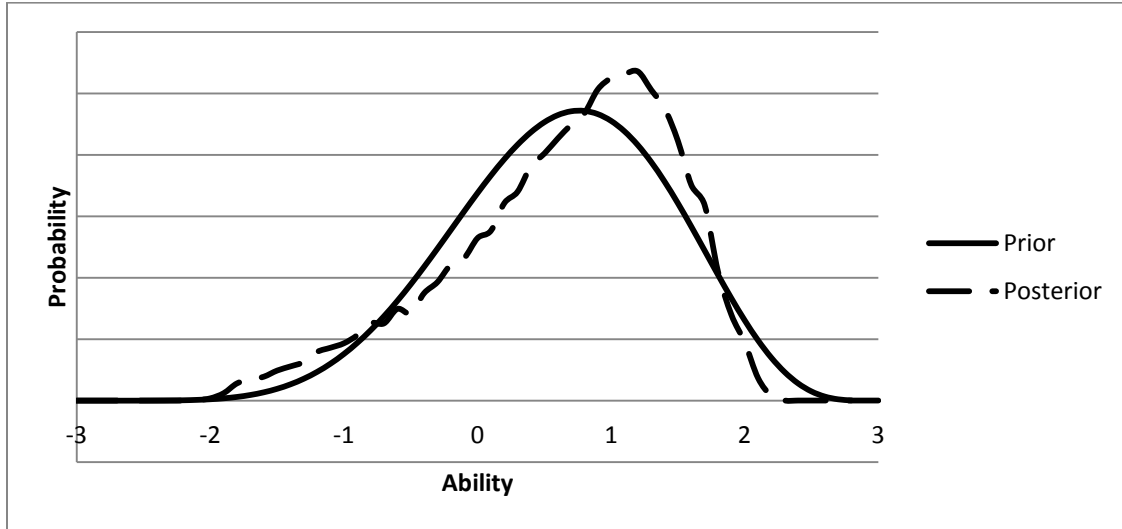
Figure 9. Bayesian analysis for Favorable condition, situation 6 – prior distribution beta (7.5, 4.73) informed by posterior distribution situation 5.

prior distribution may affect the accuracy of the estimation by the mean's proximity to the upper limit of the normal ability range (3) and latent person ability parameters may not be estimated as accurately as it was in the Baseline condition. Therefore, the skew value of -0.24 found in the beta (7.5, 4.73) distribution is deemed sufficient for the prior distribution. Likewise the excess kurtosis of the posterior distribution is nearly 0 at -0.033 and the beta (7.5, 4.73) distribution excess kurtosis is -0.31 which is more platykurtic, and therefore more acceptable in a prior distribution.

The fit statistic has very little information to contribute to selecting a suitable fit for a prior distribution with the fit statistic over 3000 in the process; however, the fit statistic did decrease when the transformed beta (7.5, 4.73) distribution for the central binned range of -1.5 to 1.5. Because the fit statistic has not been a productive analysis in the selecting of the subsequent prior, the use of the fit statistic is discontinued in this condition.

The Favorable to RBA condition, situation 6 prior distribution is beta (7.5, 4.73) for WinBUGS and N (0, 1) (the default) for BILOG-MG. The RMSE when estimating the latent ability or trait ranges from 0.71 to 0.85 for BILOG-MG and from 0.32 to 0.48 for WinBUGS across the 50 replications. The means of the RMSE for θ estimation are not similar with 0.79 for BILOG-MG and 0.37 for WinBUGS, which indicates more precision for WinBUGS; in this case the accuracy of the θ estimation indicates a 95% credible interval for WinBUGS at ± 0.73 and the BILOG-MG 95% credible interval at ±1.57. In this simulation, the accuracy in estimating θ is improved by using an informative prior in the Favorable to RBA condition, situation 6.

The simulated sample for situation 6 is drawn from a negative skewed distribution with mean = 1.0, standard deviation = 1, skew = -1.2, and excess kurtosis = 1.4. The dichotomous response matrix is also generated from the comparison of the probability of a correct answer with a random uniform variable and the total simulated sample of true θ values for 25,000 users (500 users and 50 replications) has a total mean of 0.63; however in a skewed distribution the median (0.87) is a better indicator of central tendency. The standard deviation of the simulated sample is 0.60 and the total skew equal to -2.28, which has an extreme negative skew. Figure 10 shows the prior, latent trait, and posterior distributions for situation 6 in the Favorable to RBA condition. In this final situation of the Favorable to RBA condition, the latent trait distribution has an extreme negative skew and the both the posterior and prior have negative skew. The latent trait (through the likelihood) influences the posterior to be much more leptokurtic than the prior and the prior influences the posterior to moderate to a less leptokurtic distribution.

Overall, in the Favorable to RBA condition, the RSME values for θ estimation were consistent between WinBUGS and BILOG-MG in situation 4 with both programs using non-informative priors. When WinBUGS was used with informative priors in situations 5 and 6, the RMSE values for θ estimation decreased to nominal levels
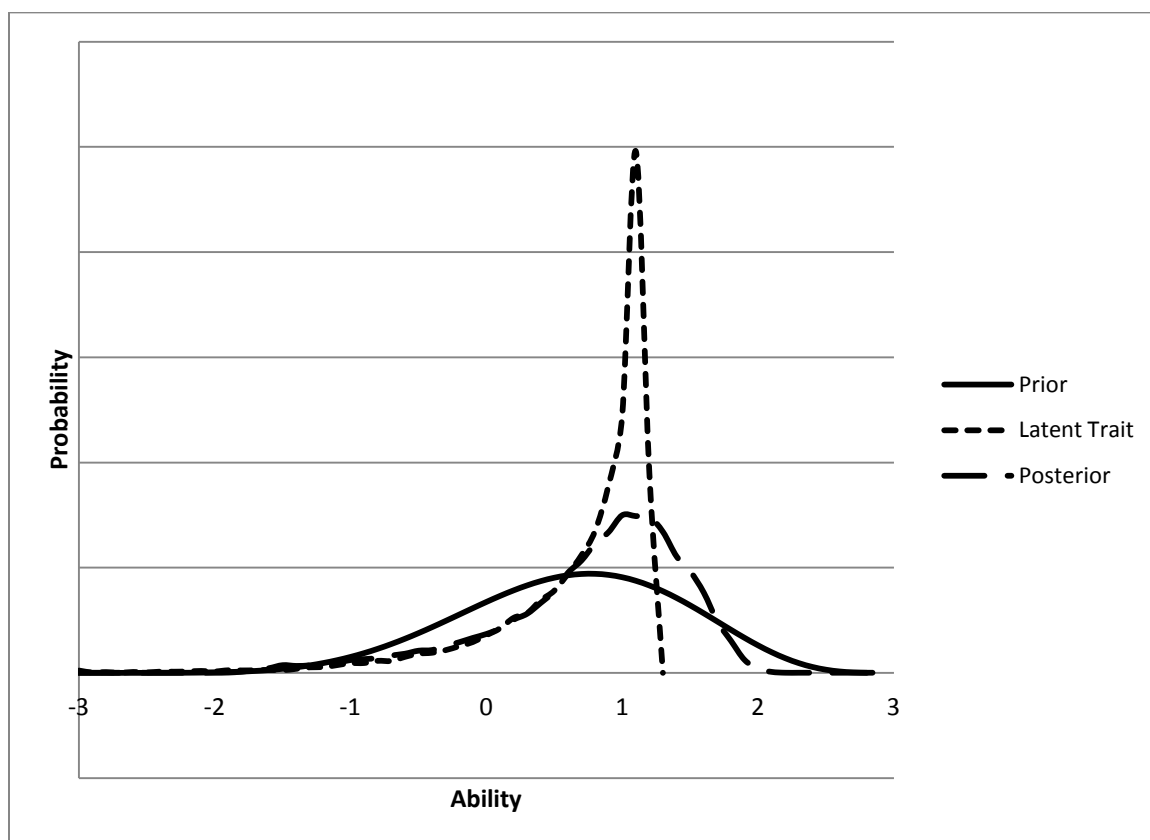
Figure 10. Bayesian analysis for Favorable condition, situation 6 – using informative prior beta (7.5, 4.73).

similar to those RMSE values previously shown in the Baseline Condition, while the RMSE values for θ estimation using BILOG-MG with non-informative prior increased as the negative skew of the sample increased. Thus, the use of RBA with WinBUGS was more accurate at θ estimation in the Favorable to RBA condition.

Though the accuracy of estimating the item parameters, *a* and *b*, are not the focus of this study, item parameter estimates by BILOG-MG and WinBUGS are compared with the true item parameters (Appendix C) using RMSE. The findings of this study show that in the Favorable to RBA condition the *a*, or discrimination parameter estimates are similar for BILOG-MG and WinBUGS as shown in table 5. The use of the informative prior with WinBUGS marginally improves the estimation of the *a* parameter over BILOG-MG as the sample's negative skew increase in situation 6. The *b*, or difficulty parameter estimates in the Favorable to RBA condition from both BILOG-MG and WinBUGS are larger in situation 4 with a RMSE of 0.62 and 0.42 respectively. However, the situation 5 and 6 *b* parameter estimation RMSEs for WinBUGS decrease to baseline levels (0.30 and 0.25), while the *b* parameter estimation RMSEs for BILOG-MG show a marked increase (0.87 and 1.16) over the same two situations. The default priors of the discrimination and difficulty parameters for BILOG-MG and the priors for the discrimination and difficulty parameters set in WinBUGS are LN(0, 0.5) and N(0, 2) respectively for all situations within the Favorable to RBA condition. The θ prior distribution is the normal default N (0,1) for all situation in BILOG-MG and N (0, 1), beta (5.4, 3.8), beta (7.5, 4.73) for WinBUGS in situations 4, 5, and 6, respectively.

<div align="center">**Unfavorable to RBA Condition Results**</div>

Each program, BILOG-MG and WinBUGS, use non-informative priors in the initial occasion of the Unfavorable to RBA condition. The RMSE when estimating the latent ability or trait ranges from 0.33 to 0.39 for BILOG-MG and from 0.32 to 0.39 for WinBUGS across the 50 replications. The means of the RMSE for θ estimation are also similar with 0.36 for BILOG-MG and 0.36 for WinBUGS (see Table 8), which indicates

that the accuracy of the θ estimation is virtually the same in situation 7. Both programs

generate similarly accurate θ estimates to the RMSE calculations as previously shown in

the Baseline Condition. The sample is drawn from a normal distribution with mean = 0,

standard deviation = 1, skew = 0, and excess kurtosis = 0, which is the same as the

sample distribution shown in the Baseline Condition. The dichotomous response matrix is

also generated from the comparison of the probability of a correct answer with a random

uniform variable and each simulated true θ value for each sample of 500 users has a mean

Table 8

Root Mean Square Error (RMSE) Calculations for Unfavorable to RBA Condition

Estimated Parameter	BILOG-MG	WinBUGS
Situation 7 (500 × 50)		
Θ	0.36	0.36
A	0.15	0.19
B	0.22	0.30
Situation 8 (500 × 50)		
Θ	0.79	0.66
A	0.32	0.34
B	1.14	0.48
Situation 9 (500 × 50)		
Θ	0.80	1.48
A	0.29	0.23
B	0.95	1.10

of 0.002. The standard deviation of the simulated sample is 1 and the total skew is equal

to 0.04. Figure 11 shows the prior, latent trait, and posterior distributions for situation 7

in the Unfavorable to RBA condition. The figure indicates that when the prior is normal

and closely fits the latent trait distribution, the posterior is also very closely fit. Though a

graphic is not included for all situations in the Baseline Condition, situation 7 in the

Unfavorable to RBA condition indicates that the when the prior distribution and latent

trait distribution match there is small indication of RSME for both programs (0.36) as

shown in Table 8 situation 7 for the θ estimation. In analyzing the situation 7 posterior

distribution to inform the situation 8 prior, the characteristics are shown in Table 9.
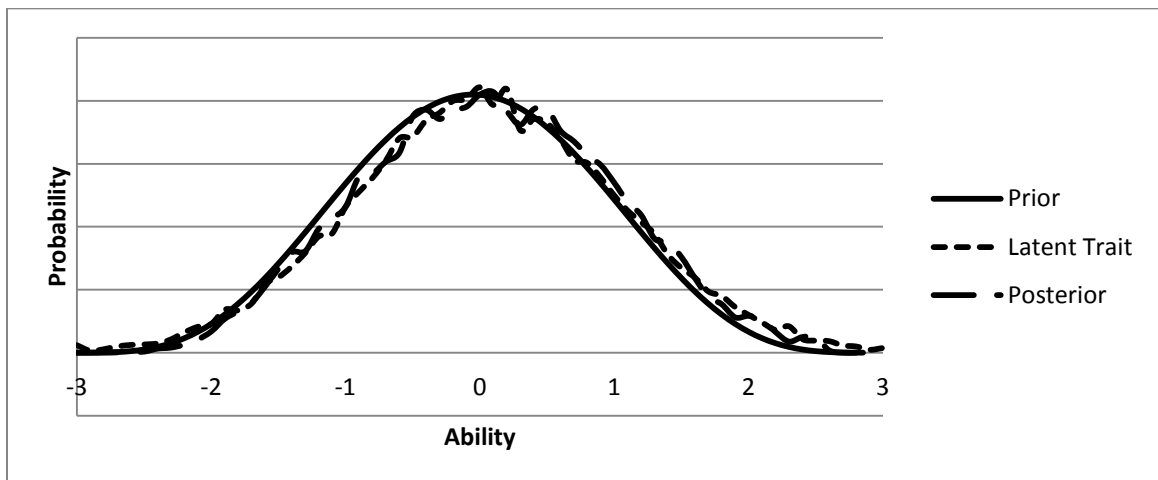
Figure 11. Bayesian analysis for Unfavorable condition, situation 7 – using non-
informative prior N(0,1).

Table 9 shows the first four central moment values of the Unfavorable to RBA

condition, situation 7 posterior distribution of the total sample size across all replications

$(500 \times 50 = 25,000)$. The transformed beta distribution α shape parameter is assigned the

value of 4.63 (because an α value of 4.63 creates a beta distribution with similar shape to

the posterior distribution) and the β parameter is calculated using formula 21. The

Table 9

Characteristics of Posterior Distribution Unfavorable Condition, Situation 7

Mean	0.011014769
Standard Error	0.005874297
Median	0.004529
Mode	1.048
Standard Deviation	0.928807967
Sample Variance	0.86268424
Excess kurtosis	-0.438219406
Skewness	0.046620227
Range	5.388
Minimum	-2.854
Maximum	2.534
Sum	275.3692247
Count	25000

mean represents the central tendency because a normal distribution is used for the latent

trait. Substituting the values for the min, max, α, and setting μ equal to the mean (0.006)

the β value is 4.61 producing a transformed distribution beta (4.63, 4.61) which has a

mean of 0.06, variance of 0.88, skew of -0.002, and excess kurtosis of -0.49. The mean,

variance, skew, and kurtosis of the transformed beta prior distribution are similar to the

corresponding central moment values of the posterior distribution shown in table 9.

The Pearson χ^2 statistic that enables comparison of the posterior distribution with

the transformed beta prior distribution is 7.23 for the central seven bins. This χ^2 statistic

has 6 degrees of freedom and is lower than the critical value of 12.59. However, as in

prior usage, when the additional three bins on either side of the central portion of the

distribution are added, the 13 equivalent intervals indicate a statistic of $\chi^2(12)=80.23$ and

the critical value of 21.03 which indicates that the shape of tails of the distribution

include points that are outside of the 95% confidence interval in comparing the two

distributions for fit.

The posterior distribution is shown in graphical format in Figure 12, with the

transformed beta (4.63, 4.61) distribution superimposed. The process of informing
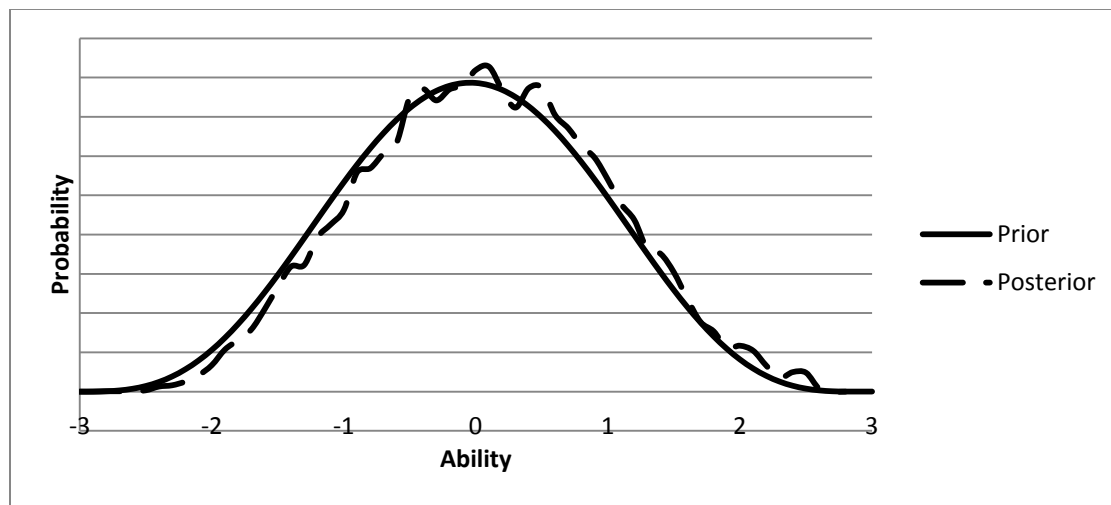
Figure 12. Bayesian analysis for Unfavorable condition, posterior situation 7 – informing prior beta (4.63, 4.61) distribution, situation 2.

this prior for the Unfavorable to RBA condition, situation 8 is similar to the informative

prior formation throughout this simulation: A beta distribution is matched to the first four

central moments of the situation 7 posterior distribution. The difference from the

Favorable to RBA condition, situation 7 is that the latent trait in the Unfavorable to RBA

condition, situation 7 is normally distributed (shown in figure 11). Thus, the posterior and

prior are virtually symmetrical with the mean of the posterior at 0.011 and the mean of

the transformed beta (4.63, 4.61) prior equal to 0.06. The variance is also similar at 0.86

and 0.88 for the posterior and prior, respectively. The skew value and kurtosis are also

similar with the skew value in both cases being near 0 and the excess kurtosis is -0.44 for

the posterior and -0.49 for the prior. Therefore, the transformed beta (4.63, 4.61) prior

will be used in the Unfavorable to RBA condition, situation 8.

The Unfavorable to RBA condition, situation 8 prior distribution is beta (4.63,

4.61) for WinBUGS and N(0,1) (the default) for BILOG-MG. The RMSE when

estimating the latent ability or trait ranges from 0.74 to 0.83 for BILOG-MG and from

0.59 to 1.08 for WinBUGS across the 50 replications. The means of the RMSE for θ

estimation are similar with 0.79 for BILOG-MG and 0.66 for WinBUGS, which indicates

slightly more precision for WinBUGS; however, in this case the accuracy of the θ

estimation indicates a 95% credible interval for WinBUGS at $\pm$ 1.32 and the BILOG-MG

95% credible interval at ± 1.58. In this simulation, the accuracy in estimating θ is similar

for WinBUGS and BILOG-MG in the Unfavorable to RBA condition, situation 8.

The simulated sample for situation 8 is drawn from an extreme negative skewed

distribution with mean = 1, standard deviation = 1, skew = -1.2, and excess kurtosis =

1.4. The dichotomous response matrix is also generated from the comparison of the

probability of a correct answer with a random uniform variable and the total simulated

sample of true θ values for 25,000 users (500 users and 50 replications) has a total mean

of 0.63; however in a skewed distribution the median (0.86) is a better indicator of central

tendency. The standard deviation of the simulated sample is 0.60 and the total skew equal

to -2.25, which is extremely skewed. Figure 13 shows the prior, latent trait, and posterior

distributions for situation 8 in the Unfavorable to RBA condition. The graph indicates

that the prior distribution is near normal and is not similar to the latent trait distribution

and the posterior clearly shows the influence of the prior distribution in that it has

characteristics of both distributions. Situation 8 in the Unfavorable to RBA condition

indicates that the when the near normal prior distribution is used to predict an extremely negative skewed latent trait distribution there is a increase of RSME (or decrease in accuracy of estimation) for both BILOG-MG at 0.79 and WinBUGS at 0.66 as shown in table 7, situation 8 for the θ estimation. The range across 50 replications is
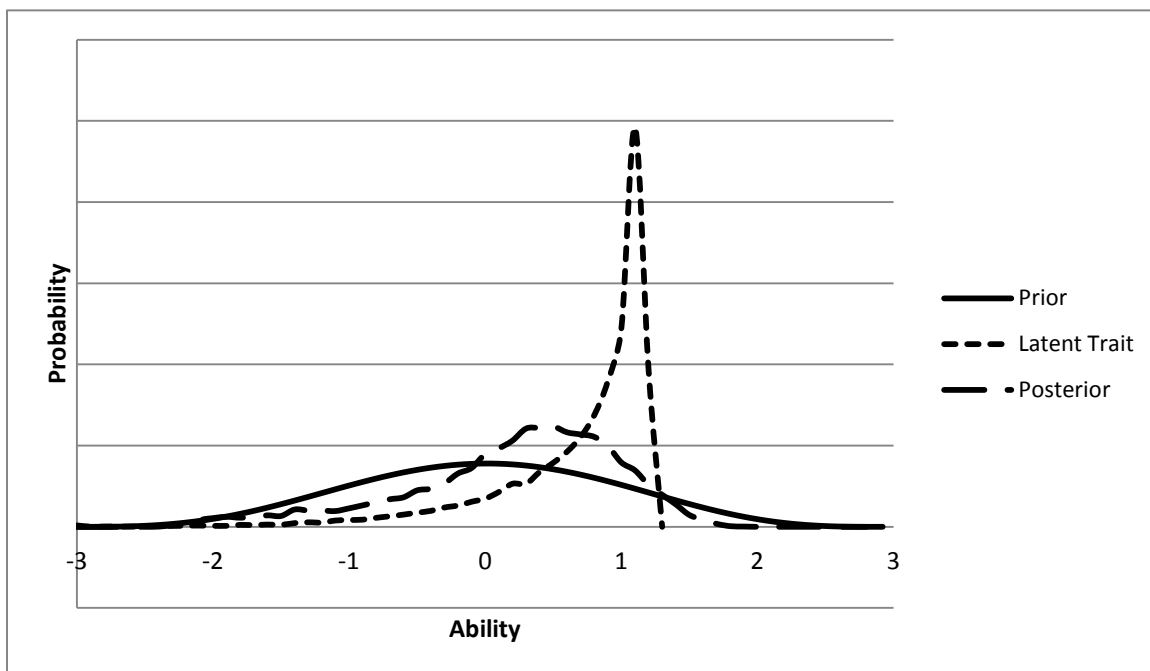
Figure 13. Bayesian analysis for Unfavorable condition, situation 8 – using informative prior beta (4.63, 4.61).

smaller for BILOG-MG (0.74 to 0.83) than for WinBUGS (0.59 to 1.08); however, because the *θ* priors for both BILOG-MG, N(0,1) and WinBUGS, beta (4.63, 4.61) are similar, the accuracy of the *θ* estimation was also similar. In analyzing the situation 8 posterior distribution to inform the situation 9 prior, the characteristics are shown in Table 10.

Table 10 shows the first four central moment values of the Unfavorable to RBA condition, situation 8 posterior distribution of the total sample size across all replications (500x50=25,000). The transformed beta distribution α shape parameter is assigned the

value of 9.3 (because an α value of 9.3 creates a beta distribution with similar shape to

the posterior distribution) and the β parameter is calculated using formula 21. Because

situation 8 in the Unfavorable to RBA condition uses an extreme negative skewed

distribution for the latent variable, the resulting posterior distribution is skewed (-0.84),

therefore the central tendency is represented by the median or mode. Substituting the

values for the min, max, α, and setting μ equal to the median (0.29) the β value is 7.84

producing a transformed distribution beta (9.3, 7.84). That beta (9.3, 7.84) distribution

Table 10

Characteristics of Posterior Distribution Unfavorable Condition, Situation 8

Mean	0.152140434
Standard Error	0.004903754
Median	0.2867
Mode	1.048
Standard Deviation	0.775351599
Sample Variance	0.601170101
Excess kurtosis	0.428286887
Skewness	-0.839853072
Range	4.298
Minimum	-2.45
Maximum	1.848
Sum	3803.510842
Count	25000

has a mean of 0.31, variance of 0.49, skew of -0.08, and excess kurtosis of -0.29. The

mean is similar to the posterior distribution's median, but the variance, skew, and kurtosis

of the transformed beta prior distribution are not similar to the corresponding central

moment values of the posterior distribution shown in Table 10.

Using the χ^2 fit statistic is of little use due to the skew of the posterior and the resulting asymmetrical distribution. The Pearson χ^2 fit statistic is calculated to be over a value of 10,000. Therefore, the use of the χ^2 fit statistic is discontinued in this condition. The posterior distribution is shown in graphical format in Figure 14, with the transformed beta (9.3, 7.84) distribution superimposed. In figure 14, the transformed beta (9.3, 7.84)
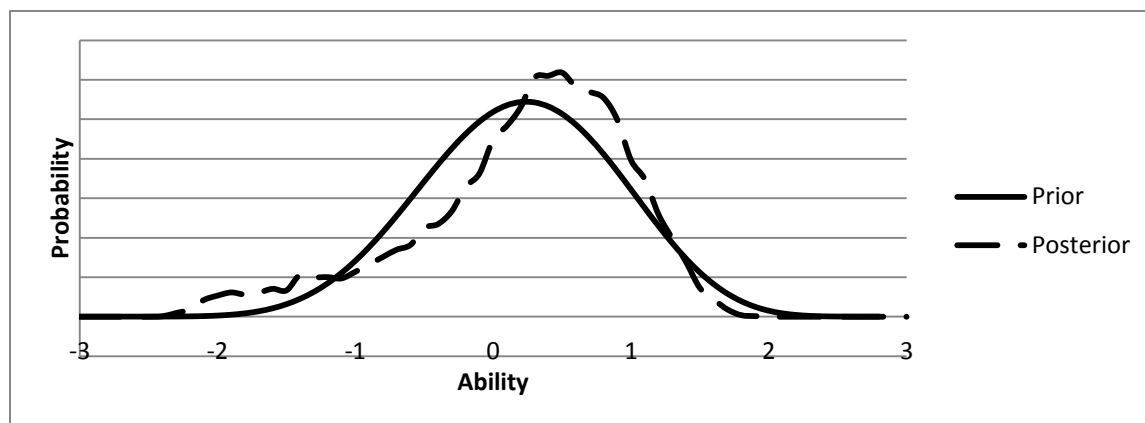
Figure 14. Bayesian analysis for Unfavorable condition, situation 9 – prior distribution beta (9.3, 7.84) informed by posterior distribution situation 8.

distribution has characteristics: μ = 0.31,   = 0.49, γ = -0.08, and *k* = -0.29. The mean of the transformed beta distribution is between the median (0.29) and mode (1.05) of the situation 8 posterior distribution which is consistent with determining the central tendency of a skewed sample. The variance is similar (0.50 and 0.60); however, the skew values are not similar (-0.11 and -0.84) and the excess kurtosis values are not similar (-0.29 and 0.42). Though these distributions indicate a similar central value for variance and an acceptable indication of mean/median comparison, the values for skew and excess kurtosis are not similar and the transformed beta distribution skew value needs adjustment.

Similar to situation 5 of the Favorable to RBA condition, the central portion of the posterior distribution in the Unfavorable to RBA, situation 8, was influenced by the non-

informative situation 7 prior and biased toward the mean of the normal distribution. This is evidenced by the elevated RSME (0.66) for θ estimation by WinBUGS, shown in table 7 for the Unfavorable to RBA condition, situation 8. The second and third central moments (variance and skew) of the posterior distribution are primarily used to inform the prior distribution because of the decreased dependence on the mean as an indicator of central tendency in a skewed distribution.

Since the posterior distribution is influenced by the non-informative prior in situation 7, the posterior distribution's skew value may indicate that the transformed beta distribution needs to increase the negative skew. Typically, the negative skew is increased by reducing the β shape parameter in a beta distribution; therefore, the informative prior is adjusted to beta (9.3, 5), thus reducing the β shape parameter from the previous beta distribution from 7.84 to 5. The transformed beta (9.3, 5) distribution has characteristics: $\mu = 0.95$, $= 0.53$, $\gamma = -0.30$, and $k = -0.22$. With this new distribution the mean is still between the median (0.29) and mode (1.05) of the posterior distribution, the variance is consistent, the skew values are closer, but not similar, and the excess kurtosis is relatively unchanged. Since a slightly more platykurtic prior distribution is desirable, the excess kurtosis of -0.22 is acceptable.

In evaluating the skew, the moderate to highly skewed (-0.83) posterior influences a change in the transformed beta skew value to -0.30. Adding additional negative skew will increase the mean indicated by the beta to be near the extremes of central tendency. Therefore, the increase in mean (0.95) caused by the adjustment of negative skew is limited to the 0.29 to 1.05 (median to mode) range. The transformed beta (9.3, 5) distribution superimposed upon the posterior distribution from situation 8 is shown in

graphic form in Figure 15. This procedure of using the skew value of the situation 8

posterior to inform the formation of the situation 9 prior is consistent with matching the

two distributions using the first four central moments of the posterior matched with the

first four central moments of the transformed beta prior.
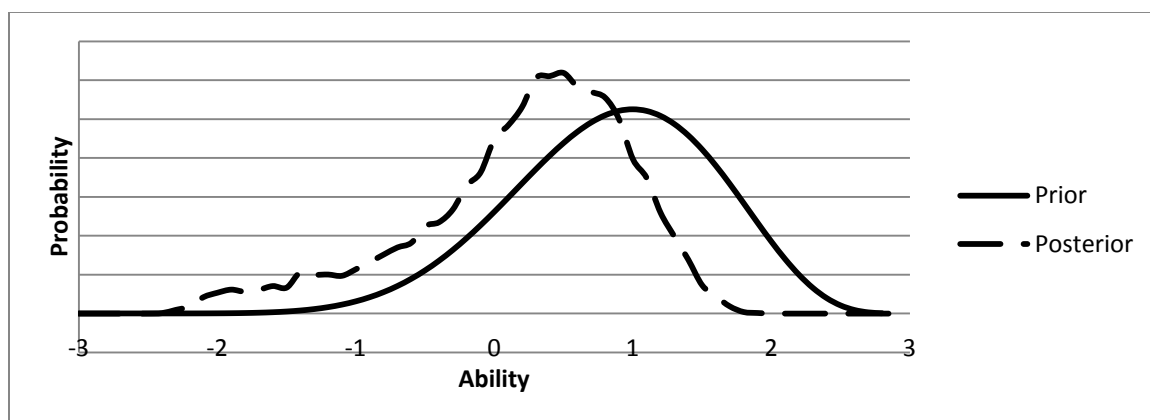
Figure 15. Bayesian analysis for Unfavorable condition, situation 9 – prior distribution beta (9.3, 5) informed by posterior distribution situation 8.

The Unfavorable to RBA condition, situation 9 prior distribution is beta (9.3, 5)

for WinBUGS and N (0, 1) (the default) for BILOG-MG. The RMSE when estimating

the latent ability or trait ranges from 0.73 to 0.86 for BILOG-MG and from 1.40 to 1.55

for WinBUGS across the 50 replications. The means of the RMSE for θ estimation are

not similar with 0.80 for BILOG-MG and 1.48 for WinBUGS, which indicates more

precision for BILOG-MG; however, in this case the accuracy of the θ estimation

indicates a 95% credible interval for WinBUGS at $\pm$ 2.96 and the BILOG-MG 95%

credible interval at ±1.60. In this simulation, the accuracy in estimating θ is poor for

WinBUGS and BILOG-MG in the Unfavorable to RBA condition, situation 9 with

WinBUGS estimates being influenced by the informative prior beta (9.3, 5). In this

situation, the prior distribution is skewed in the opposite direction from the latent trait distribution which produces a greater RMSE in the MCMC estimation of the latent ability or trait.

The simulated sample for situation 9 is drawn from an extreme positive skewed distribution with mean = -1, standard deviation = 1, skew = 1.2, and excess kurtosis = 1.4 as shown in chapter 3, Table 2. The dichotomous response matrix is also generated from the comparison of the probability of a correct answer with a random uniform variable and the total simulated sample of true θ values for 25,000 users (500 users and 50 replications) has a total mean of -0.64; however in a skewed distribution the median (-0.87) is a better indicator of central tendency. The standard deviation of the simulated sample is 0.60 and the total skew equal to 2.32, which is extremely skewed. Figure 16 shows the prior, latent trait, and posterior distributions for situation 9 in the Unfavorable to RBA condition.

Figure 16 indicates that the prior distribution has a negative skew and is not similar to the latent trait distribution, which has an extreme positive skew. The posterior clearly shows the influence of the prior distribution in that it is biased toward the mean of the prior distribution. Situation 9 in the Unfavorable to RBA condition indicates that when the negative skewed prior distribution is used to predict θ values from an extremely positive skewed latent trait distribution there is an increase of RSME (or decrease in accuracy of estimation) for WinBUGS at 1.48 as shown in table 7 situation 9
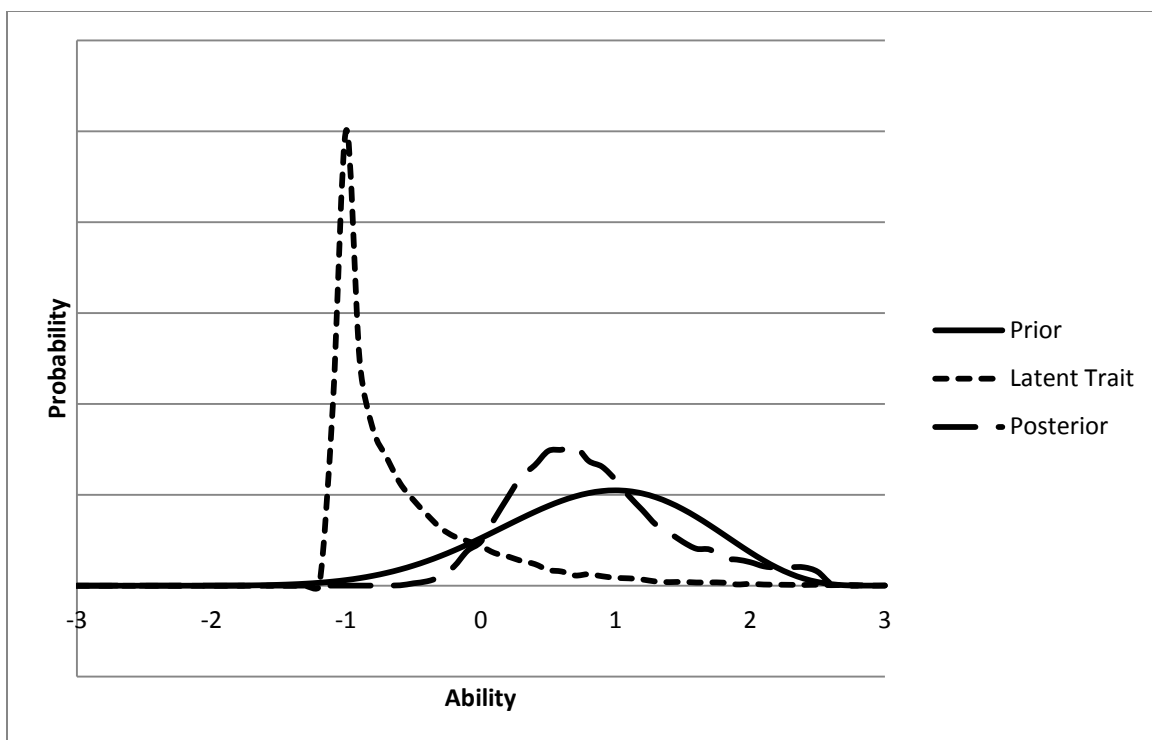
Figure 16. Bayesian analysis for Unfavorable condition, situation 9 – using informative prior beta (9.3, 5).

for the θ estimation. The RMSE range across 50 replications in situations 8 and 9 are similar for BILOG-MG (0.74 to 0.83 and 0.73 to 0.86) because the normal non-informative prior is maintained, N(0, 1), as the latent trait distribution changes from an extreme negative skew to an extreme positive skew. Though the RSME remains similar for the BILOG-MG estimation of θ, the accuracy of that estimation is poor at approximately ±1.60 for situations 8 and 9. The application of using an informative prior distribution for WinBUGS is the least accurate in situation 9. The case of using an informative prior distribution that has a negative skew with a latent trait distribution that has extreme positive skew produced an RMSE value of 1.48 (±2.96) across 25,000 simulated users.

Though the accuracy of estimating the item parameters, *a* and *b*, are not the focus of this study, item parameter estimates by BILOG-MG and WinBUGS are compared with

the true item parameters (Appendix C) using RMSE. The findings of this study show that

in the Unfavorable to RBA condition the *a*, or discrimination parameter estimates are

similar for BILOG-MG and WinBUGS within each situation as shown in table 7. The *b*,

or difficulty parameter estimates in the Unfavorable to RBA condition from both BILOG-

MG and WinBUGS are larger in situations 8 and 9 with a RMSE of 1.14, 0.95 and 0.48,

1.10 respectively. However, the situation 7 *b* parameter estimation RMSEs for both

BILOG-MG and WinBUGS decreased to baseline levels (0.22 and 0.30). The default

priors of the discrimination and difficulty parameters for BILOG-MG and the priors for

the discrimination and difficulty parameters set in WinBUGS are LN(0, 0.5) and N(0, 2)

respectively for all situations within the Unfavorable to RBA condition. The θ prior is the

normal default N(0, 1) for all situations in BILOG-MG and N(0, 1), beta (4.63, 4.61),

beta (9.3, 5) for WinBUGS in situations 7, 8, and 9 respectively.

In summary, there were three conditions with three occasions that describe the

situations within each condition in the simulation (3 x 3 = 9). The RSME is calculated by

comparing the estimated θ parameters to the true θ parameters of the sample of 500

simulated users over 50 replications, which total 25,000 estimations of the θ person

parameter. Also estimated *a* and *b* item parameters are compared to the true item

parameters using RSME.

The Baseline condition uses a normal distribution for the latent trait and prior

distributions in each situation. The analyses result in a normal distribution for the

posterior in the case of both BILOG-MG and WinBUGS. The RSME calculations for θ

estimation for both programs across all situations (3) are consistently at the baseline level

of 0.35 (as seen in Table 3), which is determined as the baseline accuracy level for this simulation.

In the Favorable to RBA condition, the RMSE values are less accurate than baseline values in situation 4 (BILOG-MG = 0.61 and WinBUGS = 0.55), but near baseline values for WinBUGS in the subsequent two situations (0.36, 0.37). While θ estimation accuracy improved for WinBUGS across the three occasions of the Favorable to RBA condition, θ estimates using BILOG-MG showed increasing RMSE (0.61, 0.76, 0.79) in the three occasions respectively, which indicates a substantial loss of accuracy for BILOG-MG using non-informative priors as the latent trait distribution became progressively more skewed. In fact, the slight negative skew of the latent trait in situation 4 produced θ estimates using BILOG-MG that have a 95% credible interval of ±1.22 on a -3 to 3 scale.

In the Unfavorable to RBA condition, situation 7 repeats the Baseline Conditions and the θ estimation results are very similar for both BILOG-MG and WinBUGS (BILOG-MG = 0.36 and WinBUGS = 0.36). For the subsequent two situations the latent trait distributions are extremely negative skewed and extremely positive skewed. In these situations, RMSE values of BILOG-MG θ estimates are consistent at 0.79 and 0.80, which produce 95% credible intervals of about ±1.6. WinBUGS using informative priors in a RBA has an RMSE of 0.66 in situation 8 and 1.48 in situation 9 showing that an informative prior distribution that is extremely different from the latent trait distribution yields a θ estimation that is very poor in accuracy with a 95% credible interval of ±2.96 on a -3 to 3 scale.

CHAPTER 5

CONCLUSIONS

As stated this dissertation sets three conditions: Baseline, Favorable to Recursive Bayesian Analysis (RBA), and Unfavorable to RBA. Within each condition are three occasions denoted as occasion 1, occasion 2, and occasion 3. There are nine situations which represent the intersections of the conditions and occasions. The latent trait, θ, is normally distributed in all three situations in the Baseline condition. In the Favorable to RBA condition, θ is distributed with a negative skew that progresses from -0.40 in the situation 4, -0.80 in situation 5, to -1.20 in situation 6. This condition is Favorable to RBA because the skew is consistently negative or in the same direction. The Unfavorable to RBA condition starts with θ distributed normally in situation 7, then extreme negative skew in situation 8 and extreme positive skew in situation 9. This condition is unfavorable because the θ distribution changes from extreme negative to extreme positive skew in adjacent situations. A chart showing the three conditions and occasions with the accompanying true θ distributions, prior distributions, and RMSE values is in figure 17.

A SAS^{TM} simulation generating 500 users over 50 replications is used for each situation. RBA and traditional analysis were compared using the same 40 item instrument in each situation. RBA is employed to estimate θ using WinBUGS in each situation and traditional analysis is employed to estimate θ using BILOG-MG in each situation. These programs were compared over three conditions and in three occasions within each condition. In each situation, BILOG-MG and WinBUGS used the same dichotomous test data. BILOG-MG used a non-informative prior for the θ estimation in each situation and

WinBUGS used a non-informative prior in the first occasion of each condition and an
informative prior in occasions 2 and 3 in each condition. In total there were 9 situations,

	Occasion 1		Replace Poor Items	Occasion 2		Replace Poor Item	Occasion 3	
Test Length All Conditions	40 Items		Replace Poor Items	40 Items		Replace Poor Item	40 Items	
	BILOG-MG	WinBUGS		BILOG-MG	WinBUGS		BILOG-MG	WinBUGS
Baseline	Situation 1			Situation 2			Situation 3	
Latent Trait	normal	normal		normal	normal		normal	normal
Prior	default	default		default	N(0.009,0.93)		default	N(0.016, 0.96)
(RSME)	(0.36)	(0.35)		(0.37)	(0.36)		(0.36)	(0.35)
Favorable	Situation 4 Skew	Skew		Situation 5 Skew	Skew		Situation 6 Skew	Skew
Latent Trait (-)	(-0.40)	(-0.40)		(-0.80)	(-0.80)		(-1.20)	(-1.20)
Prior	default	default		default	Beta(5.4, 3.8)		default	Beta(7.5,4.73)
(RSME)	(0.61)	(0.55)		(0.76)	(0.36)		(0.79)	(0.37)
Unfavorable	Situation 7			Situation 8 Skew	Skew		Situation 9 Skew	Skew
Latent Trait (- +)	normal	normal		(-1.20)	(-1.20)		(+1.20)	(+1.20)
Prior	default	default		default	Beta(4.63,4.61)		default	Beta(9.3, 5)
(RSME)	(0.36)	(0.36)		(0.79)	(0.66)		(0.80)	(1.48)

TIME →

Figure 17. Latent Trait and Prior Distributions for All Conditions and Occasions with RMSE Indicated for Each Situation.

(3 x 3 = 9) in which the estimation of θ by BILOG-MG and WinBUGS were compared
with the true simulated θ value.

The Baseline was the first condition in which there was little difference between
RBA and traditional analysis in estimation of θ. In the three situations of the Baseline the
RMSE found in comparing the θ estimates to the true θ were consistently at about 0.35
for both programs, indicating that using RBA did not increase the accuracy of estimation
of θ when the latent trait had a normal distribution. The Baseline condition also indicated

the RSME value (0.35) in this simulation for the estimation of θ when the latent trait distribution is normal.

The Favorable to RBA or second condition examined the accuracy of θ estimation when the latent trait distribution has a progressive negative skew across three occasions. In situation 4 of the Favorable to RBA condition, both BILOG-MG and WinBUGS used a non-informative prior and had a slightly negative skew (-0.40) for the latent trait. The RMSE of the θ estimation compared to the true θ value increased to 0.61 and 0.55 for BILOG-MG and WinBUGS, respectively. The increase in RMSE from 0.35 to 0.61 indicates a decrease in the precision of the θ estimate. The increase in RMSE continued across situations 5 and 6 for the Favorable to RBA for BILOG-MG with values of 0.76 and 0.79, respectively, indicating that as the skew of the latent trait increased the RMSE also increased. In contrast, WinBUGS using an informed prior through RBA decreased the RMSE to 0.36 and 0.37 in situations 5 and 6 when a progressively negative skew latent trait distribution was used. This RMSE value is near the previously measured baseline level and improves the accuracy of estimation by roughly a factor of two (0.76 to 0.36).This finding supports the alternative hypothesis that using RBA increases the accuracy of θ estimation in WinBUGS when the latent trait has a progressively negative skew distribution. Based on the simulation, this finding supports using RBA when the latent trait is skewed.

The Unfavorable to RBA or third condition was the final condition of the simulation. In this condition the researcher examined the accuracy of θ estimation when the latent trait distribution has a normal distribution as a baseline condition in situation 7, followed by an extreme negative skew in situation 8, and followed by an extreme positive

skew in situation 9. Though this is an extreme simulation and not often seen in actual practice, it is valuable to test the extreme cases to try and determine the θ estimation related to an informative prior when the latent trait distribution is dissimilar. In situation 7 both BILOG-MG and WinBUGS used a non-informative prior and a normal distribution for the latent trait. The RMSE of the θ estimation compared to the true θ value was 0.36 for both BILOG-MG and WinBUGS, which similar to the Baseline condition previously measured. BILOG-MG using an uninformative prior through traditional analysis maintained a RMSE of 0.79 and 0.80 in situation 8 and 9 with an extreme change from a negative skew to a positive skew latent trait distribution. The RMSE value increased across situations 8 and 9 for WinBUGS with values of 0.66 and 1.48. The extreme changes in the latent trait distribution produced a negative skew informed prior distribution that was actually opposite the extreme positive skew of the latent trait distribution. In the Unfavorable to RBA condition, the use of RBA resulted in decreased accuracy in the θ estimation when there were extreme changes in the latent trait distribution. This finding supports the null hypothesis for the research question because using RBA decreased the accuracy of θ estimation when the latent trait had an extreme change from a negative skew to a positive skew distribution. The researcher also maintains that the extreme change simulated in the Unfavorable to RBA condition is believed to be very rare when real data are analyzed rather than simulation data.

Applications

The analysis of skewed distributions may be productive in particular situations (Samejima, 1997, 2000). The Georgia education system places great emphasis on Criterion-Referenced Competency Tests (CRCT) to indicate student achievement,

instructional quality, and school efficacy (Georgia Department of Education; Governor's

Office of Student Achievement, 2013). In the State of Georgia, students in middle

schools who meet or exceed the standards on the CRCT , total 96.8% of students in

reading, 94.3% of students in English /language arts, and 83% of students in mathematics

(Georgia Department of Education, 2013). The CRCT results are not only used for

student achievement indicators, but the College and Career Ready Performance Index

(CCRPI) uses the CRCT mean in all subjects to assign a point value to the school. CRCT

means that meet the state target for meets or exceeds numbers will earn the school up to

70 CCRPI points which is the maximum number of points for student achievement

(Georgia Department of Education; Governor's Office of Student Achievement, 2013).

For example, a distribution of the CCRPI scores for all middle schools within Georgia is

shown in Figure 18. The extreme negative skew distribution of the graph shows the

CCRPI middle school scores and may be a rough indication of the ability distribution of

that group (Samejima, 1997). This actual student distribution mirrors the negative skew

distribution used in the simulation to represent the latent trait. The simulation described

in this dissertation indicates that when a skewed distribution is examined through IRT

processes, θ estimation is less accurate when a non-informative prior is used by either

BILOG-MG or WinBUGS as evidenced by an RMSE of 0.61 and 0.55 in the Favorable

to RBA condition, situation 4, respectively. Seen in the subsequent situations (5 and 6) of

the Favorable to RBA condition, using RBA produces lower RMSE values (0.37 and

0.36) yielding more accurate θ estimation than using a non-informative prior (0.66 and

0.72). If the latent trait distribution in subsequent situations is similar to or is skewed in

the same direction consistently, then an informative prior which is a transformed beta

distribution informed by the previous posterior distribution improves the θ estimation accuracy roughly by a factor of two. This process may be applicable to determining
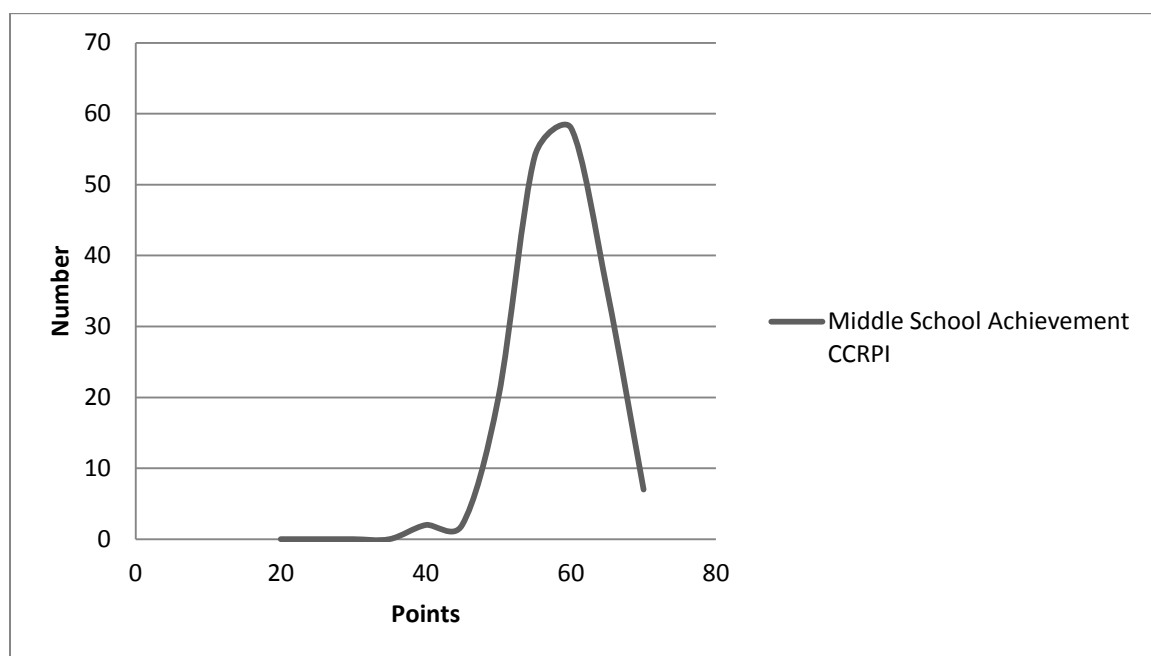
Figure 18. Middle school achievement distribution CCRPI.

the latent trait associated with the CRCT tests, which likely produce a negative skew in the latent trait distribution.

Rather than examining the assessment or items within the assessment similar to Bazan, Bronco, and Bolfarine (2003) in determining skewness, the emphasis of this dissertation is on the assessed population which is saturated with instruction specific to the assessment questions in the time prior to the assessment. More specifically, this dissertation examines an application of RBA typically used in target tracking, robot geo-location, or navigation in which the posterior informs the subsequent prior. Typically in Bayesian analysis, estimation of θ is based upon the prior belief about the distribution of the assessed population multiplied by the likelihood or the observed scores and the posterior is proportional to that product. A similar process is completed in the RBA in

which the process is recursive and the posterior informs the subsequent prior. This

recursive process is used when there are seconds or milliseconds between Bayesian

processes such as observed in real-time navigation of robotic drones; however, this

dissertation applies RBA to the IRT process which does not have similar time demands.

RBA also retains the independence of the data and does not resample, rather, the prior is

informed by the former posterior as the belief about a specific population's assessment

scores may change.

A hypothetical example of applying RBA to actual data can be demonstrated in as

follows in five steps.

Step 1: WinBUGS uses a default θ prior distribution of N(0, 1) to estimate

θ in a large population.

Step 2: The researcher analyzes the posterior distribution generated in Step

1 and defines the first four central moments of the posterior distribution. For this

example, those values can be $\mu = .04$, $\sigma^2 = 0.886$, $\gamma = -0.47$, and $k = -0.45$, where μ

is the mean, σ^2 is the variance, γ is the skew, and k is the excess kurtosis. The

median may be a better indication of central tendency because the posterior

distribution is skewed. In this case, the median is 0.53.

Step 3: The initial formation of the beta shape parameters α and β are

created. An α value of 5 is set to have a similar shape to the posterior distribution

and the β value is calculated using equation 21 substituting the median for the

mean. The result for β is 3.8 and the equations 17 through 20 are used to calculate

the first four central moments of the beta (5, 3.8) distribution. Those calculations

are $\mu = 0.46$, $\sigma2 = 0.90$, $\gamma = -0.16$, and $k = -0.47$.

Step 4: The first four moments of the posterior and transformed beta prior distributions are compared. The prior mean (0.46) is near the posterior median (0.53), the variance is similar (0.886 and 0.90), the skew is not similar (-0.47 and -0.16); however it is in the same direction, and the excess kurtosis is similar (-0.45 and -0.47). Because the skew value has a large standard deviation compared to the mean and variance, the transformed beta distribution is accepted and applied as the informed prior in the next IRT analysis.

Step 5: When the subsequent IRT analysis is complete, evaluate the accuracy of θ estimation through increased model fit compared to the Step 1 IRT analysis. A better model fit can be an indication of a more accurate θ estimation.

Limitations

A possible limitation in this dissertation is relaxing the assumption that the latent trait has a normal distribution in the Favorable and Unfavorable to RBA conditions. Typically, the distribution of the latent trait is assumed to be N (0,1) as is simulated in the Baseline condition. The negative skew distribution in the Favorable to RBA condition and extreme negative and positive skew distributions in the Unfavorable to RBA condition was manipulated to simulate latent trait distributions similar to CRCT results published by the State of Georgia. While the general practice of transforming skewed distributions to normal distributions is sometimes employed in IRT applications, the asymmetrical distribution may provide latent trait indicators that are lost in data transformation; thus, the relaxation of the normal assumption was applied in the Favorable and Unfavorable to RBA conditions.

Implications for Future Research

This is a computer simulation using SAS^{TM} programming to call WinBUGS and

BILOG-MG and the θ distribution is manipulated using Fleishman's (1978) skew

transformation and a linear transformation of the mean to simulate real CRCT data. The

latent trait distributions were simulated, therefore an extension of this study may use

actual latent trait distributions. Distributions of the latent trait typically are not known in

the actual application of the 3PL IRT model; however, the first four central moments of

the posterior distribution can be used to inform the subsequent prior even when latent

trait distributions are not observed or manipulated artificially. The accuracy of the

estimation process may be evaluated using a model fit statistic.

The use of the Pearson χ^2 fit statistic may give more productive information if a

transformed beta distribution is not used as the informative prior. It is possible that an

equation can be derived from the posterior distribution that can be used to inform the

subsequent prior. This is similar to Bayesian techniques which have been used since the

1990s in navigation and tracking applications. Additional research in this area may show

an additional increase in the accuracy of IRT parameter estimation.

The analysis of skew distributions is discussed by Samejima (1997, 2000); Bazan,

Bronco, and Bolfarine (2003); and Broccolli and Cavrini (2007) and a similar approach to

the estimation of θ within a skew distribution is used by these researchers. Samejima

(2000) introduced an "item complexity" parameter in addition to the difficulty and

discrimination parameters that is "based on a skewed conditional distribution of the item

response tendency" (p. 325). The skew is positive or negative depending on the

characteristics of the model and the assessed population. Samejima (2000) established the

Logistic Positive Exponent (LPE) family of distributions and Bolfarine and Bazan (2008) added another negative skew distribution called Reflection Logistic Positive Exponent (RLPE). Closely related to these developments, Bazan et al. (2003) and Broccoli and Cavrini (2007) also used a skew hyperparameter to examine skewed ICCs that would lead to a skewed distribution of the latent trait. Common to all of these approaches, is developing a hyperparameter at the item level that relates to the skew seen in the distribution of the latent trait. Further exploration of a Bayesian approach to gather information from the conditional distribution of the posterior may be used to inform the subsequent prior at the item level for some tests.

BILOG-MG uses Bayesian techniques and can apply informative priors that may increase the accuracy of θ estimation also. WinBUGS employs an informative prior in a different manner, therefore, an investigation of the creation and use of informative priors for BILOG-MG and any subsequent change in the θ estimation accuracy is an area that future research may examine.

This dissertation used the IRT 3PL model with RBA; however, other examples of dichotomous, polytomous, or multidimensional IRT models may be evaluated using the RBA methods if a manner to elicit information for the prior is established. This process may be straight forward with other dichotomous and polytomous models, but can be extremely complicated when working with a multidimensional IRT model. An extension of this research would be applying RBA in other IRT models.

REFERENCES

Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). Retrieved from

http://ericae.net/irt/baker

Bazan, J. L., Branco, M. D., & Bolfarine, H. (2006). A model of skew item response

theory. *Bayesian Analysis, 1*(4), 861-892.

Bolfarine, H., & Bazán, J. (2010). Bayesian estimation of the logistic positive exponent

model. *Journal of Educational and Behavior Statistics, 35*(6), 693-713.

Broccoli, S., & Cavrini, G. (2010, June) *A Bayesian IRT model for mixed responses and

skew latent variable.* Paper presented at the SIS 2010 Scientific Meeting - 45th

Scientific Meeting of the Italian Statistical Society, University of Padua, Italy.

Congdon, P. D. (2010). *Applied Bayesian hierarchical methods.* Chapman Hall/CRC,

Boca Raton, FL.

Curtis, S. M. (2010). BUGS code for item response theory. *Journal of Statistical

Software, 36*(1), Retrieved from http://www.jstatsoft.org

de Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY:

Guliford.

Fan, Xitao, Felsovalyi, A., Sivo, S.A., Keenan, S.C. (2001). *SAS^{TM} for Monte Carlo

Studies: A Guide for Quantitative Researchers*. Cary, NC: SAS Institute Inc.

Fleishman, A. I. (1978). A method for simulating non-normal distributions.

Psychometrika, 43(4), pp. 521-532.

Foley, B. P. (2010). *Improving IRT parameter estimates with small sample sizes:

Evaluating the efficacy of a new data augmentation technique.* University of

Nebraska—Lincoln: Open Access Theses and Dissertations from the College of

Education and Human Sciences. Paper 75. Retrieved from

http://digitalcommons.unl.edu/cehsdiss/75

Fox, J. P. (2010). *Bayesian item response modeling: Theory and applications.* New York: Springer.

Fox, J., & Glas, C. A. W. (2001). Bayesian estimation of a multi-level IRT model using Gibbs sampling. *Psychometrika, 66*(2), 271-288.

Fukuhara, H. (2009). *A differential item functioning model for testlet-based items using a bi-factor multidimensional item response theory model: A Bayesian approach* (Doctoral dissertation). Retrieved from ProQuest. (UMI 3373995)

Georgia Department of Education, Governor's Office of Student Achievement. (2013). *CCRPI supplementary tables & figures* (May 2013 E-Bulletin). Retrieved from GADOE website: http:/www.gadoe.org

Hambleton, R. K., Swaminathan, H., & Rogers, J. H. (1991). *Fundamentals of item response theory.* Newbury Park, CA: Sage.

Harwell, M., Stone, C. A.,Hsu,T.,Kirisci, L. (1996). Monte Carlo studies in item response theory. *Applied Psychological Measurement*, 20(2), 101-125.

Hsu, Y., Ackerman, T. A., & Fan, M. (1999). The relationship between the Bock-Aitkin procedure and the EM algorithm for IRT model estimation. *ACT Research Report Series 99*-7, 1-35.

Jones, D. H., & Nediak, M. (2000). Item parameter calibration of LSAT items using MCMC approximation of Bayes posterior distributions, *RRR 7-2000*, Piscataway, NJ: RUTCOR.

Kieftenbeld, V., & Natesan, P. (2012). Recovery of graded response model parameters: A

comparison of marginal maximum likelihood and Markov Chain Monte Carlo

estimation. *Applied Psychological Measurement, 36*(5), 399-419.

Lord, F. M. (1983). Statistical bias in maximum likelihood estimators of item parameters.

Psychometrika, 48(3), 425-435.

Liu, C., & Gambrell, J. (2009). Item parameter estimation: Does WinBUGS do better

than BILOG-MG? [Bayesian Class Presentation]. Retrieved from

http://homepage.stat.uiowa.edu/~kcowles/s138_2009/

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS: A Bayesian

modeling framework: concepts, structure, and extensibility. *Statistics and

Computing*, *10*, 325-337.

Mead, A. D., Morris, S. B., & Blitz, D. L. (2007). *Open source IRT: A comparison of

BILOG-MG and ICL features and item parameter recovery* [Unpublished

manuscript]. Chicago: Institute of Psychology, Illinois Institute of Technology.

Mislevy, R. J. (1986). Bayes modal estimation in item response models. *Psychometrika,

51*(2), 177-195.

Mislevy, R.J. & Bock, R.D. (1990). *BILOG-MG 3: Item analysis and test scoring with

binary logistic models.* Chicago: Scientific Software, Inc.

Mislevy, R. J., & Stocking, M. L. (1989). A consumer's guide to LOGIST and BILOG-

MG. *Applied Psychological Measurement, 13*, 57-75.

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of

Mathematical Psychology, 3*(1), pp. 1-18.

Patz, R. J., & Junker, B. W. (1999). A straightforward approach to Markov Chain Monte Carlo methods for item response models. *Journal of Educational and Behavioral Statistics, 24*(2), 146-178.

Proctor, T., Teo, K., Hou, J., & Hsieh, M. (2005). *Comparison of parameter recovery in a 2-parameter logistic item response model using MLE and Bayesian MCMC methods* [Unpublished manuscript]. Iowa City: The University of Iowa.

Rupp, A. A. (2003). Item response modeling with BILOG-MG and MULTILOG for windows. *International Journal of Testing, 3*(4), 365-384.

Samejima, F. (1997). Departure from normal assumptions: A promise for future psychometrics with substantive mathematical modeling. *Psychometrika, 62*(4), 471-493.

Samejima, F. (2000). Logistic positive exponent family of models: Virtue of asymmetric item characteristics curves. *Psychometrika, 65*(3), 319-335.

Sheng, Y. (2005). *Bayesian analysis of hierarchical IRT models: Comparing and combining the uni-dimensional & multi-dimensional IRT models* (Doctoral dissertation). Retrieved from ProQuest. (UMI 3322166)

Sinharay, S. (2006). Bayesian item fit analysis for unidimensional item response theory models. *British Journal of Mathematical and Statistical Psychology 59*, 429-449.

Suh, H. (2010). *A study of Bayesian estimation and comparison of response time models in item response theory* (Doctoral dissertation). Retrieved from ProQuest. (UMI 3404361)

Traub, R. (1997). Classical Test Theory in historical perspective. *Educational Measurement: Issues and Practice, 16*(4), 8-14. doi:10.1111/j.1745-3992.1997.tb00603.x

van der Linden, W. J., & Hambleton, R. K. (2010). *Handbook of modern item response theory.* New York: Springer.

Wheeler, D. J. (2011, August). Problems with skewness and kurtosis: What do the shape parameters do? *Quality Digest Daily.* Retrieved from http:/www.spcpress.com/pdf/DJW231.pdf

APPENDIXES

APPENDIX A

Fleishman's SAS^{TM} Code for Non-normal Distributions

```
/* This program calculates the coefficients for Fleisman's power
transformation in order to obtain univariate non-normal variables */

PROC IML;
/* in the following matrix 'SKEWKURT', specify the skewness and
kurtosis for each
    variable. Each row represents one variable. In each row, the ith
number is the
    skewness and the 2nd number is the kurtosis of the variable;
*/

SKEWKURT={-.40 .4,
          -.80 .7,
          -1.20 1.4,
           1.20 1.4};

START NEWTON;
      RUN FUN;
      DO ITER = 1 to MAXITER
      WHILE(MAX(ABS(F))>CONVERGE);
            RUN DERIV;
            DELTA=-SOLVE(J,F);
            COEF=COEF+DELTA;
            RUN FUN;
      END;
FINISH NEWTON;
MAXITER=25;
CONVERGE=.000001;
START FUN;
      X1=COEF[1];
      X2=COEF[2];
      X3=COEF[3];
      F=(X1**2+6*X1*X3+2*X2**2+15*X3**2-
1)//(2*X2*(X1**2+24*X1*X3+105*X3**2+2)-
SKEWNESS)//(24*(X1*X3+X2**2*(1+X1**2+28*X1*X3)+X3**2*(12+48*X1*X3+141*X
2**2+225*X3**2))-KURTOSIS);
FINISH FUN;
START DERIV;
      J=((2*X1+6*X3)||(4*X2)||(6*X1+30*X3))//((4*X2*(X1+12*X3))||(2*(X1
**2+24*X1*X3+105*X3**2+2))||(4*X2*(12*X1+105*X3)))//((24*(X3+X2**2*(2*X
1+28*X3)+48*X3**3))||(48*X2*(1+X1**2+28*X1*X3+141*X3**2))||(24*(X1+28*X
1*X2**2+2*X3*(12+48*X1*X3+141*X2**2+225*X3**2)+X3**2*(48*X1+450*X3))));
FINISH DERIV;
DO;
NUM = NROW(SKEWKURT);
DO VAR = 1 TO NUM;
      SKEWNESS=SKEWKURT[VAR,1];
      KURTOSIS=SKEWKURT[VAR,2];
      COEF={1.0, 0.0, 0.0};
      RUN NEWTON;
```

```
        COEF=COEF`;
        SK_KUR=SKEWKURT[VAR,];
        COMBINE=SK_KUR || COEF;
        IF VAR=1 THEN RESULT=COMBINE;
        ELSE IF VAR>1 THEN RESULT=RESULT // COMBINE;
END;
        PRINT "COEFICIENTS OF B, C, D FOR FLEISMAN'S POWER
TRANSFORMATION" ;
        PRINT "Y = A + BX + CX^2 +DX^3" ;
        PRINT "A = -C" ;
        MATTRIB RESULT COLNAME=({SKEWNESS KURTORSIS B C D})
                          FORMAT=12.9;
        PRINT RESULT;
END;
QUIT;
```

Output of SAS[TM] Program

```
        The SAS System          14:37 Wednesday, March 5, 2014

        COEFICIENTS OF B, C, D FOR FLEISMAN'S POWER TRANSFORMATION


                        Y = A + BX + CX^2 +DX^3


                                A = -C


                                RESULT
          SKEWNESS      KURTOSIS           B            C            D


       -0.400000000   0.400000000   0.973646447  -0.063985169   0.007362577
       -0.800000000   0.700000000   1.002054839  -0.141323322  -0.007467128
       -1.200000000   1.400000000   1.032194765  -0.257844888  -0.035001562
        1.200000000   1.400000000   1.032194765   0.257844888  -0.035001562
```

APPENDIX B

SAS^{TM} Code for the Simulation Study

Baseline Condition, Situation 1

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Baseline Condition, Situation 1*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel1.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                   p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                   y[i,k] ~ dbern(p[i,k])
          }

      theta[i] ~ dnorm(0,1)
    }

    for (k in 1:T) {
        a[k] ~ dlnorm(0,0.5)
        b[k] ~ dnorm(0,2)
        c[k] <- 0.17
    }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini1.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1),"
```

```
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0)))," 
RUN;

/*Scripts to run WinBUGS*/
FILENAME runirt1 'c:\Winbugs\WinBUGS\runirt1.txt';
DATA _NULL_;
  FILE runirt1;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel1.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata1.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini1.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog1.txt')";
  PUT@1  "quit()";
RUN;


DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt1.txt';
PUT 'exit';
RUN;
options linesize=72;
%macro runsimirt;
    %let n=1;
        %do %while(&n <= 50);

%macro gen(sz1=, length=,SIM=);

        libname temp "C:\";

        %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
        data SIM_IRT1;
                array item item1-item&length;
                array a a1-a&length;   /*tlength= test length*/
                array b b1-b&length;
                array c c1-c&length;
                array p p1-p&length;
                array x x1-x&length;
                array y y1-y&length;
                infile 'G:\IRT HW\SAS\NM_SPE_40Occasion1.dat'; /*parameter
a b and c reference*/

                    do over a;
                       input item a b c;
                    end;
```

```
                    do k=1 to &sz1;    /* set sample size*/
                        t1=rannor(0);   /*Used in Baseline condition for
occasion one for BILOG-MG and WinBUGS (1X1X2= 2 situations) */
                    SIM=&SIM;

                do over p;
                        x=ranuni(0);
                        p=c+((1-c)/(1+exp(-1.7*a*(t1-b))));
                        if x le p then y = 1; else y=0;
                    end; output;
                    end;
         run;

        data temp.SIM&SIM&i; set SIM_IRT1; run;

data Work.simcsv1;
        set Work.SIM_IRT1 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;


  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=1);
proc export data=work.sim_irt1 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt1&n..csv"
   dbms=csv
   replace;
%_sexport(data=Simcsv1,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata1.txt',
var =y1-y40);
run;


/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;

/*Read in the log file */

DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog1.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;


DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults1&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults1&n..log"';

/*Analyze the results*/
```

```
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults1&n..txt" TRUNCOVER ;
INPUT all $90.;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp1&n..txt";
PUT all;
RUN;

%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The original file was
replicated 100 times. */

Data _NULL_ ;
     SET  Work.&DataIn ;
     ID = 5 ;
     FILE "C:\B1&n..dat" ;
     PUT   @1  ID @2  (y1-y40) (1.) ;
run ;



%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT1)
%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


   /* Create BLM file for BILOGMG */
data _null_ ;
     file "C:\Program Files\BILOGMG\BASE1&n..blm ";
put
%str(">TITLE   Baseline Condition;" )    /
"      Occasion one  ;      "       /
%str(">GLOBAL DFName= 'C:\B1&n..dat', NPArm=3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = Base1 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)  ") /
%str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;  ") ;
run;
%end;
%mend RepBILOG1;
%RepBILOG1 (Cond=BASE,Occ=1,Rep=1)
%macro RepBILOG2 (Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


/*****    End Algina
***/
data _null_ ;
         file "C:\SIMIRTBLM\Baseline1&n..bat";
    put
```

```
      %str("cd\program files\bilogmg" )      /
      %str("blm1 BASE1&n")/
      %str("blm2 BASE1&n" )/
      %str("blm3 BASE1&n") /
      %str("exit") /;
run ;
/*                  Call Bilog
*/ ;
DATA _NULL_;
dm " x C:\SIMIRTBLM\Baseline1&n..bat" ;

/*****            End BILOG-MG call
*/ ;


run ;

%end;
%mend RepBILOG2;
%RepBILOG2 (Cond=BASE,Occ=1,Rep=1);

%let n=%eval(&n+1);
    %end;
%mend runsimirt;
%runsimirt;
```

Baseline Condition, Situation 2

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Baseline Condition, Situation 2*/
/* specify theta prior by analysis of Situation 1 posterior*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel2.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                  p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                  y[i,k] ~ dbern(p[i,k])
        }

      theta[i] ~ dnorm(0.00898,0.93014)
    }

    for (k in 1:T) {
      a[k] ~ dlnorm(0,0.5)
      b[k] ~ dnorm(0,2)
      c[k] <- 0.17
```

```
      }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini2.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1,1),"
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0)))),"

RUN;

/*Scripts to run WinBUGS*/
FILENAME runirt2 'c:\Winbugs\WinBUGS\runirt2.txt';
DATA _NULL_;
  FILE runirt2;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel2.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata2.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini2.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog2.txt')";
  PUT@1  "quit()";
RUN;

DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt2.txt';
PUT 'exit';
RUN;
options linesize=72;
```

```sas
%macro runsimirt;
    %let n=1;
       %do %while(&n <= 50);


%macro gen(sz1=, length=,SIM=);

       libname temp "C:\";

       %do i=1 %to 5; *number of draws from random normal;

/* SIM Data Generation*/
       data SIM_IRT2;
              array item item1-item&length;
              array a a1-a&length;  /*tlength= test length*/
              array b b1-b&length;
              array c c1-c&length;
              array p p1-p&length;
              array x x1-x&length;
              array y y1-y&length;
              infile 'G:\IRT HW\SAS\NM_SPE_40Occasion2.dat'; /*parameter
a b and c reference*/

                     do over a;
                        input item a b c;
                     end;

                     do k=1 to &sz1;   /* set sample size*/
                        t1=rannor(0);  /*Used in Baseline condition for
occasion one for BILOG-MG and WinBUGS (1X1X2= 2 situations) */
                     SIM=&SIM;

                  do over p;
                          x=ranuni(0);
                          p=c+((1-c)/(1+exp(-1.7*a*(t1-b))));
                          if x le p then y = 1; else y=0;
                     end; output;
                     end;
        run;

       data temp.SIM&SIM&i; set SIM_IRT2; run;

data Work.simcsv2;
       set Work.SIM_IRT2 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;


  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=2);
proc export data=work.sim_irt2 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt2&n..csv"
   dbms=csv
   replace;
%_sexport(data=Simcsv2,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata2.txt',
```

```
var =y1-y40);
run;



/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;



DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog2.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;



DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults2&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults2&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults2&n..txt" TRUNCOVER ;
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp2&n..txt";
PUT all;
RUN;

%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The original file was
replicated 100 times. */

Data _NULL_ ;
      SET  Work.&DataIn ;
      ID = 5 ;
      FILE "C:\B2&n..dat" ;
      PUT    @1  ID  @2  (y1-y40) (1.) ;
run ;



%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT2)

%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;
```

```
    /* Create BLM file for BILOGMG */
data _null_ ;
      file "C:\Program Files\BILOGMG\BASE2&n..blm ";
put
%str(">TITLE    Baseline Condition;" )    /
"        Occasion two  ;        "       /
%str(">GLOBAL DFName= 'C:\B2&n..dat', NPArm = 3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER =(1(1)40), INAMES =(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = Base2 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)  ") /
%str(">CALIB ACCel = 1.000, cycles = 25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;  ") ;
run;
%end;
%mend RepBILOG1;
%RepBILOG1 (Cond=BASE,Occ=2,Rep=1)
%macro RepBILOG2 (Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


/*****    End Algina
***/
data _null_ ;
         file "C:\SIMIRTBLM\Baseline2&n..bat";
    put
    %str("cd\program files\bilogmg" )     /
    %str("blm1 BASE2&n")/
    %str("blm2 BASE2&n" )/
    %str("blm3 BASE2&n") /
    %str("exit") /;
run ;
/*                 Call Bilog
*/ ;
DATA _NULL_;
dm " x C:\SIMIRTBLM\Baseline2&n..bat" ;

/*****           End BILOG-MG call
*/ ;


run ;

%end;
%mend RepBILOG2;
%RepBILOG2 (Cond=BASE,Occ=2,Rep=1);

%let n=%eval(&n+1);
   %end;
%mend runsimirt;
%runsimirt;
```

Baseline Condition, Situation 3

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Baseline Condition, Situation 3*/
/* specify theta prior by analysis of Situation 2 posterior*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel3.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
       for (k in 1:T) {
                   p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                   y[i,k] ~ dbern(p[i,k])
         }

       theta[i] ~ dnorm(0.0162,0.9642)
    }

    for (k in 1:T) {
       a[k] ~ dlnorm(0,0.5)
       b[k] ~ dnorm(0,2)
       c[k] <- 0.17
    }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini3.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1),"
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0))),"

RUN;
```

```
/*Scripts to run WinBUGS*/
FILENAME runirt3 'c:\Winbugs\WinBUGS\runirt3.txt';
DATA _NULL_;
  FILE runirt3;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel3.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata3.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini3.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog3.txt')";
  PUT@1  "quit()";
RUN;

DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt3.txt';
PUT 'exit';
RUN;
options linesize=72;
%macro runsimirt;
   %let n=1;
      %do %while(&n <= 50);

%macro gen(sz1=, length=,SIM=);

      libname temp "C:\";

      %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
      data SIM_IRT3;
            array item item1-item&length;
            array a a1-a&length;  /*tlength= test length*/
            array b b1-b&length;
            array c c1-c&length;
            array p p1-p&length;
            array x x1-x&length;
            array y y1-y&length;
            infile 'G:\IRT HW\SAS\NM_SPE_40Occasion3.dat'; /*parameter
a b and c reference*/

                  do over a;
                     input item a b c;
                  end;

                  do k=1 to &sz1;   /* set sample size*/
                     t1=rannor(0);  /*Used in Baseline condition for
occasion one for BILOG-MG and WinBUGS (1X1X2= 2 situations) */
                  SIM=&SIM;
```

```
          do over p;
                  x=ranuni(0);
                  p=c+((1-c)/(1+exp(-1.7*a*(t1-b)))));
                  if x le p then y = 1; else y=0;
              end; output;
              end;
      run;

      data temp.SIM&SIM&i; set SIM_IRT3; run;

data Work.simcsv3;
      set Work.SIM_IRT3 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;


  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=3);
proc export data=work.sim_irt3 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt3&n..csv"
   dbms=csv
   replace;
%_sexport(data=Simcsv3,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata3.txt',
var =y1-y40);
run;


/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;

/*Read in the log file to view the DIC*/

DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog3.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;

DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults3&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults3&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults3&n..txt" TRUNCOVER ;
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp3&n..txt";
```

```
PUT all;
RUN;


%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The original file was
replicated 100 times. */

Data _NULL_ ;
     SET  Work.&DataIn ;
     ID = 5 ;
     FILE "C:\B3&n..dat" ;
     PUT   @1  ID  @2  (y1-y40) (1.) ;
run ;



%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT3)

%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


   /* Create BLM file for BILOGMG */
data _null_ ;
     file "C:\Program Files\BILOGMG\BASE3&n..blm ";
put
%str(">TITLE    Baseline Condition;" )    /
"        Situation one  ;       "        /
%str(">GLOBAL DFName= 'C:\B3&n..dat', NPArm=3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = Base3 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)  ") /
%str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;  ") ;
run;
%end;
%mend RepBILOG1;
%RepBILOG1 (Cond=BASE,Occ=3,Rep=1)
%macro RepBILOG2 (Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


/*****     End Algina
***/
data _null_ ;
         file "C:\SIMIRTBLM\Baseline3&n..bat";
    put
    %str("cd\program files\bilogmg" )    /
    %str("blm1 BASE3&n")/
    %str("blm2 BASE3&n" )/
```

```
    %str("blm3 BASE3&n") /
    %str("exit") /;
run ;
/*                 Call Bilog
*/ ;
DATA _NULL_;
dm " x C:\SIMIRTBLM\Baseline3&n..bat" ;

/*****          End BILOG-MG call
*/ ;


run ;

%end;
%mend RepBILOG2;
%RepBILOG2 (Cond=BASE,Occ=3,Rep=1);

%let n=%eval(&n+1);
    %end;
%mend runsimirt;
%runsimirt;
```

Favorable to RBA Condition, Situation 4

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Favorable Condition, Situation 4*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel4.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                 p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                 y[i,k] ~ dbern(p[i,k])
        }

      theta[i] ~ dnorm(0,1)
   }

   for (k in 1:T) {
      a[k] ~ dlnorm(0,0.5)
      b[k] ~ dnorm(0,2)
      c[k] <- 0.17
   }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1
```

```
#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
   SET model;
   FILE model;
   PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini4.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1,1),"
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0)))),"

RUN;

/*Scripts to run WinBUGS*/
FILENAME runirt4 'c:\Winbugs\WinBUGS\runirt4.txt';
DATA _NULL_;
   FILE runirt4;
   PUT@1 "display('log')";
   PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel4.txt')" ;
   PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata4.txt')";
   PUT@1 "compile(1)";
   PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini4.txt')";
   PUT@1 "gen.inits()";
   PUT@1 "update(1000)";
   PUT@1 "set(a)";
   PUT@1 "set(b)";
   PUT@1 "set(theta)";
   PUT@1 "update(5000)";
   PUT@1 "stats(*)";
   PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog4.txt')";
   PUT@1  "quit()";
RUN;

DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt4.txt';
PUT 'exit';
RUN;
options linesize=72;
%macro runsimirt;
   %let n=1;
      %do %while(&n <= 50);

%macro gen(sz1=, length=, SIM=);

      libname temp "C:\";
```

```
      %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
      data SIM_IRT4;
            array item item1-item&length;
            array a a1-a&length;   /*tlength= test length*/
            array b b1-b&length;
            array c c1-c&length;
            array p p1-p&length;
            array x x1-x&length;
            array y y1-y&length;
            infile 'G:\IRT HW\SAS\NM_SPE_40Occasion1.dat'; /*parameter
a b and c reference*/

                  do over a;
                     input item a b c;
                  end;

                  do k=1 to &sz1;   /* set sample size*/
                     t1=.5+sqrt(1)*rannor(0);  /*Used in FAVORABLE
condition for occasion one for BILOG-MG and WinBUGS (1X1X2= 2
situations) */
              SIM=&SIM;
                     t1 = .063985169 + .973646447*t1 + -
.063985169*t1**2 + .007362577*t1**3; /* Fleisman non-normality
transformation skew = -.4, kurtosis = .4 */
            do over p;
                     x=ranuni(0);
                     p=c+((1-c)/(1+exp(-1.7*a*(t1-b))));
                     if x le p then y = 1; else y=0;
                  end; output;
                  end;
       run;

      data temp.SIM&SIM&i; set SIM_IRT4; run;

data Work.simcsv4;
      set Work.SIM_IRT4 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;


  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=4);
proc export data=work.sim_irt4 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt4&n..csv"
   dbms=csv
   replace;
run;
%_sexport(data=Simcsv4,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata4.txt',
var =y1-y40);
run;
```

```
/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;



DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog4.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;



DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults4&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults4&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults4&n..txt" TRUNCOVER ;
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp4&n..txt";
PUT all;
RUN;

%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The original file was
replicated 50 times. */

Data _NULL_ ;
     SET  Work.&DataIn ;
     ID = 5 ;
     FILE "C:\F4&n..dat" ;
     PUT    @1  ID  @2  (y1-y40) (1.) ;
run ;



%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT4)
%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


   /* Create BLM file for BILOGMG */
data _null_ ;
     file "C:\Program Files\BILOGMG\FAV4&n..blm ";
```

```
put
%str(">TITLE    Favorable Condition;" )    /
"       Situation one  ;       "       /
%str(">GLOBAL DFName= 'C:\F4&n..dat', NPArm=3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = FAV4 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)  ") /
%str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;  ") ;
run;
%end;
%mend RepBILOG1;
%RepBILOG1 (Cond=BASE,Occ=4,Rep=1)
%macro RepBILOG2 (Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


/*****     End Algina
***/
data _null_ ;
         file "C:\SIMIRTBLM\Favorable4&n..bat";
    put
    %str("cd\program files\bilogmg" )     /
    %str("blm1 FAV4&n")/
    %str("blm2 FAV4&n" )/
    %str("blm3 FAV4&n") /
    %str("exit") /;
run ;
/*                  Call Bilog
*/ ;
DATA _NULL_;
dm " x C:\SIMIRTBLM\Favorable4&n..bat" ;

/*****            End BILOG-MG call
*/ ;


run ;

%end;
%mend RepBILOG2;
%RepBILOG2 (Cond=BASE,Occ=4,Rep=1);

%let n=%eval(&n+1);
   %end;
%mend runsimirt;
%runsimirt;
```

Favorable to RBA Condition, Situation 5

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';
```

```
/*WinBUGS program for 3PL IRT Favorable Condition, Situation 5*/
/*Prior will be set by obtaining information from Situation 4
posterior*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel5.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                  p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                  y[i,k] ~ dbern(p[i,k])
         }

      theta1[i] ~ dbeta(5.4,3.8)
             }
             for (i in 1:N){
                  theta[i] <- ((theta1[i]*6)-3)
    }

    for (k in 1:T) {
        a[k] ~ dlnorm(0,0.5)
        b[k] ~ dnorm(0,2)
        c[k] <- 0.17
      }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini5.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1,1),"
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0)))"

RUN;
```

```
/*Scripts to run WinBUGS*/
FILENAME runirt5 'c:\Winbugs\WinBUGS\runirt5.txt';
DATA _NULL_;
  FILE runirt5;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel5.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata5.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini5.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog5.txt')";
  PUT@1  "quit()";
RUN;

DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt5.txt';
PUT 'exit';
RUN;
options linesize=72;
%macro runsimirt;
   %let n=1;
      %do %while(&n <= 50);

%macro gen(sz1=, length=, SIM=);

      libname temp "C:\";

      %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
      data SIM_IRT5;
            array item item1-item&length;
            array a a1-a&length;  /*tlength= test length*/
            array b b1-b&length;
            array c c1-c&length;
            array p p1-p&length;
            array x x1-x&length;
            array y y1-y&length;
            infile 'G:\IRT HW\SAS\NM_SPE_40Occasion2.dat'; /*parameter
a b and c reference*/

                  do over a;
                     input item a b c;
                  end;

                  do k=1 to &sz1;   /* set sample size*/
                     t1=.75+sqrt(1)*rannor(0);  /*Used in FAVORABLE
condition for occasion two for BILOG-MG and WinBUGS (1X1X2= 2
situations) */
                  SIM=&SIM;
```

```
                              t1 = .141323322 + 1.002054839*t1 + -
.141323322*t1**2 + -.007467128*t1**3; /* Fleisman non-normality
transformation skew = -.8, kurtosis = .7 */
                do over p;
                        x=ranuni(0);
                        p=c+((1-c)/(1+exp(-1.7*a*(t1-b)))));
                        if x le p then y = 1; else y=0;
                    end; output;
                    end;
        run;

        data temp.SIM&SIM&i; set SIM_IRT5; run;


data Work.simcsv5;
      set Work.SIM_IRT5 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;

  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=5);
proc export data=work.sim_irt5 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt5&n..csv"
   dbms=csv
   replace;
run;
%_sexport(data=Simcsv5,
file ="C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata5.txt",
var =y1-y40);
run;



/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;


/*Read in the log file to view the DIC*/

DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog5.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;

DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults5&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults5&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults5&n..txt" TRUNCOVER ;
```

```
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp5&n..txt";
PUT all;
RUN;

%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The situation is repeated 50
times for WinBUGS and 50 times for BILOGMG. */

Data _NULL_ ;
     SET  Work.&DataIn ;
     ID = 5 ;
     FILE "C:\F5&n..dat" ;
     PUT    @1  ID  @2  (y1-y40) (1.) ;
run ;


%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT5)
%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


   /* Create BLM file for BILOGMG */
data _null_ ;
     file "C:\Program Files\BILOGMG\FAV5&n..blm ";
put
%str(">TITLE    Favorable Condition;" )   /
"        Situation two  ;        "        /
%str(">GLOBAL DFName= 'C:\F5&n..dat', NPArm=3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = FAV5 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)  ") /
%str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;  ") ;
run;
%end;
%mend RepBILOG1;
%RepBILOG1 (Cond=FAV,Occ=5,Rep=1)
%macro RepBILOG2 (Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


/*****    End Algina
***/
data _null_ ;
         file "C:\SIMIRTBLM\Favorable5&n..bat";
    put
    %str("cd\program files\bilogmg" )     /
```

```
        %str("blm1 FAV5&n")/
        %str("blm2 FAV5&n" )/
        %str("blm3 FAV5&n") /
        %str("exit") /;
run ;
/*                    Call Bilog
*/ ;
DATA _NULL_;
dm " x C:\SIMIRTBLM\Favorable5&n..bat" ;

/*****            End BILOG-MG call
*/ ;


run ;

%end;
%mend RepBILOG2;
%RepBILOG2 (Cond=FAV,Occ=5,Rep=1);

%let n=%eval(&n+1);
    %end;
%mend runsimirt;
%runsimirt;
```

Favorable to RBA Condition, Situation 6

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Favorable Condition, Situation 6*/
/*Prior will be set by obtaining information from Situation 5
posterior*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel6.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                  p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                  y[i,k] ~ dbern(p[i,k])
          }

      theta1[i] ~ dbeta(7.5,4.73)
            }
            for (i in 1:N){
                  theta[i] <- ((theta1[i]*6)-3)
    }

    for (k in 1:T) {
      a[k] ~ dlnorm(0,0.5)
      b[k] ~ dnorm(0,2)
      c[k] <- 0.17
```

```
      }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini6.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1,1),"
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0)))"

RUN;

/*Scripts to run WinBUGS*/
FILENAME runirt6 'c:\Winbugs\WinBUGS\runirt6.txt';
DATA _NULL_;
  FILE runirt6;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel6.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata6.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini6.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog6.txt')";
  PUT@1  "quit()";
RUN;

DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt6.txt';
PUT 'exit';
RUN;
options linesize=72;
```

```
%macro runsimirt;
    %let n=1;
        %do %while(&n <= 50);

%macro gen(sz1=, length=, SIM=);

        libname temp "C:\";

        %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
        data SIM_IRT6;
                array item item1-item&length;
                array a a1-a&length;  /*tlength= test length*/
                array b b1-b&length;
                array c c1-c&length;
                array p p1-p&length;
                array x x1-x&length;
                array y y1-y&length;
                infile 'G:\IRT HW\SAS\NM_SPE_40Occasion3.dat'; /*parameter
a b and c reference*/

                    do over a;
                       input item a b c;
                    end;

                    do k=1 to &sz1;   /* set sample size*/
                       t1=1+sqrt(1)*rannor(0);  /*Used in FAVORABLE
condition for occasion three for BILOG-MG and WinBUGS (1X1X2= 2
situations) */
                   SIM=&SIM;
                       t1 = .257844888 + 1.032194765*t1 + -
.257844888*t1**2 + -.035001562*t1**3; /* Fleisman non-normality
transformation skew = -1.2, kurtosis = 1.4 */
             do over p;
                    x=ranuni(0);
                    p=c+((1-c)/(1+exp(-1.7*a*(t1-b))));
                    if x le p then y = 1; else y=0;
                 end; output;
                 end;
        run;

        data temp.SIM&SIM&i; set SIM_IRT6; run;

data Work.simcsv6;
        set Work.SIM_IRT6 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;

  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=6);
proc export data=work.sim_irt6 (keep=k t1 SIM)
    outfile="c:\SIMIRTBLM\sim_irt6&n..csv"
    dbms=csv
    replace;
```

```
run;
%_sexport(data=Simcsv6,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata6.txt',
var =y1-y40);
run;


/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;


DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog6.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;


DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults6&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults6&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults6&n..txt" TRUNCOVER ;
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp6&n..txt";
PUT all;
RUN;

%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The original file was
replicated 100 times. */

Data _NULL_ ;
     SET  Work.&DataIn ;
     ID = 5 ;
     FILE "C:\F6&n..dat" ;
     PUT    @1  ID  @2  (y1-y40) (1.) ;
run ;


%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT6)
%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;
```

```
      /* Create BLM file for BILOGMG */
data _null_ ;
      file "C:\Program Files\BILOGMG\FAV6&n..blm ";
put
%str(">TITLE    Favorable Condition;" )    /
"        Situation three  ;       "        /
%str(">GLOBAL DFName= 'C:\F6&n..dat', NPArm=3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = FAV6 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)  ") /
%str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;  ") ;
run;
%end;
%mend RepBILOG1;
%RepBILOG1 (Cond=BASE,Occ=6,Rep=1)
%macro RepBILOG2 (Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


/*****    End Algina
***/
data _null_ ;
         file "C:\SIMIRTBLM\Favorable6&n..bat";
   put
   %str("cd\program files\bilogmg" )    /
   %str("blm1 FAV6&n")/
   %str("blm2 FAV6&n" )/
   %str("blm3 FAV6&n") /
   %str("exit") /;
run ;
/*               Call Bilog
*/ ;
DATA _NULL_;
dm " x C:\SIMIRTBLM\Favorable6&n..bat" ;

/*****           End BILOG-MG call
*/ ;


run ;

%end;
%mend RepBILOG2;
%RepBILOG2 (Cond=BASE,Occ=6,Rep=1);

%let n=%eval(&n+1);
   %end;
%mend runsimirt;
%runsimirt;
```

Unfavorable to RBA Condition, Situation 7

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Unfavorable Condition, Situation 7*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel7.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                  p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                  y[i,k] ~ dbern(p[i,k])
        }

      theta[i] ~ dnorm(0,1)
    }

   for (k in 1:T) {
      a[k] ~ dlnorm(0,0.5)
      b[k] ~ dnorm(0,2)
      c[k] <- 0.17
    }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini7.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1),"
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0)))"

RUN;
```

```
/*Scripts to run WinBUGS*/
FILENAME runirt7 'c:\Winbugs\WinBUGS\runirt7.txt';
DATA _NULL_;
  FILE runirt7;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel7.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata7.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini7.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog7.txt')";
  PUT@1  "quit()";
RUN;

DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt7.txt';
PUT 'exit';
RUN;
options linesize=72;
%macro runsimirt;
   %let n=1;
      %do %while(&n <= 50);
%macro gen(sz1=, length=,SIM=);

      libname temp "C:\";

      %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
      data SIM_IRT7;
            array item item1-item&length;
            array a a1-a&length;  /*tlength= test length*/
            array b b1-b&length;
            array c c1-c&length;
            array p p1-p&length;
            array x x1-x&length;
            array y y1-y&length;
            infile 'G:\IRT HW\SAS\NM_SPE_40Occasion1.dat'; /*parameter
a b and c reference*/

                  do over a;
                     input item a b c;
                  end;

                  do k=1 to &sz1;   /* set sample size*/
                     t1=rannor(0);  /*Used in Unfavorable condition for
occasion one for BILOG-MG and WinBUGS (1X1X2= 2 situations) */
                SIM=&SIM;

              do over p;
```

```
                        x=ranuni(0);
                        p=c+((1-c)/(1+exp(-1.7*a*(t1-b))));
                        if x le p then y = 1; else y=0;
                  end; output;
                  end;
        run;

      data temp.SIM&SIM&i; set SIM_IRT7; run;

data Work.simcsv7;
      set Work.SIM_IRT7 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;


  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=7);
proc export data=work.sim_irt7 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt7&n..csv"
   dbms=csv
   replace;
run;
%_sexport(data=Simcsv7,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata7.txt',
var =y1-y40);
run;




/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;

/*Read in the log file to view the DIC*/
DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog7.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;


DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults7&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults7&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults7&n..txt" TRUNCOVER ;
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp7&n..txt";
```

```
     PUT all;
     RUN;


     %macro MakeDatFiles (Datain= );
     /*  The original file, Sim_IRT&Sim is used in each situation; therefore
     the data given
        to BILOG and WinBUGS is the same within each situation. The format
     of the file is changed
        to accomodate the different programs. The original file was
     replicated 100 times. */


     Data _NULL_ ;
          SET  Work.&DataIn ;
          ID = 5 ;
          FILE "C:\U7&n..dat" ;
          PUT   @1  ID  @2  (y1-y40) (1.) ;
     run ;



     %mend  MakeDatFiles ;
     %MakeDatFiles (datain=Sim_IRT7)
     %macro RepBILOG1(Cond= ,Occ= , Rep= );
        %do q=0 %to &Rep;


        /* Create BLM file for BILOGMG */
     data _null_ ;
          file "C:\Program Files\BILOGMG\UNFAV7&n..blm ";
     put
     %str(">TITLE    Unfavorable Condition;" ) /
     "        Situation one  ;       "        /
     %str(">GLOBAL DFName= 'C:\U7&n..dat', NPArm=3; ")/
     %str(">LENGTH NITems = (40);")/
     %str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
     %str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
     %str(">TEST1 TNAme = UNFAV7 , GUESS = (0.1700(0)40);")/
     %str("(1A1,40A1)  ") /
     %str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
     %str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
     %str(">SCORE ;   ") ;
     run;
     %end;
     %mend RepBILOG1;
     %RepBILOG1 (Cond=UNFAV,Occ=7,Rep=1)
     %macro RepBILOG2 (Cond= ,Occ= , Rep= );
        %do q=0 %to &Rep;


     /*****    End Algina
     ***/
     data _null_ ;
             file "C:\SIMIRTBLM\Unfavorable7&n..bat";
         put
         %str("cd\program files\bilogmg" )     /
         %str("blm1 UNFAV7&n")/
         %str("blm2 UNFAV7&n" )/
         %str("blm3 UNFAV7&n") /
```

```
    %str("exit") /;
run ;
/*                  Call Bilog
*/ ;
DATA _NULL_;
dm " x C:\SIMIRTBLM\Unfavorable7&n..bat" ;

/*****          End BILOG-MG call
*/ ;


run ;

%end;
%mend RepBILOG2;
%RepBILOG2 (Cond=UNFAV,Occ=7,Rep=1);

%let n=%eval(&n+1);
    %end;
%mend runsimirt;
%runsimirt;
```

Unfavorable to RBA Condition, Situation 8

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Unfavorable Condition, Situation 8*/
/*Prior will be set by obtaining information from Situation 7
posterior*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel8.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                 p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                 y[i,k] ~ dbern(p[i,k])
           }

      theta1[i] ~ dbeta(4.63,4.61)
              }
              for (i in 1:N){
                 theta[i] <- ((theta1[i]*6)-3)
      }

    for (k in 1:T) {
       a[k] ~ dlnorm(0,0.5)
       b[k] ~ dnorm(0,2)
       c[k] <- 0.17
    }

#b1 ~ dpois(2)
```

```
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini8.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1),"
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0)))," 

RUN;

/*Scripts to run WinBUGS*/
FILENAME runirt8 'c:\Winbugs\WinBUGS\runirt8.txt';
DATA _NULL_;
  FILE runirt8;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel8.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata8.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini8.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog8.txt')";
  PUT@1  "quit()";
RUN;

DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt8.txt';
PUT 'exit';
RUN;
options linesize=72;
%macro runsimirt;
   %let n=1;
      %do %while(&n <= 50);
```

```sas
%macro gen(sz1=, length=, SIM=);

       libname temp "C:\";

       %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
       data SIM_IRT8;
               array item item1-item&length;
               array a a1-a&length;  /*tlength= test length*/
               array b b1-b&length;
               array c c1-c&length;
               array p p1-p&length;
               array x x1-x&length;
               array y y1-y&length;
               infile 'G:\IRT HW\SAS\NM_SPE_40Occasion2.dat'; /*parameter
a b and c reference*/

                       do over a;
                          input item a b c;
                       end;

                       do k=1 to &sz1;   /* set sample size*/
                          t1=1+sqrt(1)*rannor(0);  /*Used in UNFAVORABLE
condition for occasion two for BILOG-MG and WinBUGS (1X1X2= 2
situations) */
                       SIM=&SIM;
                          t1 = .257844888 + 1.032194765*t1 + -
.257844888*t1**2 + -.035001562*t1**3; /* Fleisman non-normality
transformation skew = -1.2, kurtosis = 1.4 */
                       do over p;
                               x=ranuni(0);
                               p=c+((1-c)/(1+exp(-1.7*a*(t1-b))));
                               if x le p then y = 1; else y=0;
                          end; output;
                          end;
          run;

       data temp.SIM&SIM&i; set SIM_IRT8; run;

data Work.simcsv8;
       set Work.SIM_IRT8 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;


  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=8);
proc export data=work.sim_irt8 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt8&n..csv"
   dbms=csv
   replace;
run;
```

```
%_sexport(data=Simcsv8,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata8.txt',
var =y1-y40);
run;

/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;


/*Read in the log file to view the DIC*/

DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog8.txt" TRUNCOVER ;
INPUT log $90.;

RUN;


PROC PRINT DATA=log;
RUN;



DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults8&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults8&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults8&n..txt" TRUNCOVER ;
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp8&n..txt";
PUT all;
RUN;



%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The original file was
replicated 50 times. */

Data _NULL_ ;
     SET  Work.&DataIn ;
     ID = 5 ;
     FILE "C:\U8&n..dat" ;
     PUT    @1  ID  @2  (y1-y40) (1.) ;
run ;


%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT8)
%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;
```

```
data _null_ ;
      file "C:\Program Files\BILOGMG\UNFAV8&n..blm ";
put
%str(">TITLE    Unfavorable Condition;" ) /
"        Situation two  ;       "      /
%str(">GLOBAL DFName= 'C:\U8&n..dat', NPArm=3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = UNFAV8 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)  ") /
%str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;   ") ;


run;
%end;
%mend RepBILOG1;
data %RepBILOG1(Cond=UNFAV,Occ=8,Rep=1);
%macro RepBILOG2(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;


     /*****     End Algina
*********/
    data _null_ ;

          file "C:\SIMIRTBLM\Unfavorable8&n..bat";
    put
    %str("cd\program files\bilogmg" )     /
    %str("blm1 UNFAV8&n")/
    %str("blm2 UNFAV8&n" )/
    %str("blm3 UNFAV8&n") /
    %str("exit") /
    run;
    /*****                 Call Bilog
****/

    dm " x C:\SIMIRTBLM\Unfavorable8&n..bat" ;

    *****           End BILOG-MG call
********* ;


    run ;
%end;
%mend RepBILOG2;
data %RepBILOG2(Cond=UNFAV,Occ=8,Rep=1);

%let n=%eval(&n+1);
   %end;
%mend runsimirt;
%runsimirt;
```

Unfavorable to RBA Condition, Situation 9

```
TITLE 'Run WinBUGS from SAS: Recursive Bayesian Analysis using 3PL IRT
(2013)';

/*WinBUGS program for 3PL IRT Unfavorable Condition, Situation 9*/
/*Prior will be set by obtaining information from Situation 8
posterior*/
FILENAME model "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtmodel9.txt";
DATA model;
INPUT model $80.;
CARDS;/*start the model*/
model
{
  for (i in 1:N) {
      for (k in 1:T) {
                  p[i,k] <- c[k]+((1-c[k])*(phi(a[k]*theta[i]-b[k])))
                  y[i,k] ~ dbern(p[i,k])
         }

      theta1[i] ~ dbeta(9.3,5)
            }
             for (i in 1:N){
                  theta[i] <- ((theta1[i]*6)-3)
    }

   for (k in 1:T) {
      a[k] ~ dlnorm(0,0.5)
      b[k] ~ dnorm(0,2)
      c[k] <- 0.17
    }

#b1 ~ dpois(2)
#b2 ~ dpois(2)
#b11 <- b1+1
#b12 <- b2+1

#b11 ~ dunif(2,100)
#b12 ~ dunif(2,100)
}

;
RUN;
DATA _NULL_;
  SET model;
  FILE model;
  PUT model;
RUN;

/*Starting values*/
DATA _NULL_;
FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtini9.txt";
PUT
"list(a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1,1),"
```

```
PUT
"(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0)))"

RUN;

/*Scripts to run WinBUGS*/
FILENAME runirt9 'c:\Winbugs\WinBUGS\runirt9.txt';
DATA _NULL_;
  FILE runirt9;
  PUT@1 "display('log')";
  PUT@1 "check('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtmodel9.txt')" ;
  PUT@1 "data('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata9.txt')";
  PUT@1 "compile(1)";
  PUT@1 "inits(1, 'C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtini9.txt')";
  PUT@1 "gen.inits()";
  PUT@1 "update(1000)";
  PUT@1 "set(a)";
  PUT@1 "set(b)";
  PUT@1 "set(theta)";
  PUT@1 "update(5000)";
  PUT@1 "stats(*)";
  PUT@1 "save('C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtlog9.txt')";
  PUT@1  "quit()";
RUN;


DATA _NULL_;
FILE "C:\Winbugs\runirt.bat";
PUT '"C:\Winbugs\WinBUGS\WinBUGS.exe" /PAR runirt9.txt';
PUT 'exit';
RUN;
options linesize=72;
%macro runsimirt;
   %let n=1;
      %do %while(&n <= 50);

%macro gen(sz1=, length=, SIM=);

      libname temp "C:\";

      %do i=1 %to 5; *number of replications;

/* SIM Data Generation*/
      data SIM_IRT9;
            array item item1-item&length;
            array a a1-a&length;  /*tlength= test length*/
            array b b1-b&length;
            array c c1-c&length;
            array p p1-p&length;
            array x x1-x&length;
            array y y1-y&length;
            infile 'G:\IRT HW\SAS\NM_SPE_40Occasion3.dat'; /*parameter
a b and c reference*/

                  do over a;
                     input item a b c;
                  end;
```

```
                        do k=1 to &sz1;    /* set sample size*/
                            t1=-1+sqrt(1)*rannor(0);   /*Used in UNFAVORABLE
condition for occasion three for BILOG-MG and WinBUGS (1X1X2= 2
situations) */
                    SIM=&SIM;
                        t1 = -.257844888 + 1.032194765*t1 +
.257844888*t1**2 + -.035001562*t1**3; /* Fleisman non-normality
transformation skew = 1.2, kurtosis = 1.4 */
                do over p;
                        x=ranuni(0);
                        p=c+((1-c)/(1+exp(-1.7*a*(t1-b))));
                        if x le p then y = 1; else y=0;
                    end; output;
                    end;
         run;

        data temp.SIM&SIM&i; set SIM_IRT9; run;

data Work.simcsv9;
      set Work.SIM_IRT9 (keep=y1-y40); /* Create dichotomous file for
BILOG and WinBUGS analysis */
run;


  proc printto; run;
%end;
%mend gen;
%gen (sz1=500, length=40, SIM=9);
proc export data=work.sim_irt9 (keep=k t1 SIM)
   outfile="c:\SIMIRTBLM\sim_irt9&n..csv"
   dbms=csv
   replace;
run;
%_sexport(data=Simcsv9,
file ='C:\Winbugs\WinBUGS\Bugs\SAStoWB/irtdata9.txt',
var =y1-y40);
run;


/*Run WinBUGS*/
DATA _NULL_;
X "C:\Winbugs\runirt.bat";
RUN;
QUIT;



DATA log;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\irtlog9.txt" TRUNCOVER ;
INPUT log $90.;

RUN;

PROC PRINT DATA=log;
RUN;
```

```
DM OUTPUT 'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults9&n..txt"';
DM LOG    'FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults9&n..log"';

/*Analyze the results*/
DATA temp;
INFILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\allresults9&n..txt" TRUNCOVER ;
INPUT all $90.;

FILE "C:\Winbugs\WinBUGS\Bugs\SAStoWB\temp9&n..txt";
PUT all;
RUN;


%macro MakeDatFiles (Datain= );
/*  The original file, Sim_IRT&Sim is used in each situation; therefore
the data given
   to BILOG and WinBUGS is the same within each situation. The format
of the file is changed
   to accomodate the different programs. The original file was
replicated 50 times. */

Data _NULL_ ;
     SET  Work.&DataIn ;
     ID = 5 ;
     FILE "C:\U9&n..dat" ;
     PUT    @1  ID  @2  (y1-y40) (1.) ;
run ;


%mend  MakeDatFiles ;
%MakeDatFiles (datain=Sim_IRT9)
%macro RepBILOG1(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;

data _null_ ;
     file "C:\Program Files\BILOGMG\UNFAV9&n..blm ";
put
%str(">TITLE    Unfavorable Condition;" ) /
"         Situation three  ;        "        /
%str(">GLOBAL DFName= 'C:\U9&n..dat', NPArm=3; ")/
%str(">LENGTH NITems = (40);")/
%str(">INPUT NTOtal = 40, NALt = 7, NIDch = 1 ; " ) /
%str(">Items INUMBER=(1(1)40), INAMES=(Y1(1)Y40);"  ) /
%str(">TEST1 TNAme = UNFAV9 , GUESS = (0.1700(0)40);")/
%str("(1A1,40A1)   ") /
%str(">CALIB ACCel=1.000, cycles =25, GPRior, READPRIOR;" ) /
%str(">PRIORS1 ALPHA = (100(0)40), BETA = (500(0)40);" ) /
%str(">SCORE ;   ") ;

run;
%end;
%mend RepBILOG1;
data %RepBILOG1(Cond=UNFAV,Occ=9,Rep=1);
%macro RepBILOG2(Cond= ,Occ= , Rep= );
   %do q=0 %to &Rep;
```

```
        /*****      End Algina
*********/
    data _null_ ;

            file "C:\SIMIRTBLM\Unfavorable9&n..bat";
    put
    %str("cd\program files\bilogmg" )      /
    %str("blm1 UNFAV9&n")/
    %str("blm2 UNFAV9&n" )/
    %str("blm3 UNFAV9&n") /
    %str("exit") /
    run;
    /*****                    Call Bilog
****/

    dm " x C:\SIMIRTBLM\Unfavorable9&n..bat" ;

    *****            End BILOG-MG call
********* ;


    run ;
%end;
%mend RepBILOG2;
data %RepBILOG2(Cond=UNFAV,Occ=9,Rep=1);

%let n=%eval(&n+1);
   %end;
%mend runsimirt;
%runsimirt;
```

One SAS Macro from Sparapani was used in the simulation to change the SAS file

format to a WinBUGS format. The SAS Macro file is below:

```
%put NOTE: You have called the macro _SEXPORT, 2004-05-19.;
%put NOTE: Copyright (c) 2004 Rodney Sparapani;
%put;

/*
Author:  Rodney Sparapani <rsparapa@mcw.edu>
Created: 2004-00-00

This file is free software; you can redistribute it and/or modify
it under the terms of the GNU General Public License as published by
the Free Software Foundation; either version 2, or (at your option)
any later version.

This file is distributed in the hope that it will be useful,
but WITHOUT ANY WARRANTY; without even the implied warranty of
MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
GNU General Public License for more details.

You should have received a copy of the GNU General Public License
```

```
along with this file; see the file COPYING.  If not, write to
the Free Software Foundation, 675 Mass Ave, Cambridge, MA 02139, USA.

In short: you may use this file any way you like, as long as you
don't charge money for it, remove this notice, or hold anyone liable
for its results.
*/

/* _SEXPORT Documentation
    Export a BUGS or R/S+ "structure" data file from a SAS Dataset.

    REQUIRED Parameters

    FILE=                   "structure" file to create
    VAR=                    list of arrays to be included
                            Ex. two arrays: VAR=one1-one5 two1-two8

    Specific OPTIONAL Parameters


    DATA=_LAST_             default SAS dataset used
    FORMAT=best12.          default format for variables
    LINESIZE=80             default line length
    LS=LINESIZE             alias
    OUT=DATA                default name of object
    R=                      by default create BUGS-style "structure"
                            if set to anything, create a S+/R-style
    SPLUS=R                 alias

    Common OPTIONAL Parameters

    LOG=                    set to /dev/null to turn off .log
*/

%macro _sexport(file=REQUIRED, var=REQUIRED, data=&syslast,
    format=best12., linesize=80, ls=&linesize, out=&data,
    r=, splus=&r, log=);

%_require(&file &var);

%let file=%scan(&file, 1, ''"");
%let splus=%length(&splus);

%if %length(&log) %then %_printto(log=&log);

%local nobs i j var0 name;
%let nobs=%_nobs(data=&data);
%let var0=%_count(&var);

proc format;
    value __na
        .='NA'
        other=[&format]
    ;
run;

data _null_;
```

```
        file "&file" linesize=&ls;

        put %if &splus %then "%lowcase(%trim(&out)) <- ";
            "list(N=&nobs, T=%_count(%_list(&&var&i)), " @;

    %do i=1 %to &var0;
        %let var&i=%scan(&var, &i, %str( ));
        %let j=%_indexc(&&var&i, 0123456789);
        %let name=%_substr(&&var&i, 1, &j-1);

        put "%_tr(&name, from=_, to=.) = structure(.Data = c(";

            do i=1 to &nobs;
                set &data(keep=&&var&i) point=i;

                put (&&var&i) (__na.-r ',') @;

                if i=&nobs then put '), ' @;
                else put ',' @;
            end;

            put %if &splus %then ".Dim = c(%_count(%_list(&&var&i)),
&nobs))";
                %else ".Dim = c(&nobs, %_count(%_list(&&var&i))))";
                %if &i=&var0 %then ')';
                %else ',';
            ;
    %end;

        stop;
    run;

    %if %length(&log) %then %_printto;

%mend _sexport;


%*VALIDATION TEST STREAM;
/* un-comment to re-validate

data matrix;
    input col1-col5;
    datalines;
    1 2 3 4 5
    6 7 8 9 10
run;

%_sexport(data=matrix, var=col1-col5, file=_sexport.txt);
%_sexport(data=matrix, var=col1-col5, file=_sexport.r, r=1);

*/
```

APPENDIX C

Item Parameters for Each Occasion (from Hambleton, Swaminathan, & Rogers, 1991, p

156)

Occasion One Item Parameters

#	a	b	c
1	.84	.58	.17
2	.91	-.51	.17
3	1.10	-.28	.17
4	.37	-1.69	.17
5	.69	-.97	.17
6	.64	1.11	.17
7	.79	-.90	.17
8	.67	-.09	.17
9	.53	-1.36	.17
10	.02	.00	.17
11	.65	-.14	.17
12	.37	-.11	.17
13	.56	.22	.17
14	.88	-.67	.17
15	1.30	.76	.17
16	1.53	-.04	.17
17	.66	-1.32	.17
18	.66	-.32	.17
19	1.23	.51	.17
20	1.00	-3.00	.17

#	a	b	c
21	.59	-.68	.17
22	.74	-.46	.17
23	.74	.63	.17
24	1.25	-.64	.17
25	.98	-.12	.17
26	.65	1.18	.17
27	1.08	-.52	.17
28	.91	-.12	.17
29	.62	-.91	.17
30	2.00	.00	.17
31	.58	.08	.17
32	.68	-.90	.17
33	.87	.78	.17
34	1.10	.03	.17
35	1.04	1.40	.17
36	.94	-.50	.17
37	.63	1.54	.17
38	.64	-1.10	.17
39	.69	-.09	.17
40	1.00	3.00	.17

"Poor items"

Occasion Two Item Parameters

#	a	b	c
1	.84	.58	.17
2	.91	-.51	.17
3	1.10	-.28	.17
4	.37	-1.69	.17
5	.69	-.97	.17
6	.64	1.11	.17
7	.79	-.90	.17
8	.67	-.09	.17
9	.53	-1.36	.17
10	.62	.19	.17
11	.65	-.14	.17
12	.37	-.11	.17
13	.56	.22	.17
14	.88	-.67	.17
15	1.30	.76	.17
16	1.53	-.04	.17
17	.66	-1.32	.17
18	.66	-.32	.17
19	1.23	.51	.17
20	.73	-.31	.17

#	a	b	c
21	.59	-.68	.17
22	.74	-.46	.17
23	.74	.63	.17
24	1.25	-.64	.17
25	.98	-.12	.17
26	.65	1.18	.17
27	1.08	-.52	.17
28	.91	-.12	.17
29	.62	-.91	.17
30	.86	-.81	.17
31	.58	.08	.17
32	.68	-.90	.17
33	.87	.78	.17
34	1.10	.03	.17
35	1.04	1.40	.17
36	.94	-.50	.17
37	.63	1.54	.17
38	.64	-1.10	.17
39	.69	-.09	.17
40	1.00	3.00	.17

"Poor Item"

Occasion Three Item Parameters

#	a	b	c
1	.84	.58	.17
2	.91	-.51	.17
3	1.10	-.28	.17
4	.37	-1.69	.17
5	.69	-.97	.17
6	.64	1.11	.17
7	.79	-.90	.17
8	.67	-.09	.17
9	.53	-1.36	.17
10	.62	.19	.17
11	.65	-.14	.17
12	.37	-.11	.17
13	.56	.22	.17
14	.88	-.67	.17
15	1.30	.76	.17
16	1.53	-.04	.17
17	.66	-1.32	.17
18	.66	-.32	.17
19	1.23	.51	.17
20	.73	-.31	.17

#	a	b	c
21	.59	-.68	.17
22	.74	-.46	.17
23	.74	.63	.17
24	1.25	-.64	.17
25	.98	-.12	.17
26	.65	1.18	.17
27	1.08	-.52	.17
28	.91	-.12	.17
29	.62	-.91	.17
30	.86	-.81	.17
31	.58	.08	.17
32	.68	-.90	.17
33	.87	.78	.17
34	1.10	.03	.17
35	1.04	1.40	.17
36	.94	-.50	.17
37	.63	1.54	.17
38	.64	-1.10	.17
39	.69	-.09	.17
40	1.50	.58	.17