ABSTRACT

Title of Document: REWEIGHTING DATA IN THE SPIRIT OF TUKEY: USING BAYESIAN POSTERIOR PROBABILITIES AS RASCH RESIDUALS FOR STUDYING MISFIT

William R. Dardick, Doctor of Philosophy, 2010

Directed By: Professor Robert J. Mislevy
Department of Measurement, Statistics & Evaluation

A new variant of the iterative "data = fit + residual" data-analytical approach described by Mosteller and Tukey is proposed and implemented in the context of item response theory psychometric models. Posterior probabilities from a Bayesian mixture model of a Rasch item response theory model and an unscalable latent class are expressed as weights for the original data. The data weighted by the units' posterior probabilities for the unscalable class is used for further exploration of structures. Data were generated in accordance with departures from the Rasch model that have been studied in the literature. Factor analysis models are compared with the original data and the data as reweighted by the posterior probabilities for the unscalable class. Eigenvalues are compared with Horn's parallel analysis corresponding to each class of factor models to determine the number of factors in a dataset. In comparing two weighted data sets, the Rasch weighted data and the data

were considered unscalable, and clear differences are manifest. Pattern types are detected for the Rasch baselines that have different patterns than that of random or systematic contamination. The Rasch baseline patterns are strongest around item difficulties that are closest to the mean generating value of θ's. Patterns in baseline conditions are weaker as they depart from a item difficulty of zero and move toward extreme values of ±6. The random contamination factor patterns are typically flat and near zero regardless of the item difficulty with which it is associated. Systematic contamination using reversed Rasch generated data produces alternate patterns to the Rasch baseline condition and in some conditions shows an opposite effect when compared to the Rasch patterns. Differences can also be detected within the residually weighted data between the Rasch generated subtest and contaminated subtest. In conditions that have identified factors, the Rasch subtest often had Rasch patterns and the contaminated subtest has some form of random/flat or systematic/reversed pattern.

REWEIGHTING DATA IN THE SPIRIT OF TUKEY: USING BAYESIAN POSTERIOR PROBABILITIES AS RASCH RESIDUALS FOR STUDYING MISFIT

By

William R. Dardick

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:
Professor Robert J. Mislevy, Chair
Professor Gregory R. Hancock
Professor George Macready
Professor Paul Hanges
Professor Jeffrey Harring

Dedication

To my parents who have encouraged me throughout my life.

Acknowledgements

My deepest appreciation to Dr. Robert J. Mislevy for allowing me to investigate my diverse interests throughout the years and for his wisdom, patience and guidance in writing this dissertation.

Table of Contents

List of Tables

List of Figures

Chapter 1: Purpose and Rationale

Background

In educational assessments and sample surveys of opinions, there are expected patterns in the data that form the basis of our hypothesis testing. We draw inference from patterns we can anticipate; however, other patterns exist in the data, which may be expected or unexpected depending on the nature of the pattern. These other patterns may arise through systematic or idiosyncratic approaches of the respondents.

Mosteller and Tukey

Mosteller and Tukey (1977) express data as the degree of fit plus residual. For most predictive statistical analysis, such as a regression model, one can obtain a value that is indicative of the fit to that model, and anything left over, positive or negative, is the residual for that case. Mosteller and Tukey discuss the examination of residuals for detecting patterns. In the case of the current investigation, Item Response Theory (IRT) is the fit which can than be expressed as the degree of model fit, the Rasch IRT model (described in Chapter 3), plus what is left over that did not fit the IRT model (e.g., Smith, 1986, 1988). In the current investigation, the idealized model is fit while the residual includes everything not consistent with that model, but with a different approach to differentiating fit and residual. Specifically, this approach is inspired by the perspective of mixture modeling. The data will be re-expressed in terms of weights to examine the data by splitting it into the weighted proportion that fits both the model and that which is misfit from the model.

In the foundational book, *Data Analysis and Regression*, Mosteller and Tukey (1977) set aside the final chapter in the book to examine regression residuals. They effectively come up with a set of guidelines for examining regression residuals that can be very helpful in exploring if a better fit exists. This method is found more explicitly in *Understanding Robust and Exploratory Data Analysis* (Hoaglin, Mosteller and Tukey, 1983) and broken up into four components: resistance, residuals, re-expression, and revelation (Hoaglin, 2003). Resistance occurs in the data when local misfit is not revealed. The residual is expressed as the data minus the fit of the specific model. Re-expression is typically some form of transformation to aid in analysis. The data is then presented through display for revelation of expected or unexpected characteristics.

Mosteller and Tukey (1977) have the idea that the model is fit to the data and that the patterns associated with this model can be removed by examining the residuals. In removing the model fit, that portion of the data is stripped away and re-exploration of what is left over in the residual portion of the data begins. The idea here is that it will be easier to examine the remainder of misfit when the portion of the data that is fit to the model is removed. Mosteller and Tukey continue to explore the data in an iterative process, examining and searching for patterns and occasionally observing recognizable patterns that can further explain what is left over. Similar to the approach marshaled by Mosteller and Tukey, the current project investigates this notion that the fitted portion of the data can be removed in order to explore the portion of the data that does not fit the model.

In the spirit of the method set forth by Mosteller and Tukey (1977), the current investigation will explore the original data using residuals to look for suspicious patterns that could indicate an alternative and better fit. If one re-expresses the original data as residual fit to the Rasch class, what is left over when the Rasch is fit will be in the unscalable class. Examining what is left over in this sense is similar to what is done in regression. We examine what is left over once the model removes the perfectly fitting portion of the data in the Rasch class.

This investigation differs from the Mosteller and Tukey approach in its notion of misfit. Their conception of misfit is that data = fit + misfit and is a residual to the model in the traditional aspects of a regression. Mosteller and Tukey fit a linear regression model to the data and calculate predicted values and the difference of the predicted value from the actual value being observed. This difference is the residual. The residual is what is left over when the regression function is fit. Along this conceptual framework of data= fit + misfit, the Rasch model may be fit to the data but a known value to generate a residual for an individual's ability does not exist. Here, we examine the model through a mixture perspective using the structure of a latent class model. An unscalable class is used as a type of residual catch all.

The present investigation extends work done by Mislevy and Verhelst (1990) that shows how individuals can employ different strategies to respond to items even though the strategy used by each individual is unknown. The strategy used for each response is inferred and a separate IRT model is fit for each strategy component in the mixture. In order to calculate a conditional estimate of the respondent's ability given the strategy they are likely to be using, Mislevy and Verhelst used probabilities

that a respondent was employing a specific strategy. One of the examples Mislevy and Verhelst used posited two classes: a Rasch model for students engaged in the tasks, and a random-guessing class.

What follows are three additional examples of analogous models. Extending work by Goodman (1975), Dayton and Macready (1980) proposed a latent class model with intrinsically unscalable subjects. Andreassen et al. (1987) used a similar strategy in a medical diagnosis example in which the states of the latent "disease" variable included several possible states of the suspected diseases, including a "normal" state, and an "unknown" state with independent probabilities. Patients with pathologies other than those anticipated would have high posterior probabilities in the unknown class. Yamamoto, (1987,1989,1995) employed a HYBRID model and extended it for diagnosing test speediness (1990, 1995) with a discrete latent class model and Rasch.

Another method for examining Rasch mixtures was proposed by Rost (1990). Rost's work with the Rasch model used latent class analysis (LCA) to conceptually split data sets and permit different parameters for a Rasch model within each latent class. The mixed Rasch model gives an alterative to testing the fit of the Rasch. Using LCA as this conceptual split for Rasch permits the latent classes to form the basis of comparison regarding parameters instead of a typical available variable (such as score, age, or gender) with known criteria that may not make a meaningful qualitative difference for research.

A brief philosophical examination into exploratory research

When we move from confirmation of theory to investigation of data in an exploratory process, we no longer have the same scientific goal. We have moved from an a priori framework of hypothesis testing in which statistical significance testing is appropriate to an exploratory method that should be used at a completely different stage of scientific investigation. According to Gorsuch (1983), if any part of a model is based on the data then we are sure to fit the data, even if it is only by chance. Probability levels should not be reported or taken seriously. Statistical testing for significance and the use of confidence intervals are only valid in the context of prior theoretical development.

When we fail to reject the null hypothesis then we are saying that the results of our analysis are not statistically significant. In basic foundations (Minium, King, & Bear, 1993) we learn that we should be find direction and significance level before we even collect the data. In an exploratory analysis one cannot attribute differences, which statistical theory indicates has probably occurred through chance, to some new hypothesis generated after seeing the results. The exploratory mode of analysis requires "greater caution" (Dayton, 1998) because we do not know what significance levels mean when we fit models to data. Even if we reject the null hypothesis, great caution is taken in statistical decision theory, stating that our alternative hypothesis is one of many possible explanations for some difference in the data. If we reason after the fact and develop a model from the data then we are stepping away from an already cautious justification of theory.

Immanuel Kant (1778), in The *Critique of Pure Reason* discusses the role prior reasoning plays in our scientific investigations:

> When Galileo caused balls…to roll down an inclined plane; when Torricelli made the air carry a weight of which he had calculated beforehand… a light broke upon all students of nature. They learned that reason has insight only into that which it produces after a plan of its own… constraining nature to give answer to questions of reason's own determining. Accidental observations…can never be made to yield a necessary law, which alone reason is concerned to discover (Kant, 1778).

This idea summarized here by Kant, is foundational in our pursuit of science. We need to have well reasoned prior theory to test against the observations we gather.

Albert Einstein certainly understood the division of exploratory and confirmatory analysis. Einstein's (1916) theory of relativity was born out of imagination and mental exploration. When it came time to test and confirm his general theory of relativity he had a very difficult time. There was the competing theory of Newtonian mechanics, which was in "far-reaching agreement" (Einstein, 1961) with his theory of relativity. Based on the general theory of relativity, it would be very hard to deduce an a priori hypothesis and then test it against another competing theory that is in such close agreement. Deflection of light by gravitational fields were accomplished and tested.

In statistics we also construct models to explain theory, just as the physicist does. Atomic theory was first proposed over 2000 years ago by Democritus. Exploration and confirmation of theory lead us to present day quantum theory (Bohm,

1980). Models of atoms are now perceived as waves/particles with probability fields. Science very reluctantly (Capra, 1982) turned to probability models to explain the atomic nature of the world. The need to turn to such models of the universe was brought about by the combinations of exploratory and confirmatory research, each playing its proper role in the construction of a theoretical framework.

When our hypotheses fail or are untestable in a traditional sense, where does this leave us, as researchers? We begin developing new ideas that may be in conflict with existing paradigms (Kuhn, 1996) of scientific knowledge, colliding with the rigorous barriers of a current theoretical framework that we would not notice during normal science. It is only in times of conflict, like those that could arise when an expected theory fails or an unexpected one prevails, that we take much notice of the rules in which we work.

The current investigation is a preliminary examination of the data with structured hypotheses that is likely to lead to additional hypotheses that can be further explored in future research. Investigating the data structures through preliminary analysis prior to running the complete research will insure that reasonable techniques are used on the data to find meaning and evidence.

Novelty of study

The goal of the current investigation is to examine residuals from a Bayesian Rasch IRT model for patterns of responses that do not fit the model, due to utilization of different underlying strategies employed by respondents. Notably, "the residual" is a reweighted facsimile of the original data set: the same response vectors, but with cases weighted in proportion to the degree to which they do not fit the posited model.

The investigation will use the Rasch model as a filter and obtain residuals to reweight the data. The residually reweighted data will then be examined for what is left behind once a model, which is unsatisfactory for the complete dataset in the sense of Mislevy and Verhelst (1990), has been employed.

Mosteller and Tukey provide a road map of exploring residuals, particularly those in the regression context. Some similarities exist between the process they follow and the variant of residual exploration in this research. The first similarity is the disbelief of perfect fit. The model will fit the data to some degree regardless of the underlying distributions. The fit of the model can be extracted and what is left behind can be explored. These ideas run parallel to this study. Data is transformed and reevaluated to explore for additional anomalous and probable patterns. Mosteller and Tukey explore the residual to determine what type of transformation of the original data will simplify and clarify the analysis. The differences arise here where in this investigation: data is reweighted by the posterior probability of class membership – specifically by the probability of a residual class rather than a class defined by the "model" class corresponding to the Rasch model – and then transformed and explored. Each respondent receives a weight from 0 to 1 and the new data is a weighted copy of the original dataset. The same data vectors exist but each individual case represents only a proportion of the original case. In the Mosteller and Tukey method data is transformed and a new data set is manifest with each case still representing the same proportion as in the original data, while each data point is the difference between the original data point and the value predicted by the model. Their transformation changes the data values while maintaining case proportionality. This

investigation changes case proportionality and maintains the original data values. If one where to apply this method to regression, instead of transforming the data by a square root or logarithmic function, the data would be weighted by the regression's residual size proportionate to the entire data.

Overview of the study

In the current investigation data will be generated in accordance with an IRT model. The baseline condition data will be generated strictly in accordance with the Rasch model, while in others conditions data will be contaminated in some way, as suggested by earlier studies in the literature about person fit analysis.

Posterior probabilities from a Bayesian mixture model of a Rasch and an unscalable latent class are expressed as weights for the original data. The data weighted by the unscalable class is used for exploration of further structure, in a new variant of the iterative "data = fit + residual" data-analytical approach described by Mosteller and Tukey. Exploratory factor analysis models using tetrachoric correlations are used to determine if the contamination can be detected in the generated data. Factor structures are evaluated to determine if, on average, systematic differences are still manifest after the data has been weighted by the unscalable class or "residually reweighted" data set.

This dissertation will present information on: the review of relevant literature, methods involved in the current investigates, preliminary analysis, results of the research, and, lastly, conclusions. Chapter 2 provides a literature review of pertinent topics. Background is presented in the literature review for the Rasch model and

departures from the Rasch model. Types of potential misfit are identified and used later in the method chapter. Models and topics relevant to generating and estimating data for this method are reviewed. Examples of topics include: Latent Class Analysis, Mixture Item Response Theory, Factor Analysis and Bayesian estimation.

Chapter 3 outlines the method for this research. The mixture Bayesian Rasch model for estimating the data is presented along with generating conditions for the configuration of the data. Fixed and manipulated factors, including the baseline test and subtests are described. The simulation process for data generation, transformations and analysis using the SAS and Winbugs computer programs is explained. Evaluation and analytical procedures are arranged to: assess residuals, transform weighted and unweighted data through the use of orthogonal principal components analysis using tetrachoric correlation, determine significance through the use of an effect size for Wilks' lambda and confidence intervals, graph comparison between Rasch and residually weighted data and visually compare subtests within the residual data. Hypotheses are generated for all analysis to theorize expectations and directionality of the results.

Chapter 4 is the preliminary research and investigation of the model used in this investigation. Winbugs is used to estimate four conditions and models are evaluated for accuracy and parsimony. From this investigation fixed number of burn-in cycles and estimation cycles are set. Priors and sample size are determined and held constant for all conditions.

Chapter 5 is the preliminary analysis to determine the viability of the research to uncover patterns for Rasch and alternative strategies. This chapter uses five

conditions from the full analysis and examines them using the Winbugs estimation model from Chapter 4 and evaluates these conditions using the method prescribed in Chapter 3.

Chapter 6 presents results from the full analysis. Residuals are examined to determine if they are of sufficient size. Eigenvalues for weighted and unweighted data are investigated using horns parallel analysis as a determinant for number of factors. MANOVA, confidence intervals and graphs compare Rasch and residually weighted data for differences in factor patterns. Residually weighted data is explored to determine visual differences in patterns for the Rasch and contaminated subtests.

Chapter 7 provides a final discussion of the results, research implication and directions for future research. Conclusions regarding the results are expounded on and direction is provided for future work into using Bayesian posteriors in a similar manner.

Chapter 2: Review of the Literature

This chapter provides a literature review related to the current investigation. The areas under review include Item Response Theory (IRT), Latent Class Analysis (LCA), Bayesian analysis, and analysis of person-fit in relation to types of departures from the Rasch model.

Item Response Theory (IRT)

Item Response theory (IRT) models the probability of a person's response to an item as a function of one or more parameters for the person's ability and one or more parameters for characteristics of the item, such as its difficulty and sensitivity to ability. Mislevy and Verhelst (1990) review the general form of the IRT model, with "the probability of response x_{ij} (*1* if correct, *0* if not) from Subject I to item j is given by":

$$p(x_{ij} \mid \theta_i, \beta_j) = \left[f(\theta_i, \beta_j)\right]^{x_{ij}} \left[1 - f(\theta_i, \beta_j)\right]^{1-x_{ij}}$$

where the person parameter, θ_i associated with subject *i* and item parameter β_j associated with item *j* and f is "a known, twice differentiable, function whose range is the unit interval" (Mislevy & Verhelst, 1990).

Assuming local independence:

$$p(x_{ij} \mid \theta_i, \beta_j) = \prod_{j=1}^{n} p(x_{ij} \mid \theta_i, \beta_j)$$

The general formula for a response vector as the product of item-by-item probabilities. The general form of a one parameter item response model:

$$P\left(x_{ij} \mid \theta_i, \beta_j(\alpha)\right) = \frac{\exp\left[x_{ij}\left(\theta_i - b_j\right)\right]}{1 + \exp\left(\theta_i - b_j\right)}.$$

In IRT, the main trait(s) or factor(s) in a model ideally accounts for the responses one would give to items that should measure that ability (Hambleton et. al, 1991). The probability that an examinee will respond correctly to an item increases as the ability increases, as represented by that factor. The assumptions of correct dimensionality and local independence are intertwined. Correct dimensionality is the assumption of all IRT models, and in a one-dimensional model that assumption is unidimensionality. Unidimensionality assumes that only one dimension of ability is being measured by the items of the test (θ).

There is always some level of departure from the unidimensionality assumption, as psychological and educational research does not take place in a vacuum. All abilities of the test takers are brought to bear when they take the assessment. Minor departure from unidimensional ability at work on a test might be motivation, anxiety, guessing tendencies, response speed, and additional cognitive skills (Hambleton et. al, 1991). The application for the unidimensionality assumption to be met is that only one dominant component is being measured on the test. However, perfect unidimensionality is not probable. The first assumption of correct dimensionality is essential in order for the second assumption of local independence to be met. When the assumption of dimensionality is met, so is the assumption of local independence.

Mislevy and Chang (2000) present local independence as the "cornerstone of Item Response Theory (IRT)". Local independence is necessary regardless of the

dimensionality of the test. The goal is to have no relationship remain between the test takers' item responses after the ability being measured has been removed. When ability is held constant, responses should be statistically independent. This will be true if the correct dimensionality has been specified in the model. When Θ is the ability of interest and it is removed, all remaining responses should be unrelated. This process is similar to looking at factors in a Factor Analysis (FA) model. When local independence occurs in the data, one dominant factor is present that can account for much of the variability in unidimensional data. Other factors are small and insignificant in comparison and are merely random fluctuation, as some minor disturbances from departure are expected in any model.

In order for local independence to hold with multiple dimensions, the entire latent space needs to be specified. There are several ways that latent space can be understood. There could be multiple dimensions at work on items or there could be subpopulations of abilities. One can view the multiple aspects of ability as multivariate space and/or mixture of classes. When the underlying dimension is continuous the multidimensional approach is appropriate. If the underlying dimension is a discrete dimension then the mixture approach is suitable.

Departures from the Rasch Model

In conjunction with ideas of alternative strategies in testing, misfit of the Rasch model has been explored in relevant literature. The statistical properties, estimation algorithms, and statistical fit of data to the Rasch model have been explored (Kelderman, 1984). Looking to examine types of data that will misfit the Rasch model (Mead, 1976; Smith, 1986) typically conforms to some form of mixing

or multidimensional disturbance. This approach is considered to be detrimental because the desirable qualities of the Rasch model are then lost. The following three types of departures from the Rasch that are explored by Mead (1976): Random guessing, also called carelessness; Practice, which is related to how quickly one finishes the test; and finally, Bias. These data all violate the model's assumptions, but still show acceptable fit when the fit statistics are used to test these models.

Smith (1986, 1991) explored distributional properties of fit statistics through investigations based on thorough investigations of simulated data. Similar types of patterns to those Mead studied for departure from the Rasch model are used by Smith (1986), including: guessing, startup (fumbling), plodding, content interaction, sloppiness (carelessness), reanalysis, and random responses. Guessing, speed and item bias are considered amongst types of misfit. Guessing is a very common type of misfit modeled throughout Rasch research (Wainer & Wright, 1990; Mislevy & Verhelst, 1990). Wright (1995) examines Low Mean Squares (LMS). Misfit is represented by LMS inspired methods with departures from the Rasch model with less randomness than expected results

Effects of change on parameters, such as time effect, curriculum effect as shifts in θ or difficulty, have been conceptualized prior to this research (Mead 1976; Mislevy 1981). Mead (1976) refers to speed as both the need to warm up and the rushing influenced by lack of speed. Changing theta or changing difficulty can be shown to be mathematically the same if done systemically. If done for the same population or sub-population, a positive one standardized shift in item difficulty as a generating parameter is mathematically equivalent to a negative one shift in theta

generating parameters. Mead also examines bias as the same type of shift in parameters. Bias in this case is typical; in population A the items are relatively harder than in population B. This is said to have the same effect as practice or speed and speed is to be a special form of bias. Curriculum effect is also just a special form of bias. It is bias to look at two subpopulations with one having increased ability over the other with respect to some subsets of items but not others, due to their educational experiences. The underlying cause is philosophically different only insofar as one group is said to have special knowledge, or an increased ability. Curriculum is a special case of Mead's bias in which the mathematics remains identical and the discussion is pointed toward an increase in ability in some subpopulation for some subset of items.

Gentner and Gentner (1983) describe models of erroneous knowledge that can serve as an inferential framework. Serial and parallel combinations of resistors and batteries are posed to participants using different analogies. People are separated based on differencing analogy for electricity and respond differently based on the analogy. Those using the flowing water model performed well on the battery section and poorly on the resistor section, while those using the moving crowd model performed better on resistor section and poorly on the battery section. The model of their performance shows a reversal effect. Given the same questions using different analogy, respondents differ depending on the interaction of the analogy and the model.

In IRT the advantages hold when the assumption, principally independence, of IRT hold. (Hambleton, Swaminathan, & Rogers, 1991) Testing the fit of the model is

crucial to insure violations of the underlying assumptions are not severe. Goodness of fit studies (Divgi, 1986; Rogers & Hattie, 1987) were flawed considering their sensitivity to sample size (Hambleton, Swaminathan, & Rogers, 1991). The literature of item and person fit is valuable here even though the model of misfit is different in the current investigation. The literature on misfit statistics is not as useful for the present research in the context of mathematical theory or statistical fit, but as the theory underlying why the misfit exists within the data. The current investigation does not use misfit literature for diagnosis but for investigating common types of departures found in the investigation of IRT and particularly Rasch data.

Hambleton, Swaminathan, and Rogers (1991) discuss an approach to assessing data model fit as "designing and conducting a variety of analyses to detect expected types of misfit" As noted above, useful mechanisms that can produce misfit have been developed within this literature of departure from Rasch data.

In the current investigation, the simulated IRT mixture data represent two alternative strategies to responding to questionnaire style survey tools or examination assessments. The idea of "person fit" as laid out by Richard Smith (1986) traces the major themes focusing on the concept of believability of a pattern of responses from a person. Person fit statistics for the Rasch are shown to provide a useful framework to test person fit and believability. In comparing the ability of two Rasch statistics to detect disturbances, Smith simulates 10 sets of responses with 100 patterns per set for four guessing level of random responses 1/3, ½, 2/3, and all items. In addition, he generated 9 sets of 100 simulated patterns to study subsets. In situations in which the power of the test was low, the impact of measurement disturbance was also

considered low. Smith goes on to state that total and between statistics are required to detect all disturbances.

From Strategies to Factors: A Martial Example

Strategies can be translated into factors through expectation of the patterns that might arise from following one strategy over another. When one strategy is followed, a certain patterns of results are likely to occur for a given set of test items. That pattern of results will yield one factor model. Alternatively a different strategy may yield different patterns of results. Gentner and Gentner (1983) provide one example of reversed strategies from analogy in studying electricity. Let us take an alternative example from martial arts. Different strategies therefore yield different patterns of interrelationships among item responses, which are revealed in different results from factor analyses of the data. When it is not known a priori what strategies respondents may be using, the revealed factor patterns can be analyzed for hypothesis generation in light of what is known about the substance of the items and respondents' plausible distinctive ways of interacting with that substance.

In training modern style martial arts such as Taekwondo, there are many choices in training strategies. Currently it is very popular to train students for fighting in the Olympics for a point style combat sport. In training, students may learn to break boards, forms and patterns, move quickly in and out of range and always keep their hands down to protect their vital scoring targets. Alternatively, students of other Taekwondo schools may focus more on the martial aspect of the art. These students also learn to break boards, forms and patterns but spar with hands raised and strike to do maximum damage. In repeated test of breaking boards and moving through formal

movements students of both styles would have very comparable abilities. If we were to design an exam of 20 items, where items were forms and boards to break, students would perform based on their ability and both strategies could lead to a similar scale, such as a Rasch scale for both strategies. If we then moved to other arenas for examination we might see where alterative strategies would give rise to alterative factors. In the formal sport of Olympic style Taekwondo, a student trained to score points would likely have very high ability compared to a student taught combat. The sport trained student would score points and move in and out with footwork and hands protecting the chest before the combat student learned the elements of the contest. The sport trained martial artist would appear to have a high ability for elements or items in this venue while the combat student would have low ability. If the same two athletes where to meet and be tested on elements of combat a different and likely opposite result would occur. The combat trained athlete would ignore points and look to deliver deadly strikes attacks. The guard of the sports trained athlete would not serve well during real confrontation. The two would seem to have reversed in ability as in the reversed Rasch case in this current investigation. This alterative strategy, tested over several iterations could yield part of the subtest. In the end each strategy would have an underlying set of factor patters that may be similar when viewed through one subtest and very different on another test of skill.

The above martial arts exam could be envisioned as a 40 item assessment where the first subtest consisting of 20 items measured the skill of breaking boards and performing martial arts patterns. The second subtest could be items relating to sparing where the two strategies would perform differently. The underlying strategies

can give rise to the alternative factor patterns in the sparring subsection of the exam. Strategy one leads to high patterns of sparing on some items and low patterns in the others while the alternative strategy reverses the effect. This patterns could be picked up in unweighted data or data that had been filtered to attempt to remove one of the two underlying strategies.

Latent Class Analysis

Categorical latent variables distinguish among respondents in a Latent Class Analysis (LCA) model. In this type of latent structure model, both the observable and latent variables are categorical (Dayton, 1998). The covariance in the manifest variables can be explained by the latent variables (McCutcheon, 1987). In LCA, a mutually exclusive set of latent classes accounts for the distribution in a cross tabulation of the observable variables. The latent variable in LCA is defined such that there is a set of classes and for people within a given class, the manifest variables are independent (i.e., local independence). In a cross tabulation table, each of the cells is equal to the sum of the expected values over classes, weighted by the class-size, or mixing, proportions.

LCA is related to discrete mixture models and factor analysis (Dayton & Macready, 2007). The basic form of LCA is a mixture of product-multinomial distributions, but can also be perceived as a factor analytic model for categorical data.

Dayton and Macready (2007) represent the formula for a latent class model given the assumption of local (conditional) independence as:

$$\Pr(Y_i) = \sum_{c=1}^{C} \theta_c \Pr(Y_i \mid c),$$

"where Pr($\boldsymbol{Y}_i$|c) is the conditional probability for response vector, $\boldsymbol{Y}$i given latent class c." (Dayton & Macready, 2007).

As addressed by Dayton and Macready (1980), the concept of an unscalable class has a history in LCA. The mixture IRT model is a combination of the IRT and LCA models, and the mixture IRT model that is the focus of the present research uses a class that is analogous to an unscalable class.

Multidimensionality and Mixture IRT

In some instances, the assumption that ability is of one dimension is violated. Or, it can be that multiple abilities are at work in an assessment. This paper will present such cases as a mixture of underlying class ability in a Bayesian framework. It is important to have an understanding of common ways of interpreting and modeling dimensionality in an IRT model. In IRT two general frameworks for dealing with multiple factors are multidimensional IRT models (MIRT) and Mixture IRT models (Davier & Carstensen, 2007). The focus of the present research will be on a Mixture IRT model of a Rasch class and an unscalable class in a Bayesian framework, as presented by Mislevy and Verhelst (1990).

In exploring the literature of Mixture IRT, it is valuable to consider MIRT an alternative way to model data that cannot be satisfactorily fit with a unidimensional model. The MIRT model differs in its theoretical framework of underlying structure within the data. In MIRT models, all the person's ability parameters are continuous. A mixture IRT model has both continuous ability parameters within classes and a discrete parameter for class membership. Mixture IRT examines the latent classes that are posited to be inherent within the data (Rost, 1990). In Mixture IRT, there is an

assumption that ability is from multiple subclasses on multivariate categorical data. A powerful model is the general mixture IRT framework (Kelderman & Macready 1990, Mislevy & Verhelst, 1990) in examining latent groups.

Factor Mixture Analysis

A framework for different types of latent variable models by Muthen (2008) configures different variations of models within the general area of mixture modeling or modeling with categorical latent variables. The model overview (Muthen, 2008) includes continuous, categorical, and hybrid latent variables for both cross-sectional and longitudinal models. Hybrid models include continuous and categorical latent variables. Cross-sectional models have one instance in time, while longitudinal models have several time periods. Central to the current investigation is the branch of mixture modeling known as Factor Mixture Analysis (FMA). FMA is nested within the cross-sectional hybrid latent variable models division of mixture modeling.

A division of FMA in the IRT literature (Muthen 2008) includes Mislevy and Verhelst (1990), Mislevy and Wilson (1996), Wilson (1989), and Yamamoto and Gitomer (1993). The current investigation can also be nested in this branch of FMA. Much like one of the models used by Mislevy and Verhelst (1990), the current investigations looks at the case of a two class mixture model, one with a 1-parameter Rasch model and the other with an unscalable class. Although the model under investigation can easily be set in the framework of FMA, the question at hand is not a best fit question within that framework. The question is about what is left behind in a given selected model when fit is determined for each person. The model selected for

investigation is amongst the most popular models used in item response theory, the Rasch model.

Principal Components Analysis or Factor Analysis

There are many different extraction methods used in Factor Analysis (Gorsuch, 1983; Tabachnick & Fidell, 1996). A Comparison of factor procedures (Gorsuch, 1983) for the current investigation is reviewed for selection of an appropriate method. The theoretical rationale in the analysis concerns whether the technique chosen serves the purpose of data reduction or if the factors will be used to draw theoretical inferences regarding the constructs (Lawrence & Hancock, 1999). Components are mathematically abstract composites, while factors explain the theoretical underpinnings of the observed data. The major consideration for extraction is whether the current investigation is dealing with factors or components.

Mathematically, the difference between PCA and the common factor analytic technique of Principal Axes Factoring (PAF) lies in the positive diagonal of the correlation matrix (Tabachnick & Fidell, 1996). FA differs from PCA in that commonalities derived through an iterative procedure are used instead of those in the diagonal of the correlation matrix. When the principal factor procedure is used on the original matrix with those in the diagonal, the result is PCA (Gorsuch 1983). When the diagonal is a commonality, such as in principal axes factoring, the result is a factor. The goal of PCA is to reduce a large number of variables to a small number of components while extracting the maximum variance in the data with each component. FA is concerned only with covariance or commonalities: the variance in observed variables shared with other observed variables.

In the case of systematic contamination of the current investigation within the data that is proportionally reweighted to conform to the misfit, contamination can be detected by examining composite or factor structures. How we classify this contamination in the current investigation may not be important. In the current study we expect to find Rasch data, systematic and random contamination, as well as random uniqueness. The concepts underpinning the ability in the current investigation are generic in that argued general ability, special abilities of contamination, and uniqueness are each examined in the data. When looking at the Rasch model as having one underlying latent variable for a given ability, the data in the systematic contamination condition is generated to have a secondary latent special ability factor. These components in the simulation are generic and could hold equal meaning in math or reading ability. In factor analysis it is valuable for the final results to be understood as a conceptually clear construct, summarizing the interrelationships among variables to be understood (Gorsuch, 1983).

There are several factor extraction procedures (Tabachnick & Fidell, 1996) besides the traditional principal factor method used in the discussion between PCA and FA and involving the solution of the characteristic roots and vectors also known as eigenvalues and eigenvectors (Gorsuch, 1983). Image analysis is another variant using principal factors that has the same vector (Gorsuch, 1983; Harris, 1964) as PCA and PAF but different roots. These three solutions should have the same pattern of high and low loadings (Gorsuch, 1983). Minimum residual analysis and alpha factoring also use the same mathematics behind principal factoring but have alternative methods of obtaining communalities. The maximum likelihood factor

extraction method differs mathematically from those methods using the solution of characteristic roots and vectors. The maximum likelihood method estimates population parameters from sample statistics.

The current investigation will use PCA and look at rotated components instead of factors. Mathematically, the same principal factor method is used in PCA as in common factor analysis but the choice is not to reduce the variance to only commonality. Under the principal factor method, PCA is often used as a first step to FA as it can help determine the number of likely factors, variables to remove and the general factorability (Tabachnick & Fidell, 1996). In the current investigation, there is no need to make the argument that components underlying the data need to draw inference, as the main research questions are not regarding content development. The main question to examine is if any of the systematic difference is still left in the data after the Rasch model has been used as a filter. Devising explanations for what might be left could be performed on a case by case basis using real data instead of simulated data.

Bayesian estimation of Mixture IRT

Bayesian inference is described by Gelman, Carlin, Stern, and Rubin (2004) as a process that fits a probability model to data sets whose results are summarized by probability distributions on both parameters of the model and unobserved quantities. Inference comes from the data using probability models to quantify uncertainty for observed and unobserved information. Probability statements are made about parameter θ (Gelman, Carlin, Stern, & Rubin, 2004) and are conditional on observed values of y: p($\theta|y$). The joint probability distribution for θ and y allows for probability

statements about θ given y. Density functions are the prior distribution p(θ) and the sampling distribution p($y|\theta$):

$$p(\theta, y) = p(\theta)p(y \mid \theta)$$

Conditioning on the known data y:

$$p(\theta \mid y) = \frac{p(\theta)p(y \mid \theta)}{p(y)}$$

Omitting p(y) yields the unnormalized posterior density:

$$p(\theta \mid y) \propto p(\theta)p(y \mid \theta)$$

Unlike likelihood functions or estimation equations Bayesian estimation, using Markov-Chain Monte-Carlo (MCMC) estimation iterates through many draws in model parameter space (Rost 1990). Priors in conjunction with the algorithm are used for each parameter to estimate. MCMC permits extensions of more complex IRT models, as the number and estimation of parameters is not limited to the more conventional likelihood function. In this study, the Winbugs program (Lunn et al., 2000) will be used to carry out MCMC estimation.

Chapter 3: Method

Research Design

The goal of this research is to determine if the posterior residual used to reweigh data can extract out from the data significant differences in factor patterns between a Rasch and unscalable class. The current study will examine the use of posterior residuals from a Bayesian mixture model comprised of a Rasch and an unscalable latent class, expressed as weights for the original data, to explore factors that may still exist. Such factors include the use of multiple strategies to answer survey questions. In order to address this issue, multiple procedures will be used in conjunction with one another to accomplish the desired method. Replications will be simulated using SAS to conform to the Rasch model in the null condition with variants on the Rasch model adding in different levels of random and systematic pattern in other conditions.

The Model

In the investigation, simulated data will be generated to provide fit and intentional misfit (i.e., contamination) to a Rasch model. Using a Bayesian procedure via the Winbugs computer program, a mixture Rasch model will be fit to each dataset. Each person will have a value ranging from zero to one, in terms of posterior probabilities that the case accords with a class defined by the Rasch model response process, as opposed to a residual class represented by the independent product of .5 probabilities.-– in essence, a class of unscalables (Dayton & Macready, 1980). An IRT mixture model will be fit to all data conditions, in which one class consists of

examinees responding in accordance with the Rasch model, and the other class represents examinees responding randomly. The following expressions are adapted from Mislevy and Verhelst (1990):

The two class mixture model:

$$P(x_i \mid \xi) = \sum_{k=1}^{2} P(x_i \mid \phi_k = 1, \xi)\pi_{k,}$$

where π is the class proportion, ξ is the full vector of item parameters for all b's and c's, The two mixture model for estimation will contain two classes. The first class defined by the Rasch model and the second by Random Guessing.

Rasch model:

$$P(x \mid \phi_1 = 1, b_1, ..., b_J) = \int \prod_j \frac{\exp[x_j(\theta - b_j)]}{1 + \exp(\theta - b_j)} p(\theta) \partial\theta \,,$$

where b_j is the difficulty parameter for Item j under the Rasch model.

Random guessing:

$$P(x \mid \phi_2 = 1, c_1, ..., c_J) = \prod_j c_j^{x_j} (1 - c_j)^{1 - x_j} \,,$$

where c_j is a prespecified guessing parameter for Item j. It follows from the definition of the model that

$$P(x_i \mid \xi) = P(x \mid \phi_1 = 1, b_1, ..., b_J)\pi_1 + P(x \mid \phi_2 = 1, c_1, ..., c_J)\pi_2$$

$$= \left[\int \prod_j \frac{\exp[x_j(\theta - b_j)]}{1 + \exp(\theta - b_j)} p(\theta) \partial\theta \right] \pi_1 + \left[\prod_j c_j^{x_j} (1 - c_j)^{1 - x_j} \right] \pi_2 .$$

The posterior probability of a given subject with response vector x_i belonging to each of the two classes is thus calculated as follows:

Rasch class:

$$= \frac{\left[\int \prod_j \frac{\exp\left[x_j\left(\theta - b_j\right)\right]}{1+\exp\left(\theta - b_j\right)} p(\theta)\partial\theta\right]\pi_1}{\left[\int \prod_j \frac{\exp\left[x_j\left(\theta - b_j\right)\right]}{1+\exp\left(\theta - b_j\right)} p(\theta)\partial\theta\right]\pi_1 + \left[\prod_j c_j^{x_j}\left(1-c_j\right)^{1-x_j}\right]\pi_2}.$$

Random class:

$$= \frac{\left[\prod_j c_j^{x_j}\left(1-c_j\right)^{1-x_j}\right]\pi_2}{\left[\int \prod_j \frac{\exp\left[x_j\left(\theta - b_j\right)\right]}{1+\exp\left(\theta - b_j\right)} p(\theta)\partial\theta\right]\pi_1 + \left[\prod_j c_j^{x_j}\left(1-c_j\right)^{1-x_j}\right]\pi_2}.$$

The posterior probability estimate can be used to determine class each response pattern likely belongs in, Rasch or Non-Rasch, and to what degree each response pattern likely belongs. In this model, we can interpret an examinee's posterior probability of belonging to the Rasch class as his fit to the Rasch model, and the posterior probability of belonging to the random guessing class as a residual. Each subject's posterior probability of being in the unscalable class will be used to reweight the original data as the "residual" data set. The implementation of this weighting is discussed later in this chapter in the section

Responses that fit perfectly to the Rasch class will have a weight near one, while those patterns that accord very poorly with the Rasch model will have a weight nearer to zero. All other response patterns will have a value between zero and one from the posterior distribution representing the degree of fit with the Rasch model. Those that fit well with the Rasch model should be those respondents that were

generated to conform to the Rasch model. Those that differ are less like the theoretical model, and exploring this residual information can yield reasons for differences.

As specified above, by using the residual weights from the posterior distribution of the Bayesian Rasch model the original data will be re-examined through Factor analytic techniques. Tetrachoric correlations will be used instead of Pearson's correlation coefficient, as the data are dichotomous. Three exploratory factor analysis (EFA) models will be constructed, one with unweighted data, the other two are based on reweighted data of posterior probabilities belonging to the unscalable or Rasch class.

The first exploratory model will be of the unweighted dataset. It may be useful to return to this unweighted model after examining the other model, and it may be very useful as a baseline in determining the number of factors in the overall dataset.

The second EFA model is the model of interest for examining response patterns to the extent that they are not in accord with the expected one-dimensional model. This model uses reweighted data, where the weights for each case are posterior probabilities of belonging to the unscalable class; that is, a weighting of the data that best represents misfit to the Rasch model.

The third EFA model is reweighted and estimated to represent fit to the Rasch model. This model uses reweighted data, where the weights for each case are posterior probabilities of belonging to the Rasch class; that is, a weighting of the data that best represents fit to the Rasch model.

Rasch and Misfit Data

Data will be generated in the current investigation to conform to three different strategies: Rasch, random, and a Rasch reversal effect. The Rasch model will represent the expected response pattern. Randomly generated responses will represent the first patterns of misfit from the Rasch. These two strategies are commonly found in mixture IRT simulations (e.g., Smith, 1986, 1988 and Mislevy and Verhelst 1990). The random effect misfit will be characterized by random chance regardless of the difficulty of the item; the reversal effect represents special training, course of study, or analogy difference. The reversal alternative strategy is based on the reversal effect shown by Gentner and Gentner (1983). Under the same domain of knowledge, one can perform strong in one section and weak in another. Gentner and Gentner (1983) illustrate this situation using tasks on electrical circuits. One kind of task was relatively easier for students thinking of electrical flow in analogy to hydraulic flow. Another kind of task was relatively easier for students thinking of electrical flow in analogy to teeming crowds trying to get through different configurations of turnstyles. Based on analogy or special knowledge, the reversal effect will be simulated by reversing the difficulties of the items in one of the subsections for those in the contaminated condition

The first strategy will be represented by the one parameter logistic model, known also as the Rasch model. The Rasch equation will be used to model the data for respondents using the first strategy:

$$P(X_i = 1 \mid \theta) = \frac{e^{\theta - b_i}}{1 + e^{\theta - b_i}}$$

The second strategy will be represented by generating random responses. This "Random" strategy is common and sometimes based on issues such as the lack of care from a participant, rushing through responses, lack of knowledge, or some other underlying concept that would manifest an apparently random set of responses.

$$P(X_i = 1 \mid \theta) = .25$$

The probability, P, in the current investigation is set to .25 because the data will be generated assuming that four responses were possible for each item, as is typical for most multiple choice exams. If a participant responds randomly to item with C possible response categories the resulting probability of guessing is:

$$P(X_i = 1 \mid \theta) = \frac{1}{C}$$

Selecting .25 as the unscalable class is analogous to Andreassen et al. (1987) for diseases that are unknown. When looking at diseases in their Bayesian system they introduce a state of "other". This "other" condition helps to avoid strong favorable statements for any of the set of prespecified disease states, for when cases fit poorly. Instead of placing the case in the least poorly fitting prespecified condition, the network places this case in the "other" condition with a high probability. This is similar to the unscalable class in the current investigation. It is not that the data fits the unscalable condition but that it is placed into this class at all regardless of the cases structure. This can be an indicator that anomalous structure not accounted for by the hypothesized model(s) is present.

The third strategy will be represented by generating responses affected by special knowledge or analogy. The reversal effect will be simulated by reversing the difficulties on the generating parameters.

The Data Generating Factors

The subsequent section will provide an overview for fixed factors and manipulated factors in the simulation. This will facilitate the kinds of misfit from the Rasch model within the current investigation. Fixed factors in the research are: Theta distribution, number of items, sample size, and number of replications per cell. The four manipulated factors and the associated levels are: Misfit (2), Proportion within items misfit (contamination) (5), Scaling factor for scaling the item difficulty parameters in the test (4), Size of subtest (2). The following are tabular summaries of the fixed and manipulated factors. The next two sections of methodological discussion focus first on the selection fixed factors and later on the selection of levels for manipulated factors.

Table 3-1: Fixed Factors

Fixed Factor	
Theta Distribution	Normal(0,1)
Number of items	40
Sample size	500
replications per cell	50

Table 3-2: Manipulated Factors

Manipulated Factors	Levels					
Type of Misfit	3(2)	No Misfit, All Rasch	Reverse effect	Random effect		
Proportion within item Misfit (contamination)	5(4)	0	0.5	0.80	0..95	1
Scaling factor for scaling the difficulties	4	1/1	3/3	1/3	3/1	
Size of subtests	2	30/10	20/20			
Number of Cells	All Rasch Baseline = 1*1*4*2 + misfit 2*4*4*2=8+64=72 Total Cells					

Fixed Factors

Several factors are fixed in this study for all conditions of the simulation. Theta for the Rasch portions of the mixture will have a normal distribution with a mean of zero and a constant standard deviation of one. The sample size is fixed to 500 simulees and is discussed in the sample and subtest size section of this chapter. The number of replications per cell is set to 50. The sample size and replications that were tested in preliminary investigations are the discussions of these are deeply ingrained in chapters four and five. The number of items will be held constant at 40 with evidence supporting this number of items or less from previous research (e.g. Smith 1988, and Wright & Tennant 1996).

While searching the literature for a rational number of items to be used in the current investigation, several standardized tests of achievement were explored. The primary sources used to determine the number of items in the current investigation include the subtests of the GRE, TOEFL, NAEP and SAT.

ETS (2009) provides test information on the GRE's. The computer based test typically has 30 verbal and 28 quantitative questions. The paper based version of the GRE has 2 sections of verbal and quantitative 38 and 30 items per section

respectively. The TOEFL (Test of English as a Foreign Language) sections of listening comprehension have 50 questions, structure and writing expression have 40 questions and reading comprehension has 50 questions. The NAEP for 2003 has a range of items used to develop its scales. In the 4th grade assessment, the subscales ranged from 19 to 75, while in the 8th Grade the subscales ranged from 30 to 51. The SAT's subsections ranged in item numbers from 18 to 25 for reading, 16 to 20 for mathematics and 14 to 35 for writing.

The range of items for a test and subtest for standardized tests vary greatly. The range of the above values is from 14 to 75 for subsections of these well established standardized tests. The typical range for number of items can be limited from the mid 20's to 40's. Based on this review of test item length and preliminary analysis of 20 and 40 item tests lengths, the longer of the two 40 items has been selected for use in the current research. It would be reasonable to select a shorter test length of 20 considering the simulation research, targeted inspections of items have been limited to 3 items (Mislevy & Verhelst, 1990). Wainer and Wright (1980) used 3 tests lengths of 10, 20 and 40 items in their simulation work. In this investigation, 40 items will increase the chances of finding effects in the data. The contamination effect will change from 5 out of 20 to 10 out of 40 in the first subtest condition, which will double the number of items in comparison and increase the overall power of the design.

The researchers' preference to select precise points instead of having them randomly generated for the test will permit the evaluation of the study to have multiple overlapping item difficulties so that comparisons within and between

subsections of the test can be made clear. To evaluate common structure, patterns will be looked at individually and aggregated by these common difficulties within subsets.

Manipulated Factors

There will be a total of four manipulated factors in the current investigation: Type of Misfit or contamination (2), Proportion of misfit or contamination (5), Scaling factor of the test (4), and Size of subtest (2). The all Rasch conditions that are represented by the no misfit condition will not be repeated as one set overlaps. There are two sets of cells for the no contamination condition: no random contamination and no reverse contamination are the same effective cell. The number of cells replicated will be 2x4x4x2+8 = 72 cells in total.

The first of the manipulated factors, type of misfit to the Rasch model, involves departure from Rasch generating data. The data will be generated by altering the proportion of strategies being mixed and the strength of the models. The Rasch versus Non-Rasch (Random, Reverse effect) condition will represent key difference in cells. These Non-Rasch conditions are considered contamination in the investigation.

The second manipulated factor, the mixture proportions of expected and alternative models, will have 5 levels. The amount of Non-Rasch, contaminated responses mixed in with the expected Rasch responses will be: 0, .5, .80, .95 and 1. The intent here is to cover many possible levels in which patterns may interact and strength of patterns may be more or less prevalent. The mixing proportion here is for the total dataset, not within an individual's response pattern. This factor for a given set of responses is considered Rasch or Contaminated between case proportions.

Based on this first factor, responses will be all Rasch or contaminated within a given response pattern.

The third manipulated factor is the scaling factor of the test difficulties. In determining the range of the difficulties, several articles were reviewed (Chyn, Tang & Way, (1994); Chen & Davis, 1991; Mislevy & Verhelst, 1990; Kingston & Durons, 1982). In the sensitivity analysis (Hambleton & Rovinelli, 1973; via Hambleton et al, 1991) carried out on the goodness of fit statistics to determine sample size, item difficulties were chosen to those commonly found in practice, between -2 and +2. Items in the current investigation, taking into account previous research and practical considerations, are constrained to have a base range of 2 through -2. This item range will interact with a change in scaling factor specifically considered to interact with the contaminated conditions.

The scaling factors of the item difficulties in the test will take on four conditions: The base range of the exam extends from ±2. This scaling factor is used to create the four conditions where the base range is multiplied by: 1 for both subtests, 3 for both subtests, 1 for the Rasch only subtest and by 3 for the contaminated subtest, and 3 for the Rasch only subtest and by 1 for the contaminated subtest. This will provide a variety of test ranges from ±2 for all items in both subtests to a relatively extreme condition with a range of ±6 for items on subtests with interaction effect of those ranges also being examined.

The fourth and final manipulated factor is the sizes of subtests. In the first condition, the Rasch only subtest will have 30 items generated by Rasch only, with the remaining 10 items in the contaminated subtest changing based on the mixing

parameters. The second subtest condition for this factor will be equal subtests of 20 items each.

The Random proportion is envisioned as the lack of care, laziness, or even an alternate rushing factor. In conforming to this notion, randomness will be adjusted within a response pattern to represent different levels of carelessness. The reverse contamination will be generated to conform to the analogy or alternative knowledge models. These two represent contamination in the second subtest.

The contaminated responses will be generated within the second subtest. This subset represents not caring responses or differencing inferential models within the non-Rasch response subtest. All individuals will conform to the Rasch model for the first subtest. This factor is the amount of randomness within each response pattern.

Sample size and subtests

The approach to investigating misfit proposed here is quite different from that of Smith (1988), who calculates residuals for each item by person combination. However, we can use his results as a guide to determining sample sizes for which it is possible to both fit the Rasch model and detect residuals from the model. Particularly of interest for number of replication is Smith's second analysis with samples ranging from 30 to 2000 for a 10 and 20 item tests. Smith states that for 10 or 20 items, the relative frequencies appear to be independent of the sample size, and that the differences in frequencies seem like random fluctuations of the 10 replication, 20 item, 1000 person sample. In further examining tables 3 and 4 of Smith's work, 100 persons in a simulated sample seem minimally adequate. Wright and Tennant (1996) suggest at least 10 items "with a reasonably targeted sample of 50 persons, there is

99% confidence that the estimated item difficulty is within ±1 logit of its stable value."30 people are suggested to be good enough for pilot studies and 200 participants brings you within ± .5 logits. Keeping all this in mind, and understanding that the Rasch model is used here for the purposes of extracting information, the current investigation will simulate a fixed number of 500 subjects. This is above the 400 minimum suggested by Kolen and Brennan (1995) who are dealing with high stakes issues of equating and falls within research suggested by Smith (1988).

This sample size was also evaluated in preliminary investigation, and the results indicate 500 people estimated parameters significantly better than 100 people but not significantly different from 2000 people per replication.

In the current simulation 4 core characteristics were manipulated. Type of Misfit or contamination (2), Proportion of misfit or contamination (5), Scaling factor of item difficulties on the test (4), and Size of subtest (2). The above description of the factors includes: three different strategies of misfit, all Rasch, random and reversed inferential knowledge; five levels of mixing contamination between Rasch and misfit; four interacting scaling factors with the subtests; two separate subtest sizes, and all with a constant number of response patterns. Overall there will be 72 mixture response patterns generated with 500 subjects per replication.

Simulation Method

While discussing the overall process, it is useful to explain the method with respect to design considerations and model approach. In order to generate, transform, format, transport, analyze and graph data, several macros and general code were developed in SAS. The model for consideration was constructed and run in Winbugs.

Moving through the simulation process step by step and understanding all decisions and generating parameters will be helpful to understand the overall research.

Modeling

The measurement model was developed as a mixture Rasch model which fits a latent class model with the first class being the Rasch model and the second class being unscalable. Estimation of residual is done using a Bayesian mixture Rasch model within each of the 72 cells.

The Rasch model is used as the basis for model specification for the mixture in this analysis. The one parametric model

$$P(x_i \mid \theta_i, \beta_j) = \frac{\exp\left[x_{ij}(\theta_i - \beta_j)\right]}{1 + \exp(\theta_i - \beta_j)}$$

The mixture uses one class as Rasch with the second class being an unscalable class. The unscalable class for this two class model is set to have a p of .25. This mixture of Rasch and unscalable assumes that multiple classes exist within the population and for a portion of them the Rasch model will hold, the other portion of the population will be more appropriately within the unscalable.

The mixture model was presented earlier in this chapter. The posterior probabilities that a subject belongs to each of the two classes can be written as

Rasch class:

$$= \frac{\left[\int \prod_j \frac{\exp\left[x_j(\theta - b_j)\right]}{1+\exp(\theta - b_j)} p(\theta)\partial\theta\right]\pi_1}{\left[\int \prod_j \frac{\exp\left[x_j(\theta - b_j)\right]}{1+\exp(\theta - b_j)} p(\theta)\partial\theta\right]\pi_1 + \left[\prod_j c_j^{x_j}(1-c_j)^{1-x_j}\right]\pi_2}.$$

Random class:

$$= \frac{\left[\prod_j c_j^{x_j}(1-c_j)^{1-x_j}\right]\pi_2}{\left[\int \prod_j \frac{\exp\left[x_j(\theta - b_j)\right]}{1+\exp(\theta - b_j)} p(\theta)\partial\theta\right]\pi_1 + \left[\prod_j c_j^{x_j}(1-c_j)^{1-x_j}\right]\pi_2}.$$

These can be used as weights for an analogy to Tukey's "fit" and residual," to split the original data set casewise in proportion to these probabilities.

In the current research a 40 item test with 500 simulees is used. In the following example Table 3-3, four items with 50 simulees are used to represent an analogous but smaller example.

Table 3-3: 4 item, 50 simulees example.

0000	Observed	Rasch	Unscalable
0000	1	1	0
0001	3	3	0
0010	2	1.5	.5
0011	6	6	0
0100	0	0	0
0101	6	5	1
0110	2	1.5	.5
0111	11	9	2
1000	0	0	0
1001	4	3	1
1010	2	1.5	.5
1011	9	8	1
1100	0	0	0
1101	2	1.6	.4
1110	0	0	0
1111	2	2	0
Total	50	43.1	6.9

If a simulee with a 1101 pattern is determined to be .80 in the Rasch class and .20 in the unscalable class, that simulee and others with the same pattern will be assigned a weight of .20 to the residual class. The posterior probability is used as a weight for each individual within each class by assigning the proportionately associated with the patter that simulee falls in. The above example would take the 6.9 responses classified in the unscalable class and use them as weights in a factor analysis for the residual class.

Winbugs model

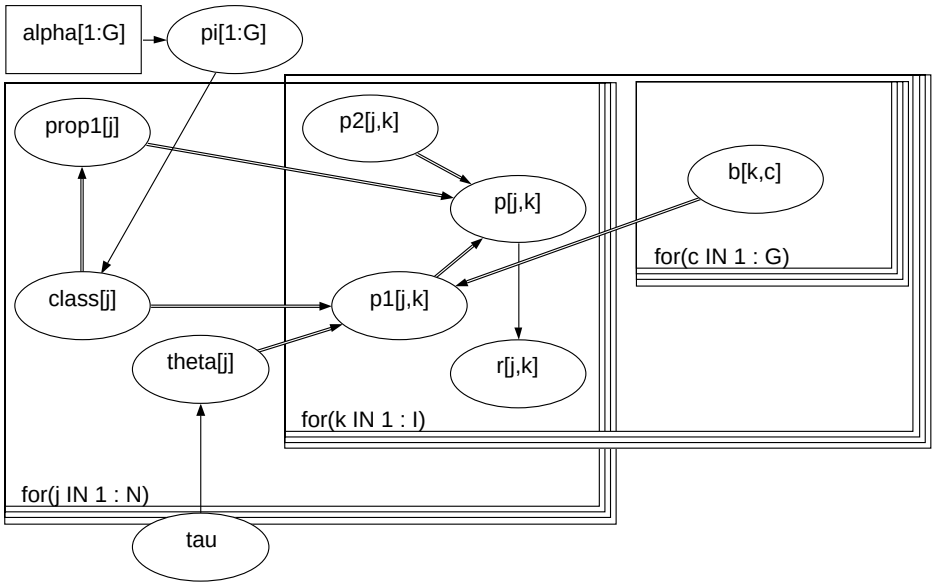
The first choice to be made in generating the data was the model to be used in the simulation. In the Preliminary investigation discussed earlier, the priors were set for each parameter and tested to make sure priors did not overwhelm the data. Data points were fixed in order to resolve the label switching indeterminacy with multiple

classes (Chung, Loken, & Schafer, 2004). The data is fixed by the first replication always being set to the first class. The model was set to remain as a fixed factor in the overall simulation. The preliminary study also set the number of burn-in and estimation cycles for the study. The end results of a set of fixed parameters for the current research are discussed sufficiently in the preliminary investigation chapters 4 and 5, so that only the brief recap of the model and its fixed factors are mentioned. The two class mixture model of Rasch and an unscalable class is used in all replications across all cells. This is the model that the generated data, which is generated under differing conditions depending on the cells manipulated factors, will be applied to like a filter to examine residuals. The burn-in cycles are set to 2000 and the estimation cycles are set to 5000.

Winbugs diagram

In developing the code to transform the above equations into Winbugs code, the intermediate step of modeling was taken. Figure 3-1 is a directed acyclic graphical representation of the model, called a doodle in the WinBUGS program. This representation aides a structural explanation of the model.

Figure 3-1: Winbugs doodle of the 2 class model



m

The WinBUGS code in figure 3-2, is modeled from the above doodle. The code and the doodle show the two class model with Rasch and unscalable classes.

Figure 3-2: Winbugs code for the 2 class model

```
model
{
for( j in 1 : N ) {
for( k in 1 : I ) {
p1[j,k] <- exp(theta[j]-b[k,class[1]])/(1+exp(theta[j]-b[k,class[1]]))
p2[j,k] <- 0.25
p[j,k] <- p2[j,k]*prop1[j]+p1[j,k]*(1-prop1[j])
r[j,k] ~ dbern(p[j , k])
}
}
for( k in 1 : I ) {
for( c in 1 : 1 ) {
b[k , c] ~ dnorm( 0.0,0.25)
}
}
for( j in 1 : N ) {
theta[j] ~ dnorm( 0.0,tau)
class[j] ~ dcat(pi[])
prop1[j] <- class[j] - 1
}
pi[1:G] ~ ddirch(alpha[])
tau ~ dgamma( 0.5,1)
                }
list(N=500, I=20, G=2, alpha=c(10,10),
class=c(
1,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
……
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA),
r=structure(.Data=c(
1,0,1,0,0,1,1,0,1,0,0,0,1,0,1,0,0,1,1,1,
0,0,1,0,0,0,0,1,1,1,1,0,0,0,1,1,0,1,0,1,
……
1,0,1,0,0,0,1,1,1,0,0,1,0,1,1,1,1,0,0,1,
), .Dim = c(500, 20)))
```

Generating data: Calculating the Probability

The simulation process for generating the data in the current investigation starts with a Monte Carlo simulation of varying mixture conditions of Rasch data and unscalable, random data. In generating one replication within one cell, several choices are made. The first decision was to set the test to be used in the simulation process.

The general base test has set points of +2, +1, 0, -1, -2 repeated 8 times in the assessment for a total of 40 items. There are 6 of each base values in the thirty item subtest, 4 of each in the twenty item subtest and 2 of each in the ten item subtest. These interact with the scaling factor for a given subtest and are either multiplied by one, meaning they do not change, or multiplied by three to show a more extreme scaling factor condition of +6, +3, 0, -3, -6.

The Rasch data is generated to conform to a normal distribution with a mean of zero and a standard deviation of 1. These theta values are randomly drawn for each respondent. Regardless, the first subtest is generated to conform entirely to the Rasch model for all respondents.

In order to get one respondent for one replication within a given cell in the first subtest, the randomly generated theta value is used to generate probabilities. The probabilities are calculated for each of the items in the subtest resulting in item response probabilities for the simulee. The model used to generate this is simply the Rasch model in SAS code.

```
PROB[P] = (EXP(THETA-ITEM[P])/(1+(EXP(THETA-ITEM[P])))
```

In the code PROB is the probability and is calculated by taking the exponential (EXP) of the difference between the generated THETA value and the ITEM difficulty

divided by one plus that same value. This is effectively the Rasch model in SAS code. The P parameter represents that this occurs over P items and for the Rasch subtest.

The remaining items in the second test represent either 25% or 50% of the items and are generated under one of several mixing conditions. The two contaminated alternatives to the above Rasch model used in the mixing condition are an unscalable class of random data and the reverse analogy effect. The Random probabilities are simply set to .25 for each of the items in the subtest. This is the same as random guessing. The Reverse effect swaps sign of the difficulties so that +2 becomes -2 and +1 becomes -1 for that portion of the subtest. The data is generated the same way as above except for those in the contaminated portion of the subtest where the code now reverses, as shown bellow.

```
PROB[P] = (EXP(THETA+ITEM[P])/(1+(EXP(THETA+ITEM[P])))
```

These items are generated as Rasch with a proportion mixing conforming to the misfit generating parameter, 0, .5, .80, .95 or 1. In this subtest zero, Rasch and all of the alternative contaminated condition can be generated, partial Rasch and partial contamination or all Rasch can be present in the subtest, which would result in the baseline, non-mixed, uncontaminated condition.

The probabilities are generated in SAS based on the Rasch and misfitting, contaminated conditions. The generation of these probabilities is based on equations mentioned earlier in the misfit from Rasch section of the paper.

Generating data: the observable zero/ones

Based on the simulee's theta and mixing condition, each simulee has 40 probabilities, one for each item in the test.. The probabilities are then transformed into

zero one responses based on a random procedure in SAS called RANBIN (SAS, 2009). This function RANBIN returns a random value for 1 or 0 from a binomial distribution with set parameters of number and probability of success taken from the probability calculated in the previous section. 500 such simulees are run per each replication of a cell.

Transformation and Transportation of Data

The response data is stored first in SAS as zeros and ones and contains 500 respondents by 40 items. The data is restructured and merged with Winbugs formatting to create a plain text document that can be used with Winbugs. In addition to the data being formatted as a text file to be used in Winbugs, the Bayesian Mixture model is also written out by SAS into a text file to be used in the analysis.

A command script is written in SAS and exported as text to run Winbugs remotely through SAS. The script contains generating parameters, locations of files to be used for the data and model, information to set parameters and collect summary statistics information from set parameters. The statistics are then saved as a text file. This script is run automatically using code within SAS. A DOS command prompt opens, runs code that starts Winbugs, runs the script, closes Winbugs, and the data is stored as a text file and then transformed in its entirety into a SAS dataset. A second SAS dataset from the raw text file generated from the Winbugs analysis contains only the probability of class membership for each simulated respondent. This information, the posterior probability of class membership, is stored to be used as the weight variable to investigate the original data. This is the residual or misfit by which the

data is weighted and reexpressed to have fit stripped away, leaving the misfitting portion to make it easier to find patterns should they still exist.

Evaluating the outcome

Exploratory in the spirit of Tukey

In examining the reproduction of factors, Gorsuch (1983) discusses replication and invariance of factors across random samples of individuals in terms of desirability of the factor model. A solution is considered invariant when the factor patterns are similar across multiple replications. Invariance is said to occur when there is a high correlation of factor scores. Slight differences in a model are attributed to chance fluctuations. In evaluating the factor models in the preliminary investigation, visual inspection as well as correlation will be used to determine if invariance exists. Invariance and replication of factor structures will be considered crucial for the final study.

Mosteller and Tukey (1977) provide an in-depth treatment of exploratory data analysis of residuals for regression models. In a similar general framework of data=fit + residual the data is explored through reweighing data to the residuals and fit of the model. A critical difference between Mosteller and Tukey's exploration of regression and the current investigation of the Rasch model is the benefit of knowing the outcome in real data situations. In regression, one uses a predictor, or set of predictors to predict an outcome, and the outcome of the model regression is known. In the Rasch model, although it too develops a model with parameters for prediction, the true values are not know and typically sufficient statistics are used to estimate a model. In regression, we have the model and a calculated residual between observed

and true values. The residual to the Rasch model in this framework is created from a two class LCA model with one class, as Rasch and everything else unscalable, which includes the residual to the Rasch model.

In the current investigation, 50 replications in each cell will be aggregated and reported in terms of summary statistics for the posterior mixture model. Summary statistics of the posterior probabilities of class membership will be evaluated using means and SE. After all cells are simulated, 3 factor models will be constructed from each individual dataset by weighing the data: unweighted, Rasch weighted, and residual weighted data. The number of factors in a condition and the structure of factors will be evaluated. Eigenvalues will be evaluated by comparison and patterns of change. Patterns will be aggregated within a cell and within each dataset type and compared using MANOVA and Hotelling's T test. Individually aggregated patterns can be compared using t-test and ANOVAs. Grouping patterns will be compared in terms of means and standard errors (SEs).

Tetrachoric correlation

Tetrachoric correlation coefficients are estimated in SAS using the PLCORR option in the Tables statement of the PROC FREQ Procedure (SAS 2004). PROC FREQ is the procedure for producing frequency tables, while the option PLCORR produces polychoric correlation coefficients. In the limiting case of a 2 by 2 contingency table, the polychoric correlation coefficient is described as the tetrachoric coefficient. In SAS, the correlation starts with the Pearson correlation coefficient as the starting point and attempts to iteratively solve for the tetrachoric solution. The model will run until it converges, or until the maximum number of

replications has been reached. This value is typically 20. In the current investigation it is crucial that each solution for the tetrachoric coefficient be solved. There are 40 items, so there are 40*39/2= 780 unique correlations to be estimated for each correlation matrix used in this analysis. The maximum number of iterations has been set to 100 and the convergence criteria have been set to .001. If a solution is not found for every coefficient in the matrix, the result cannot be used in a factor analysis model. If unsuccessful, the Pearson correlation will be substituted into the model, because when using tetrachoric correlations or Pearson correlations, no difference could be detected in results in preliminary analysis.

Incorporating the weights for a tetrachoric correlation

In SAS the frequency proc freq procedure which produces the tetra correlation was an option to weigh the data by some other variable (SAS, 2009). In the current investigation, the weight used is the posterior probability of class membership, specifically the probability of being in the unscalable class. This is the concept of residual discussed earlier. When this type of weighted structure is applied to tetrachoric correlations there is a modification of the construction of the correlation by using the weighted value for each case instead of counting each case as one. Instead of frequency of response being entered into each of the 2x2 contingency table cell, 11,10,01,00; the weighted frequency is entered. Each individual counts once in the original data so the summed weight for each condition is effectively entered into the cell. One could also take the average weight times N to show the value as well. For 500 people, when all responses have the same frequency of relation there is 125 in each cell combination as shown in Table 3-4:

Table 3-4: Equal frequency N = 500

Raw frequency	1	0
1	125	125
0	125	125

Weights sum to: average weight times N Total N=131.25

11=.476, 10=.238, 01=.095, 00=.190

Table 3-5: Unequal frequency N=131.25

Weighted frequency	1	0
1	62.5	31.25
0	12.5	25

Eigenvalues and Horn's parallel analysis

Horns parallel analysis (HPA), was initially evaluated using Pearson correlation coefficients and tetrachoric correlation coefficients. The two performed equally in discriminating eigenvalues. To determine the number of factors for the unweighted factor models, 50 replications are used in the modified HPA to stay consistent with the 50 replications in each cell. Each replication is run on random data with the same structure as the data used in the current analysis: 40 items and 500 simulees in the unweighted model, and the correct proportion when data is weighted.

Eigenvalues from the appropriate 50 replication HPA will be evaluated using this as the cutoff criteria for selecting the number of eigenvalues. Eigenvalues with

values greater than their counterparts from the random modified HPA eigenvalues will be quantified as being justifiable factor.

The weighted factor model comparison will have a similar method with the additional step of using posterior weights of the unscalable condition, as is also used in the cells weighted condition. The randomly generated data will be estimated using the same mixture model used on analyzed data. The posterior proportion for the unscalable class will be used as a weight and tetrachoric correlations will be run to generate factor models.

Analyzing the data

Analysis in the current investigation will address the first two factors unless results of the HPA investigation determine indicate inquiry beyond these two factors. The initial examination of the conditions will be to determine what proportion of times respondents who were generated to have misfit data are judged to be in the residual class. Essentially, how well does the mixture model do at categorizing these respondents into the unscalable class? Examination of the residuals will be comparative to the baseline Rasch model. The average and standard error of correct classification of misfit will be compared for each cell to the baseline Rasch model. A 95% confidence interval will be used. A condition will be determined to be categorizing misfitting respondents if the portion of correctly identified misfitting data is outside of the 95% confidence interval of the baseline Rasch model average.

Conditions with residuals that fall outside this value will be explored in further analysis when possible. In order to explore the unscalable class weighted data,

the residual must be of sufficient size. Practical considerations are considered when determining what might make a meaningful residual.

A practical exploration value of 1.25% is a reasonable guide to consider when exploring residuals in the current investigation. In the smallest misclassification generating parameters, 5% of the data in 25% of the items is contaminated. Although this does not translate directly to the expected size of the residual, it is a good guide in the current investigation to suggest a residually weighted dataset that is large enough to explore regardless of other concerns with the residual.

In the current investigation, if a residual value meets or exceeds 1.25% of the data it will be considered for further analysis to determine what types of patterns are left in the weighted data. This value also includes the residuals for the Rasch only baseline conditions should they be 1.25% or larger. This is not the smallest size of residuals possible for further analysis in the weighted data, but value for preliminary screening to suggest that a residual is large enough to explore.

Depending on the results of analysis in the baseline condition and what is feasible given the analyses to be conducted, a smaller residual value may still be set. The residual must be of sufficient size to work with the mechanics of the analysis procedures. When analyzing the results, smaller residuals will be explored provided they work with the factor analysis procedure and are reasonably larger than the baseline Rasch models. Unless baseline Rasch conditions increase greatly, the smallest condition explored will be .4%, as it is both feasible for analysis and just outside the 2SE band from the Rasch preliminary investigation.

Hypotheses surrounding the expectations of the size of the residual are based on the manipulated factors and will be evaluated based on the proportionate size relative to the percent contaminated condition.

1. It is expected that the proportionate residual size will:
    a. Increase as the percent of contamination generated into the data decreases
    b. Increase as the contaminated subtest range increases from +/-2 to +/-6
    c. Increase as the contaminated subtest size increases from 10 to 20 items
    d. Be larger for the systematic contamination conditions when compared to the random contamination conditions.

The residual weights will then be used to reweigh the data into three separate datasets. The first is the unweighted initial dataset, and no weight is applied. The second uses the weighted fit portion of the data to proportionately create a new dataset that removes unscalable effects, and theoretically increases the Rasch proportion of the data. The third dataset is weighted with the residual data and is the proportion of data categorized into the unscalable class for that replication. These three datasets,--unweighted, Rasch weighted, and unscaled--will be used in the next set of analysis.

In evaluating the patterns within the factors, all Rasch generated cells with no contamination will serve as a baseline. This uncontaminated Rasch generated condition is generated with patterns that would be expected for the model. There is a second factor of contamination being introduced in the reverse effect condition. When that factor is systematic variation, it should manifest in the unweighted factor

analysis and have a stronger secondary factor than when no systematic variation is present in the uncontaminated condition. The patterns in the uncontaminated condition will load heavily on the first factor. The second factor in this unweighted condition will be the best of what is left over when the first factor is removed. The contaminated second factor should be greater than those in the uncontaminated baseline second factor and the patterns within that factor should also be large on average in an absolute sense. On average, the patterns in the second factor in the contaminated conditions should be significantly greater than the patterns in the uncontaminated condition for the unweighted factor analysis.

There are several hypotheses to test surrounding the expectation of the number of factors in the data. HPA will be used as a threshold to determine the number of factors for all hypotheses regarding the number of factors in a given condition.

2. In the eight Rasch only generated baseline conditions, one factor will be present in the unweighted data.
3. All other 64 systematic or random contamination conditions will have present a second factor in the data for all unweighted datasets.

An exploratory factor analysis (EFA) model will be estimated on the unweighted data. The eigenvalues will be compared to a Horn's parallel analysis to determine the number of factors in the unweighted data. It is expected that models conforming to the Rasch model will have one factor, and that models with contamination will have two factors in the unweighted data. The two factors in the contaminated condition are expected to be enhanced when the data is reweighted.

EFA will be conducted for the two remaining weighted datasets, Rasch and unscalable data, to determine if significant patterns exist in a given cell and to evaluate eigenvalues. The eigenvalues will first be examined prior to exploring the loadings of the two weighted datasets.

The Rasch weighted data is examined for expected eigenvalues patterns. In general, it is expected that the Rasch weighted data will have more clearly one dominant factor compared to the unweighted data. This is because some of the systematic and random fluctuation from the unscalable class should be reduced. Furthermore, this will interact with the proportion correctly into Rasch and unscalable classes such that the greater the proportion of contaminated data correctly classified, the smaller the secondary factor should be proportionate to the unweighted data. It also should be easier to extract information that is systematic instead of random. The related hypothesis surrounding the Rasch weighted data are thus as follows:

4. On average, in the Rasch weighted there will be fewer factors extracted from the data when compared to the unweighted data.

The eigenvalues of the unscalable weighted data are also examined. The residual data should reweight the data such that Rasch effects are reduced in factor exploration. One important question with regards to the contamination in the data is as follows: can the contamination still be detected, or are the remaining factors no better than what would occur by random chance alone? The unscalable weighted data will have a different proportion of Rasch versus contamination.

The unscalable weighted data should have a reduced Rasch factor and an increased random or systematic factor. The related hypotheses surrounding the

residually weighted data are as follows and assume the residual is large enough, over .4 to be used in analysis:

5. On average it is expected that the systematically contaminated conditions will have more factors than the random contamination conditions when the residual is detectable.
    a. When factors are found for the systematically contaminated conditions there will be two factors: one Rasch and one Reversed Rasch
    b. When factors are found for the random contaminated conditions there will be only one factor in the data which is a suppressed Rasch factor.

MANOVA

The patterns of the factor models will be compared first using a MANOVA or Hotelling's T-test to examine the multivariate differnec between the Rasch weighted data and the residually weighted data across all items on the exam. The structures of factor patterns will be evaluated using the all Rasch generating conditions as the baseline. The Rasch condition will serve as a baseline under the assumption that the data is generated to fit the Rasch model and only random fluctuation from the Rasch model will occur in these generated conditions. The posterior weights for the unscalable condition should have no substantial meaning in terms of modeling contamination, as only random variation lead to its categorization in the unscalable class.

The baseline Rasch has 8 cell conditions: The four scaling factor conditions interacting with the two subtest sizes; the three baseline factor models derived from these Rasch conditions that will not use the posterior residual, as this would be too

small and not comparative; instead a random assignment of respondents to one of the two classes will be used to split the data and determine if the MANOVA structure is significant as a baseline. The data will be randomly weighted in each cell with 4 different weight sets as appropriate for comparison: .5 represents 250 Rasch 250 unscalable, .8 represents 400 Rasch, 100 unscalable, .95 represent 475 Rasch and 25 unscalable, and .97 485 Rasch and 15 unscalable. This represent the 3 generated contaminated conditions of .05, .8 and .95 as well as a smaller 3% condition (15 respondents) condition for the expected residuals around 2%-4% as found in some preliminary analysis.

In order to conduct the MANOVA, the two weighted factor models for the all Rasch weighted data and for the unscalable weighted data will be used. The factor patterns from the first two factors will be used in conducting two MANOVAs, one on factor one, and a second on factor two. In a multivariate analysis the two sets of patterns, one from the Rasch weighted data and the other from the unscalable data, will be compared to determine if they are different. In the first MANOVA Factor one, patterns for all items will be used as the dependent variables. The independent variable is the Rasch weighted factor patterns compared to the unscalable weighted factor patterns. The second MANOVA will follow the same method but use factor two for both Rasch and unscalable patterns.

These models are all expected to come up as not significant. The same logic for the baseline will be used for the other conditions except the splitting of the data will be the posterior classification of class membership.

In all other cells having contamination, the two weighted patterns for the Rasch factor model and unscalable factor model will be compared. A MANOVA will be conducted comparing the matrix of patterns across all items as the dependent measures. The Rasch factor patterns and unscalable factor patters are the independent measure, just as in the baseline condition. The sample sizes will be equal in all cases even though factor models are estimated using a different proportion of respondents. If the multivariate model is different, it will be a statement that there is a difference in the dimensional space of patterns between the two sets of patterns. This analysis will only be conducted for the first two factors within each cell.

The results of the Wilks' lambda F value for each condition will be compared to the baseline F value in the Rasch conditions. For comparison reasons, any MANOVA F value that is not over the Rasch Baseline value, regardless of significance will not be considered a large enough effect. The F value of the baseline model will be the comparison value of interest.

Instead of relying just on significance and one F value, 100 MANOVAs will be used and a mean F and SE will be derived. F values outside of 2SE will be identified in a results table.

The hypotheses surrounding the comparison of patterns between Rasch weighted and residually weighted data involve the manipulated factors in the current research.

6. Through the use of MANOVA it is expected that when residual misfit is extracted there will be a significant difference between the Rasch weighted patterns and residually weighted patterns in the first and second factors.

a. It is expected that differences will be more detectable when the contamination is stronger. Specifically, stronger contamination is measured by: an increase in scaling factor in the contaminated subtest and an increase in the number of items from 10 to 20 items.
b. It is expected that as the proportion of contamination increases, less residual effects will be significant. The contamination will overwhelm the data in both the residual and Rasch conditions and cancel out differences between the two weighted datasets.

Continued investigation: MANOVA and Confidence Intervals to support visual graphs

Once it is determined that the proportion in the residual is large enough to explore for a given cell, and that the multidimensional structure has been investigated, the individual differences will then be examined. Each pattern will be examined to determine if the unscalable pattern is different than the Rasch pattern. This will be done using confidence intervals. The overall goal is not to be able to speak to each of the 40 individual differences between patterns per cell, but to lend support when necessary to the graphical models used to visualize differences between the overall patterns in the weighted datasets. Furthermore differences clarified by CI and graphical representation will help to articulate where separation occurred in the 40 dimensional multivariate model used in MANOVA.

In the residual dataset, the data is weighted to maximize unscalable data. This approach should have the effect of reducing the Rasch factor from the unweighted dataset and leaving more of the secondary factors from the unweighted dataset. The

Rasch weighted data should have the contrary effect of reducing the contamination and clarifying the Rasch factor. In comparing the two sets of patterns from the Rasch weighted data and the residual weighted data several patterns should be apparent. There is expected to be differences in both random and reversed contamination conditions but for different reasons. It is expected that random conditions will have suppressed weak patterns while the reversed condition will have strong patterns in a different structure than the Rasch weighted data. The hypotheses about the difference in patterns stated above will be investigated through the use of CI, graphs and Wilks' lambda.

7. When the Rasch only, baseline data are randomly split, there should be no visible difference between the two sets of patterns for the first or second factor.
8. In the random contamination conditions, the residually weighted dataset should have suppressed patterns for the first and second factor. The Rasch weighted data should still have strong Rasch patterns. The differences should be captured, with CI differences attributed to residual weighted values close to 0 and Rasch weighted data following Rasch type patterns. These differences should be apparent in graphs comparing the two weighted datasets The Wilks' lambda should be significant and larger than the baseline F.
9. In the reversed contamination condition, the residually weighted dataset should have strong patterns similar to Rasch weighted data but in a different graphic structure for both the first and second factor. The differences should be captured with CI differences attributed to the differences in patterns

particularly in the contaminated subtest. The Wilks' lambda should be significant and larger than the baseline F.

Examining the residual patterns of subtests

Finally, the patterns will be examined by groups of like items. Within each subtest, items with the same difficulty base value, +2, +1, 0, -1, -2 will be averaged and the patterns will be examined. Each cell will have five averaged patterns in subtest one compared to five average patterns in subtest two. Descriptive comparison through graphical representation is the primary form of comparison. The all Rasch condition will be used as a baseline condition.

There are several hypotheses to test surrounding the expectation of the patterns.

10. It is expected that the subtest in the Rasch baseline conditions will look like the remainder of the exam. Both subtests will have Rasch patterns
11. In the random condition it is expected that the contaminated subtest will have significantly smaller pattern values, close to zero, than the all Rasch subtest.
12. It is expected that patterns in the reverse effect condition should show a reversed pattern in the subtests which becomes more prominent as the strength of the effect increases.

Chapter 4: Preliminary Investigations into Model Justification

Logic of preliminary investigation

The test was comprised of forty items, with a base range in difficulty values of ±2 This size and range was selected after an extensive literature review. All of the analysis at this stage of investigation was initially evaluated using a 20 item model with a test range of +1 to - 1. The 20 item analysis was used as guidance in structuring the current method in this preliminary analysis and a baseline for expected times for cycles. It is important to reevaluate the model under the 40 item condition to substantiate a good model.

The base condition to be evaluated for the simulated test is forty items with a range of +/-2. The test has two subsections: (a) 30 item difficulties all generated as Rasch and (b) 10 items with varying degrees of contamination. This test is identical to the model test that will be used in the final analysis. In the current investigation, five individual datasets are reviewed to explore the robustness of conditions and help provide a rational for burn-in cycles, estimation cycles, sample size, effect size and finally number of replications. The investigation process is not intended to exhaustively explore all possible conditions but instead to iteratively select a high-quality model from amongst reasonable models to be used in the final analysis.

The initial burn-in and estimation cycles are examined using a sample size of 500 simulees. The foci of the study of the number of burn-in cycles are to first insure sufficient cycles are removed and that the remaining data are stable. Once the burn-in condition is set, the estimation cycles will beare evaluated to determine if 2500, 5000

or 7500 cycles are used in the final simulation. When the model has a stable number of cycles for burn-in and estimation cycles, the sample size is examined further to determine if 500 is a reasonable sample size that reproduces expected parameters by comparing it to sample sizes of 100 and 2000. Finally, the effects of reversing the Rasch model are examined.

The preliminary investigation has several phases that are based on issues of correct technique from the method section in chapter three. In the first phase of investigation, a rationale is demonstrated for selection of the number of burn-in cycles and estimation cycles along with a brief review of priors. This first step in this investigation is necessary to be confident in the accurate estimation of the Bayesian mixture model. The criteria for selection of burn-in cycles and estimation cycles are balanced between two main factors, accuracy and overall estimation time. The Gelman-Rubin statistic, cycle history and density plots are used to select accuracy level for burn in criteria. The Gelman- Rubin statistic must approach the value of 1.05, the history must show convergence of the three disperse chains, and finally, the density plot for the three chains must overlap and begin to look like one distribution instead of three.

The second criterion is parsimony. In most cases, for accuracy, more is better: the more cycles used, the more accurate the outcome. However, the model could be run into infinity, or at least until the computer stops running. The lowest number of burn-in and estimation cycles are used to meet the criteria set for convergence across five datasets. In addition to the convergence criteria, reproduction of the difficulty parameters is examined. The parameters should be reasonably close to the generating

parameters when the Rasch model is expected to fit the data. In selecting a larger cycle size over a smaller size, the difficulty parameters will show change greater than expected by chance.

Once the burn-in and estimation cycles are set, the sample size is adjusted to see if it would be reasonable to lower the sample size or if it needs to be increased. Samples sizes of 100, 500 and 2000 are evaluated.

The next phase uses the selected burn-in, estimation cycles, and sample size from the first phase to investigate cells from study and determine the final simulation design and analysis plan using the trial cells from this second Stage of investigation. The all-Rasch generated model is used in the main analysis as a baseline condition while the data is generated to fit the Rasch model. It is expected that if a sufficient effect in the contamination condition occurs there should be some difference in the amount of contamination.

Burn-in and Estimation Cycle Investigation: Preliminary cycle investigation

In order to begin running the full scale simulation using MCMC estimation in Winbugs, preliminary investigations of parameters were conducted to present evidence that parameters are being estimated with the correct number of burn-in and estimation cycles with a reasonable number sample size. The Gelman-Rubin (modified by Brooks & Gelman (1998)) statistic in Winbugs (Spiegelhalter et. al., 2003), is used to examine burn-in cycles. In addition, statistics of posterior parameters, history of cycles and densities of the posterior data are examined under several conditions to add evidence and cover the rational for selecting the number of

burn in cycles and adequate estimation cycles computed per replication in the final simulation analysis.

Four datasets were examined to be falling under differing conditions within the main simulation study. The intention here is not a comprehensive examination of all datasets in the analysis, which will occur during the main investigation, but an exploration of a sub-sampling of datasets from different conditions to justify cycles for use in the main investigation, as well as to provide some additional evidence that the model being used is a reasonable model.

Limiting the burn-in cycle

This initial exploration at this stage of the investigation is to provide practical limitations for the second stage of the burn-in and estimation cycle study. Extreme burn-in cycles such as 50 or 8000 are examined alongside of several reasonable conditions to insure thorough review of limits on the size conditions in the second stage of the burn-in cycles are apparent. The conditions of 50 and 8000 burn-in cycles are principally for comparison purposes in the first stage of investigation to the more practical conditions between 250 and 4000, and are expected to be the boundaries considered extreme conditions of very poor estimation at 50 and loss of efficiency at 8000. This groundwork also provides some examination of the estimation cycles, and if it is worth examining, small estimation cycles in the next stage. The estimation sample size of 50 is useful in investigating initial estimation densities, history and summary statistics, but not for the Gelman-Rubin statistic which was investigated under the 500 burn-in cycle condition. Before running a full scale MCMC simulation, it is essential in preliminary investigations to consider good estimation along with

prudence of cycles. The goals of the first stage of the investigation were to narrow down the number of burn-in conditions to those that showed good quality convergence in the initial data and to determine if any could be eliminated for the second stage of investigation as superfluous.

Manipulated Factors

In the first stage of this part of the investigation, burn-in is initially examined in only one of five of the datasets under consideration. Two estimation conditions, 50 and 500 cycles are run past the initial burn-in cycles to determine which number of iterations for burn-in cycles would be worth further investigation in the other four datasets. Initially 50, 250, 500, 1000, 2000, 4000, and 8000 burn-in iterations were examined for a total of seven burn-in conditions. This initial study covers fourteen conditions, seven burn-in cycle conditions by two estimation conditions. The purpose here is simply to remove the extreme conditions of burn-in cycles, should they not meet the accuracy criteria or are unnecessarily time intensive.

Initial Data Structure and Model

The factors from the main analysis for which this dataset falls under in the first stage of investigation are a Mixture of Rasch and unscalable data. A total of 40 items were used as the test structure and 500 respondents were simulated. The initial 250 response in the dataset are simulated as Rasch with N(0,1) with the remaining 250 response are a mixture of 75% Rasch N(0,1) and 25% guessing or unscalable with P=.25. The data was analyzed in Winbugs.

Winbugs Model:

```
Model
{
for( j in 1 : N ) {
for( k in 1 : I ) {
p1[j,k] <- exp(theta[j]-b[k,class[1]])/(1+exp(theta[j]-b[k,class[1]]))
p2[j,k] <- 0.25
p[j,k] <- p2[j,k]*prop1[j]+p1[j,k]*(1-prop1[j])
r[j,k] ~ dbern(p[j , k])
}
}
for( k in 1 : I ) {
for( c in 1 : 1 ) {
b[k , c] ~ dnorm( 0.0,0.25)
}
}
for( j in 1 : N ) {
theta[j] ~ dnorm( 0.0,tau)
class[j] ~ dcat(pi[])
prop1[j] <- class[j] - 1
}
pi[1:G] ~ ddirch(alpha[])
tau ~ dgamma( 0.5,1)
            }
```

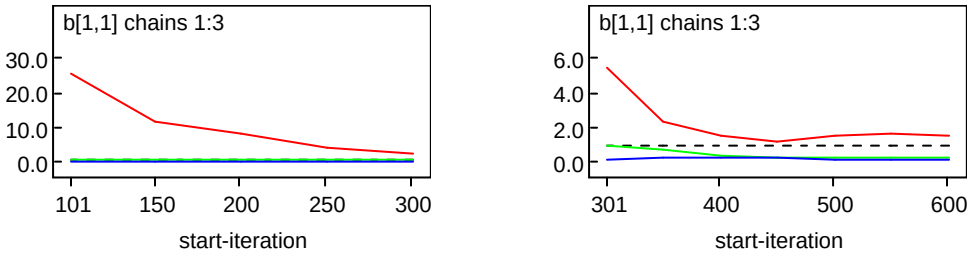
Three chains were run with every analysis used in the preliminary investigation. The starting values were set to be dispersed with convergence being examined with the Gelman-Rubin statistic as suggested for use in Winbugs. Class mixing proportions were intentionally set to have diverse starting points (.9,.1), (.5,.5), and (.1,.9), while the remainder of starting values were permitted to be generated by Winbugs. In addition to reviewing the model with the Gelman-Rubin statistic, the summary statistics, history and density of each analysis were investigated to assist in determining the best number of burn-in cycles to investigate in other datasets for the next stage of investigation.

Results for limiting burn-in cycle

The value for the Gelman-Rubin statistic suggested to be good is below 1.05. The two smaller conditions of 50 and 250 burn-in cycles are examined here. The Gelman-Rubin statistic initial value changed within the burn-in cycle condition of 50, from a range of 5 to 100 for item difficulties with the occasional extreme value

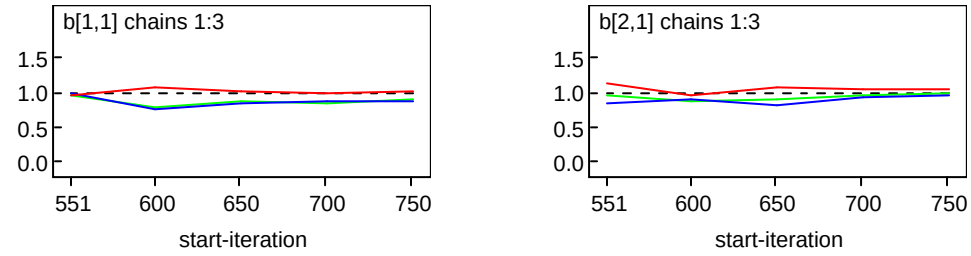
spiking to 200 in ability estimations In the 250 burn-in cycle condition, more stable estimation were revealed with ranges of approximately 1 to 3, with ability estimates spiking to no more than values in the twenties. However, both of these conditions are eliminated from further consideration as they do not fall with-in reasonable values for the Gelman-Rubin statistic.

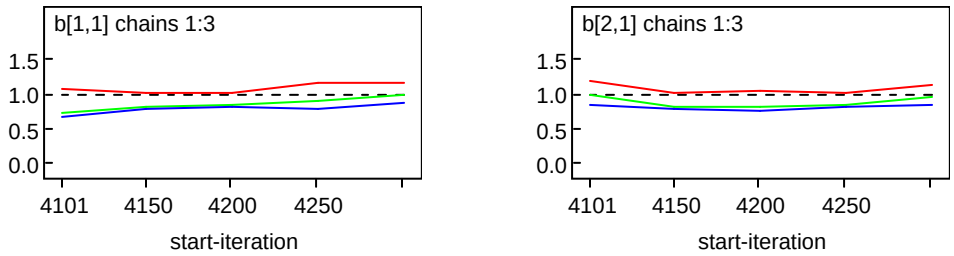
Figures 4-1: Examples of Gelman-Rubin graph for 50 and 250 burn-in cycles of item difficulty 1



All other conditions 500 through 8000 have very similar Gelman-Rubin statistics and are considered to have met the criteria of accuracy for further investigation.

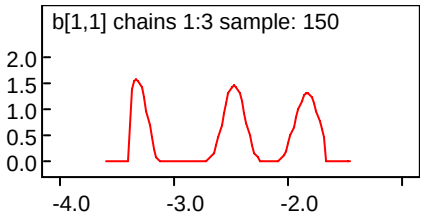
Figures 4-2: Examples of Gelman-Rubin graph for 500 and 8000 burn-in cycles of item difficulty 1 and 2.





Density patterns are similar for all graphs to those displayed below for preliminary analysis. The graphs are identified by the number of burn in cycles and estimation cycles. The 50 and 250 conditions are not under commiseration after evaluation of the Gelman-Rubin statistic and Graph 6-7 shows and example of poor convergence in a density plot.

Figure 4-3: Poor convergence density plot



The 500 burn in condition for 50 estimation cycles is the first visual smoothing of a distribution and by 500 estimation cycles the density plot is not distinguishable from the 8000 burn-in cycle condition

Figure 4-4: 500 burn-in 50 estimation cycles per chain.

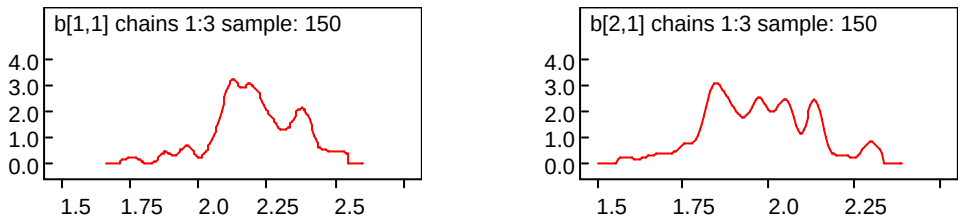


Figure 4-5: 500 burn-in 500 estimation cycles per chain.

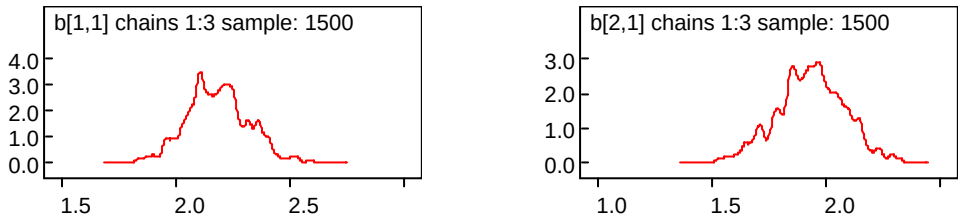


Figure 4-6: 8000 burn-in 50 estimation cycles per chain

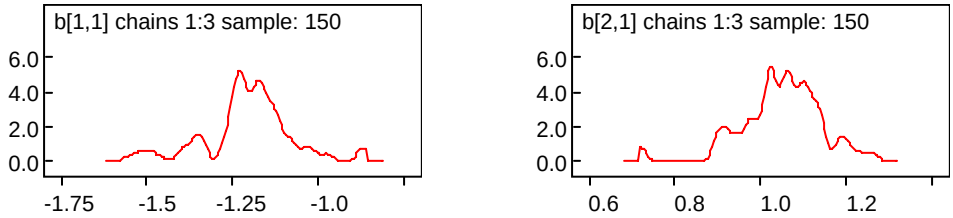
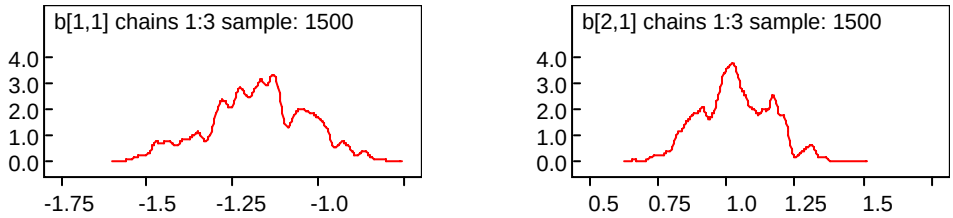


Figure 4-7: 8000 burn-in 500 estimation cycles per chain



Using the 500 burn-in condition shows cycles form this burn-in estimation have the chains converging, which does not occur in the earlier conditions.

500 burn-in, 500 estimation: The conditions greater than 500 of 1000, 2000, 4000 and 8,000 simply add support to the chains converging prior to 500 cycles.

Figure 4-8: Three chain convergence for item difficulty

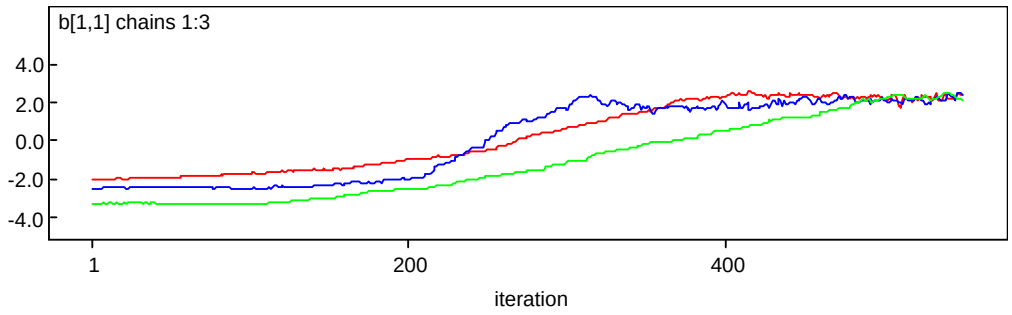


Figure 4-9: Three chain convergence for class

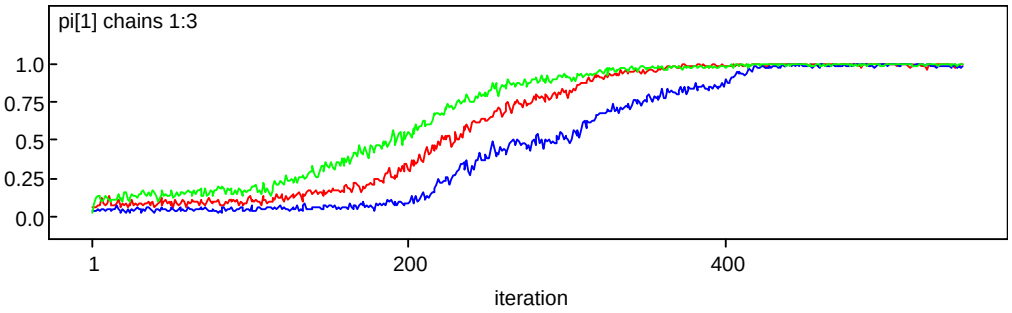
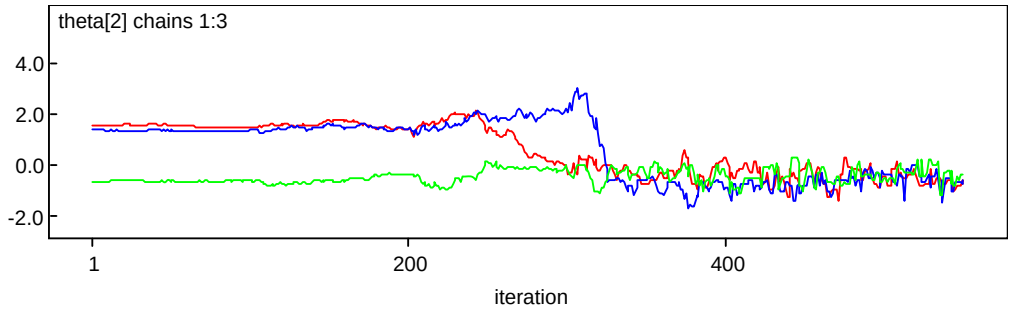


Figure 4-10: Three chain convergence for Theta



Decisions for next stage of analysis

The conditions of 50 and 250 burn-in cycles for both levels of estimation cycles, 50 and 500, were considered unsatisfactory for further investigation in the next stage of the study but were examined thoroughly in the first stage to assist the overall goal of quality estimation. The 8000 burn-in cycle condition was also not

considered useful in further investigation after the examining the preliminary dataset because it yielded no distinguished improvement in Gelman-Rubin statistics even when compared with the modest 500 burn-in condition. The history convergence and density plots also did not show any reason to keep it over other more concise conditions. The removal of the 8000 cycle conditions reflects the undesirability of using extreme burn-in when unnecessary for large scale simulations.

The 50 and 250 burn-in cycle conditions served as a baseline for unwanted Gelman-Rubin statistics, providing visual example of what is not desirable for this model. These two conditions displayed statistics that did not level off or come near to the value of 1.05. The Gelman-Rubin statistic was greatly increased in the two conditions than in other conditions, noticeably the 50 cycle condition to the 250 burn-in cycle condition was a large jump toward a good convergence. This change from 50 to 250 burn-in cycles represents a marked improvement, but remains unsatisfactory for the next stage of investigation.

All other conditions for further consideration in the first stage of analysis are so similar to the most extreme and time consuming condition of 8000 burn-in cycles that it will not be considered further in the second stage of investigation--unless the Gelman-Rubin statistics indicate in the other burn-in conditions that the more cycles are needed and this condition needs to be reinvestigated. It is considered from preliminary results that 500 estimation cycles used to estimate the Gelman-Rubin statistics in the current stage of investigation should be increased to provide more stable conditions in the estimation cycles.

The investigation of the density for these conditions was also informative in helping to provide further evidence for the selection of the number of burn-in cycles, as well as total estimation cycles. Despite the extremely low number of replications of 50 burn-in cycles, the density never fully overlaps in the estimation cycle examples provided. Moving up the number of burn-in iterations to the 250 condition the density already displayed some overlapping in the chains, even when only 50 estimation cycles are examined though not condensed enough to be used in further analysis. Showing the density with low frequency, as in the first 50 cycles, show estimation cycles that have high disturbance. The history of cycles shows three chains are easy to distinguish in the 50 burn-in condition regardless of the number of estimations afterward, while in the 500 burn-in cycles with 500 replications start to produce a useful mix of chains. The exploration of the densities support the Gelman-Rubin statistic in that 500 burn-in cycles with 500 estimation cycles look to converge well even with dispersed initial values on the three chains.

The history of the multiple chains shows what is going on in the burn-in cycles as well as in the estimation cycles. In order to have thorough exploration of values, the history is a tool to provide evidence in evaluating the burn-in cycles, as well as what was being kept afterward in the estimation cycles. Throwing out too many cycles can be very wasteful, as in the 8000 burn-in case, and overly time-consuming in simulations. The primary focus of the history review as with the density is on the early cycles to estimate good burn-in for the next stage of the investigation.

It is apparent that for item difficulty and person ability that very little effect occurs in the burn-in cycles until somewhere around 400 to 450 cycles are discarded..

This is representative of all item difficulties. Theta and class proportion values followed a similar pattern with visual convergence for most ability estimates beginning to converge around 350 cycles and the class proportion converging just after 400 cycles.

The history also indicates evidence of support for the Gelman-Rubin statistic as very little if any improvement is seen after 450 or so estimations are used as burn-in. The conclusion from the history graphs support a 500 burn-in cycle stage for the next leg of investigation.

Burn-in and estimation cycle selection

In this stage of analysis, four datasets were analyzed including the dataset under from stage one of the burn-in and estimation cycle investigation. Both the estimation cycles and the burn-in cycles are explored here. Four burn-in conditions and 3 estimation cycle conditions are further investigated. The burn-in cycle values from the first stage have been narrowed down to 500, 1000, 2000, and 4000. 2500, 5000, and 7500 estimation cycles are scrutinized to choose a relatively fast convergence that still yields high-accuracy estimation. The Gelman-Rubin statistic, history and density plots are explored for all four datasets under all 12 conditions.

These four datasets are used as a sampling of datasets from the final set of cells for analysis without having to test each cell. The sample size for this portion of the analysis is limited to 500, but is explored further to ensure a reasonable sample size. The datasets explored:

Dataset 1 is generated to conform to the Rasch model with Theta values of N(0,1)

Dataset 2 is a mixture of Rasch and random data. Replication's 1 through 250 are generated to conform to the Rasch model and have Theta N(0,1). Replication's 251 through 500 are generated as a mixture with Rasch Theta N(0,1) for the first 30 items and randomly generated response for the remaining 10 items.

Dataset 3 is a mixture of Rasch and random data for all replications. Replication's 1 through 500 are generated as a mixture with Rasch Theta N(0,1) for the first 30 items and randomly generated response for the remaining 10 items.

Dataset 4 is a mixture of Rasch and reverse Rasch. Replication's 1 through 250 are generated to conform to the Rasch model and have Theta N(0,1). Replication's 251 through 500 are generated as a mixture with Rasch Theta N(0,1) for the first 30 items. The final 10 items are generated to have the reverse Rasch

Evaluation of accuracy for burn-in cycles

The Gelman-Rubin statistics for all conditions are examined. All Gelman-Rubin statistics for the data conditions are similar and meet the 1.05 criteria for all parameters of difficulty theta and class proportion. In addition to the Gelman-Rubin statistic, the generated conditions of Rasch and the random responses mixed with Rasch meet all criteria for evaluation of history of estimation cycles and density plot convergence.

These results for the worst of the divergent Gelman-Rubin statistic are displayed here and the improvement is displayed in the 1000 and 2000 burn-in cycle conditions.

Figure 4-11: 500 burn-in cycles condition

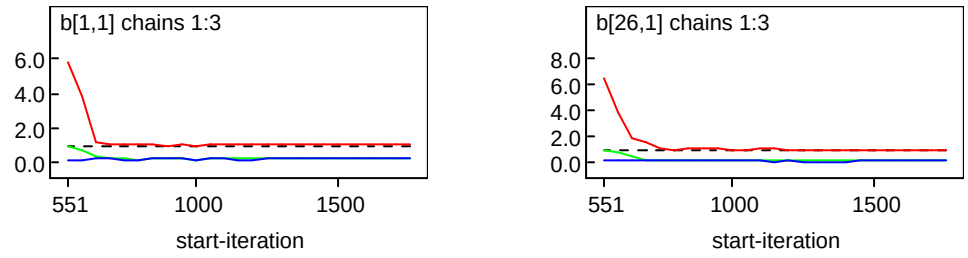


Figure 4-12: 1000 burn-in cycle condition

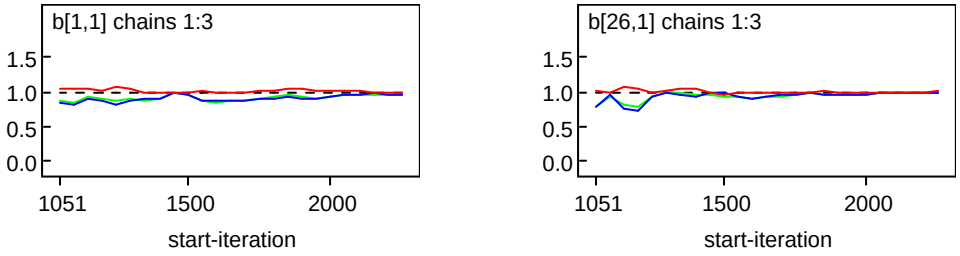
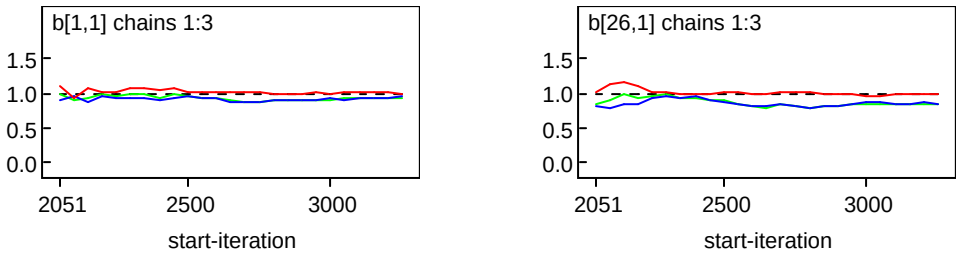
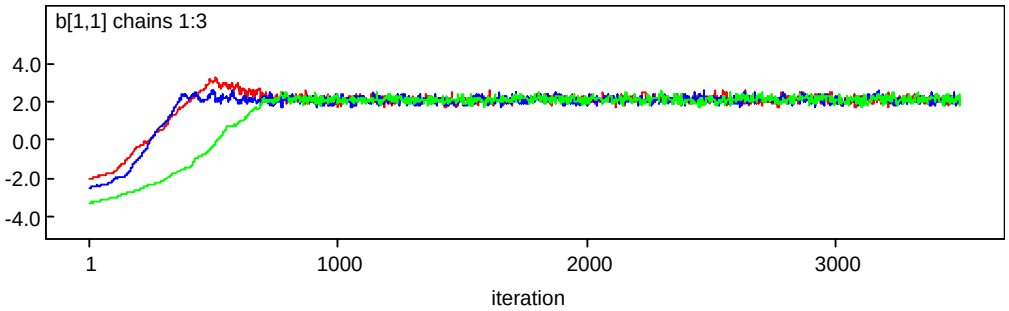


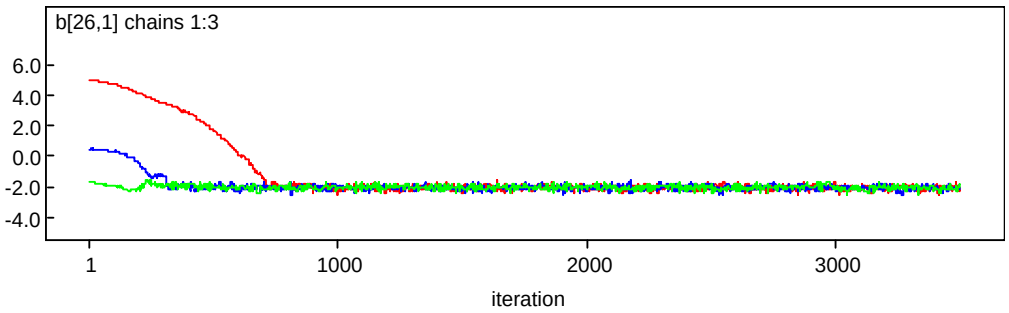
Figure 4-13: 2000 burn-in cycle condition



In supporting the Gelman-Rubin statistic the history of cycles Figure 4-14, shows convergence in all parameters prior to 1000 cycles.

Figure 4-14: Three chain convergence for multiple item difficulties





The ability and class parameters follow the same pattern as the difficulty parameters displayed above. The following in Figures 4-15 and 4-16 show the improvement from 500 burn-in cycles to 1000 burn-in cycles for a theta example and class proportion examples.

Figure 4-15: Improvement in Theta from 500 to 1000 burn-in cycles

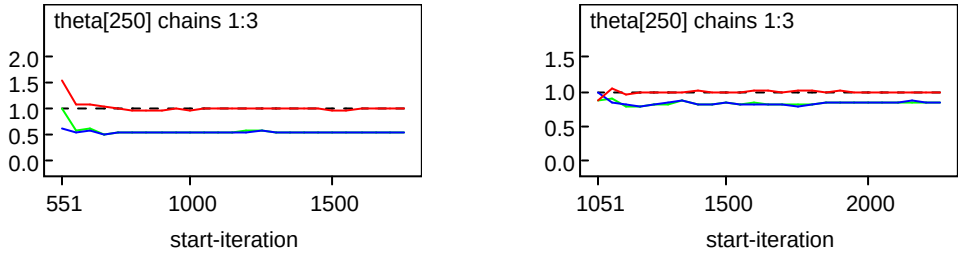
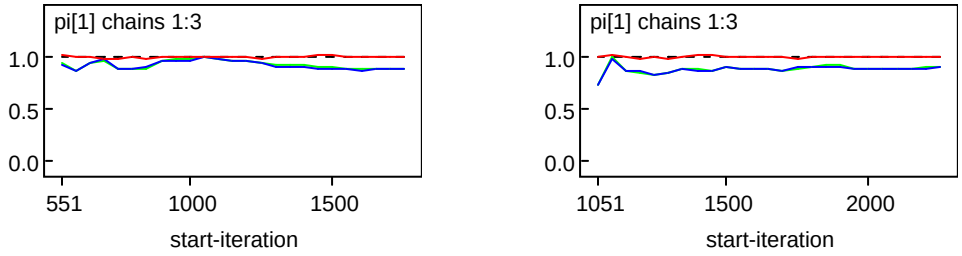


Figure 4-16: Improvement in class proportion from 500 to 1000 burn-in cycles



Final burn-in selection

The 1000, 2000 and 4000 burn-in conditions seem sufficiently well behaved for all parameters. Considering parsimony, the 4000 burn-in cycle case is ruled out of selection. This case adds a few thousand wasted cycles being thrown out without

justification. The number of estimation cycles beyond 2500 does not change the conclusion that 1000 burn-in cycles would probably be likely to be acceptable. There are additional considerations besides the empirical evidence for selection. The default in Winbugs for the adaptive burn-in stage is 2000 cycles. This is commonly used as the number of cycles thrown out in a given analysis by default. The datasets are representative of the datasets to be used in the final research; however, they do not cover all cell conditions. Also, effect sizes may be implemented that are stronger than those presented here. Regardless, the conservative selection is to choose the 2000 burn-in condition. In further preliminary analysis, the cycles are checked to make sure they still hold up to inspection.

Selection of estimation cycle size

The selection focus for estimation cycle selection is different than the burn in cycles from the previous section. The 2000 burn-in cycle condition has been chosen for accuracy using the Gelman-Rubin statistic and supporting graphical evidence, under several data conditions. All three estimation conditions were tested: 2500, 5000 and 7500. The criteria for estimation cycle selection are accurate reproduction of parameters in posterior statistics and an evaluation of those differences in parameters given the size of the cycle. The goal here is to choose estimation size that accurately but prudently yields good estimation. Reproduction of parameters should occur when it is expected, such as in the Rasch-only condition. Parameters will also be examined and selected on the basis of the difference between the parameters given the estimation cycle size. All things being equal, the 7500 condition is run twice and used as the baseline for differences in parameters.

Statistics for Rasch only condition

Inspection of the Rasch condition and expected replication of parameters is examined using the posterior means of the data. The posterior means are compared to the generating parameters for item difficulties to show how closely parameters are being reproduced in each condition.

When looking across all item difficulties, the absolute average difference remains the same for all three conditions with a mean of .102. The average MC error is also very low for all conditions of the Rasch generated data: 2500 cycles = .043, 5000. The difference in theta value between 7500 cycles and 2500 cycles is .002. The difference between the 7500 and 5000 cycles case is .0003. The class proportion difference amongst all groups is .0001. These values are effectively the same under the Rasch condition. When estimating the parameters a second time using the same data, the difference between theta values is .0004 slightly greater than the difference between the 5000 and 7500 cycle condition. The proportion was estimated to be the same to the 6^{th} digit.

Statistics for the remaining datasets

The remaining datasets are evaluated and described earlier in the burn-in and estimation cycle selection. Table 4-1 contains three comparisons of item difficulties. The first comparison is a Rasch generated dataset estimated twice. The second data comparison is of Rasch generated data compared to a mixture of Rasch and random data with 50% Rasch 50% contaminated data. The third data comparison is of Rasch generated data compared to a mixture of Rasch and random with 100% contamination.

Table 4-1: Comparisons of difficulty differences

	Repeated Rasch N(0,1)	Half Rasch Random	Full Random
b[1,1]	0.001	0.005	0.005
b[2,1]	0.0076	0	0.007
b[3,1]	0.0051	0.002	0.0106
b[4,1]	0.0047	0	0.0114
b[5,1]	0.0063	0.002	0.0084
b[6,1]	0.0062	0.0005	0.0023
b[7,1]	0.0023	0.0033	0.0068
b[8,1]	0.0056	0.0008	0.0085
b[9,1]	0.0022	0.005	0.0081
b[10,1]	0.0007	0.0026	0.0001
b[11,1]	0.0014	0.0017	0.0015
b[12,1]	0.0008	0.0022	0.0061
b[13,1]	0.0069	0.0001	0.0039
b[14,1]	0.00213	0.00079	0.0044
b[15,1]	0.0042	0.00005	0.0013
b[16,1]	0.00173	0.0053	0.0065
b[17,1]	0.0021	0.0026	0.0024
b[18,1]	0.0033	0.0032	0.0026
b[19,1]	0.0018	0.003	0.0047
b[20,1]	0.0048	0.00214	0.0023
Mean difference	0.003543	0.002114	0.005195

Less stable parameters such as person ability are examined in table 4-2 using average absolute value differences.

Table 4-2: Comparisons of ability parameters

	Repeated, same condition 2000 burn-in 7500 estimation cycles	2000 burn-in 7500 estimation cycles compared to 2000 burn-in 2500 estimation cycles	2000 burn-in 7500 estimation cycles compared to 2000 burn-in 5000 estimation cycles
Average Theta difference	0.010221	0.014425	0.011574

Final Conclusions of Burn-in and Estimation Cycle preliminary study

In the most volatile of datasets the model is being judged on the behavior of the worst Gelman-Rubin statistics. The Gelman-Rubin statistics found in the 1000

conditions on average looked relatively stable for the five datasets examined. The 1000 burn-in cycle condition, a reasonable candidate for selection based solely on empirical evidence of the five dataset, could have been a reasonable selection. The 2000 burn-in condition was selected when considering unexamined datasets and possibly more extreme effect sizes that could require more cycles to become stable. As stated earlier, Winbugs also uses the 2000 cycles as the adaptive stage as default.

The overall class proportion parameter 'pi' and the individual class parameter 'class' converge rapidly over conditions examined. The Gelman-Rubin statistic for Pi can be calculated for each dataset and is very stable. In the example above the history clearly shows stability in small overall fluctuations in class proportion, SD of .014. The Gelman-Rubin statistic for the individual class parameter can not be calculated for each respondent in a dataset. There is, more often than not, nothing to calculate. If all chains converge to consistently place a condition in the Rasch class there is no variability to be had for the statistic. There is nothing to converge, as it is effectively a constant, and all chains and all replications are falling in the same class, nearly always Rasch. In some replications such as 214 and 313 there is enough mixing of the two classes in that instance that convergence can be calculated and graphed, allowing the Gelman-Rubin statistic to be calculated. The quick convergence and stability in the two types of class parameters makes them far less interesting for investigating the data for irregularity. However, they are of interest concerning the overall issue of convergence of the model and essential to the overall investigation.

When examining the summary statistics for item difficulty and person ability under the condition of replications for 4000 burn-in cycles and 7500 estimation

cycles, one dataset run twice, and there is a fluctuation in parameters with an average parameter difference of .00354 in this repeated Rasch condition. The condition of 2000 burn-in cycles and 5000 estimation cycles differs no more than is reasonable, with an average deviation of .00362, compared to the repeated Rasch condition of .00354. This difference provides some support for the selection of the 2000 burn-in 5000 estimation cycles condition for use in the full scale analysis. In general, most conditions have some minor discrepancy between 2500 and 75000 estimation cycles, but appear to be no more than the difference seen in Rasch under two runs of the same data. Person ability differs no more than is visible in the repeated analysis.

The value added in being able to use less burn-in estimations is important when considering estimation time is cut significantly when 500 burn-in cycles is used with 2500 estimation cycles. Unfortunately the statistics and visual inspection does not warrant this decision. The decision to go with more burn-in and estimation cycles, as would seem advisable from inspection, leads the way to more stable and accurate parameter estimates. Estimation precision is also balanced with unnecessary discarding useful estimations cycles leading to 2000 burn-in cycles being chosen over 4000. 1000 burn-in cycles might be considered sufficient if the tested datasets were the only data in consideration and were certain to span the entire study. Rather, it is understood that using 2000 burn-in cycles should be robust to account for other dataset conversion issues without grossly throwing out useful estimation cycles and without having to test every cell which would be an entire research venture in and of itself. 5000 estimation cycles, even under the worst of conditions, reproduces results

that are no more biased then rerunning the best of conditions, the 7500 cycles, and comparing the two groups.

The results of the preliminary investigation indicate reasonable reproduction of parameters such as person ability and item difficulty and it can be inferred that they should reproduce to the original generating parameters. A more thorough investigation of parameter estimation is carried out in the primary investigation.

Test Range

The preliminary investigation permitted a chance to examine the test to be used in the simulation as well as priors set on the distributions of the mixture model. As discussed earlier in the literature review, the test ranging from +2 to -2 item difficulties seems to be producing desired results and is in accordance with ranges found more commonly in practice. The explored parameters in the preliminary investigation reproduce generating parameters.

Priors

The choice to have a non-informative prior is intentional in that the research will be testing variety of contaminations to the Rasch model. One of the goals of the current investigation is to determine if one static model with one set of parameters for item difficulties and mixing parameters will find contamination of a variety of types. In future investigations this assumption may change as other researchers use alternative hypothesis underlying prior information or gain knowledge and insight into their populations that is not assumed in the current study.

In the current analysis a normal distribution (e.g. Mislevy, 1986) is used for the prior of item difficulties. In most uninformed situations the mean of the normal for item difficulties is set to 0 and 1 (Zimowski, et al., 2003) is typically used as a variance. The normal distributions variances as a weak prior were set to 3 for item difficulties by Rupp (2003). The choice of priors does not have a strong effect on the parameters as long as the sample size is sufficient. If too strong a prior is placed on a small sample size, the item difficulties may drift toward the mean (Rupp 2003). A weak prior shown to be empirically stable in reproducing item difficulties from this preliminary analysis is used with a normal distribution with mean of 0 and variance of 4 as the item difficulties in some conditions are very disperse. In the full investigation a scaling condition yields item difficulties as large as ±6 which wererecovered well under the prior N (0, 4).

Priors are tested under conditions containing 500 and 2000 burn-in cycles by 2500 and 5000 estimation cycles. When the item difficulty prior was changed from N(0, .25) to N(0, .05), no change in difficulty parameters were noticeable in the first two decimal places. (Note: WinBugs characterizes the normal distribution in terms of mean and precision rather than mean and variance, so N(0, .25) is a normal distribution with mean 0 and variance = 1/dispersion = 1/.25 = 4.)

However when the Dirichlet prior on class proportion with the vector of parameters (labeled as alpha in the Winbugs code shown earlier) was altered, comparing values of (10,10), (5,5) and (2,2), it became apparent that the weight of the prior was more than expected in the (10,10) condition; that is, as the amount of information in the prior increased from (2,2) to (10,10), the size of the unscalable

class increased as shown in the following examples. In the condition of 2000 burn-in cycles 5000 estimation cycles that are used in the study, the proportion in the all Rasch only class dropped from P= .931 Rasch, P= .069 Random, in the alpha(10,10) to P=.975 Rasch P=.025 alpha(2,2). The data is entirely generated to be in the Rasch condition. The half unscalable mixing condition from above changed from P=.852 Rasch, P=.148 Random, in the alpha (10, 10) to P=.892 Rasch P=.108 random with alpha (2,2). Due to a test length of forty items it was expected that most simulees would be estimated as being in the Rasch latent class when the data is modeled with Rasch only data. It was this researcher's judgment that nearly 7% of the cases were too large a proportion to be in the unscalable class given that the only thing being changes was the prior and the test is of a reasonable length, for the following reasons. Both of these datasets show change in proportion expected when the prior is relaxed. In the Rasch only generated condition, the unscalable class should be very small, since all data is simulated to be Rasch with only random variation in the forty items giving any class expectancy in the unscalable class. However, for the half mixed condition with θ centered at 0, the value for the unscalable class might be as large as 12.5% or P=.125. This probable value is derived from 50% of the data being contaminated in 25% of the items. The general sentiment is that when the data is contaminated it is more likely to have a larger proportion in the unscalable class. Prior to running the full scale simulation, these results of loosening alpha seem to be cautious estimation. The researcher does not want class proportions overwhelmed by prior information with loss of reasonable class proportion. The outcome of this loosened alpha on class determination seems to improve class proportion expected

values but leaves a very small prior of effectively 2 cases (alpha (2, 2)), which seems to not greatly impact the model.

Estimation time

Finally, it should be noted that time is an important component of any simulation study. The researcher has two fully dedicated machines for simulation. The first has a 2.4 Ghz processor with 512 MB of Ram, the second computer is specked at Intel Core duo 2.50 GHz with 4 GB of RAM, Initially the model was tested on a third computer with an Intel Core duo cpu T7500 @2.20 GHz and 4 GB of RAM. Using the current selection of 2000 burn-in cycles, 5000 estimation cycles with a sample size of 500 it will take approximate 15 minutes per replication using the Core duo computer. It takes almost twice as long on the single processor computers. The analysis portion of the data once simulated will take several days to complete, about 2 hours per cell, given that 50 replications are run in each cell and tetrachoric correlations and FA computations in the full macro takes up real time in simulation. With each replication, one cell is now expected to take just under 15 minutes each; 50 replications per cell will likely take somewhere between 12 hours to complete not including Winbugs traps that may occur. It is this researcher's overall goal to limit the replications to a reasonable timeframe. The current investigation proposes 65 cells. In preliminary tests on the two dedicated machines, the faster machine is estimated at completing a cycle in less than 12 hours with the second computer taking twice as long. This number of cells will put the simulation computation time around 32 computer days on the faster machine; this estimated time does not include Winbugs

errors, tetrachoric correlation, factor analysis, graphing or any descriptive analysis that will take place.

Sample size

Three sample sizes are tested to insure that a reasonable sample size is used, which reproduces expected parameters but is not excessively large. In determining the sample size, the first criterion is the accurate reproduction of parameters. The item difficulty values in each dataset are evaluated compared to generating parameters when those values are expected to replicate. This will only be done for the first 30 items in the dataset except in the entire dataset with Rasch only data where all 40 items are reviewed. Items are summed in each of the sample size conditions and the values compared. The second criteria to evaluate which sample size to use are time to complete estimation. This is a critical balancing point because even though a larger sample size will lead to a more precise measure, in Bayesian estimation the larger sample size takes proportionately more time. When examined together, the sample size used will accurately produce a model while minimizing the length of time in the model.

The first factor used for evaluating the data is precision. The average deviation from the generating parameters in the item difficulties was similar within each sample size condition. Across all five datasets, the first 30 item parameters were reproduced with reasonable accuracy. The average for all five datasets difficulty values, as well as the standard error of the mean for the five datasets is reported in Table 4-3.

Table 4-3: Average and S.E for difficulty values

Sample Size	100	500	2000
Mean	-0.225	-0.0776	-0.066
S.E.	0.084	0.0286	0.0244

Improvement in the accuracy toward reproducing the generating parameters is as expected. The parameters are reproduced more accurately as sample size increases. The trend is not linear. The average increased precision from 100 to 500 is 0.15. The average increase in precision from 500 to 2000 is .01. In addition, the mean difference between generating parameters and the estimated parameters shows that the mean of the 100 sample size does not fall within 2 SE of the 500 or 200 sample size conditions. The mean of the 500 sample size condition falls well within the first half SE of the mean of the 2000 condition. The 100 sample size condition falls well outside the %5 CI around the 500 and 2000 conditions and is significantly different than these two conditions. The 500 and 2000 sample size conditions are not significantly different than one another.

The second factor used for evaluating the data is estimation time. Time is constant across datasets within each sample size condition and increases proportionately with and increase in sample size. All data used had the same time within several seconds of each other for a given sample size condition according to Winbugs updates. Times for each condition across all datasets are reported in table 4-4 using seconds for precision as well as minutes for ease of discussion.

Table 4-4: Average estimation time in seconds and minutes

Sample Size	100	500	2000
Seconds	140	730	2925
Minutes	2.5	12.2	49.3

Time is nearly identical across all datasets within a given sample size and is a proportional linear transformation using sample size to predict the outcome of time.

When considering the two criteria used to evaluate sample size it is clear that the greater the sample size, the more precise the parameters become for reproduction of expected generating parameters. Also, time increased along with the increased sample size. In addition, the time trend is linear while the sample size is a curved function with diminishing returns. The trade off in precision from the 100 sample size condition to 500 sample size condition of .15 seems relatively important with respect to the time increase of approximately 10 minutes. The mean is also outside of the %5 CI of the other conditions. This increase in time is large but is still manageable with modest replications per cell. The increase in precision from the 500 sample size condition to 2000 sample size condition of .01 does not hold the same proportional value and adds 37 minutes to each replication within a cell. In addition, the two conditions are not significantly different from one another. Given the balancing of these factors, the middle condition of 500 sample size is used.

Final Model

Amongst the competing models, the final model for the full study is selected to be among the useful models to examine. 2000 burn in and 5000 estimation cycles

are used to gain stable parameters. Each replication in a cell will have a sample size of 500. In the next chapter, the model is tested to determine the number of replications per cell and evaluate a method for determining meaning in the factor patterns.

Chapter 5: Preliminary Investigations: Results of pilot study using current method

The goal of the current preliminary investigation is to determine whether patterns are clear enough to be aggregated, or if an alternative descriptive technique will need to be used. 50 replications per cell are used and the aggregates of factor patterns are examined and discussed in the main investigation

Exploratory Method: Stage 1

The general method from the main study is used in the preliminary investigation. This method will determine the soundness of the method and allow for changes based on empirical evidence, if necessary. The central purpose here as well as in the main study is to review and interpret patterns specifically built in to the data generating process. The preliminary investigation will determine if patterns can be aggregated and interpreted as a whole within a cell, or if each individual case will need to be examined to determine which pattern it fits.

Five cells, including the Rasch generated baseline that is appropriate for the other four cells, are investigated prior to reviewing all cells in the main study. Preliminary exploration will involve running 50 replications for each of these cells. 50 replications are considered sufficiently large if overall contamination effect is more the two SE's from the baseline model. 50 replications are compared to 5 replications to discriminate between posterior classifications of class membership.

The conditions are selected to represent a variety of manipulated factors thought to conform to aggregation conditions described in the method section. The

conditions are referred to by the mixing proportions and type of mixing. All conditions involved are centered at a theta value of zero and have two equally sized subtests of twenty items each. The first subtest for all conditions is generated to conform to the Rasch model, and the second subtest is mixed as described herein: Condition one is all Rasch with no mixing of data, the baseline condition. Condition two is 80% Rasch in the first subtest and 20% unscalable in the second subtest, Condition three is 95% Rasch in the first subtest and 5% unscalable in the second subtest. Condition four is 80% Rasch in the first subtest and 20% reversed effect in the second subtest. Condition five is 95% Rasch in the first subtest and 5% reversed effect in the second subtest.

As described in the methods section, the number of meaningful factors for each model within each cell is examined. It is expected that data generated to conform to the Rasch model is one factor for the unweighted data, and contamination is a different model. Each FA model is compared to an appropriate Horn's parallel analysis to determine the number of eigenvalues that are greater than chance. The number of factors for each of the four models for weighted and unweighted data within each cell is evaluated.

Results and evaluation of posterior classification for 5 or 50 replications.

The Baseline Rasch condition posterior classification was compared to the other four conditions in two conditions, the 5 replications per cell and the 50 replications per cell. A confidence interval with 2 SE was constructed for each condition and the values compared.

Table 5-1: Rasch classification using 5 replications per cell

5 replications	%Rasch	Contamination Type	%Average Rasch Classification	-2SE	+2SE
	80	Random	98.41	97.00	99.82
	80	Reverse	96.11	92.87	99.36
	95	Random	99.84	98.18	99.70
	95	Reverse	96.88	95.42	98.33
	100	None	99.80	99.24	100

Table 5-2: Rasch classification using 50 replications per cell

50 replications	%Rasch	Contamination Type	%Average Rasch Classification	-2SE	+2SE
	80	Random	98.30	97.18	99.40
	80	Reverse	96.10	92.80	99.37
	95	Random	99.17	98.81	99.53
	95	Reverse	97.50	96.24	98.76
	100	None	99.85	99.61	100

In the 5 replication condition all 4 average classification falls outside of the 2SE CI from the baseline, however all 4 of the CI cross with the baseline. In the 50 replication condition, all 2SE CI ban are separated from the Baseline condition. This result/finding/outcome gives confidence that in all 4 contaminated conditions; significantly more contamination is being classified than in the baseline condition.

In the 95% Rasch, 5% contaminated condition 2.21% and 2.88% are correctly classified out of the 5% possible. These pilot effects are not meant to be the strongest effects in the main analysis, and it is highly likely that in some of the stronger effects nearly all of the generated contamination is extracted. Given the comparison, the 50 replication condition is more favorable than the 5 replication condition because it

holds stronger evidence that the classifications is significantly different than the baseline condition.

Pilot Results:

Eigenvalues for all three datasets across all cells are compared to the appropriate HPA value for weighted conditions.

Table 5-3: Unweighted number eigenvalues comparison

Unweighted	%Rasch	Contamination Type	F1	F2	F3
	80	Random	*5.76	*2.42	1.37
	80	Reverse	*5.92	*4.55	1.30
	95	Random	*6.12	*1.67	1.35
	95	Reverse	*6.17	*2.46	1.31
	100	None	*6.36	1.43	1.32
* above the threshold for HPA					

Table 5-4: Residually weighted number eigenvalues comparison

Residual	%Rasch	Contamination Type	F1	F2	F3
	80	Random	*6.60	5.16	4.19
	80	Reverse	*4.82	3.82	2.99
	95	Random	9.23	6.82	4.98
	95	Reverse	*6.38	4.84	4.11
	99.85	None	NA	NA	NA
* above the threshold for HPA					

Table 5-5: Rasch weighted number eigenvalues comparison

Rasch	%Rasch	Contamination Type	F1	F2	F3
	80	Random	*5.71	*2.40	1.35
	80	Reverse	*5.95	*4.14	1.30
	95	Random	*6.11	*1.60	1.35
	95	Reverse	*6.23	*1.94	1.32
	100	None	*6.32	1.40	1.28
* above the threshold for HPA					

In the baseline condition, one factor is established for the unweighted and Rasch weighted datasets. The residual data for the baseline condition is too small to explore with a weighted sample size of .15% of the entire data or the equivalent of 3/4 of one person. Two factors are present in the four contaminated conditions for the unweighted data.. In the residual weighted data, the reversed contamination has one factor for both levels of contamination, while the random contamination has one factor for 20% contamination but none for the 5% condition.

MANOVA results

The multivariate analysis is conducted for all 5 conditions, comparing the factor patterns from the Rasch only weighted data to the residual weighted data.

Table 5-6: F-values for factor 1

MANOVA Factor 1	%Rasch	Contamination Type	F Value	DF	Sig.
	80	Random	68.88	59	<.0001
	80	Reverse	66.99	59	<.0001
	95	Random	35.60	59	<.0001
	95	Reverse	19.56	59	<.0001
	100	None	1.38 (.99)	59	.1265

Table 5-7: F-values for factor 2

MANOVA Factor 2	%Rasch	Contamination Type	F Value	DF	Sig.
	80	Random	81.41	59	<.0001
	80	Reverse	6.76	59	<.0001
	95	Random	40.56	59	<.0001
	95	Reverse	29.41	59	<.0001
	100	None	1.55 (1.45)	59	.0628

Except for the baseline, all factor comparisons are significantly different in the multivariate analysis.. The baseline was tested at 3% and with a 50% split in

parentheses. Overall, when the data is reweighted by Rasch and residual classes the patterns are different.

Comparing patterns with CI

In checking each pattern, a table is developed to show where the CI using 2 S.E. indicates separation of patterns. As shown in table 4-12, if a residually weighted average factor pattern falls above 2S.E. from the average Rasch weighted pattern, a value of 1 is assigned to the cell. If it is below the 2 S.E. s value of -1 is assigned to the cell and if the value falls within the 2SE band a *0* is assigned to the cell. This is used to help explain overall pattern differences.

Table 5-8: CI for factor 1 patterns

Univariate Factor 1 Patterns 1-10	%Rasch	Type	1	2	3	4	5	6	7	8	9	10
	80	Random	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	80	Reverse	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	95	Random	-1	-1	-1	-1	-1	-1	-1	0	-1	-1
	95	Reverse	-1	-1	-1	-1	-1	-1	-1	-1	1	0
	100	None	0	0	0	0	0	0	0	0	0	0
11-20			11	12	13	14	15	16	17	18	19	20
	80	Random	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	80	Reverse	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	95	Random	-1	-1	-1	-1	-1	0	-1	-1	-1	0
	95	Reverse	-1	-1	-1	-1	-1	-1	-1	0	0	0
	100	None	0	0	0	0	0	0	0	0	0	0
21-30			21	22	23	24	25	26	27	28	29	30
	80	Random	-1	-1	-1	-1	-1	-1	-1	-1	0	0
	80	Reverse	-1	0	-1	-1	-1	-1	-1	-1	-1	-1
	95	Random	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	95	Reverse	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	100	None	0	0	0	0	0	0	0	0	0	0
31-40			31	32	33	34	35	36	37	38	39	40
	80	Random	-1	-1	-1	-1	-1	-1	-1	-1	0	0
	80	Reverse	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	95	Random	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	95	Reverse	0	0	-1	0	-1	-1	-1	-1	-1	-1
	100	None	0	0	0	0	0	0	0	0	0	0

Table 5-9: CI for factor 2 patterns

Univariate Factor 2 Patterns 1-10	%Rasch	Type	1	2	3	4	5	6	7	8	9	10
	80	Random	0	0	0	0	0	0	0	0	0	-1
	80	Reverse	0	0	1	0	0	0	0	0	1	0
	95	Random	0	0	0	-1	-1	0	-1	-1	0	-1
	95	Reverse	-1	0	-1	-1	-1	-1	-1	-1	-1	-1
	100	None	0	0	0	0	0	0	0	0	0	0
11-20			11	12	13	14	15	16	17	18	19	20
	80	Random	0	0	-1	0	0	0	0	0	0	0
	80	Reverse	0	0	0	0	0	0	0	0	0	0
	95	Random	0	0	-1	0	-1	-1	-1	0	-1	-1
	95	Reverse	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	100	None	0	0	0	0	0	0	0	0	0	0
21-30			21	22	23	24	25	26	27	28	29	30
	80	Random	0	0	-1	0	-1	-1	-1	-1	-1	-1
	80	Reverse	1	1	1	1	0	0	-1	-1	-1	-1
	95	Random	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	95	Reverse	1	0	-1	0	-1	-1	-1	-1	-1	-1
	100	None	0	0	0	0	0	0	0	0	0	0
31-40			31	32	33	34	35	36	37	38	39	40
	80	Random	0	0	0	-1	-1	-1	-1	-1	-1	-1
	80	Reverse	1	1	1	1	0	0	-1	-1	-1	-1
	95	Random	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	95	Reverse	1	1	0	0	-1	-1	-1	-1	-1	-1
	100	None	0	0	0	0	0	0	0	0	0	0

In both factors, the CI comparisons in the baseline have no significant difference amongst all 40 patterns in both factors. The other four cells have many significant effects mostly in a negative direction. These effects start to show patterns of suppression in the random and reversed Rasch contaminated conditions for the first factor. The second factor pattern differences may simply be attributed to the factor being extracted out for the Rasch condition and not for the residually weighted data. These patterns and differences are further explored in the results and discussion of the overall research.

In order to better understand what is left over in the unscalable class, the next comparison focuses on the residual weighted data and compares the Rasch subtest to the contaminated subtest.

Rasch subtest patterns compared with Residual subtest patterns. These values are graphically compared and aggregated for like difficulties within each subtest.

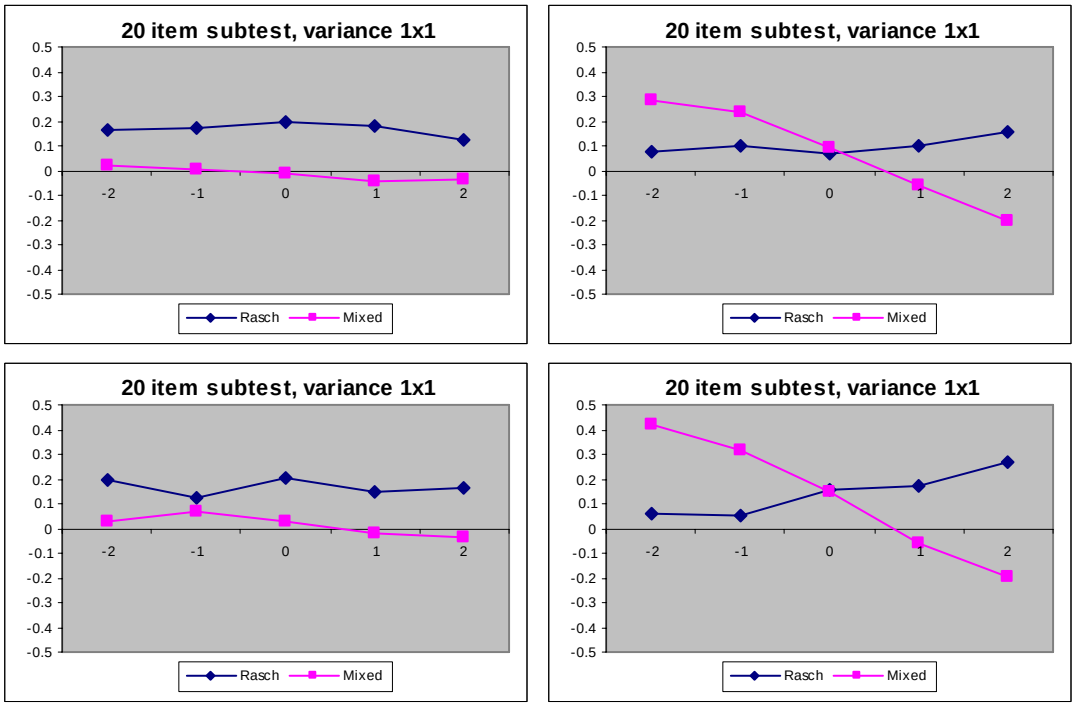
Table 5-10: Factor 1 Subtest patterns

%Rasch	Type	Subtest 1					Subtest 2				
		2	1	0	-1	-2	2	1	0	-1	-2
80	Random	0.17	0.17	0.20	0.18	0.12	0.02	0.00	-0.01	-0.05	-0.04
80	Reverse	0.08	0.10	0.07	0.10	0.16	0.29	0.24	0.09	-0.06	-0.20
95	Random	0.20	0.13	0.20	0.15	0.16	0.02	0.07	0.03	-0.02	-0.04
95	Reverse	0.06	0.05	0.16	0.18	0.27	0.42	0.31	0.15	-0.06	-0.19

Table 5-11: Factor 2 Subtest patterns

%Rasch	Type	Subtest 1					Subtest 2				
		2	1	0	-1	-2	2	1	0	-1	-2
80	Random	0.12	0.11	0.09	0.09	0.07	0.00	0.02	-0.01	0.00	-0.01
80	Reverse	0.08	0.12	0.12	0.09	0.10	0.12	0.09	0.09	0.02	-0.04
95	Random	0.13	0.08	0.10	0.14	0.13	0.07	-0.02	0.00	-0.06	0.02
95	Reverse	0.12	0.08	0.10	0.07	0.08	0.14	0.12	0.09	0.04	-0.03

Figure 5-1: Factor 1 for 80% Random, 80% Reversed, 5% Random and 5% reversed



In the random contamination condition in the residual weighted data, the random contamination looks flat compared to the Rasch generated portion of the data in both the 20% and 5% conditions.

In the reversed Rasch contamination condition in the residual weighted data, the reversed portion of the subtest has a backwards pattern when compared to the Rasch portion of the data in both the 20% and 5% conditions.

These patterns are explored further in the results chapter of this document and have been shown here as preliminary work to make support further investigation. In general, the patterns in the random contaminated condition for the contaminated portion of the subtest have smaller absolute values compared to the Rasch generated subtest in those conditions. In the first factor, the patterns in the reverse effect contaminated subtest have flipped nearly perfectly, but the effect looks stronger in the

5% conditions than in the 20% conditions. The patterns are large when the patterns are small in the all Rasch subtest, and they are small or negative when the all Rasch subtest is large. The zero difficulty has not changed in either subtest and is reflected here as well. The second factor seems to be similar to the random effects condition.

Chapter 6: Results

The first stage of evaluation is to examine the size of the unscalable class from the two-class model. The second stage is the determination of the number of factors in those conditions from which it is reasonable to proceed. The third stage is an evaluation of the multivariate effect within each cell amongst all items between the two level conditions of Rasch and residually weighted data. The fourth stage is an exploration of the univariate effects on each item. The final stage is a series of graphs, first between the Rasch and residually weighted factor patterns, then the exploration of the residually weighted data condition.

In this investigation 72 conditions were used with 50 replications per condition for a total number of 3600 replications. There was very little if any convergence issues in this analysis that could not e contributed to computer operational errors. Once the appropriate settings for each computer used in this study were set only 11 nonconvergence replications occurred. These minimal problems where spread out over conditions with no clear pattern and may have been caused by interactions with the computer during simulation. Overall convergence was not an issue with these generating parameters.

The class proportions for the class are shown in a series of tables as a percentage classified into the unscalable condition. The remainder is classified into the Rasch condition. First, Table 6-1 shows the average unscalable percent produced from the eight Rasch only baseline conditions. Table 6-2 shows the average unscalable percent for all other conditions.

Table 6-1: Average percent unscalable for Rasch, Baseline, conditions

Scaling factor	Subtest size	Mean %	Standard Deviation
1X1	10	0.1434	0.1174
	20	**0.1487**	**0.1203**
3X3	10	0.0007	0.0048
	20	0.0000	0.0002
3X1	10	0.0009	0.0050
	20	0.0052	0.0178
1X3	10	0.0199	0.0428
	20	0.0020	0.0077

The largest of these unscalable class percentages is μ=.1482% σ=.1203. The Rasch class has just over 99.85% classified as Rasch. The smallest of the Rasch only unscalable class is nearly perfectly scaled as Rasch, with μ=.0000% σ=.0002 being the residual or unscalable class. On average, the amount classified into the unscalable class across all eight Rasch conditions was μ=.0401% σ=.0657. All of these conditions are baseline conditions for the other 64 cells. The size of the unscalable class for the Rasch baseline conditions are too small to have any meaningful investigation of the residually weighted data, and typically the weights in most cases are simply representing a random anomalous case or two where a simulee's vector of scores is weighted as partially unscalable.

Table 6-2 shows the percentage classified into the unscalable class for the random and reverse contamination conditions across all levels of contamination, subtest size and subtest variability.

Table 6-2: Average proportion unscalable for the random and reverse contamination conditions

			Random		Reverse	
% Contaminated	Range	Subtest	Mean	Standard Deviation	Mean	Standard Deviation
100	1X1	10	**2.5083**	0.8827	0.1227	0.0858
		20	**6.3493**	1.7142	0.1645	0.1543
	3X3	10	0.0111	0.0374	0.0001	0.0005
		20	0.0542	0.0768	0.0001	0.0002
	3X1	10	0.0117	0.0504	0.0006	0.0025
		20	0.0477	0.0737	0.0006	0.0021
	1X3	10	**2.6328**	0.9885	0.0275	0.0630
		20	**6.5922**	1.5231	0.0064	0.0263
50	1X1	10	**0.6445**	0.3489	0.2633	0.1193
		20	**1.9236**	0.7773	**0.5315**	0.2436
	3X3	10	0.0072	0.0274	0.0120	0.0374
		20	0.0635	0.0807	**44.0916**	16.4291
	3X1	10	0.0063	0.0225	0.0003	0.0011
		20	0.0274	0.0556	0.0038	0.0091
	1X3	10	**1.1385**	0.5103	**0.5642**	0.3255
		20	**18.4359**	18.0213	**50.0769**	0.1119
20	1X1	10	0.3034	0.1914	0.3791	0.2168
		20	**1.7076**	0.5548	**3.9066**	1.6428
	3X3	10	0.0093	0.0227	0.0099	0.0385
		20	**1.4665**	0.6717	**20.0404**	0.0016
	3X1	10	0.0043	0.0265	0.0004	0.0015
		20	0.0135	0.0352	0.0257	0.0435
	1X3	10	**1.5743**	0.7086	**7.8835**	4.2881
		20	**19.1810**	0.5119	**20.0722**	0.0437
5	1X1	10	0.1667	0.1144	**0.4030**	0.2984
		20	**0.8253**	0.1819	**2.4993**	0.6300
	3X3	10	0.0419	0.0744	0.2420	0.2288
		20	**3.4653**	0.5816	**5.0104**	0.0015
	3X1	10	0.0001	0.0003	0.0002	0.0005
		20	0.0130	0.0367	0.0954	0.1239
	1X3	10	**1.2520**	0.4837	**4.6536**	0.3724
		20	**4.8105**	0.1657	**5.0303**	0.0483

A threshold of .4% for the residual class was used to determine further investigation feasibility. The .4% threshold is over 2SE apart from the largest of the Rasch baseline conditions and holds value from an exploratory perspective,

particularly when investigating a small sample size that could be evaluated graphically, as seen in later graphs. Values smaller than .4% are now deemed too close to the Rasch only conditions and systematically had class sizes that were not significantly different than the baseline. Equally of importance is that for conditions with residual class size less than .4% the meaning of factor patterns become unclear as the values could sometimes be weighted to one or two cases and analysis could not be conducted. To follow this section is an examination of eigenvalues, MANOVA, CI, and the examination of the residual condition using the .4% cut point for examining the residual class.

In the condition in which the Rasch subtest had the smaller scaling factor of 1 for the base range of items and the contaminated subtest had a scaling factor of 3, the residual was always large enough to explore except in the reverse Rasch condition. When the scaling ranges are switched, the mixed scaling factor contamination conditions with range of ±6 for the Rasch subtest, and ±2 for the contaminated subtest were always too small to be examined. When the Range of the subtest was equal, the size of the contaminated subtest affected the size of the residual. When the subtest with both ranges had 10 items for the 50%, 20% and 5%, only one of the 6 residuals crossed the .4 threshold, for the 5% condition and even then only with a value of .403. In contrast, all 6 of the 20 item subtest condition exceeded the .4 cut point. The same pattern was duplicated for the 12 values in the Random condition. When both ranges had 10 items for the 50%, 20% and 5%, only one of the 6 residuals crossed the .4 threshold this time for the 50% condition. In contrast, all 6 of the 20 item subtest condition exceeded the .4 cut point.

Overall, the size of the contaminated subtest had a large impact on the magnitude of the residual. In rank order, it was always larger as a residual for true contaminated conditions, and often by an exponential magnitude.

Relative to the size of the generated contamination and what was estimated into the unscalable class, the success of the model was inversely proportionate to the size of the residual. In the 100% condition, the largest residual was 6.5%; in the 50% contamination condition, there was one residual around 50%, one around 45% and one around 18.5%.In the 20% residual, there were three residuals at approximately 20% and one around 8%. In the 5% contamination condition there were 4 residual around 5%, one at 3.5% and one at 2.5%. Proportionally, it was clear that as the size of the residual decreased, the proportion classified more accurately appreciated the maximum number of contaminated simulees.

Eigenvalues

Eigenvalues are examined for all conditions with a class value over .4%. Suggested Factors are Graphical inspection of these conditions, as shown in the residual section of the results chapter. Horn's parallel analysis (HPA), as well as visual inspection for confirmation of the 72 conditions, was used to determine the number of factors. The eigenvalues selected as exceeding the threshold value are highlighted in bold.

In the eight Rasch only baseline conditions, Table 6-3, one dominant factor is visually evident across all eight conditions for the unweighted data. It was expected from hypothesis 2 regarding the baseline Rasch condition that only one factor would be present when HPA was used to determine the number of factors. However, in two

of the baseline conditions for the 3x3 scaling factor condition with 10 and 20 items, a second factor is just above the threshold value.

Table 6-3: The first five unweighted eigenvalues for the Rasch only condition

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.3444**	1.4270	1.3298	1.2729	1.2361
	20	**6.3646**	1.4294	1.3192	1.2647	1.2287
3X3	10	**3.2942**	**1.6473**	1.4581	1.3812	1.3251
	20	**3.2971**	**1.5898**	1.4478	1.3724	1.3210
3X1	10	**4.0615**	1.5508	1.4091	1.3333	1.2882
	20	**4.8274**	1.4905	1.3757	1.3176	1.2763
1X3	10	**5.5334**	1.4607	1.3531	1.3003	1.2572
	20	**4.7254**	1.4821	1.3750	1.3175	1.2741

The Rasch weighted condition is nearly identical to the above unweighted condition, as the sample size is nearly identical to only a small fraction of a percentage point being extracted from each condition.

Table 6-4: The first five Rasch weighted eigenvalues for the 100% Rasch only condition

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.3366**	1.4209	1.3301	1.2728	1.2358
	20	**6.3554**	1.4227	1.3196	1.2646	1.2289
3X3	10	**3.2937**	**1.6453**	1.4582	1.3812	1.3251
	20	**3.2971**	**1.5898**	1.4478	1.3724	1.3210
3X1	10	**4.0613**	1.5489	1.4091	1.3333	1.2882
	20	**4.8269**	1.4871	1.3754	1.3176	1.2764
1X3	10	**5.5311**	1.4572	1.3533	1.3004	1.2572
	20	**4.7251**	1.4812	1.3746	1.3175	1.2741

A table was not created for the first five residual weighted eigenvalues for the 100% Rasch only condition, as there were not sufficiently large residuals to examine.

According to hypothesis 3 a second factor would be generated for all 64 systematic and randomly contaminated conditions when HPA was used to determine the number of factors in the data. The results of the unweighted and Rasch weighted

data for the 100% random and reversed contamination conditions are shown in tables 6-5 through 6-9. Only one dominant factor was represented with no substantial changing of the eigenvalues patterns in the Rasch weighted conditions with only two exceptions. In the reversed Rasch condition for the 3x3 scaling factor condition with 10 and 20 items, there is again a second factor just above the threshold value. In most other contaminated conditions, a second factor was extracted with the exception of the 5% random contamination condition, in which some conditions had only one factor.

In all other unweighted conditions explored with a residual value over .4%, there is evidence of a dominant first factor along with secondary factor, with the exception being the 5% random contamination condition with a scaling factor of 1 for the contaminated subtest. Table 6-10 through 6-15 display the unweighted eigenvalues.

Table 6-5: The first five unweighted eigenvalues for the 100% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**4.9966**	1.4497	1.3813	1.3278	1.2884
	20	**3.6212**	1.4985	1.4232	1.3777	1.3368
3X3	10	**2.7061**	1.5510	1.4472	1.3934	1.3414
	20	**2.1673**	1.5401	1.4592	1.4057	1.3634
3X1	10	**2.7098**	1.5381	1.4323	1.3813	1.3427
	20	**2.1519**	1.5485	1.4650	1.4095	1.3657
1X3	10	**4.9756**	1.4652	1.3911	1.3334	1.2923
	20	**3.6461**	1.4946	1.4308	1.3799	1.3384

Table 6-6: The first five Rasch weighted eigenvalues for the 100% random contamination condition

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**4.7300**	1.4518	1.3853	1.3348	1.2919
	20	**3.1934**	1.5147	1.4439	1.3948	1.3518
3X3	10	**2.7021**	1.5465	1.4470	1.3929	1.3416
	20	**2.1534**	1.5288	1.4549	1.4033	1.3610
3X1	10	**2.7064**	1.5352	1.4325	1.3816	1.3417
	20	**2.1407**	1.5431	1.4611	1.4078	1.3643
1X3	10	**4.7011**	1.4635	1.3905	1.3388	1.3023
	20	**3.1802**	1.5169	1.4501	1.3996	1.3550

Table 6-7: The first five unweighted eigenvalues for the 100% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.2834**	1.4438	1.3320	1.2762	1.2317
	20	**6.3489**	1.4386	1.3270	1.2746	1.2290
3X3	10	**3.2529**	**1.6212**	1.4538	1.3642	1.3118
	20	**3.2633**	**1.6094**	1.4601	1.3734	1.3234
3X1	10	**4.0404**	1.5495	1.4158	1.3479	1.2985
	20	**4.7820**	1.5067	1.3769	1.3186	1.2785
1X3	10	**5.4926**	1.4575	1.3583	1.3019	1.2580
	20	**4.7987**	1.5063	1.3846	1.3238	1.2722

Table 6-8: The first five Rasch weighted eigenvalues for the 100% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.2743**	1.4386	1.3318	1.2766	1.2321
	20	**6.3297**	1.4320	1.3282	1.2754	1.2306
3X3	10	**3.2529**	**1.6211**	1.4538	1.3642	1.3118
	20	**3.2633**	**1.6094**	1.4601	1.3734	1.3234
3X1	10	**4.0403**	1.5491	1.4157	1.3479	1.2985
	20	**4.7892**	1.5073	1.3765	1.3187	1.2789
1X3	10	**5.4944**	1.4537	1.3533	1.3019	1.2576
	20	**4.7973**	1.5013	1.3846	1.3237	1.2722

Table 6-9 shows the residual eigenvalues for the random condition. A table was not created for the first five residual weighted eigenvalues for the 100%

Reversed contamination condition as there were not sufficiently large residuals to examine.

Table 6-9: The first five residual weighted eigenvalues for the 100% random contamination condition

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	4.9537	4.0693	3.4752	3.0436	2.7023
	20	3.3837	3.0013	2.7183	2.4879	2.2820
3X3	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
3X1	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
1X3	10	4.9635	4.0566	3.4747	3.0904	2.6794
	20	3.2828	2.9278	2.6756	2.4476	2.2587

Examining the size of the residually weighted data set and comparing it to HPA value, as well as visual inspection, results in nothing better than random factors manifest in this residual data for the 100% random contamination conditions.

Table 6-10: The first five unweighted eigenvalues for the 50% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**5.3097**	**2.1779**	1.3911	1.3115	1.2617
	20	**4.9584**	**3.0156**	1.4324	1.3163	1.2548
3X3	10	**3.3961**	**2.7139**	1.4995	1.3846	1.3219
	20	**6.0490**	**2.3391**	1.4257	1.3344	1.2817
3X1	10	**3.2277**	**2.0170**	1.5063	1.3921	1.3329
	20	**4.2464**	**2.2639**	1.4871	1.3671	1.3057
1X3	10	**5.1930**	**3.2936**	1.3599	1.2942	1.2406
	20	**6.0540**	**3.7422**	1.3537	1.2851	1.2388

Table 6-11: The first five unweighted eigenvalues for the 50% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.0143**	**3.1909**	1.3578	1.2664	1.2189
	20	**5.9617**	**5.4927**	1.2639	1.2112	1.1673
3X3	10	**7.0989**	**3.0363**	1.4797	1.3511	1.2847
	20	**13.9823**	**2.8496**	1.3845	1.2424	1.1825
3X1	10	**3.7310**	**3.1335**	1.5241	1.3685	1.2932
	20	**5.7133**	**4.1207**	1.3594	1.2739	1.2234
1X3	10	**7.1292**	**5.3537**	1.3415	1.2434	1.1912
	20	**14.0150**	**4.3651**	1.2284	1.1606	1.1183

Table 6-12: The first five unweighted eigenvalues for the 20% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**5.8968**	**1.8162**	1.3854	1.2986	1.2565
	20	**5.7550**	**2.4171**	1.3693	1.2954	1.2458
3X3	10	**3.5639**	**2.8822**	1.4957	1.3770	1.3191
	20	**6.3703**	**2.6892**	1.4362	1.3361	1.2768
3X1	10	**3.6771**	**1.7518**	1.4884	1.3848	1.3231
	20	**4.6330**	**2.1005**	1.4205	1.3351	1.2866
1X3	10	**5.3210**	**3.4116**	1.3763	1.2895	1.2358
	20	**6.3740**	**4.1357**	1.3403	1.2707	1.2171

Table 6-13: The first five unweighted eigenvalues for the 20% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.0388**	**2.6310**	1.3536	1.2777	1.2277
	20	**5.9211**	**4.5453**	1.2960	1.2248	1.1816
3X3	10	**6.7404**	**3.0632**	1.4751	1.3544	1.2796
	20	**13.2091**	**2.8474**	1.3314	1.2397	1.1897
3X1	10	**3.7561**	**2.6341**	1.5258	1.3801	1.3079
	20	**4.6862**	**4.1970**	1.4005	1.2875	1.2256
1X3	10	**6.7681**	**5.3414**	1.3372	1.2535	1.2068
	20	**13.2131**	**4.4206**	1.2457	1.1694	1.1205

Table 6-14: The first five unweighted eigenvalues for the 5% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.1851**	1.4878	1.3394	1.2857	1.2463
	20	**6.1155**	**1.6660**	1.3502	1.2845	1.2454
3X3	10	**3.3104**	**2.7825**	1.5198	1.3940	1.3176
	20	**5.2100**	**2.9793**	1.5255	1.3844	1.3123
3X1	10	**3.9355**	1.5502	1.4350	1.3620	1.3110
	20	**4.7449**	**1.6343**	1.4282	1.3333	1.2823
1X3	10	**5.4496**	**2.8437**	1.3786	1.3024	1.2485
	20	**5.3952**	**4.3060**	1.4098	1.3209	1.2555

Table 6-15: The first five unweighted eigenvalues for the 5% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.2424**	**1.7370**	1.3610	1.2870	1.2392
	20	**6.1724**	**2.4635**	1.3094	1.2490	1.2027
3X3	10	**5.4173**	**3.1296**	1.4647	1.3568	1.2859
	20	**10.5617**	**3.0188**	1.3718	1.2590	1.1997
3X1	10	**3.9337**	**1.7463**	1.4485	1.3642	1.3049
	20	**4.6246**	**2.5018**	1.4305	1.3158	1.2721
1X3	10	**5.6837**	**5.1510**	1.3403	1.2610	1.2093
	20	**10.5669**	**4.5524**	1.2617	1.1888	1.1416

Table 6-16 through 6-21 display the Rasch weighted eigenvalues. In some conditions the secondary factor has been greatly reduced or eliminated. It was expected from hypothesis 4 that the Rasch weighted data would have fewer factors extracted when compared to the unweighted data when HPA was used to determine the number of factors. Overall 3 less second factors were extracted in the Rasch weighted condition. Eigenvalues selected as factors are highlighted in bold.

Table 6-16: The first five Rasch weighted eigenvalues for the 50% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**5.2563**	**2.1838**	1.3808	1.3101	1.2594
	20	**4.8807**	**3.0450**	1.4028	1.3124	1.2543
3X3	10	**3.3949**	**2.7131**	1.4906	1.3838	1.3213
	20	**6.0471**	**2.3343**	1.4098	1.3280	1.2785
3X1	10	**3.2261**	**2.0161**	1.5002	1.3906	1.3323
	20	**4.2455**	**2.2632**	1.4758	1.3659	1.3041
1X3	10	**5.0900**	**3.3006**	1.3558	1.2915	1.2418
	20	**5.6532**	**3.4170**	1.4315	1.3381	1.2901

Table 6-17: The first five Rasch weighted eigenvalues for the 50% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**5.9933**	**3.1923**	1.3518	1.2665	1.2190
	20	**5.9486**	**5.4736**	1.2624	1.2104	1.1669
3X3	10	**7.0993**	**3.0304**	1.4751	1.3443	1.2829
	20	**4.7205**	**2.0373**	1.5639	1.4598	1.3899
3X1	10	**3.7309**	**3.1335**	1.5240	1.3685	1.2932
	20	**5.7135**	**4.1201**	1.3590	1.2735	1.2233
1X3	10	**7.1261**	**5.2899**	1.3311	1.2452	1.1939
	20	**5.1578**	**2.6176**	1.5468	1.4604	1.3953

Table 6-18: The first five Rasch weighted eigenvalues for the 20% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**5.8782**	**1.8177**	1.3778	1.2971	1.2565
	20	**5.7096**	**2.4029**	1.3471	1.2896	1.2444
3X3	10	**3.5611**	**2.8815**	1.4915	1.3766	1.3188
	20	**6.1252**	**2.7122**	1.4259	1.3190	1.2670
3X1	10	**3.6751**	**1.7467**	1.4867	1.3847	1.3220
	20	**4.6326**	**2.0996**	1.4178	1.3348	1.2856
1X3	10	**5.2670**	**3.3322**	1.3739	1.2947	1.2402
	20	**4.7857**	**2.0083**	1.5045	1.4123	1.3398

Table 6-19: The first five Rasch weighted eigenvalues for the 20% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.0143**	**2.6230**	1.3483	1.2767	1.2282
	20	**5.9453**	**4.1393**	1.3048	1.2371	1.1927
3X3	10	**6.7399**	**3.0611**	1.4719	1.3541	1.2786
	20	**3.2680**	**1.6681**	1.4812	1.3986	1.3396
3X1	10	**3.7560**	**2.6341**	1.5255	1.3798	1.3079
	20	**4.6830**	**4.1951**	1.3955	1.2854	1.2247
1X3	10	**6.9567**	**4.5907**	1.3578	1.2747	1.2279
	20	**4.7957**	1.5518	1.4362	1.3630	1.3126

Table 6-20: The first five Rasch weighted eigenvalues for the 5% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.1770**	1.4819	1.3390	1.2860	1.2461
	20	**6.1091**	**1.5971**	1.3501	1.2879	1.2470
3X3	10	**3.2984**	**2.7684**	1.5085	1.3918	1.3161
	20	**3.4979**	**2.5589**	1.6080	1.4293	1.3536
3X1	10	**3.9345**	1.5531	1.4361	1.3634	1.3117
	20	**4.7444**	**1.6292**	1.4263	1.3339	1.2825
1X3	10	**5.4562**	**2.4894**	1.3883	1.3124	1.2566
	20	**4.8211**	**1.6651**	1.4506	1.3552	1.2917

Table 6-21: The first five Rasch weighted eigenvalues for the 5% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	**6.2232**	**1.7041**	1.3615	1.2880	1.2399
	20	**6.2276**	**1.9389**	1.3239	1.2628	1.2205
3X3	10	**5.3526**	**3.1253**	1.4694	1.3581	1.2860
	20	**3.3042**	**1.6349**	1.4747	1.3749	1.3190
3X1	10	**3.9336**	**1.7462**	1.4484	1.3642	1.3049
	20	**4.6256**	**2.4730**	1.4287	1.3154	1.2703
1X3	10	**5.6154**	**2.6609**	1.3838	1.3028	1.2526
	20	**4.7954**	1.5165	1.3896	1.3363	1.2908

Tables 6-22 through 6-27 show eigenvalues from the residually weighted data. It was projected in hypothesis 5, concerning residually weighted data that if HPA was used to determine the number of factors, then systematic contamination would have more factors than random contamination. Overall, two first factors were found in the random contamination condition with no secondary factors. In the systematic contamination condition, eight first factors and 6 secondary factors were extracted from the data supporting the hypothesis. These findings also support the more specific hypotheses 5a and 5b, that states if factors are found in the systematically contaminated data then there would be two factors, and when they are found in the random contaminated condition there would be one factor. In later graphs, further inspection of these factors will occur to help determine if: the factors in the systemic condition are the Rasch and reversed Rasch factor, and if the factor in the random contamination condition is a suppressed Rasch factor. Eigenvalues selected as factors are highlighted in bold.

Table 6-22: The first five residual weighted eigenvalues for the 50% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	7.7466	5.5577	4.1923	3.4137	2.8107
	20	5.3749	4.3438	3.6645	3.1735	2.7209
3X3	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
3X1	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
1X3	10	6.9340	5.2255	4.2073	3.4230	2.8611
	20	**3.6531**	2.6700	2.4070	2.2147	2.0425

Table 6-23: The first five residual weighted eigenvalues for the 50% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	NA	NA	NA	NA	NA
	20	7.6171	4.9744	3.7229	3.0868	2.5250
3X3	10	NA	NA	NA	NA	NA
	20	**4.9363**	**2.4714**	1.6372	1.4326	1.3229
3X1	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
1X3	10	9.1968	5.1176	3.7410	3.0684	2.5425
	20	**5.0409**	**3.5082**	1.6213	1.5106	1.4391

Table 6-24: The first five residual weighted eigenvalues for the 20% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	NA	NA	NA	NA	NA
	20	6.5832	5.1559	4.2184	3.5107	2.9490
3X3	10	NA	NA	NA	NA	NA
	20	6.7149	4.8974	3.9177	3.2062	2.6443
3X1	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
1X3	10	7.0682	5.6018	4.5789	3.7169	3.1274
	20	**3.8658**	2.2997	2.1267	1.9981	1.8803

Table 6-25: The first five residual weighted eigenvalues for the 20% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	NA	NA	NA	NA	NA
	20	4.8211	3.8164	3.3637	2.9936	2.6402
3X3	10	NA	NA	NA	NA	NA
	20	**4.2375**	**2.7367**	1.9371	1.7467	1.6246
3X1	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
1X3	10	**5.1941**	**3.6205**	2.9582	2.6414	2.4009
	20	**5.9762**	**4.5741**	2.0601	1.8481	1.7236

Table 6-26: The first five residual weighted eigenvalues for the 5% random contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	NA	NA	NA	NA	NA
	20	9.2327	6.8199	4.9846	3.7678	2.8354
3X3	10	NA	NA	NA	NA	NA
	20	4.7926	3.9659	3.3870	2.9499	2.6110
3X1	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
1X3	10	8.7898	6.7594	5.2235	4.2986	3.3996
	20	**5.3898**	4.1212	3.5888	3.2529	2.8652

Table 6-27: The first five residual weighted eigenvalues for the 5% reverse contamination condition.

RANGE	SUBTEST	F1	F2	F3	F4	F5
1X1	10	10.0521	6.2315	4.5616	3.4005	2.6963
	20	6.3871	4.8404	4.1082	3.5727	3.1016
3X3	10	NA	NA	NA	NA	NA
	20	**5.4787**	3.1466	2.4618	2.1411	1.8626
3X1	10	NA	NA	NA	NA	NA
	20	NA	NA	NA	NA	NA
1X3	10	**6.5605**	4.4109	3.5651	3.0622	2.7671
	20	**8.0057**	**5.4538**	3.3460	2.8577	2.5078

MANOVA

Multivariate Analysis of Variance (MANOVA) was used to evaluate the differences amongst patterns sets between the Rasch weighted dataset and the residually weighted dataset. MANOVA was used to test if in the 40 item space would display a multivariate difference in patterns on the two levels of the independent variable, Rasch and residual. The number of significant values for MANOVA supports the expectations of multivariate targeted hypothesis 6. There were a substantial number of differences. Across all levels of contamination explored, significant differences were between the Rasch weighted patterns and residually

weighted patterns are evident for most of the first factor as reported in Table 6-28. The second factor for random contamination in the residual weighted condition patterns was significantly different as reported in table 6-28. The second factor for reversed contamination in the residual weighted condition patterns had fewer significant differences as reported.

Hypothesis 6a expects an increase in scaling factor in the contaminated subtest and that an increase in the number of items from 10 to 20 items will generate more differences. There was a significantly larger number for the 20 item subtests, 15, than 10 item subtest, 8, but this finding interacts with the residual selection. This is also true of the scaling factor: 14 significant values for a scaling factor of 3 and 9 for a scaling factor of 1.

Hypothesis 6b expects that as the proportion of contamination increases, fewer residual effects are significant. In the reversed Rasch conditions, significant differences were more frequent as the proportion of contamination decreased. All values selected for the random contamination condition were significant except the second factor with 100% contamination.

Although the number of patterns for all conditions was the same, the conditions of how they were created differed from condition to condition. In order to provide a relatively fair comparison, the Wilks' lambda F value was not evaluated for significance, but instead compared to an aggregated F value from the base condition. The Baseline Wilks' lambda F values were constructed from 100 MANOVAS. Each MANOVA baseline condition was generated from a random set of weights applied to the Rasch only data. A random weighted variable was created and used to parse out

the Rasch data into the two datasets. Table D-1 and D-2 show the average Wilks' lambda F-Value for the baseline Rasch condition to compare to cells with similar unscalable class proportions. Each of the proportional splits was constructed by randomly assigning simulees to one of the two conditions, unscalable or Rasch. The MANOVA values for 50%, 20% 5% and 2% have maximized separation in that the percentages represent whole cases. In order to both correctly represent the smaller residual class weights in the data in the 1% and .4% condition, and to still be able to conduct the Factor analysis, the cases in the unscalable class are proportionately distributed in that class as follows: In the 1% condition, 10 randomly assigned simulees are randomly assigned to a an equal number of proportions of .8 and .2. In the .4% case 10 randomly assigned simulees with one simulee being assigned to a proportion of .92 and the remainder are .12. These cases represent the types of weights found in cells with proportions of those sizes. The following table, table 6-xx, shows which F-values are most meaningful. Those values that are larger than 2 SE than the average baseline F-value are highlighted in bold. Values which are significant are italicized even when they are not larger than the 2 SE from the baseline F-value.

Table 6-28: Wilks' Lambda F-values for first and second factor

			Factor 1		Factor 2	
% Contaminated	Range	Subtest	Random	Reverse	Random	Reverse
100	1X1	10	***46.21***	NA	1.12	NA
		20	***74.59***	NA	0.77	NA
	3X3	10	NA	NA	NA	NA
		20	NA	NA	NA	NA
	3X1	10	NA	NA	NA	NA
		20	NA	NA	NA	NA
	1X3	10	***51.96***	NA	1.32	NA
		20	***51.37***	NA	1.3	NA
50	1X1	10	***54.78***	NA	***27.75***	NA
		20	***30.54***	3.04	**43.83**	2.85
	3X3	10	NA	NA	NA	NA
		20	NA	1.31	NA	1.35
	3X1	10	NA	NA	NA	NA
		20	NA	NA	NA	NA
	1X3	10	***28.2***	*2.55*	**54.43**	**41.78**
		20	***65.01***	1.61	**44.1**	1.36
20	1X1	10	NA	NA	NA	NA
		20	***68.88***	***66.99***	***81.41***	***6.76***
	3X3	10	NA	NA	NA	NA
		20	***173.15***	***2.78***	***28.39***	1.47
	3X1	10	NA	NA	NA	NA
		20	NA	NA	NA	NA
	1X3	10	***52.23***	*3.24*	***65.56***	*3.29*
		20	***121.02***	*2.07*	***15.47***	1.03
5	1X1	10	NA	***18.38***	NA	***8.35***
		20	***35.6***	***19.56***	***40.56***	***29.41***
	3X3	10	NA	NA	NA	NA
		20	***19.42***	***3.62***	***19.6***	1.37
	3X1	10	NA	NA	NA	NA
		20	NA	NA	NA	NA
	1X3	10	***41.23***	***6.76***	***89.42***	1.85
		20	***17.84***	***10.64***	*3.03*	*1.93*

In the 16 random contamination conditions inspected, all 16 had F-values for the first factor that exceeded the baseline F-value for comparison well beyond the 2 SE set for comparative purposes. 11 of the 16 random contamination conditions had significant F-value for the second. The four 100% random contamination and the 5%

20 item subset with 1x3 scaling factor were no longer significant for the second factor.

In the 13 reversed Rasch contamination conditions inspected, 7 of the 13 F-values, the first factor exceeded the baseline F-value for comparison and was significant in comparison to the 2 SE set from the Rasch only condition. Only 4 of the 13 were also significant for the second factor.

It was expected that an increase in scaling factor and an increase from 10 to 20 items for the contaminated subtest would cause more detectable contamination. This type of strength seemed to have been more of an indication of whether a residual could be detected at all, not necessarily if it would be significant. It was also expected that as the proportion of contamination increased, the significant values would decrease. This expectation seems to be supported with 14 exceeding the baseline in the 5% condition, 11 exceeding the baseline in the 20% condition, 9 exceeding the baseline in the 50% condition, 4 exceeding the baseline in the 100% condition. All first factor random contamination replications that were detected were significant and exceeded the baseline. In the systematic contamination, the pattern of significance was detectable in the first factor. All five detectable residual were above the baseline F-values for the 5% condition, while only two of the four were above the baseline but all were significant in the 20% condition. Also, none of the four in the 50% condition were above the baseline but 2 were significant, and there were no usable residuals in the 100% condition.

Several Wilks' lambda values were significant beyond the .05 level but did not exceeded the threshold value of 2 SE's above the average Wilks' lambda for the

baseline Rasch generated data. These conditions will still be explored further with CI and graphically. The whole comparison is explored further in the discussion.

Confidence Intervals

Each set of patterns for all items within a condition is tested to determine if the unscalable class patterns fall outside of a two SE CI developed for each set of Rasch patterns for the associated item. Effectively, a two SE CI around the Rasch values for every item within each condition is the threshold value to determine if the unscalable value is different. If the unscalable class average pattern for an item falls above the two SE mark, the item for that condition is given a positive value of +1. If the average pattern for the unscalable class falls below the two SE mark, then the value receives a -1 value. If the average pattern for the unscalable class on the item falls within the two SE's, then the item is given a zero and is not considered different from the Rasch condition.

The CI's were conducted for all conditions in which the percentage of data used was greater than .4% as discussed in the prior sections. Many significance tests are generated, so it is expected to find some significant differences in patterns by chance alone. These tests are in conjunction with observational pattern differences in graphs at the end of this section. They additionally tell a distinguishing comparison amongst the item patterns that may be lost when explored from a multivariate perspective. These significance values are used to help provide a baseline, and are intended along with the multivariate analysis to help support the explanation of the visual inspection of the patterns graphs.

In the following CI and graphic examination of patterns, it is evident that the Rasch condition analyses show no visible difference between the two sets of patterns for the first or second factor.

In the random contamination conditions, when the residual is of sufficient size, the residually weighted dataset are suppressed patterns for the first and second factor. The Rasch weighted data has a wave-like pattern and is greatly suppressed in the residually weighted data for the random conditions as shown in the graphs to follow. The CI supports the visual representation of the data. As shown earlier, the Wilks' lambdas are significant and represent the set of patterns as a whole being different.

In the reversed contamination condition, when the residual is of sufficient size, the residually weighted dataset patterns are similar to Rasch-like patterns but have various differences depending on the condition. CI help clarify visual differences and are displayed in tables to follow The Wilks' lambda significance supports visual graphs with relatively small values for the 50% contamination condition, larger values which exceed the threshold limits for the 20% condition where visual differences are manifest, as well as the 5% condition.

The results of CI's and the MANOVAs from the previous section support the visual differences displayed in the graphs to follow the CI tables. In the CI Table 6-29 through 6-35 the entries -1, 0 1 and NA are used. -1 represents the average residual weighted data value outside and bellow the CI of the Rasch weighted data. 0 represents the average residual weighted data value is within the CI of the Rasch weighted data. 1 represents the average residual weighted data value outside and above the CI of the Rasch weighted data. NA indicates that a CI was not calculated

for these values. The -1, 0 and 1 values are to assist the visual representation of the graphs and show evidence of matching visual patterns.

Table 6-29: First factor average patterns for 100% contamination conditions

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I2	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I3	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I4	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I5	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I6	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I7	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I8	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I9	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I10	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I11	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I12	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I13	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I14	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I15	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I16	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I17	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I18	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I19	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I20	-1	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I21	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I22	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I23	-1	1	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I24	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I25	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I26	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I27	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I28	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I29	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I30	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I31	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I32	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I33	1	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I34	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I35	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I36	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I37	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I38	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I39	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I40	0	0	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA

Table 6-30: First factor average patterns for 50% contamination conditions

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I2	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	-1	0
I3	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I4	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I5	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	-1	1
I6	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I7	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I8	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I9	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	1	0
I10	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	1	0
I11	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	1	1
I12	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	1	0
I13	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	1	1
I14	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I15	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I16	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	1	0
I17	-1	0	NA	NA	NA	NA	-1	0	NA	-1	NA	0	NA	NA	1	0
I18	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I19	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I20	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	0
I21	-1	0	NA	NA	NA	NA	-1	1	NA	0	NA	0	NA	NA	1	1
I22	-1	0	NA	NA	NA	NA	-1	1	NA	0	NA	0	NA	NA	1	1
I23	-1	-1	NA	NA	NA	NA	-1	0	NA	-1	NA	-1	NA	NA	0	1
I24	-1	-1	NA	NA	NA	NA	-1	1	NA	-1	NA	0	NA	NA	1	1
I25	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	1	0
I26	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	0	0
I27	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	-1	0
I28	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	1	0
I29	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	-1	-1
I30	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	1	-1
I31	-1	0	NA	NA	NA	NA	-1	1	NA	0	NA	0	NA	NA	-1	1
I32	-1	0	NA	NA	NA	NA	0	1	NA	-1	NA	0	NA	NA	-1	1
I33	-1	-1	NA	NA	NA	NA	-1	1	NA	1	NA	-1	NA	NA	0	1
I34	-1	0	NA	NA	NA	NA	-1	1	NA	-1	NA	-1	NA	NA	1	0
I35	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	1	0
I36	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	1	0
I37	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	1	0
I38	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	1	0
I39	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	1	-1
I40	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	1	-1

Table 6-31: First factor average patterns for 20% contamination conditions

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	0	-1
I2	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	0	-1
I3	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I4	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	0	-1
I5	NA	-1	NA	1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I6	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I7	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I8	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I9	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I10	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I11	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	0	-1
I12	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	0	-1
I13	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I14	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	0	-1
I15	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I16	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I17	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I18	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I19	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I20	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I21	NA	-1	NA	1	NA	NA	-1	0	-1	-1	NA	0	NA	NA	0	1
I22	NA	-1	NA	1	NA	NA	-1	0	-1	0	NA	1	NA	NA	-1	1
I23	NA	-1	NA	1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	0
I24	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	0
I25	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I26	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I27	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	0	NA	NA	0	-1
I28	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I29	NA	0	NA	-1	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I30	NA	0	NA	-1	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I31	NA	-1	NA	1	NA	NA	0	-1	-1	-1	NA	0	NA	NA	-1	1
I32	NA	-1	NA	1	NA	NA	0	0	-1	-1	NA	0	NA	NA	-1	1
I33	NA	-1	NA	1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	0
I34	NA	-1	NA	1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	0
I35	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I36	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I37	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	1	-1
I38	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	1	-1
I39	NA	0	NA	-1	NA	NA	-1	0	-1	-1	NA	0	NA	NA	1	-1
I40	NA	0	NA	-1	NA	NA	0	0	-1	-1	NA	0	NA	NA	1	-1

Table 6-32: First factor average patterns for 5% contamination conditions

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	0	NA	NA	-1	-1
I2	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	0	NA	NA	-1	-1
I3	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I4	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I5	NA	-1	NA	-1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I6	NA	-1	NA	-1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I7	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I8	NA	0	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I9	NA	-1	NA	0	NA	NA	-1	-1	-1	1	NA	0	NA	NA	-1	-1
I10	NA	-1	NA	0	NA	NA	-1	0	0	0	NA	0	NA	NA	-1	-1
I11	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I12	NA	-1	NA	0	NA	NA	-1	-1	-1	-1	NA	0	NA	NA	-1	-1
I13	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I14	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I15	NA	-1	NA	-1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I16	NA	0	NA	-1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1
I17	NA	-1	NA	0	NA	NA	-1	0	-1	-1	NA	0	NA	NA	-1	-1
I18	NA	-1	NA	-1	NA	NA	-1	-1	-1	0	NA	0	NA	NA	-1	-1
I19	NA	-1	NA	0	NA	NA	-1	0	-1	0	NA	0	NA	NA	0	-1
I20	NA	0	NA	0	NA	NA	-1	0	-1	0	NA	0	NA	NA	0	-1
I21	NA	-1	NA	0	NA	NA	-1	0	-1	0	NA	1	NA	NA	-1	1
I22	NA	-1	NA	1	NA	NA	-1	0	-1	0	NA	0	NA	NA	-1	1
I23	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	0	NA	NA	-1	0
I24	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	0	NA	NA	-1	0
I25	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I26	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I27	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I28	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I29	NA	-1	NA	-1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	0	-1
I30	NA	-1	NA	-1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	0	-1
I31	NA	-1	NA	1	NA	NA	0	0	0	0	NA	1	NA	NA	1	1
I32	NA	-1	NA	1	NA	NA	0	0	1	0	NA	1	NA	NA	1	1
I33	NA	-1	NA	0	NA	NA	-1	-1	1	-1	NA	0	NA	NA	0	0
I34	NA	-1	NA	-1	NA	NA	-1	-1	-1	0	NA	-1	NA	NA	0	-1
I35	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I36	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I37	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I38	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1
I39	NA	-1	NA	-1	NA	NA	0	0	-1	-1	NA	0	NA	NA	-1	-1
I40	NA	-1	NA	-1	NA	NA	-1	0	-1	-1	NA	-1	NA	NA	-1	-1

Table 6-33: Second factor average patterns for 100% contamination conditions

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	-1	0	NA	NA	NA	NA	0	1	NA	NA	NA	NA	NA	NA	NA	NA
I2	0	-1	NA	NA	NA	NA	1	1	NA	NA	NA	NA	NA	NA	NA	NA
I3	1	0	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I4	-1	0	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I5	1	1	NA	NA	NA	NA	0	1	NA	NA	NA	NA	NA	NA	NA	NA
I6	1	1	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I7	0	1	NA	NA	NA	NA	1	1	NA	NA	NA	NA	NA	NA	NA	NA
I8	0	1	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I9	-1	-1	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I10	1	1	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I11	0	1	NA	NA	NA	NA	1	1	NA	NA	NA	NA	NA	NA	NA	NA
I12	1	1	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I13	-1	1	NA	NA	NA	NA	1	1	NA	NA	NA	NA	NA	NA	NA	NA
I14	0	1	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I15	1	1	NA	NA	NA	NA	1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I16	1	1	NA	NA	NA	NA	1	1	NA	NA	NA	NA	NA	NA	NA	NA
I17	1	1	NA	NA	NA	NA	1	1	NA	NA	NA	NA	NA	NA	NA	NA
I18	1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I19	0	1	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I20	-1	0	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I21	1	0	NA	NA	NA	NA	0	-1	NA	NA	NA	NA	NA	NA	NA	NA
I22	1	-1	NA	NA	NA	NA	0	-1	NA	NA	NA	NA	NA	NA	NA	NA
I23	1	1	NA	NA	NA	NA	0	1	NA	NA	NA	NA	NA	NA	NA	NA
I24	1	1	NA	NA	NA	NA	1	0	NA	NA	NA	NA	NA	NA	NA	NA
I25	-1	-1	NA	NA	NA	NA	1	0	NA	NA	NA	NA	NA	NA	NA	NA
I26	1	-1	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA	NA	NA	NA
I27	-1	1	NA	NA	NA	NA	1	0	NA	NA	NA	NA	NA	NA	NA	NA
I28	-1	-1	NA	NA	NA	NA	1	1	NA	NA	NA	NA	NA	NA	NA	NA
I29	1	0	NA	NA	NA	NA	1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I30	1	-1	NA	NA	NA	NA	1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I31	-1	0	NA	NA	NA	NA	0	1	NA	NA	NA	NA	NA	NA	NA	NA
I32	-1	0	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I33	-1	1	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I34	1	-1	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I35	-1	1	NA	NA	NA	NA	0	-1	NA	NA	NA	NA	NA	NA	NA	NA
I36	-1	0	NA	NA	NA	NA	-1	1	NA	NA	NA	NA	NA	NA	NA	NA
I37	0	-1	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I38	-1	0	NA	NA	NA	NA	-1	-1	NA	NA	NA	NA	NA	NA	NA	NA
I39	1	-1	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA
I40	-1	0	NA	NA	NA	NA	-1	0	NA	NA	NA	NA	NA	NA	NA	NA

Table 6-34: Second factor average patterns for 50% contamination conditions

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I2	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	1	NA	NA	-1	-1
I3	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I4	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	-1
I5	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I6	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I7	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	-1
I8	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I9	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I10	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	1	NA	NA	-1	-1
I11	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I12	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I13	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I14	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I15	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I16	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	-1
I17	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I18	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	-1
I19	0	-1	NA	NA	NA	NA	1	-1	NA	-1	NA	0	NA	NA	-1	-1
I20	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I21	0	-1	NA	NA	NA	NA	0	0	NA	-1	NA	0	NA	NA	-1	1
I22	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	1	NA	NA	-1	1
I23	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I24	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	1	NA	NA	-1	-1
I25	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	-1
I26	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I27	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I28	0	-1	NA	NA	NA	NA	1	-1	NA	-1	NA	-1	NA	NA	-1	-1
I29	0	-1	NA	NA	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	-1
I30	0	0	NA	NA	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	0
I31	1	-1	NA	NA	NA	NA	1	0	NA	-1	NA	0	NA	NA	0	1
I32	0	-1	NA	NA	NA	NA	1	-1	NA	-1	NA	0	NA	NA	0	0
I33	-1	-1	NA	NA	NA	NA	1	-1	NA	-1	NA	1	NA	NA	0	-1
I34	-1	-1	NA	NA	NA	NA	1	0	NA	-1	NA	0	NA	NA	-1	-1
I35	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	-1	NA	NA	-1	-1
I36	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	-1	-1
I37	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	-1	NA	NA	0	-1
I38	-1	-1	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	0	-1
I39	-1	0	NA	NA	NA	NA	-1	-1	NA	-1	NA	-1	NA	NA	0	0
I40	-1	0	NA	NA	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	0	0

Table 6-35: Second factor average patterns for 20% contamination conditions

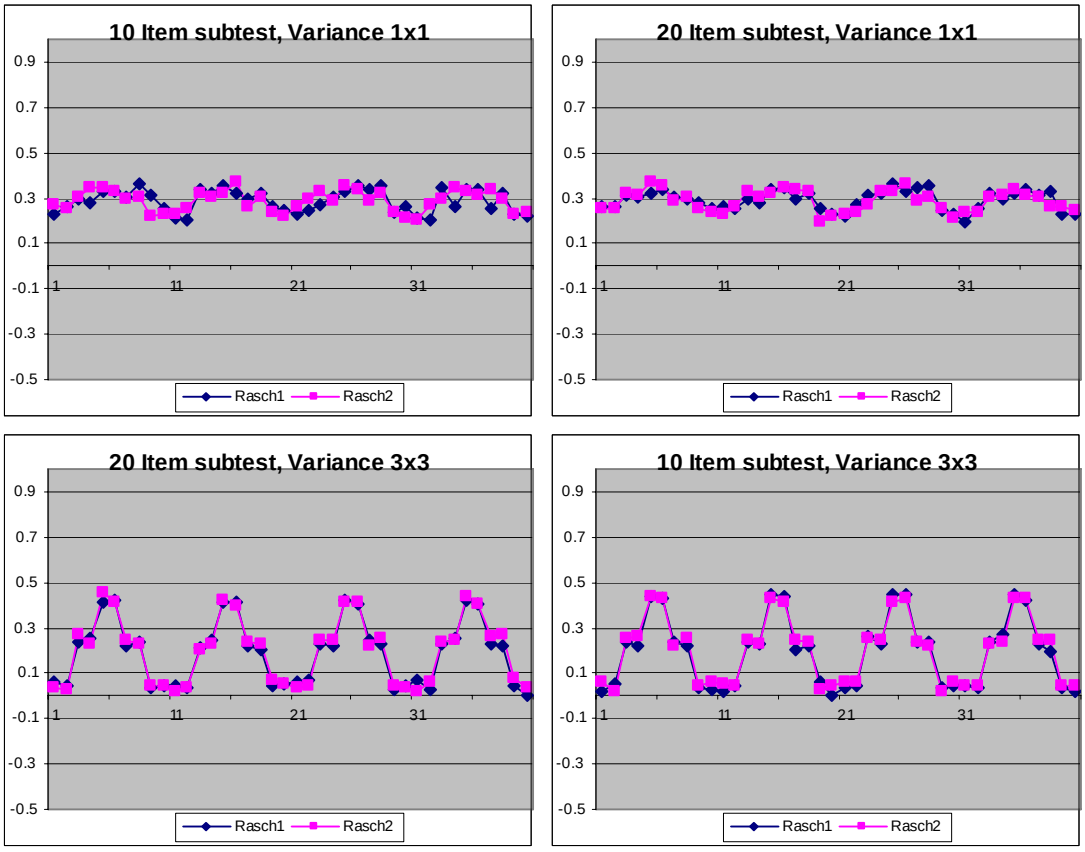
	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	0	NA	0	NA	NA	0	1	NA	0	NA	0	NA	NA	-1	1
I2	NA	0	NA	0	NA	NA	0	0	NA	0	NA	0	NA	NA	-1	0
I3	NA	0	NA	-1	NA	NA	0	1	NA	1	NA	1	NA	NA	-1	0
I4	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	1
I5	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	1
I6	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	0
I7	NA	0	NA	0	NA	NA	1	0	NA	0	NA	1	NA	NA	-1	1
I8	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	1
I9	NA	0	NA	0	NA	NA	1	0	NA	1	NA	0	NA	NA	-1	1
I10	NA	-1	NA	0	NA	NA	1	0	NA	0	NA	-1	NA	NA	-1	0
I11	NA	0	NA	0	NA	NA	0	0	NA	0	NA	-1	NA	NA	-1	1
I12	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	-1	NA	NA	-1	0
I13	NA	-1	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	0
I14	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	1
I15	NA	0	NA	-1	NA	NA	1	0	NA	0	NA	1	NA	NA	-1	1
I16	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	1
I17	NA	0	NA	-1	NA	NA	0	0	NA	0	NA	1	NA	NA	-1	1
I18	NA	0	NA	-1	NA	NA	1	0	NA	0	NA	1	NA	NA	-1	1
I19	NA	0	NA	0	NA	NA	0	0	NA	0	NA	-1	NA	NA	-1	1
I20	NA	0	NA	0	NA	NA	0	0	NA	0	NA	0	NA	NA	-1	1
I21	NA	0	NA	0	NA	NA	1	1	NA	1	NA	-1	NA	NA	-1	1
I22	NA	0	NA	0	NA	NA	0	1	NA	1	NA	-1	NA	NA	-1	1
I23	NA	-1	NA	-1	NA	NA	0	1	NA	1	NA	1	NA	NA	-1	0
I24	NA	0	NA	-1	NA	NA	0	0	NA	1	NA	1	NA	NA	-1	1
I25	NA	-1	NA	-1	NA	NA	1	-1	NA	0	NA	1	NA	NA	-1	1
I26	NA	-1	NA	-1	NA	NA	1	0	NA	0	NA	1	NA	NA	-1	1
I27	NA	-1	NA	0	NA	NA	0	-1	NA	-1	NA	1	NA	NA	-1	0
I28	NA	-1	NA	0	NA	NA	0	-1	NA	-1	NA	1	NA	NA	-1	0
I29	NA	-1	NA	0	NA	NA	0	-1	NA	-1	NA	-1	NA	NA	-1	-1
I30	NA	-1	NA	0	NA	NA	0	-1	NA	-1	NA	0	NA	NA	-1	-1
I31	NA	0	NA	0	NA	NA	1	1	NA	1	NA	1	NA	NA	-1	1
I32	NA	0	NA	0	NA	NA	1	1	NA	1	NA	-1	NA	NA	-1	0
I33	NA	0	NA	-1	NA	NA	1	0	NA	1	NA	1	NA	NA	-1	0
I34	NA	-1	NA	-1	NA	NA	1	0	NA	1	NA	1	NA	NA	-1	0
I35	NA	-1	NA	-1	NA	NA	-1	0	NA	0	NA	1	NA	NA	-1	1
I36	NA	-1	NA	-1	NA	NA	-1	0	NA	0	NA	1	NA	NA	-1	1
I37	NA	-1	NA	-1	NA	NA	-1	-1	NA	-1	NA	1	NA	NA	0	0
I38	NA	-1	NA	0	NA	NA	-1	-1	NA	-1	NA	1	NA	NA	0	0
I39	NA	-1	NA	0	NA	NA	-1	-1	NA	-1	NA	0	NA	NA	1	-1
I40	NA	-1	NA	0	NA	NA	-1	-1	NA	-1	NA	-1	NA	NA	1	-1

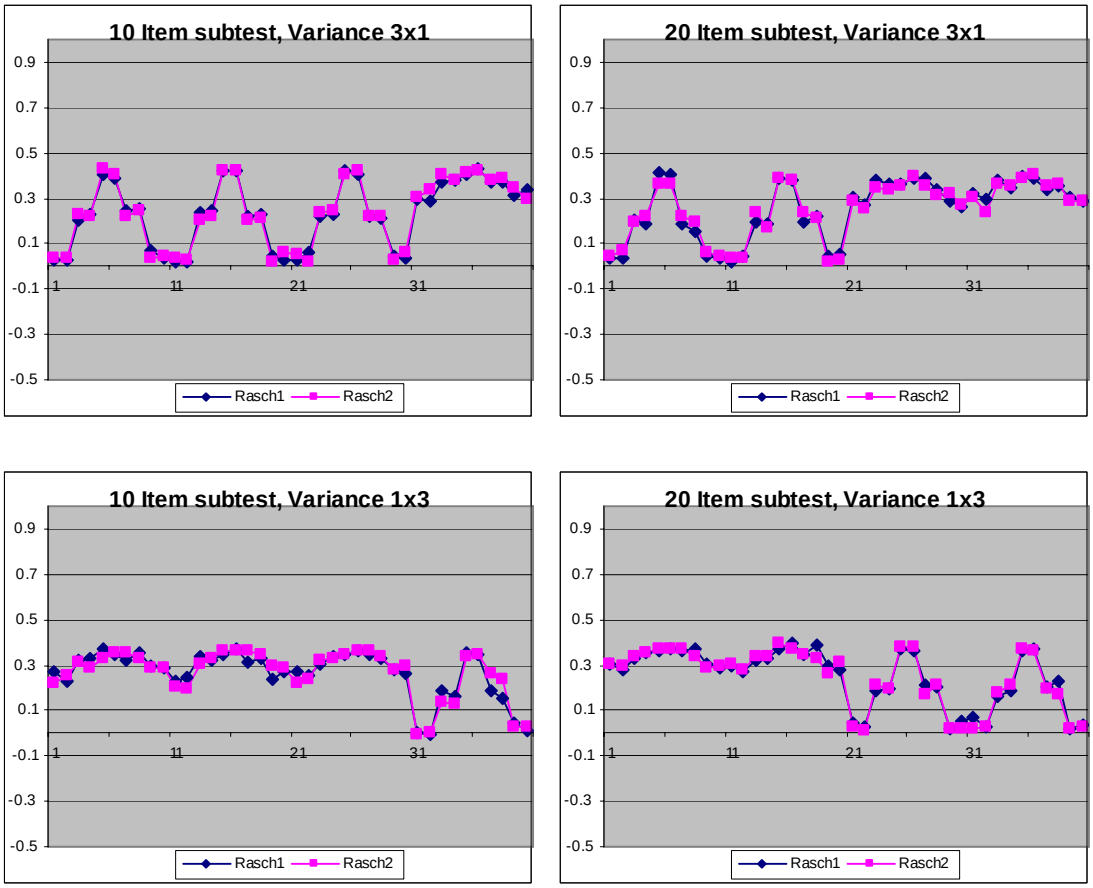
Table 6-36: Second factor average patterns for 5% contamination conditions

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	0	NA	0	NA	NA	0	1	-1	-1	NA	-1	NA	NA	1	1
I2	NA	0	NA	0	NA	NA	0	1	-1	0	NA	0	NA	NA	0	1
I3	NA	0	NA	0	NA	NA	0	0	-1	-1	NA	1	NA	NA	1	1
I4	NA	-1	NA	0	NA	NA	0	1	-1	-1	NA	0	NA	NA	1	1
I5	NA	-1	NA	0	NA	NA	0	1	-1	-1	NA	1	NA	NA	1	1
I6	NA	0	NA	-1	NA	NA	0	1	-1	-1	NA	1	NA	NA	1	1
I7	NA	-1	NA	0	NA	NA	1	1	-1	-1	NA	1	NA	NA	1	1
I8	NA	-1	NA	0	NA	NA	0	1	-1	-1	NA	0	NA	NA	1	1
I9	NA	0	NA	0	NA	NA	0	0	-1	-1	NA	-1	NA	NA	1	1
I10	NA	-1	NA	0	NA	NA	1	1	0	-1	NA	-1	NA	NA	1	1
I11	NA	0	NA	0	NA	NA	1	0	-1	0	NA	0	NA	NA	1	1
I12	NA	0	NA	0	NA	NA	0	0	-1	0	NA	-1	NA	NA	1	0
I13	NA	-1	NA	0	NA	NA	0	0	-1	-1	NA	-1	NA	NA	1	1
I14	NA	0	NA	0	NA	NA	0	1	-1	-1	NA	0	NA	NA	1	1
I15	NA	-1	NA	0	NA	NA	1	1	-1	-1	NA	1	NA	NA	1	1
I16	NA	-1	NA	0	NA	NA	1	0	-1	-1	NA	1	NA	NA	1	1
I17	NA	-1	NA	0	NA	NA	0	0	-1	-1	NA	0	NA	NA	1	1
I18	NA	0	NA	0	NA	NA	1	1	-1	-1	NA	-1	NA	NA	1	1
I19	NA	-1	NA	0	NA	NA	0	-1	-1	-1	NA	-1	NA	NA	1	1
I20	NA	-1	NA	0	NA	NA	0	1	-1	-1	NA	-1	NA	NA	1	1
I21	NA	0	NA	0	NA	NA	1	0	-1	1	NA	-1	NA	NA	1	0
I22	NA	0	NA	1	NA	NA	0	0	-1	0	NA	0	NA	NA	1	0
I23	NA	-1	NA	0	NA	NA	0	-1	-1	-1	NA	0	NA	NA	1	1
I24	NA	-1	NA	0	NA	NA	0	1	-1	0	NA	0	NA	NA	1	0
I25	NA	-1	NA	-1	NA	NA	0	-1	-1	-1	NA	1	NA	NA	1	1
I26	NA	-1	NA	-1	NA	NA	1	-1	-1	-1	NA	1	NA	NA	1	1
I27	NA	-1	NA	-1	NA	NA	1	-1	-1	-1	NA	1	NA	NA	1	0
I28	NA	-1	NA	-1	NA	NA	0	-1	-1	-1	NA	0	NA	NA	1	1
I29	NA	-1	NA	-1	NA	NA	0	-1	-1	-1	NA	-1	NA	NA	1	-1
I30	NA	-1	NA	-1	NA	NA	1	-1	-1	-1	NA	-1	NA	NA	1	-1
I31	NA	0	NA	1	NA	NA	1	0	-1	1	NA	0	NA	NA	1	0
I32	NA	0	NA	1	NA	NA	1	0	-1	1	NA	0	NA	NA	1	0
I33	NA	-1	NA	-1	NA	NA	1	0	-1	0	NA	0	NA	NA	1	1
I34	NA	-1	NA	0	NA	NA	1	-1	-1	0	NA	0	NA	NA	1	0
I35	NA	-1	NA	-1	NA	NA	0	0	-1	-1	NA	1	NA	NA	1	1
I36	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	1	NA	NA	1	1
I37	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	1	NA	NA	1	1
I38	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	1	NA	NA	-1	1
I39	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	0	NA	NA	-1	-1
I40	NA	-1	NA	-1	NA	NA	-1	-1	-1	-1	NA	-1	NA	NA	-1	-1

The Graphs to follow show what the above CI's look like and the multivariate differences as a graphic. The Rasch weighted data patterns are shown in comparison to the residually weighted data. The Rasch graphs are displayed first in order to determine what the Rasch generated data should look like under the baseline conditions, and to show what no difference looks like from a visual perspective. The graphs for the first and second factor show no visible differences supporting hypothesis 7.

Figure 6-1: Rasch patterns as baseline factor



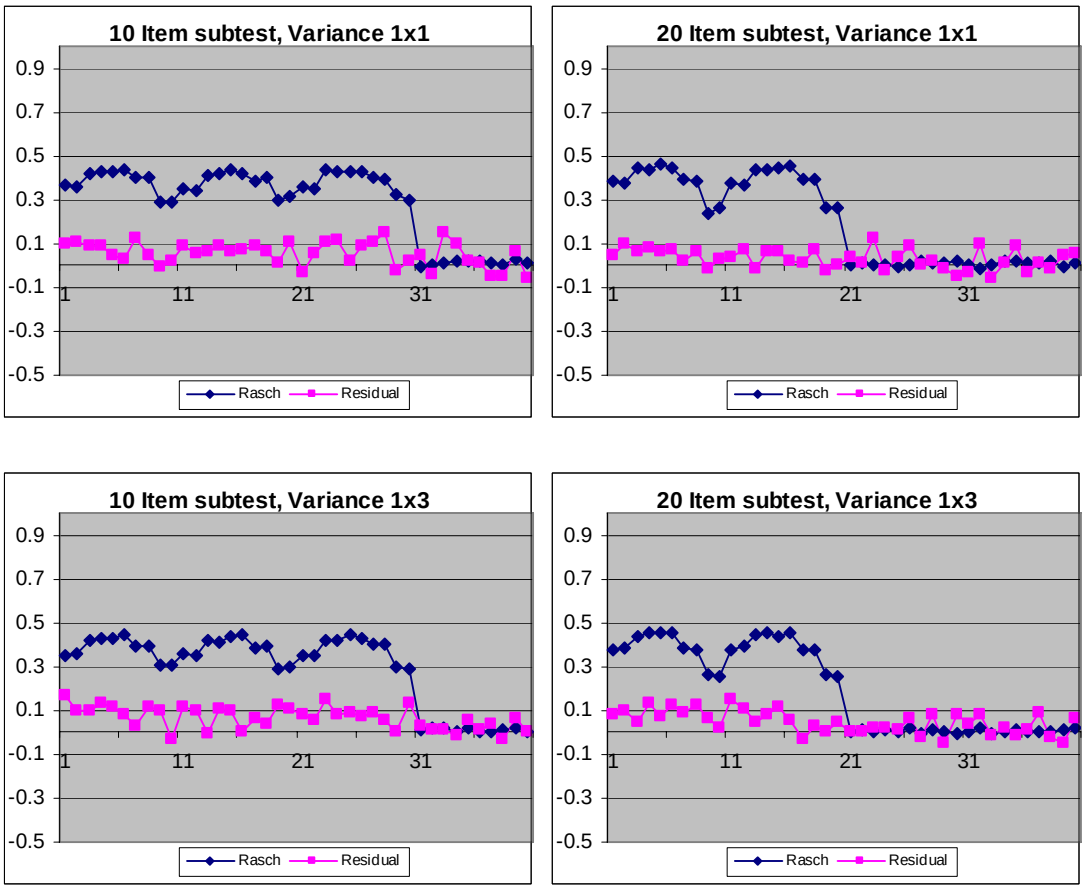


The preceding Rasch values graphically display no difference between the patterns. It was expected that the baseline data would look the same. They do, however, show what expected Rasch patterns look like for each of the conditions. The baselines above and in the second factor show that the scaling factor of 1 has a relatively flat pattern but hovers around .3, while the increase in range creates a wave live pattern with the extreme values having values close to zero, and the items close to zero have patterns near .3. In the graphs to follow, Rasch-like patterns are prevalent, particularly in the Rasch weighted data and can be visually compared to the residually weighted data.

It was generally expected in Hypotheses 8 and 9 that weighted Rasch generated data looks like the Rasch baseline data and that data generated with

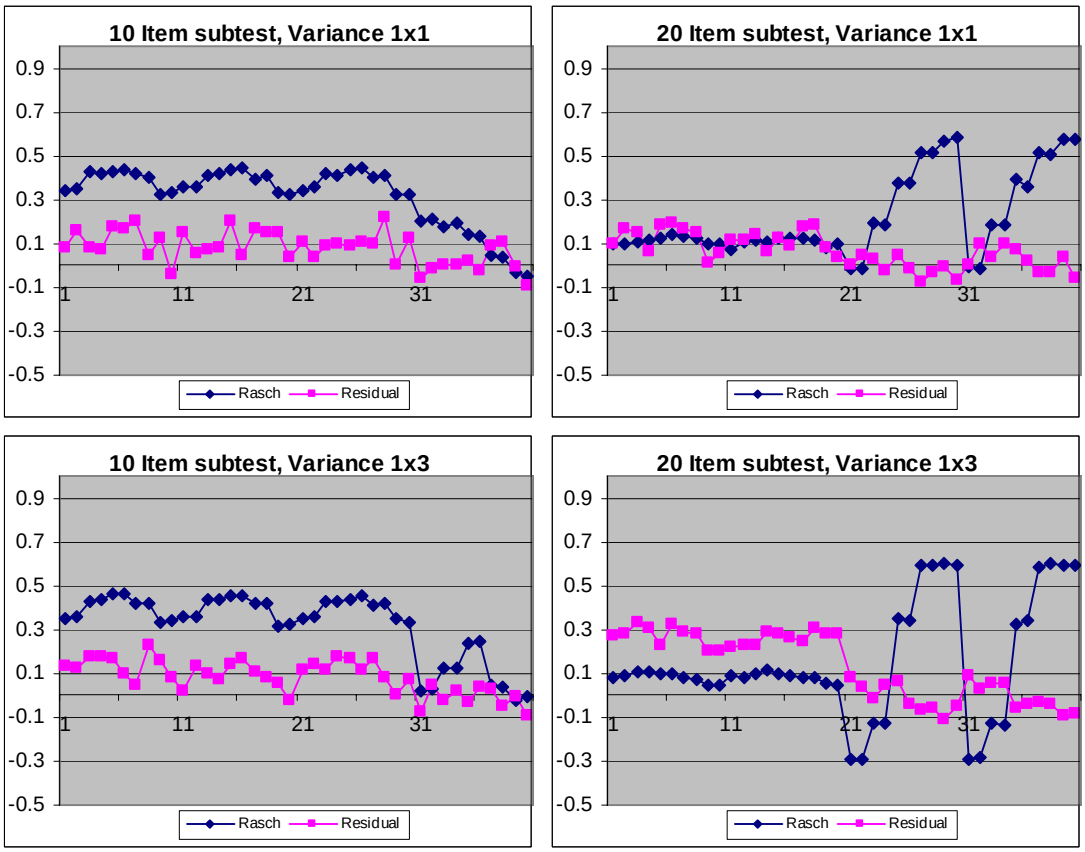
contamination would look different. In most cases, the following sets of graphs support the hypotheses. The data generated in the random contamination conditions looks suppressed. In the reversed Rasch contamination condition, offset patterns to the Rasch data are apparent.

Figure 6-2: Factor 1 residual patterns over .4% for 100% random contamination



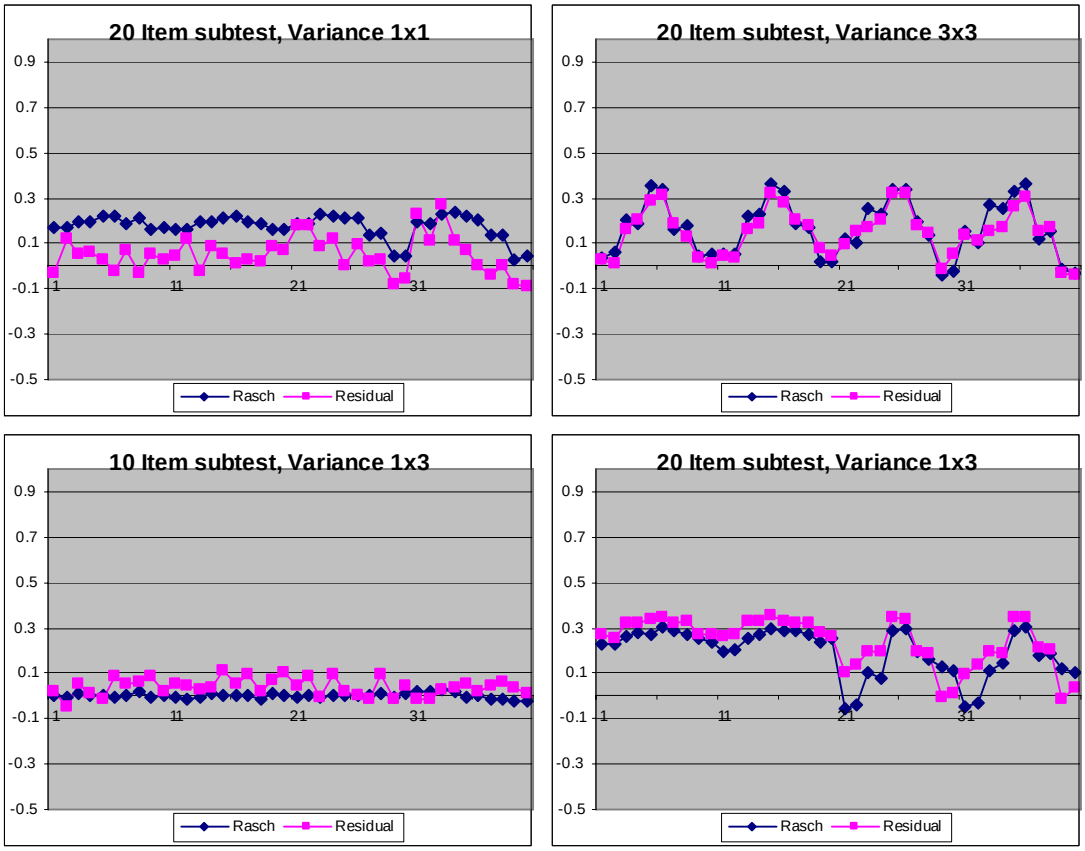
The residually weighted data condition and the contaminated portion of the subtest for the Rasch weighted data are all hovering around zero. The Rasch generated subtest for the Rasch weighted data look like Rasch baseline patterns.

Figure 6-3: Factor 1 residual patterns over .4% for 50% random contamination



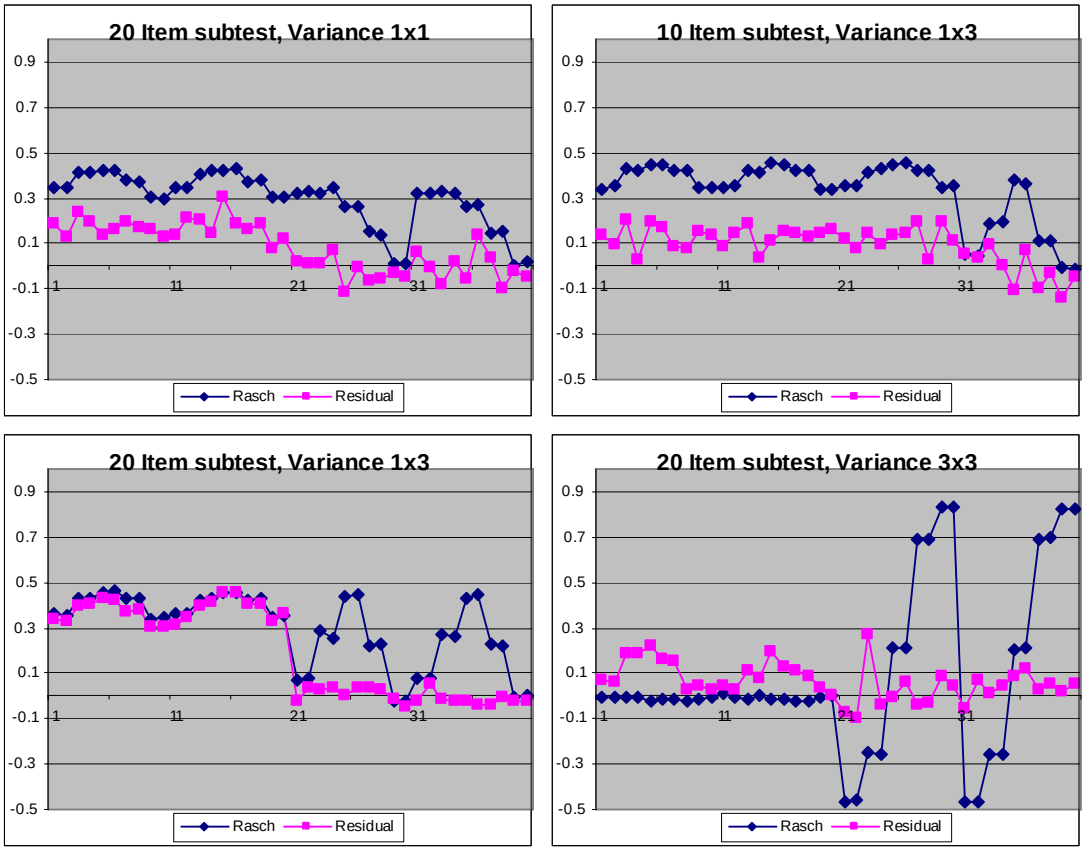
The first two graphs are similar to the 100% contamination condition. In the 20 item, scaling factor 1x1 the random residual pattern looks like previous graphs, but the Rasch condition now has wave like pattern with patterns increasing as the item increases. This is an increasing item-pattern wave pattern. In the 20 item, scaling factor of 1x3 the random condition looks similar to a somewhat suppressed Rasch condition, while the Rasch condition seemed to have a more drastic increasing pattern similar to the previously discussed graph.

Figure 6-4: Factor 1 residual patterns over .4% for 50% reversed contamination



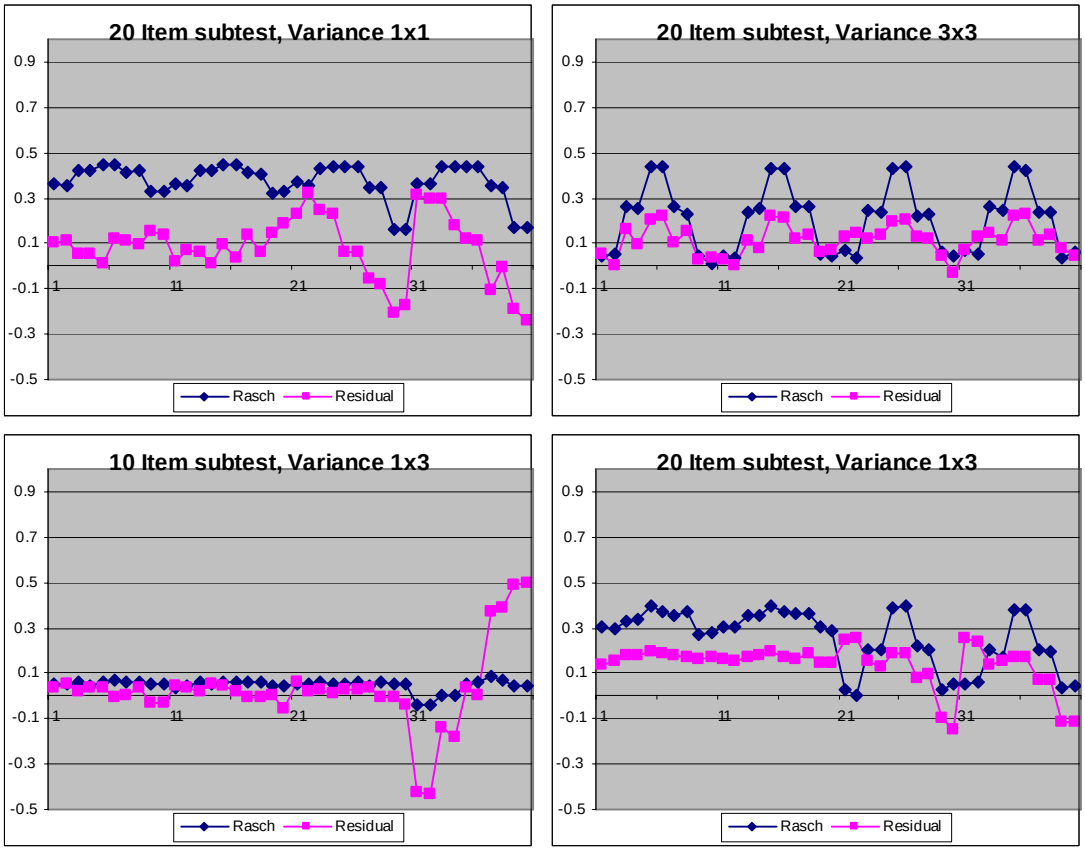
In the 1x1 scaling factor condition the residual patterns look somewhat suppressed with a spike in patterns in the residual subtest. In the 20 item 1x3 graph there appears to be some separation, but in general these graphs do not show much separation between the two weighted datasets.

Figure 6-5: Factor 1 residual patterns over .4% for 20% random contamination



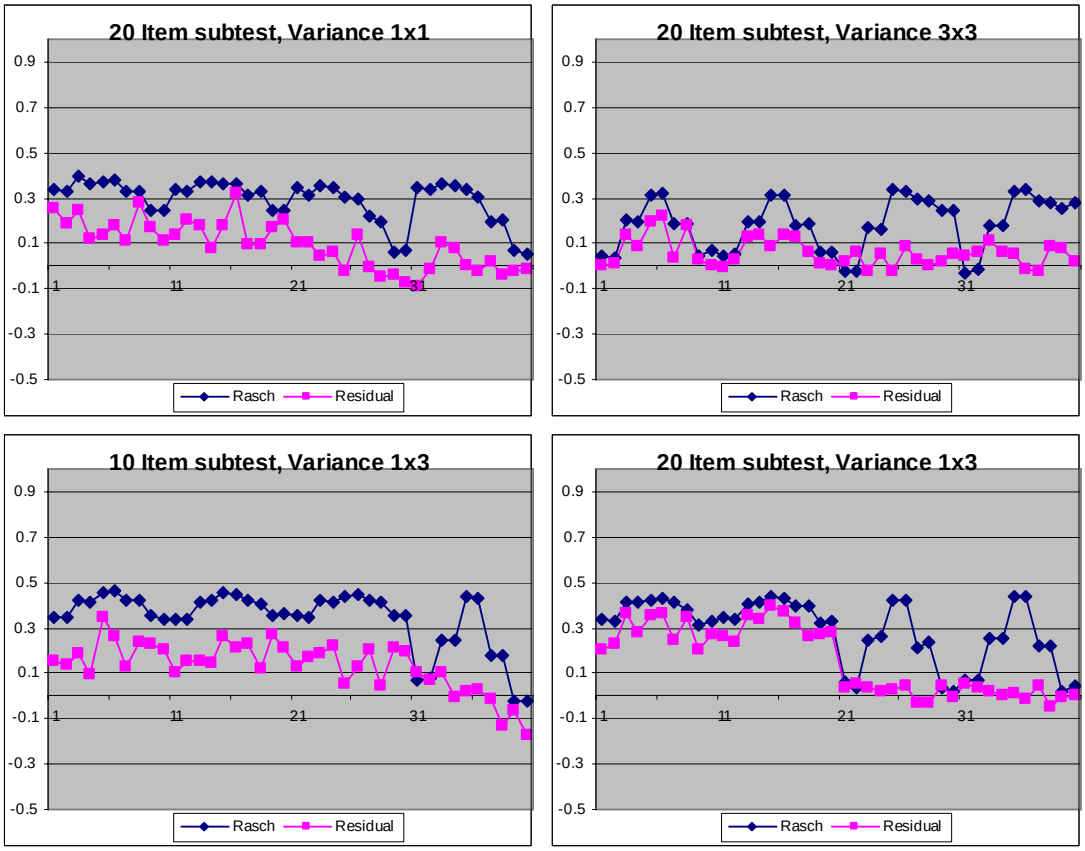
In the first two graphs, the pattern is again suppressed for residually weighted dataset and Rasch like for the Rasch weighted dataset. In the 20 item 1x3 scaling factor, the residual has again a Rasch pattern for the Rasch generated subtest and a random pattern for the random generated pattern. This time, the Rasch weighted dataset has a very distinct pattern found in the baseline condition. In the final graph, the random residually weighted dataset looks random for most patterns, while the wavelike pattern from the previous 50% condition seems extended and more extreme.

Figure 6-6: Factor 1 residual patterns over .4% for 20% reversed contamination



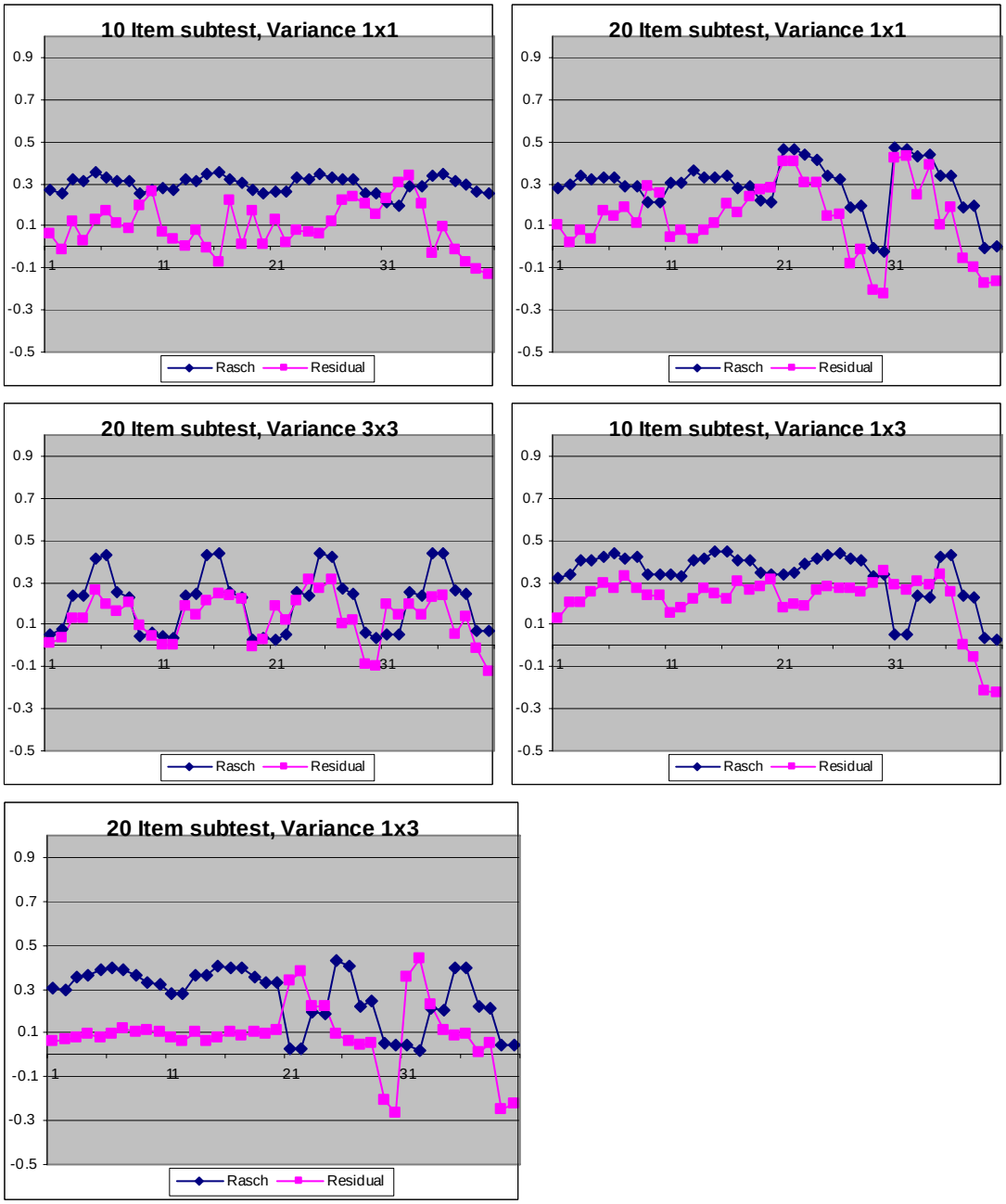
The first graph seems to show a distinct Rasch pattern compared to a reversed version of the Rasch increasing item-pattern wave pattern seen in some of the random conditions. The item-pattern is actually the same direction because the reversed Rasch condition has positive values were negative items should be, and vice versa. In the second graph, a distinct Rasch graph is compared to a suppressed Rasch graph. In the third graph, the Rasch data is suppressed, and the reversed Rasch data has the item-pattern. In the fourth graph, the Rasch data is again distinct and the reversed condition has a suppressed and off centered Rasch-like pattern.

Figure 6-7: Factor 1 residual patterns over .4% for 5% random contamination



In all conditions, the Rasch pattern is distinct. There is less suppression overall in the residually weighted data, but the suppression is still operative. In the final graph for the residually weighted data, there is a Rasch pattern for Rasch subtest and a random pattern for the contaminated subtest.

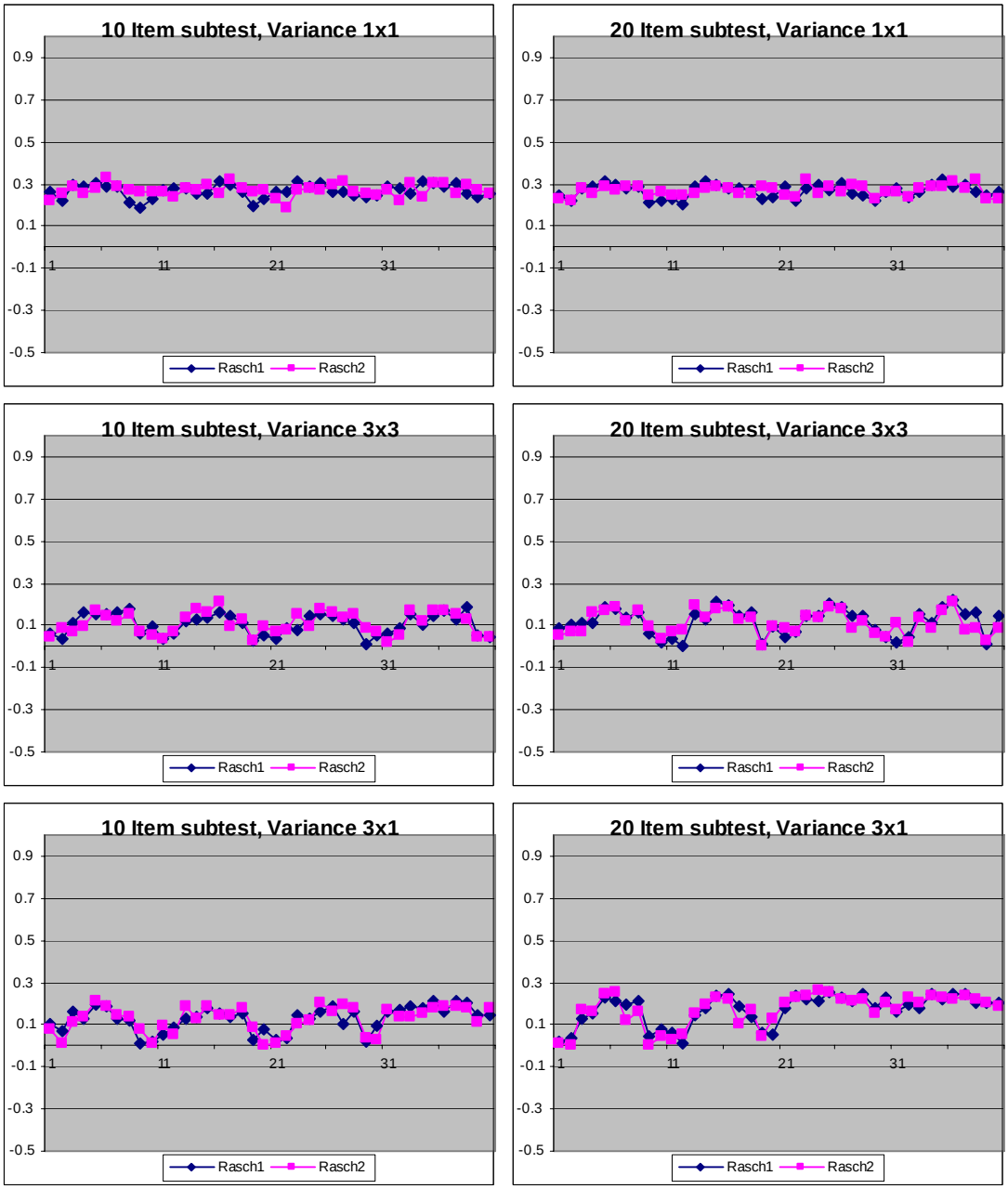
Figure 6-8: Factor 1 residual patterns over .4% for 5% reversed contamination

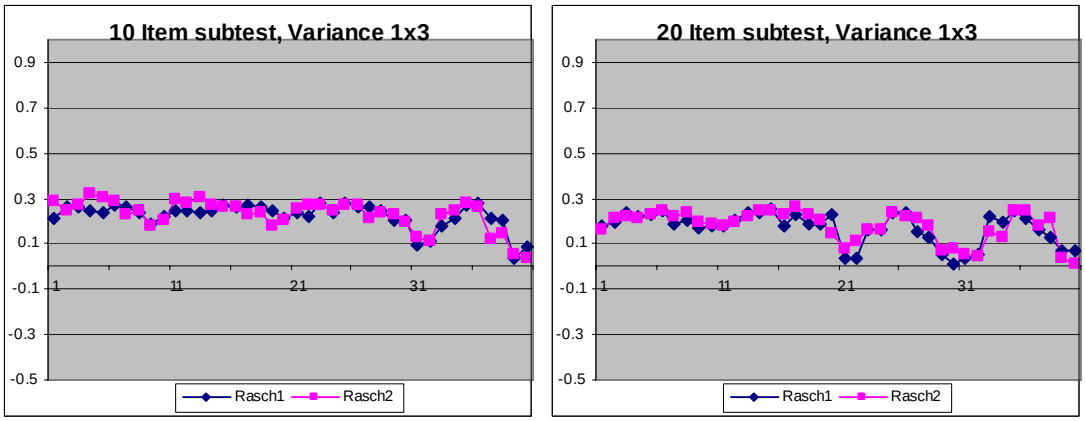


All of the Rasch weighted datasets pattern in the 5% reversed Rasch graph look like the appropriate Rasch baseline.

In the 20 item, 1x3 scaling factor condition in the Rasch generated subtest is suppressed, but the item-pattern is manifest in the contaminated subtest.

Figure 6-9: Rasch patterns as baseline factor 2





A suppressed version of Rasch patterns is apparent in the second factors.

Graph 6-10: Factor 2 residual patterns over .4% for 100% random contamination

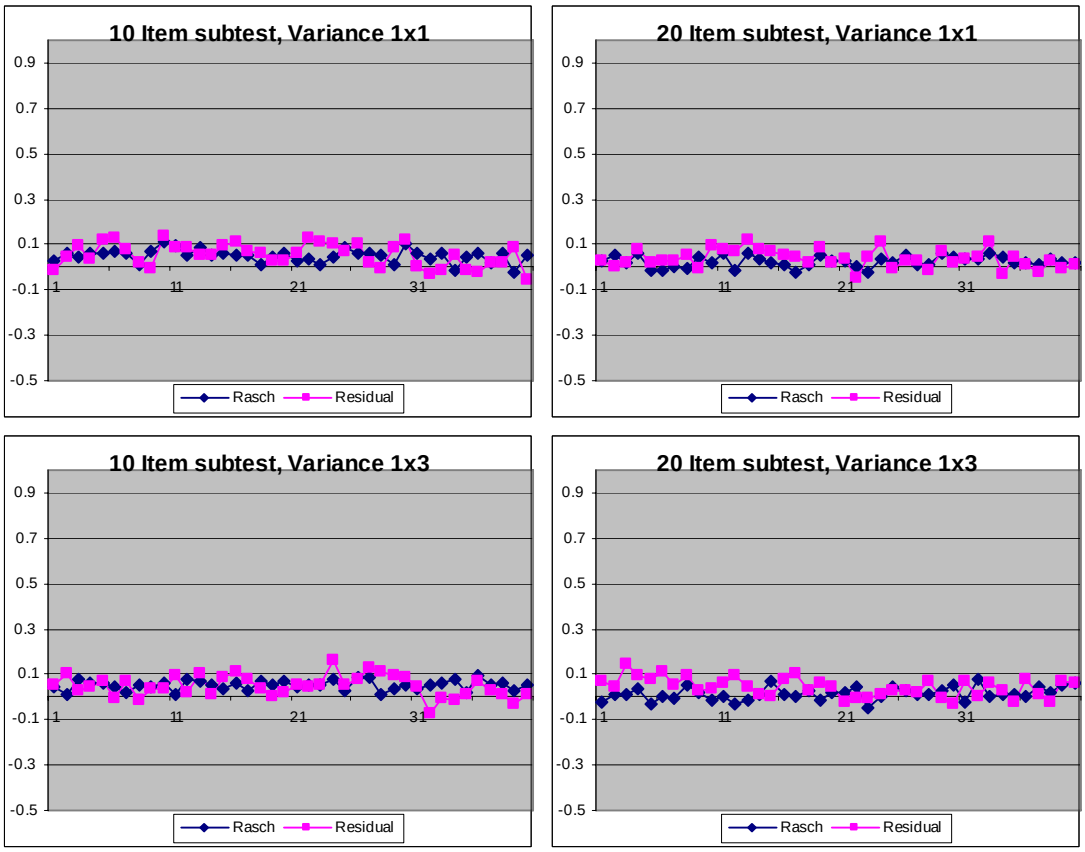
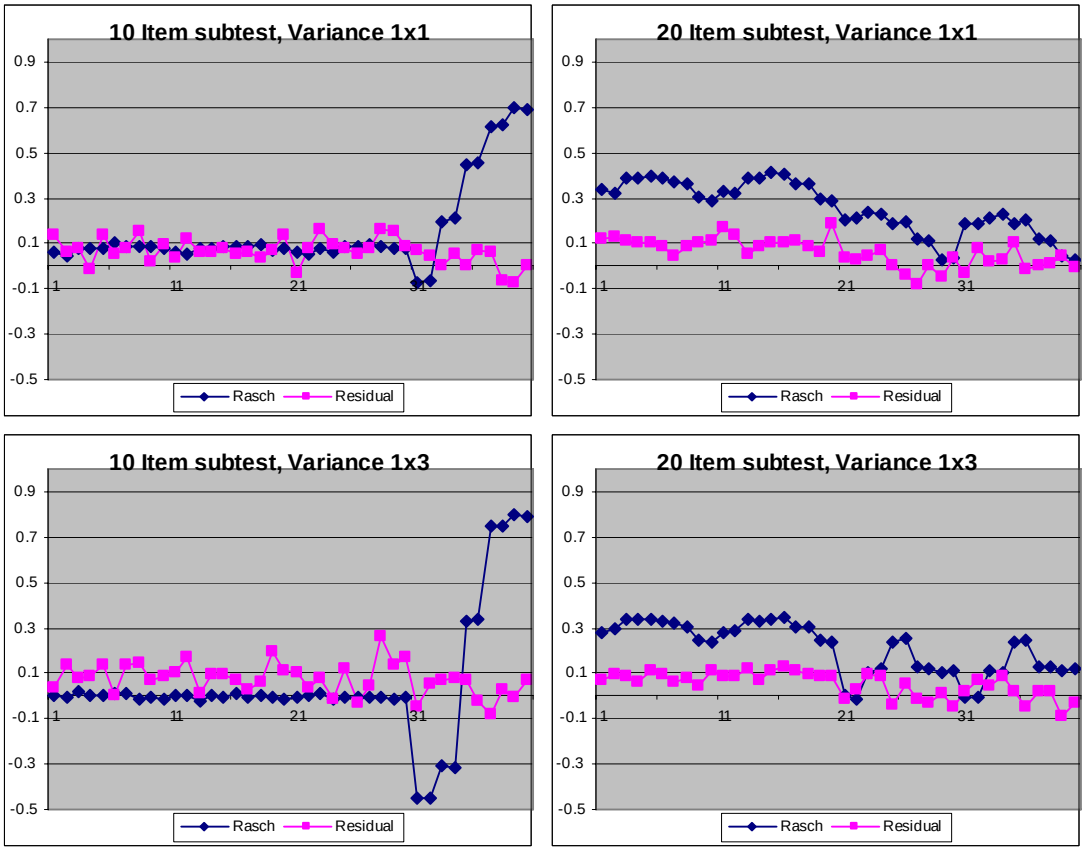
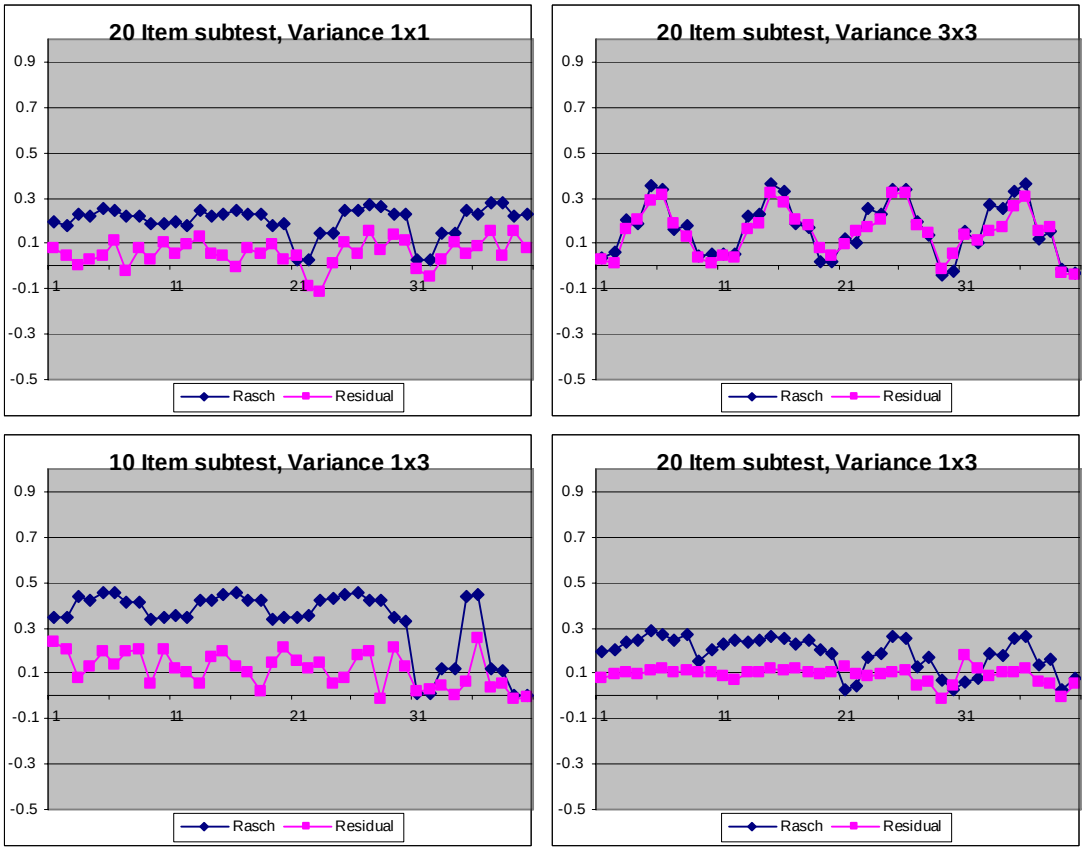


Figure 6-11: Factor 2 residual patterns over .4% for 50% random contamination



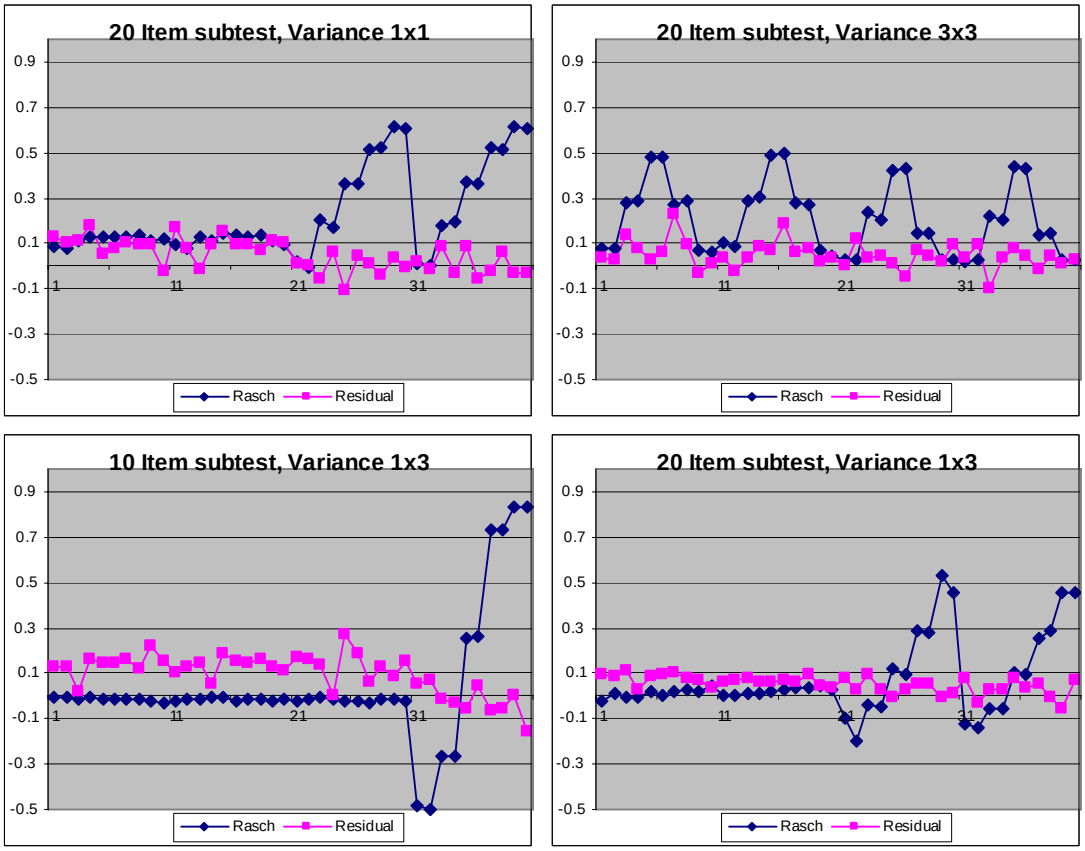
In all cases, the residually weighted data have a suppressed pattern. The secondary factors are now the Rasch like patterns not shown on the first factor. The item-pattern pattern is clear in the 10 item test, and the Rasch pattern is manifest in the 20 item graph.

Figure 6-12: Factor 2 residual patterns over .4% for 50% reversed contamination



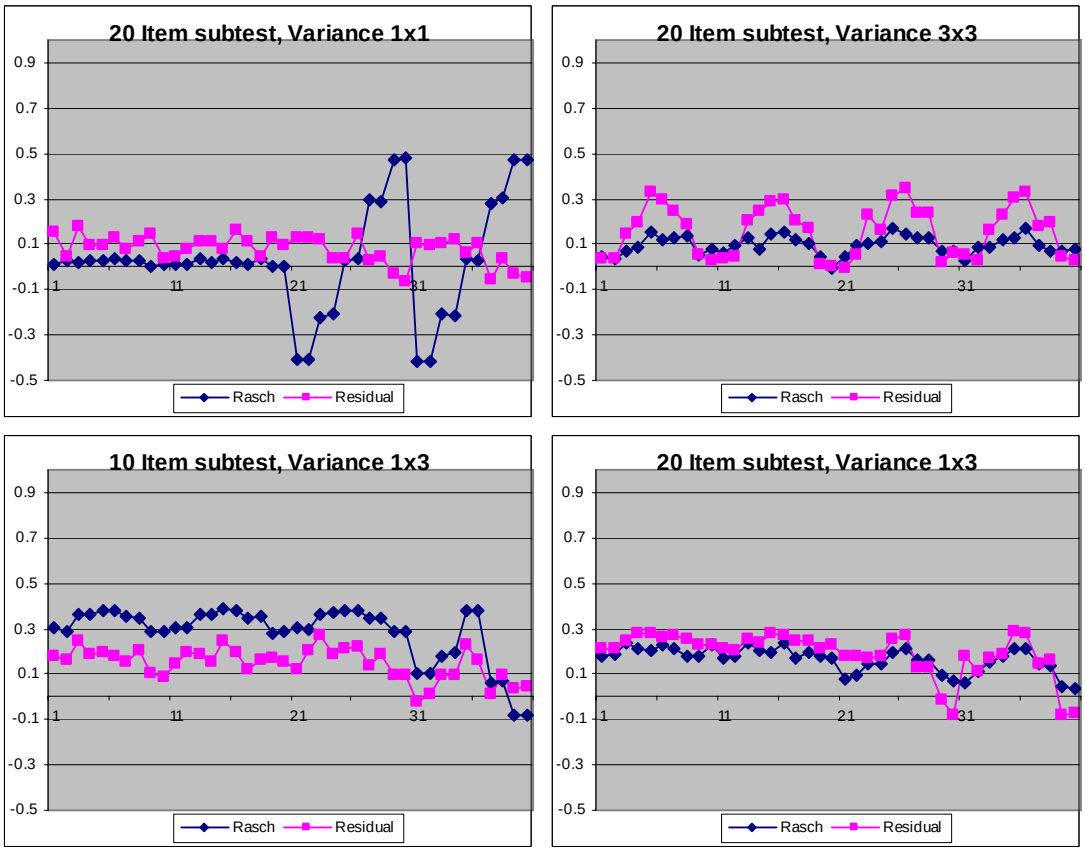
The Rasch weighted data still have Rasch patterns. Though the residual weighted data are suppressed, the 20 item 3x3 condition show a clear Rasch pattern.

Figure 6-13: Factor 2 residual patterns over .4% for 20% random contamination



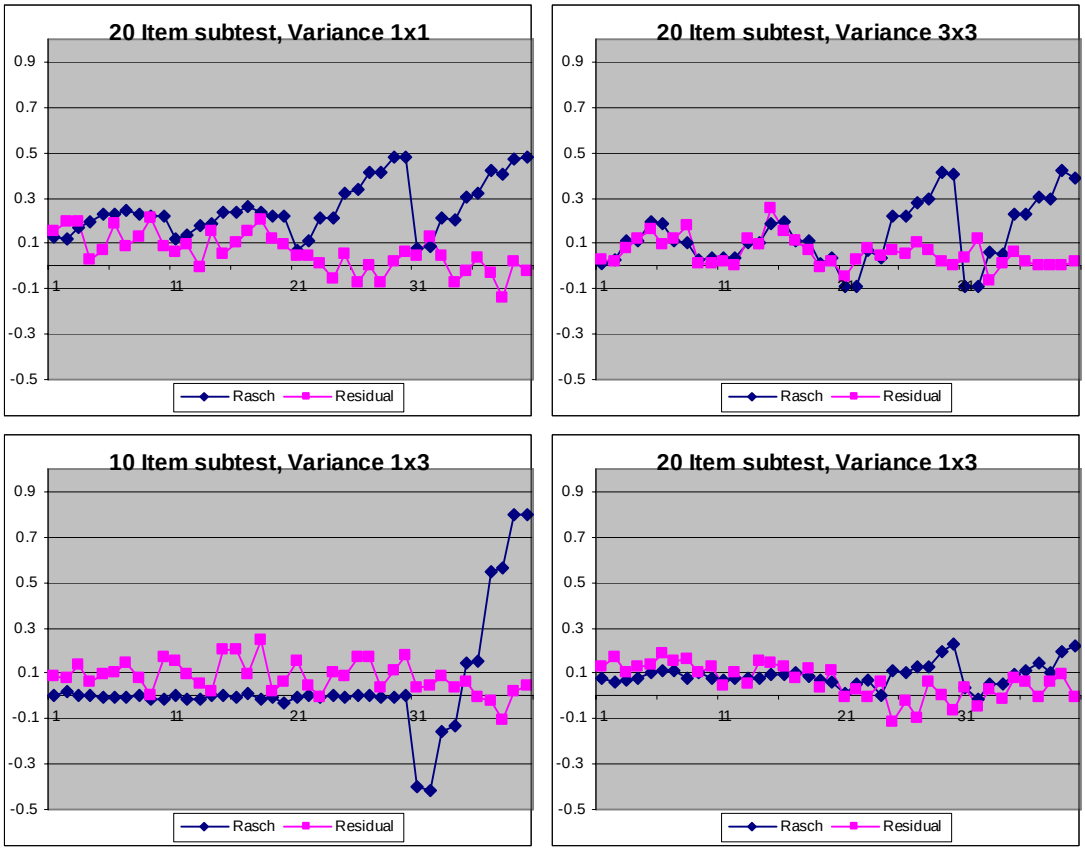
The residually weighted data is completely suppressed. Again, the Rasch weighted data shows the others pattern not manifest in the first factor.

Figure 6-14: Factor 2 residual patterns over .4% for 20% reversed contamination



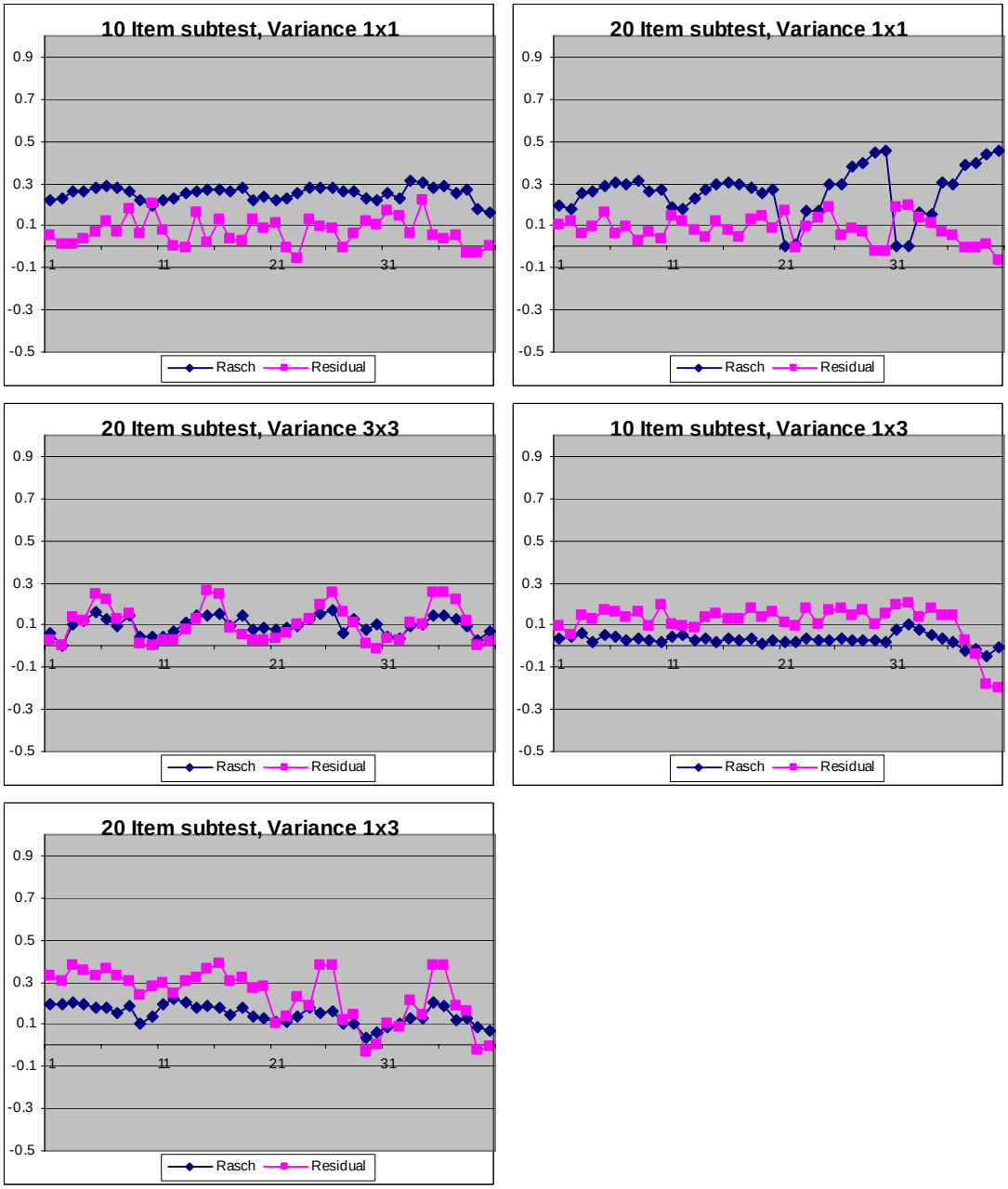
The Rasch weighted data in the 10 item conditions still has Rasch patterns. However, they are suppressed in the 20 item conditions. In the 20 item conditions, the reversed Rasch has Rasch-like patterns but is suppressed in the 10 item conditions.

Figure 6-15: Factor 2 residual patterns over .4% for 5% random contamination



The residual weighted data is suppressed for most patterns across all 5% conditions. There are some clear Rasch item-patterns in the Rasch weighted data. The stair-like and bulging structures are reminiscent of some previous Rasch baseline and some first factor patterns that were weighted to be Rasch.

Figure 6-16: Factor 2 residual patterns over .4% for 5% reversed contamination



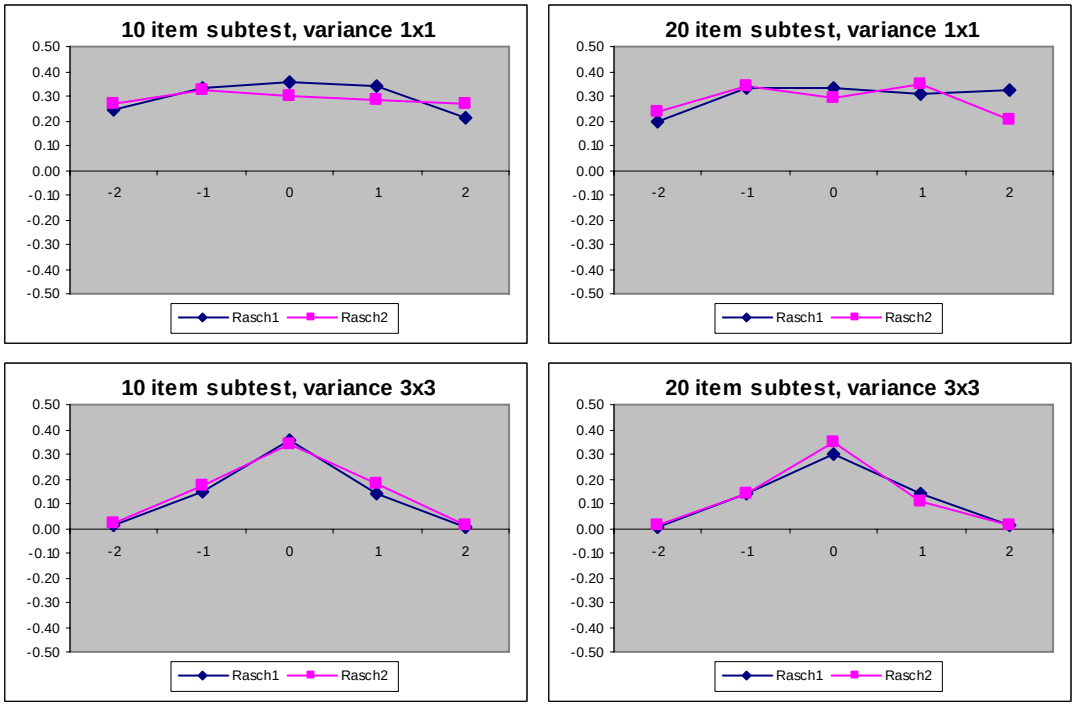
As seen above, there are still some Rasch patterns in the first two graphs for the Rasch weighted data, but very little for the last three. Though mostly suppressed

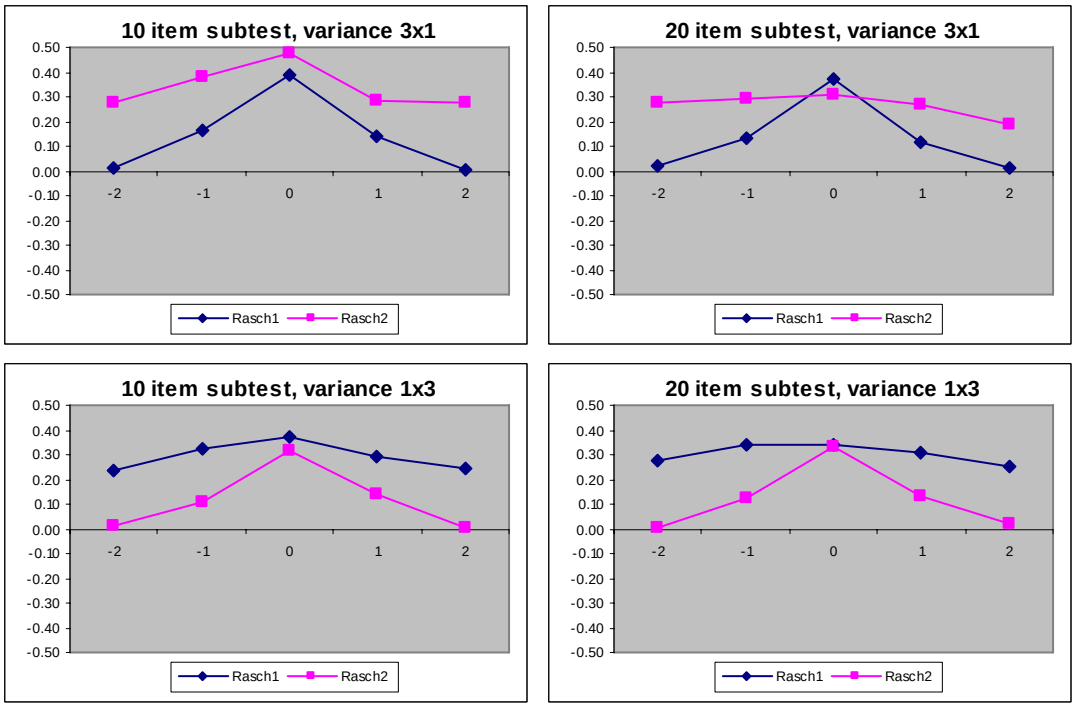
in the graphs, the residually weighted data has some Rasch type patterns visible in the last graph.

Residual patterns

This section shows graphs of the residual patterns aggregated by difficulties within subtest. The values from both subtests are shown of the same scale access from -2 to +2. The comparisons are made within the residual conditions explored in the previous sections. The visual representation shows patterns on the Rasch subtest compared to the contaminated subtest. The baseline Rasch patterns are generated using a random 2% condition. Conditions with greater percentages conform to the same pattern.

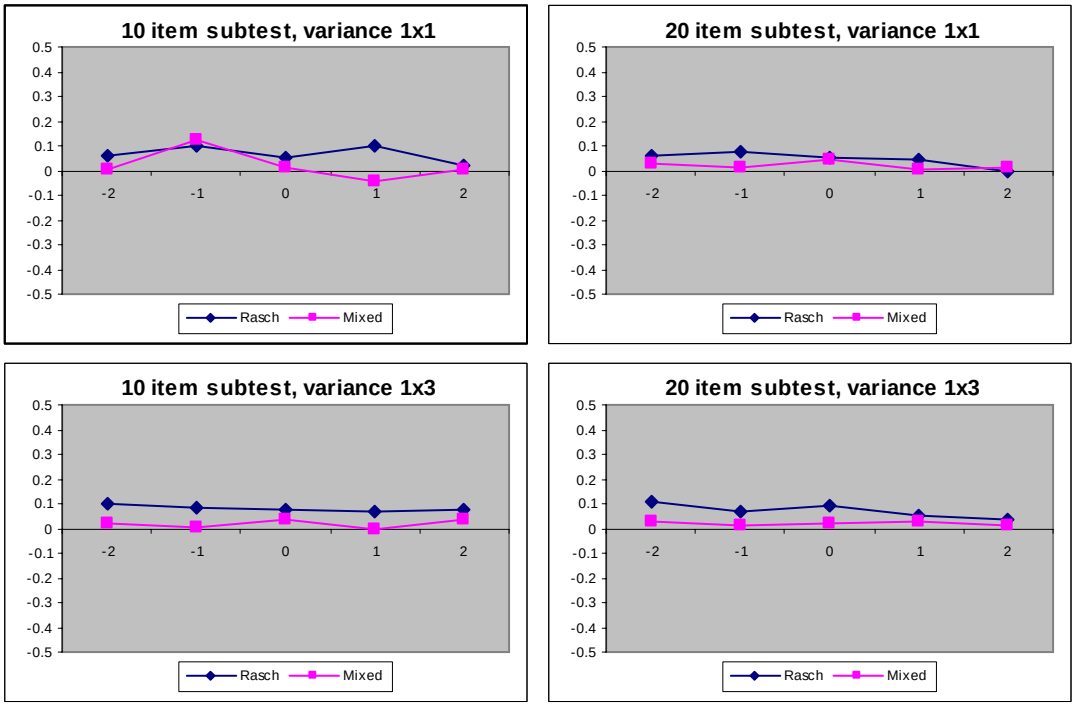
Figure 6-17: Rasch patterns as baseline factor 1





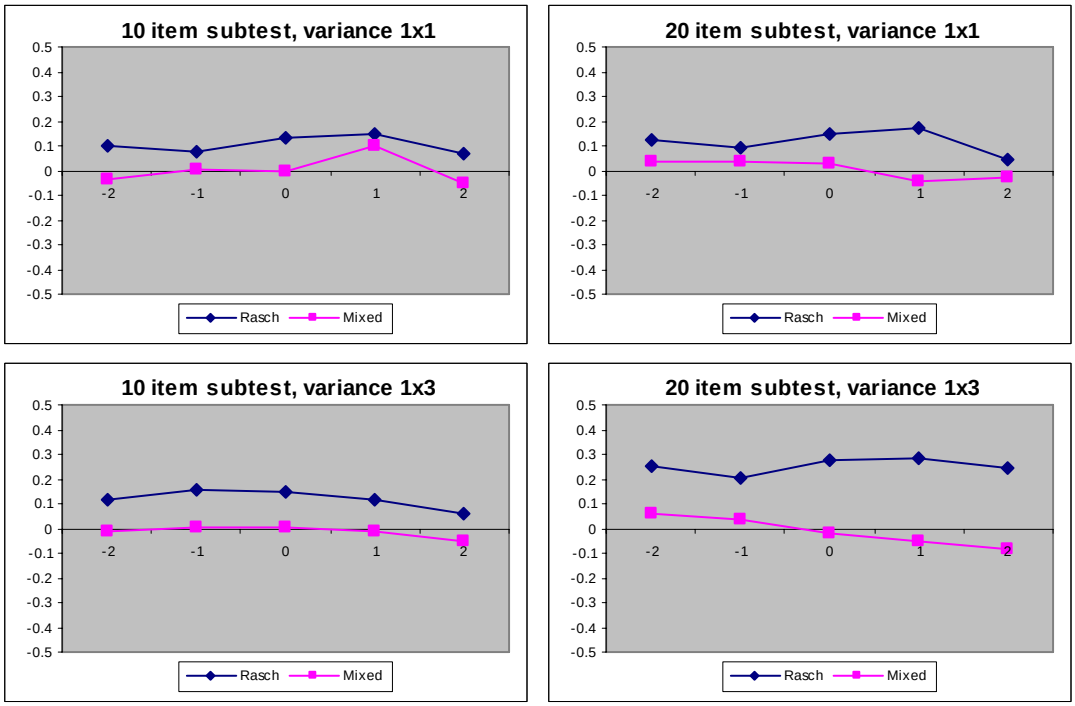
It was expected from hypothesis 10 that the subtest in the Rasch baseline would look like the second subtest. In the first four graphs this is true. In those remaining, four different patterns exist for the subtest. However, the differences are simply due to the change in scaling factor. The Rasch pattern is the same across all conditions from the same scaling factors. The scaling factor of one has a relatively flat set of pattern around .3, whereas the scaling factor of three has a mountain or wave like pattern with the extreme values going to zero and the items around zero have patterns around .3. These two stable patterns can be used as a baseline.

Figure 6-18: Factor 1 residual patterns over .4% for 100% random contamination



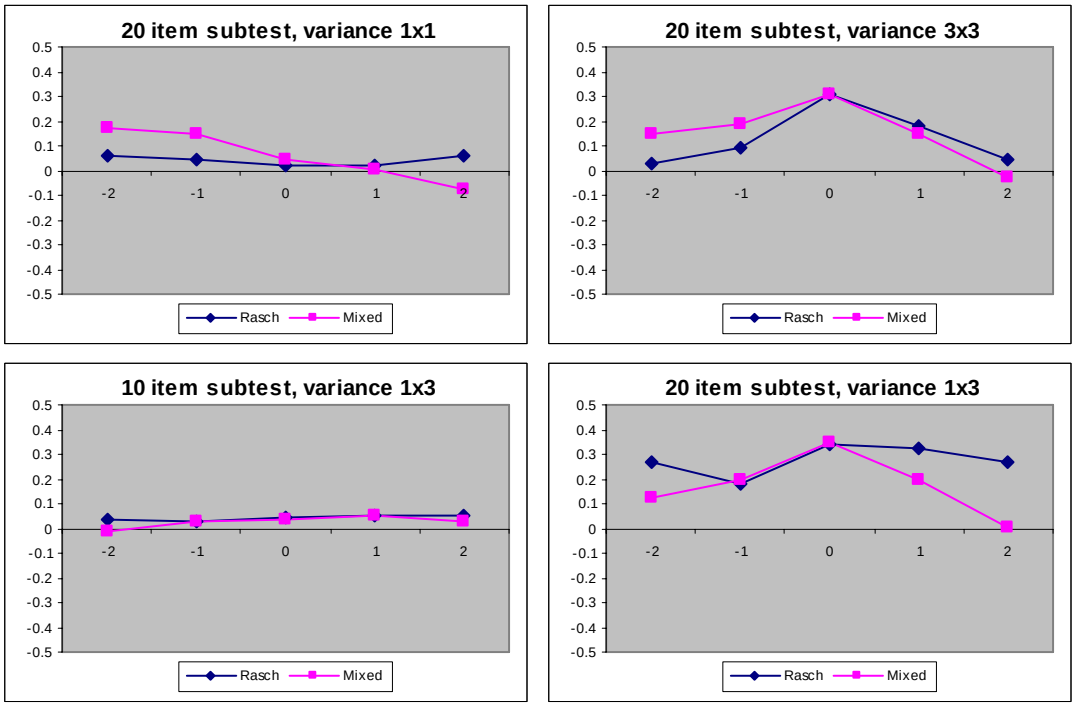
These patterns are small and have no distinct features.

Figure 6-19: Factor 1 residual patterns over .4% for 50% random contamination



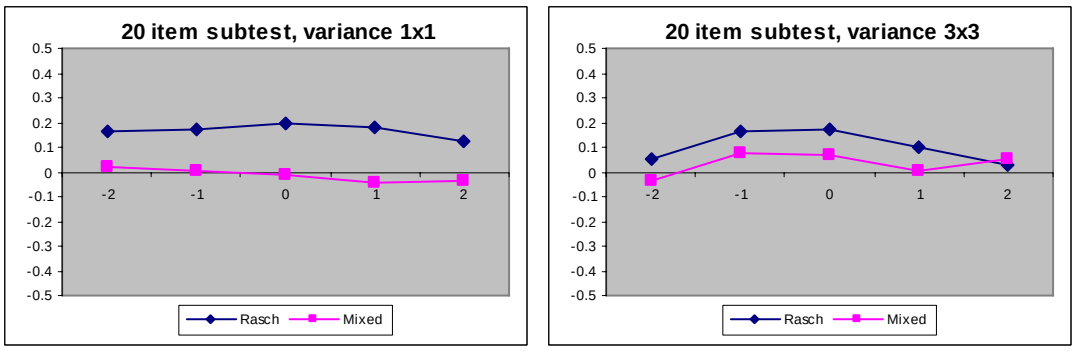
These graphs start to separate the subtests with the Rasch subtest being distinctly larger than the random contamination subtest.

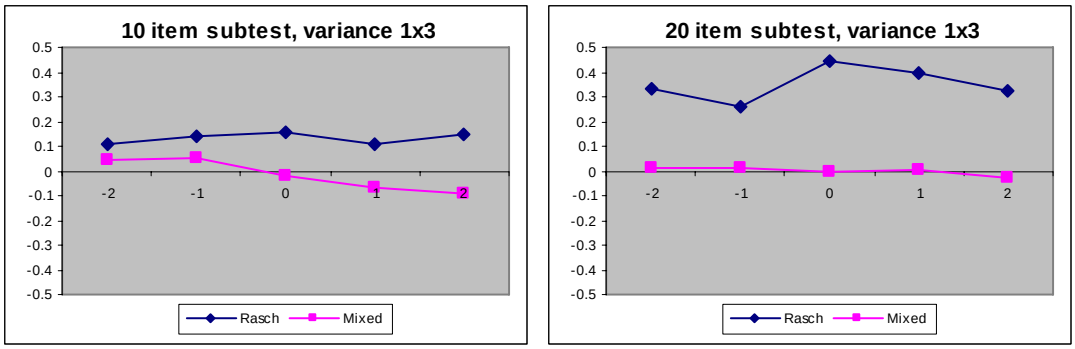
Figure 6-20: Factor 1 residual patterns over .4% for 50% reversed contamination



The two twenty-item subtest conditions show patterns with distinct features of some of the baseline Rasch condition. There may be some crossing patterns in the graphs but not anything with clear evidence.

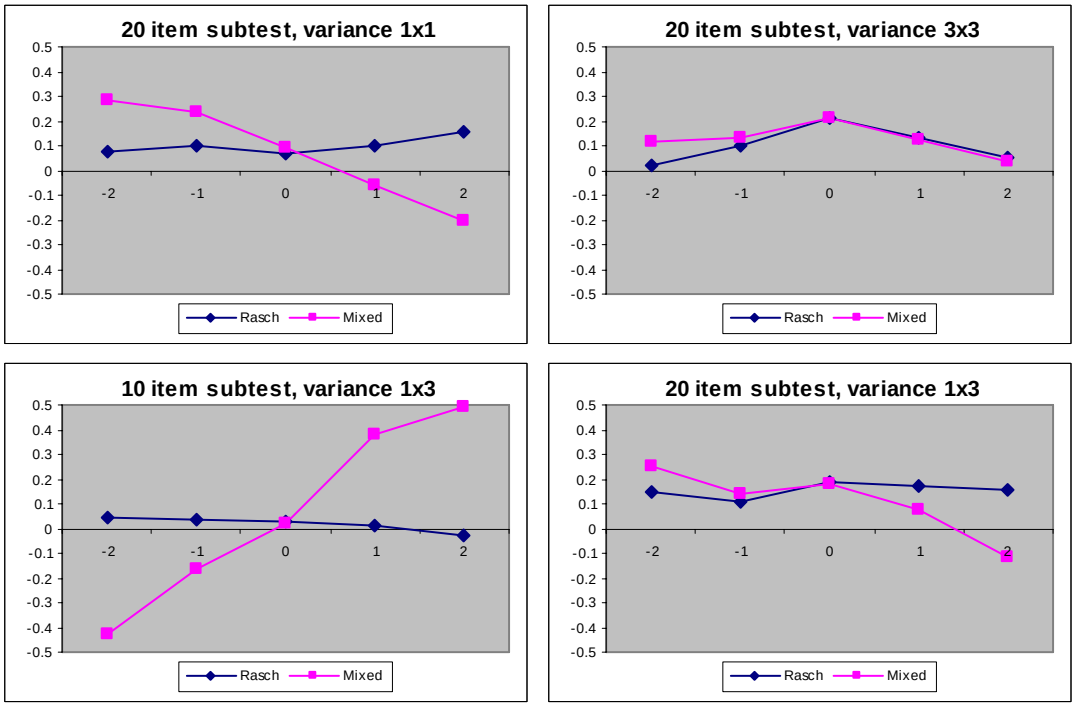
Figure 6-21: Factor 1 residual patterns over .4% for 20% random contamination





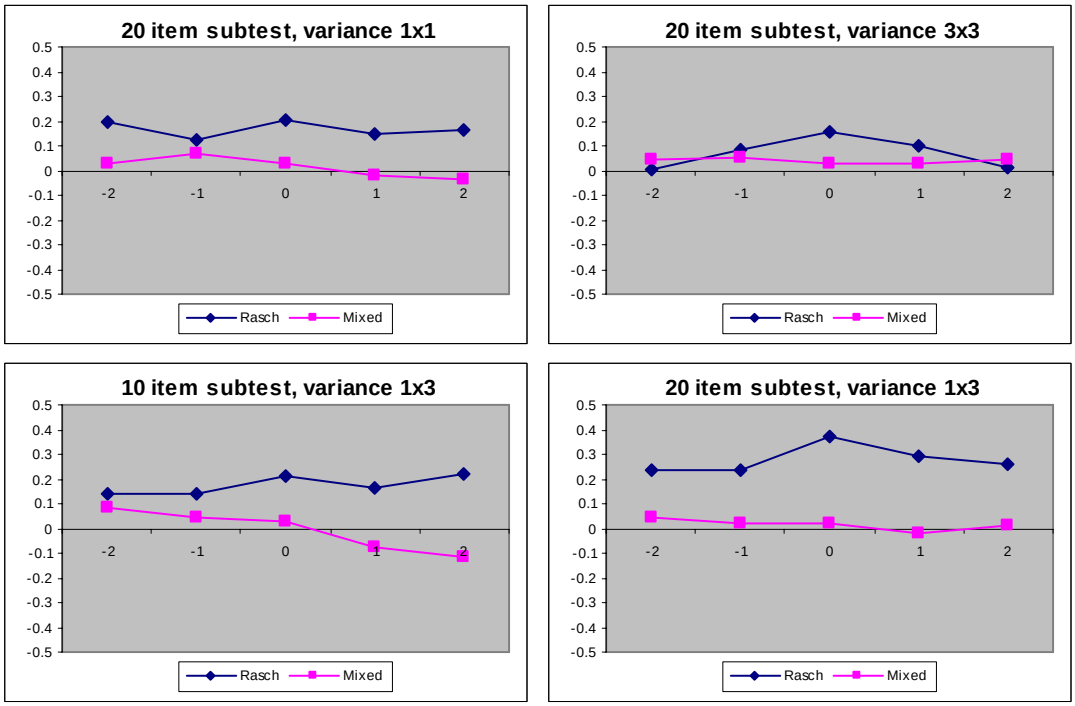
The 20% random contamination condition further displays the distinct pattern of greater values for the Rasch subtest portions of the test.

Figure 6-22: Factor 1 residual patterns over .4% for 20% reversed contamination



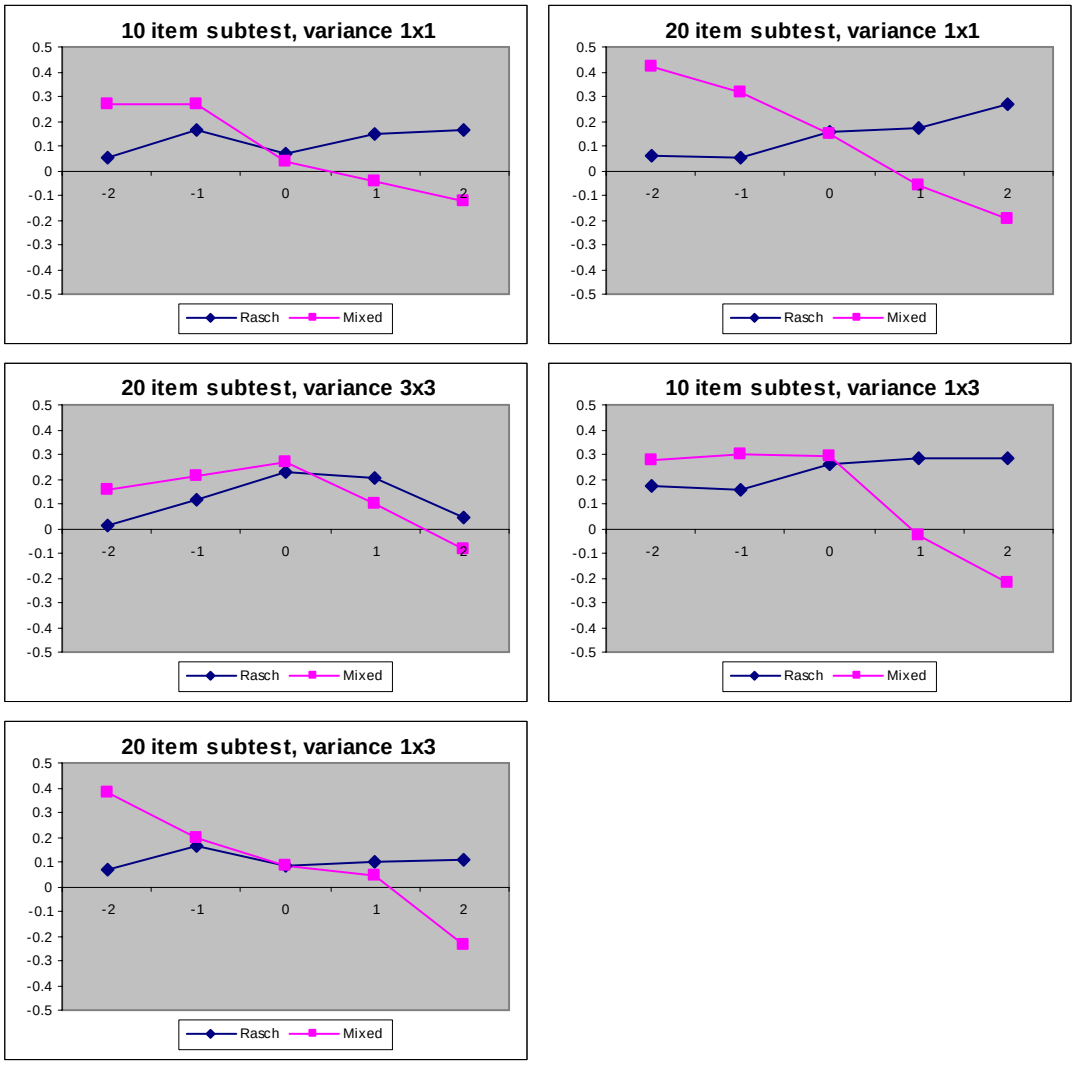
A clear crossing pattern is visible in the 20 item equal scaling factor of 1 and 10 item scaling factor of 1 for Rasch and 3 for contaminated. The 20 item, equal scaling factor of 3 condition displays a Rasch pattern. The 20 item scaling factor of 1 for Rasch and 3 for contaminated seems to be a mixture of the crossing pattern and Rasch.

Figure 6-23: Factor 1 residual patterns over .4% for 5% random contamination



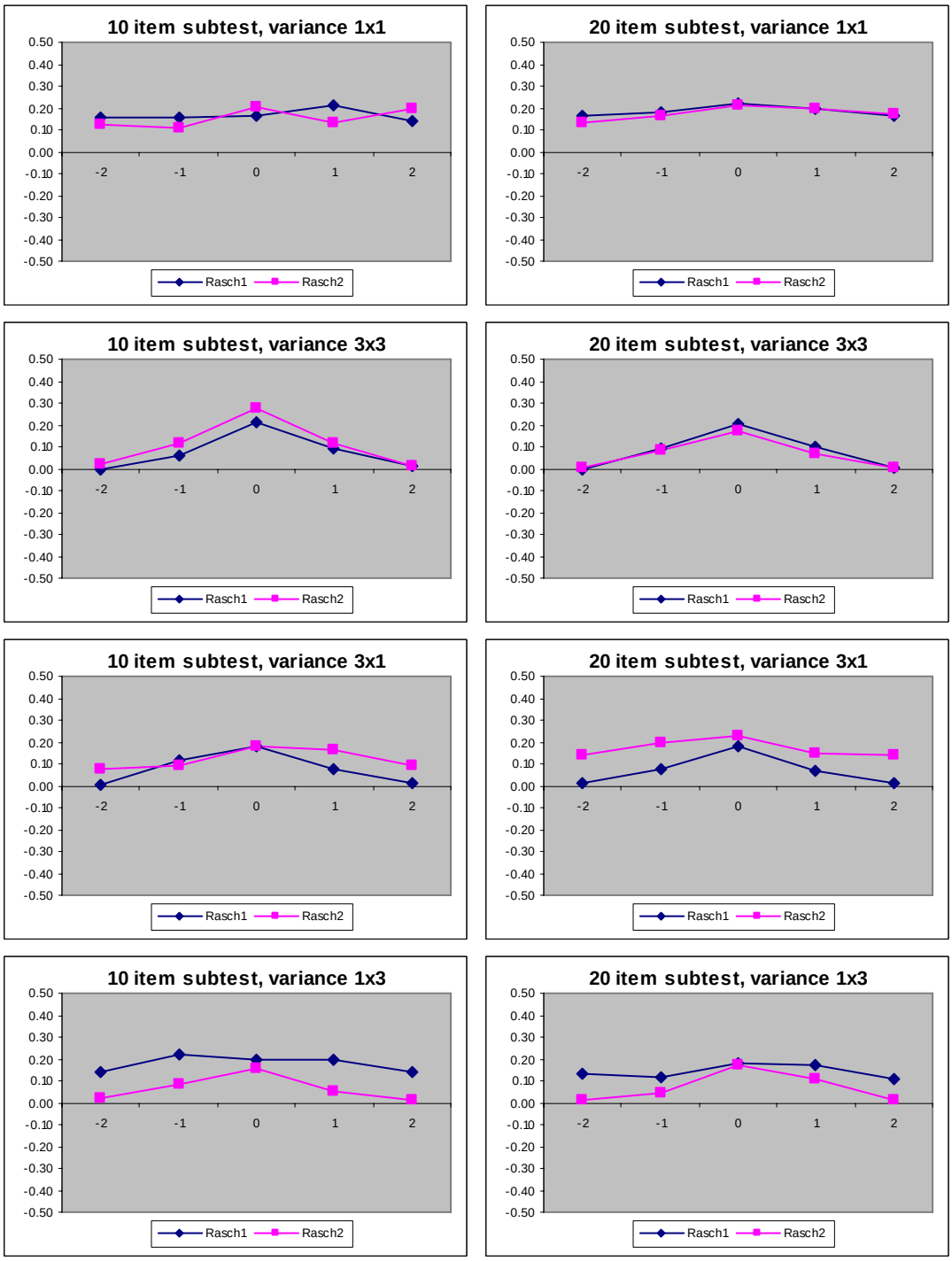
The same pattern is still apparent in the 5% random condition with some more Rasch visible in some of the graphs for the Rasch generated subtests.

Figure 6-24: Factor 1 residual patterns over .4% for 5% reversed contamination



The crossing pattern is manifest in all of the graphs for the 5% reversed contamination condition and is indicative of the expected type of pattern hypothesized in the data.

Figure 6-25: Rasch patterns as baseline factor 2



Although there are patterns recognizable from the first factors, the sizes of the patterns in the baseline condition are small. Many of the following secondary factors have no pattern or magnitude of pattern to discuss.

Figure 6-26: Factor 2 residual patterns over .4% for 100% random contamination

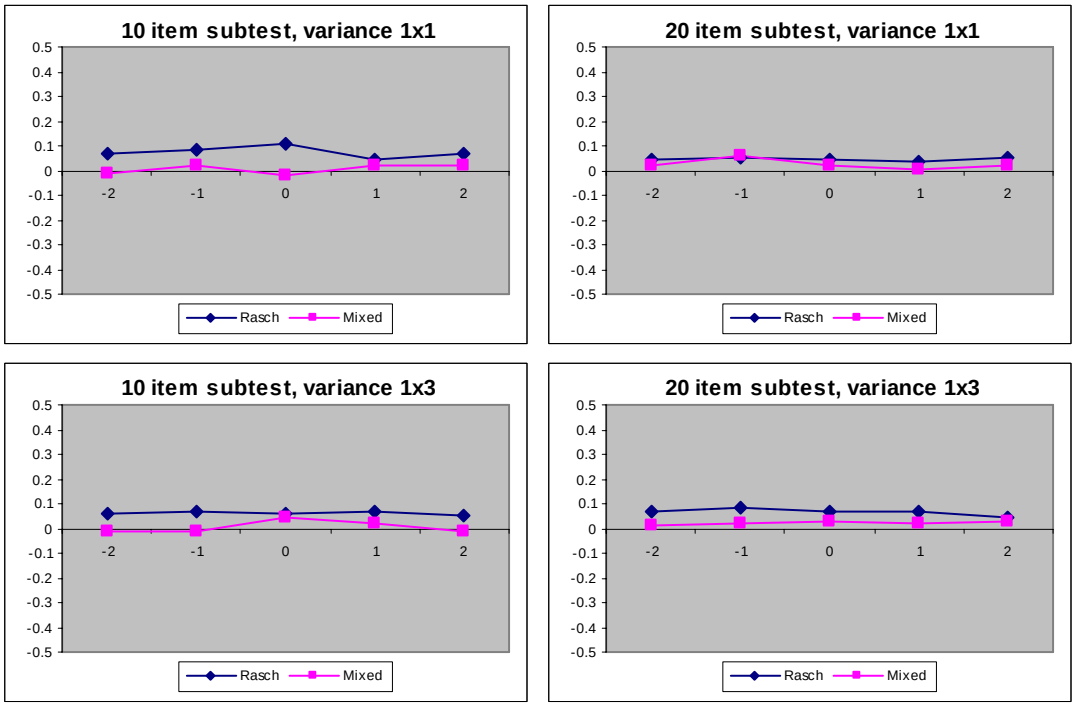


Figure 6-27: Factor 2 residual patterns over .4% for 50% random contamination

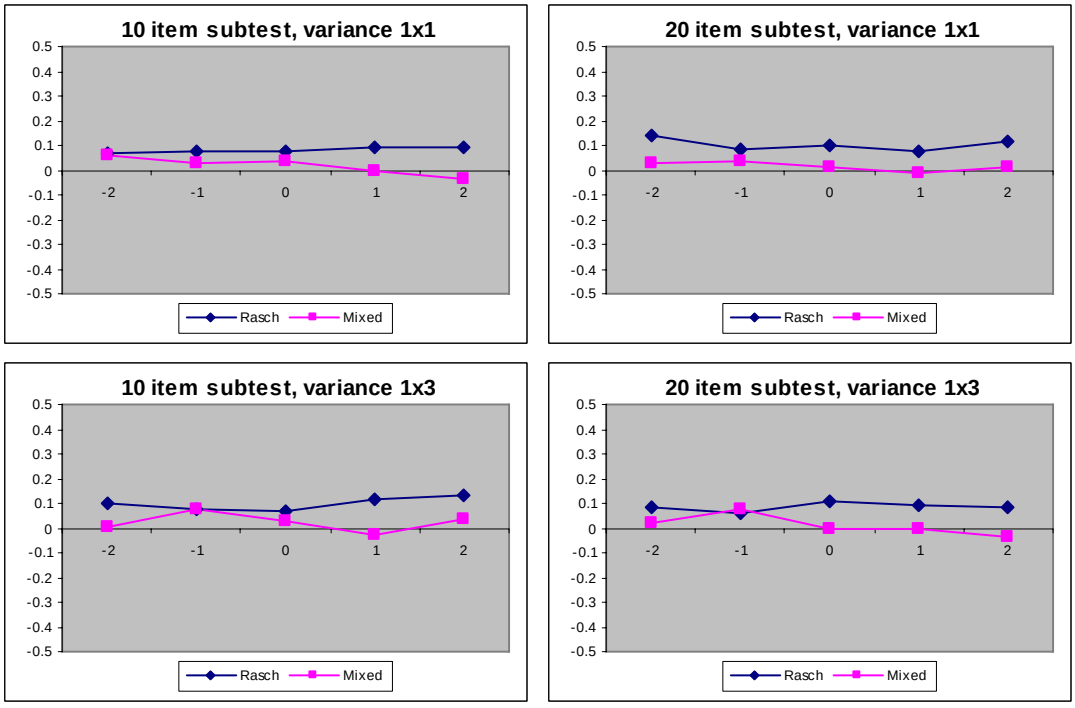
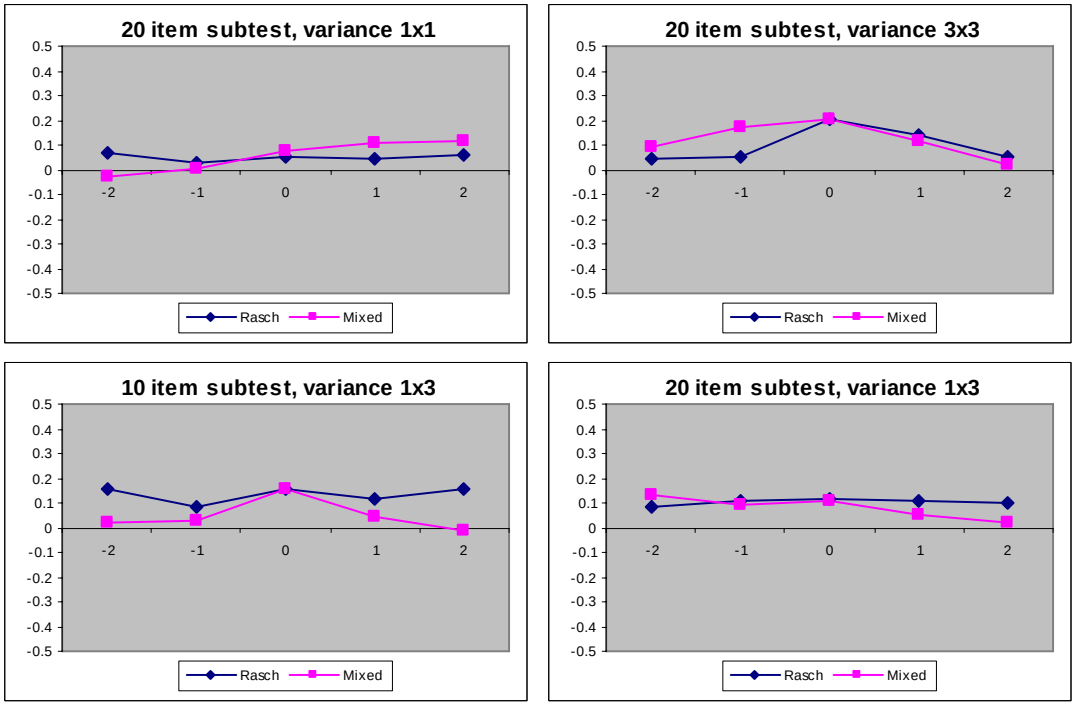
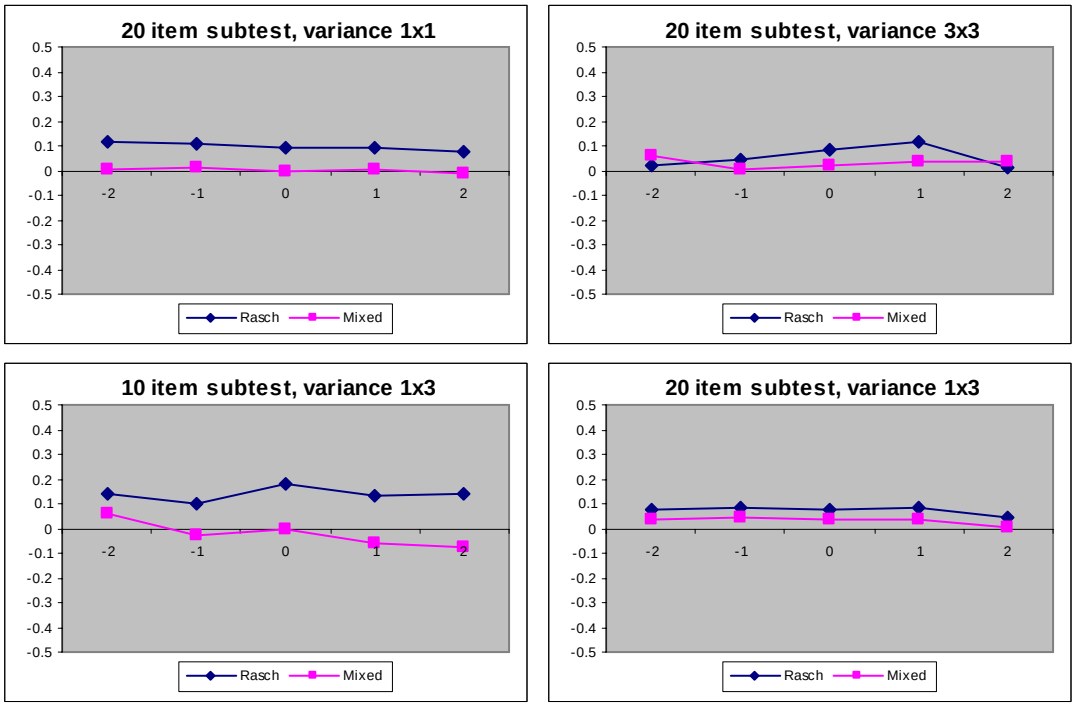


Figure 6-28: Factor 2 residual patterns over .4% for 50% reversed contamination



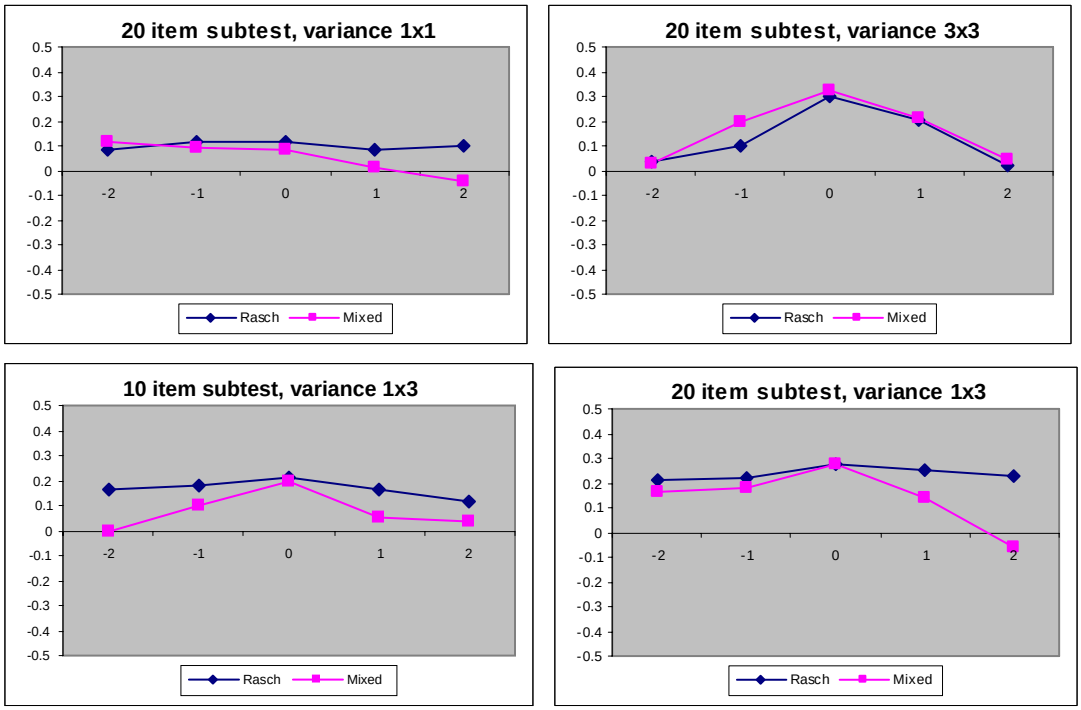
The 20 item scaling factor of 3 condition, the 10 item Rasch scaling factor 1, and contaminated scaling factor 3 show some residual pattern similar to the baseline conditions, but the magnitude of the patterns are very small.

Figure 6-29: Factor 2 residual patterns over .4% for 20% random contamination



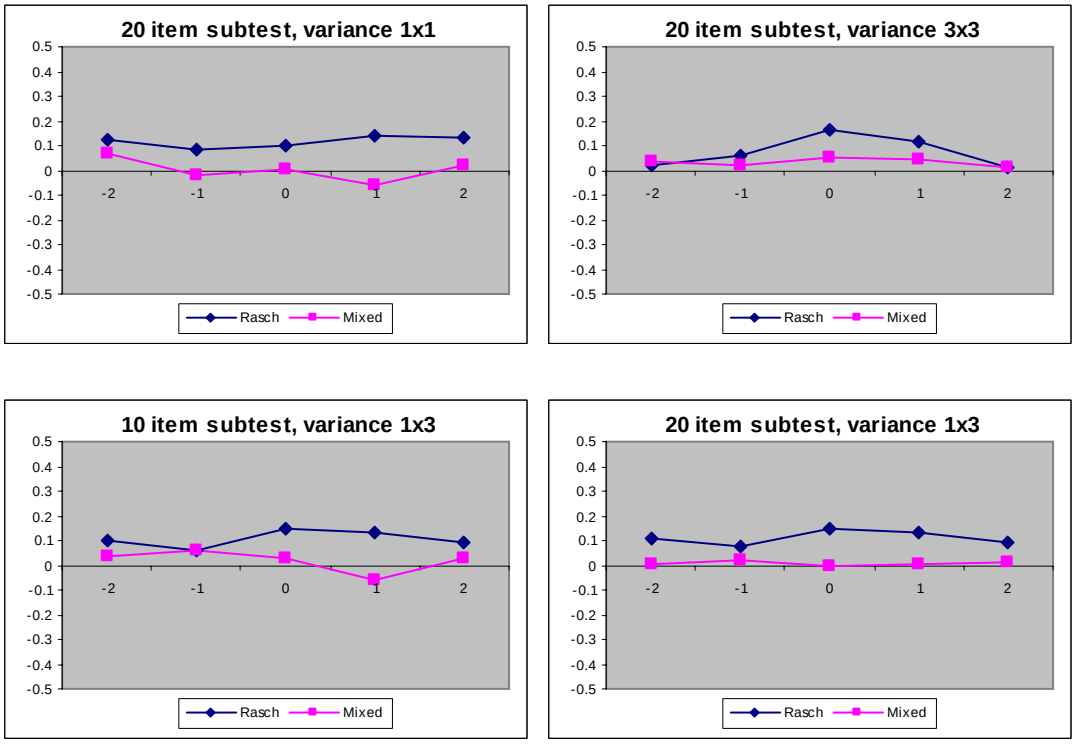
Some of the separation from the first factor is possibly present but the magnitudes are too small to interpret.

Figure 6-30: Factor 2 residual patterns over .4% for 20% reversed contamination



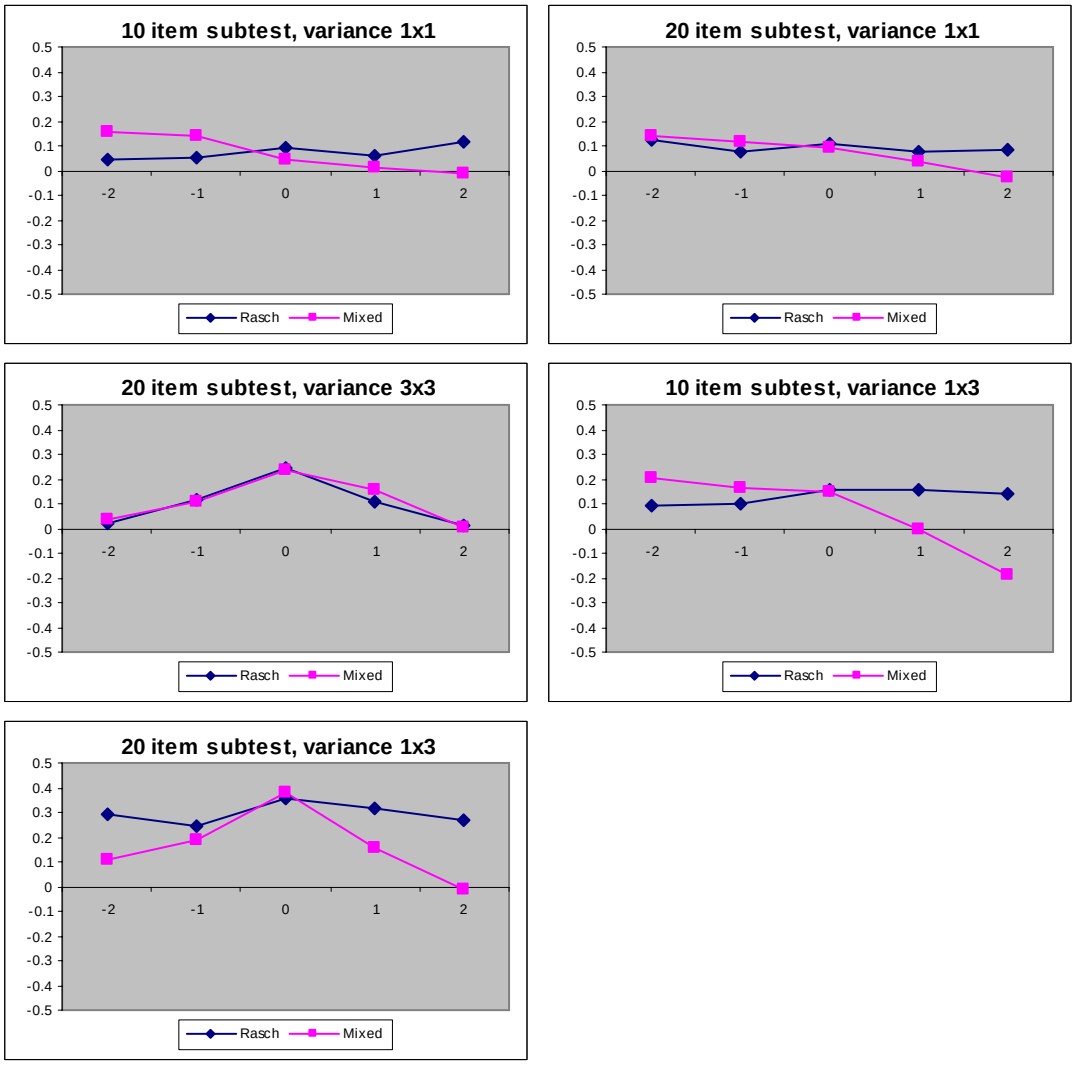
There are some clear patterns with some pattern values approaching and exceeding .3. The 20 item equal scaling factor of 3 patterns look very similar to the Rasch first factor as do the rest of the patterns in this section.

Figure 6-31: Factor 2 residual patterns over .4% for 5% random contamination



There is again some potential separation here but nothing with enough difference attribute any real findings.

Figure 6-32: Factor 2 residual patterns over .4% for 5% reversed contamination



The Patterns with a scaling factor of 3 in the contaminated subtest are reminiscent of the Rasch patterns in the first factor.

Summary of residuals and eigenvalues

This study explored the unscalable class as a residual to the Rasch class, in which an unscalable class proportion of reasonable size was discovered. A meaningful size of the unscalable class was established as .4% classification. In exploring the residuals, weighted factor analyses were conducted and the patterns of items on factors were explored visually and supported extensively through the use of MANOVA and CI. Hypotheses regarding effects support visual representation and MANOVA and CI are tools to help explain meaningful and unique patterns within a condition. The objective of the significance testing and graphing was to show a more thorough picture of what was left over in residuals for each of the 29 conditions in which the residuals were large enough to inspect.

Patterns of effects were observed for the manipulated factors in the study of subtest size, scaling factor of the subtests, % contamination and type of contamination.

The size of the residual increased on average from the random condition μ=2.35 σ=4.67 to the reversed μ=5.19 σ=12.12condition. As percent contamination condition decreases, the residual increases 100% μ= .01 σ=.02, 50% μ= .15 σ=.32, 20% μ= .24 σ=.39, 5% μ= .36 σ=.42. The size of the subtest increases from μ= .77% σ= 1.55 in the 10 item condition to μ= 6.77 σ= 11.14 in the 20 item subtest condition. An increase in the range of the subtest increased from μ= 0.72 σ=1.40 when the range was -2 to m+2 to μ= 6.83 σ=12.31 for the -6 to +6 condition. In general, the residual size is largest proportionately in the 20 subtest with a larger range. The one exception is the reverse Rasch 100% contamination condition, which is effectively a Rasch

baseline condition. The pattern of residuals is small and represents less than one simulee.

In investigating the number of factors by evaluating the size of the eigenvalues for the unweighted data, there is predominantly one main factor in the baseline conditions and the 100% contamination conditions with some minor secondary factors being found in the Rasch and Reversed Rasch condition. The reversed Rash is nearly indistinguishable from the Rasch baseline condition. The random noise added to the Rasch data in the 100% random contamination condition was systematic for all simulees and did not add a large enough factor to the data. In the unweighted data for the reversed Rasch contamination conditions for 50%, 20% and 5% and for the random contamination of 50% and 20%, two factors were present. In the 5% random contamination condition, only a second factor existed for the conditions with the increased subtest range by a factor of 3.

When the data is weighted to the Rasch class and compared to the unweighted data, a change in the number of factors is apparent in the data for the 20% contamination condition for both Random and Reversed Rasch, 20 items subtest, and a range of ±6. In the 5% random and reversed Rasch contamination condition, the change occurs in the same subtest conditions as well as in the 20 item subtest with equal smaller ranges. Across the conditions, the secondary factor decreases when the unscalable class is reduced by weighing the data to the Rasch class. In some conditions, such as a larger subtest, smaller percent contaminated and large range for a subtest these factors seem to be removed more efficiently.

In the residually unscalable class, weighted data for the 50%, 20% and 5% conditions across both random and reversed Rasch contamination, the 20 item subtest with a larger range than the Rasch subtest has eigenvalues larger than HPA. In the reversed Rasch 50%, 20% and 5% conditions, eigenvalues are larger than the HPA values for the 20 item subtest with equal but large ranges and for the 20% and 5% 10 item subtests with a larger range than the Rasch only subtest. When a factor was detected for the systematic contamination six of the eight times there was a secondary factor. Visual graphs support the idea in these tests that one factor is Rasch and the other is a reversed Rasch factor. However aspects of both can be found in the other. A factor is only found clearly twice in the weighted data for random contamination and looks like a Rasch factor, with some suppression.

The Wilks' lambda F-test in conjunction with CI supports the overall evaluation of the 29 conditions large enough to be explored thoroughly. The patterns of importance have been displayed in the result section along with graphical representation of comparison between Rasch and unscalable weighted data, and within the unscalable class by itself. In general, the visual graphics support the statistical evaluation and vice versa. While the factor is significant in the residual condition for the multivariate analysis, a visual separation or crossing of the Rasch and unscalable data is both apparent and strong. When the multivariate analysis is not significant, the patterns are small or overlapping and the graphic much less telling of any visual omnibus separation in patterns. The individual CI tend to support separation of the Rasch and unscalable data item by item. The univariate procedure

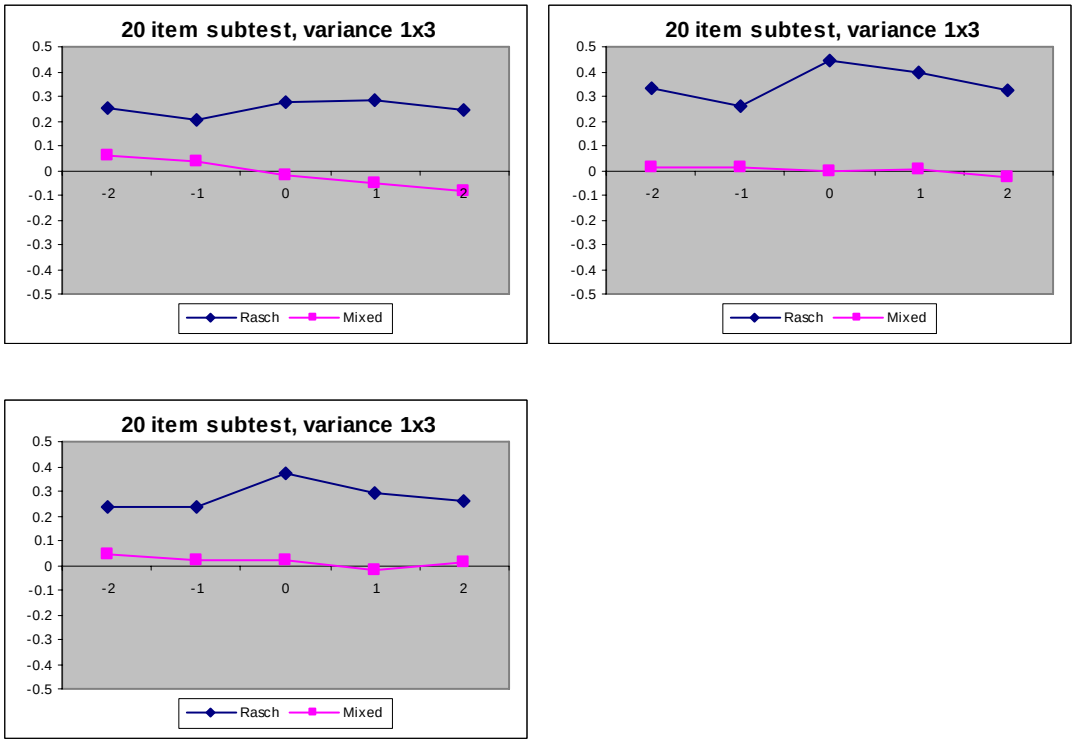
supports the presence of the patterns revealed in the visual display, which are discussed in the following section.

Central findings residuals

The results examined residual size, the number of factors extracted, comparative MANOVAs between weighted data, CI's and graphical evidence. The most interesting cells in which to explore the residual data are those that have a residual size of .4% or greater and have eigenvalues over the HPA threshold. This means there is a large enough residual extracted and factor or two, which is still apparent in the residual. In many conditions from the results section, Significant F values and CI supports show visual differences between the Rasch and residually weighted data. These differences between the two datasets do not fully explain what is going on within the residual data. The focus of this section is to more thoroughly investigate residuals whose characteristics exhibit something left over when the Rasch partitioning of the data is removed.

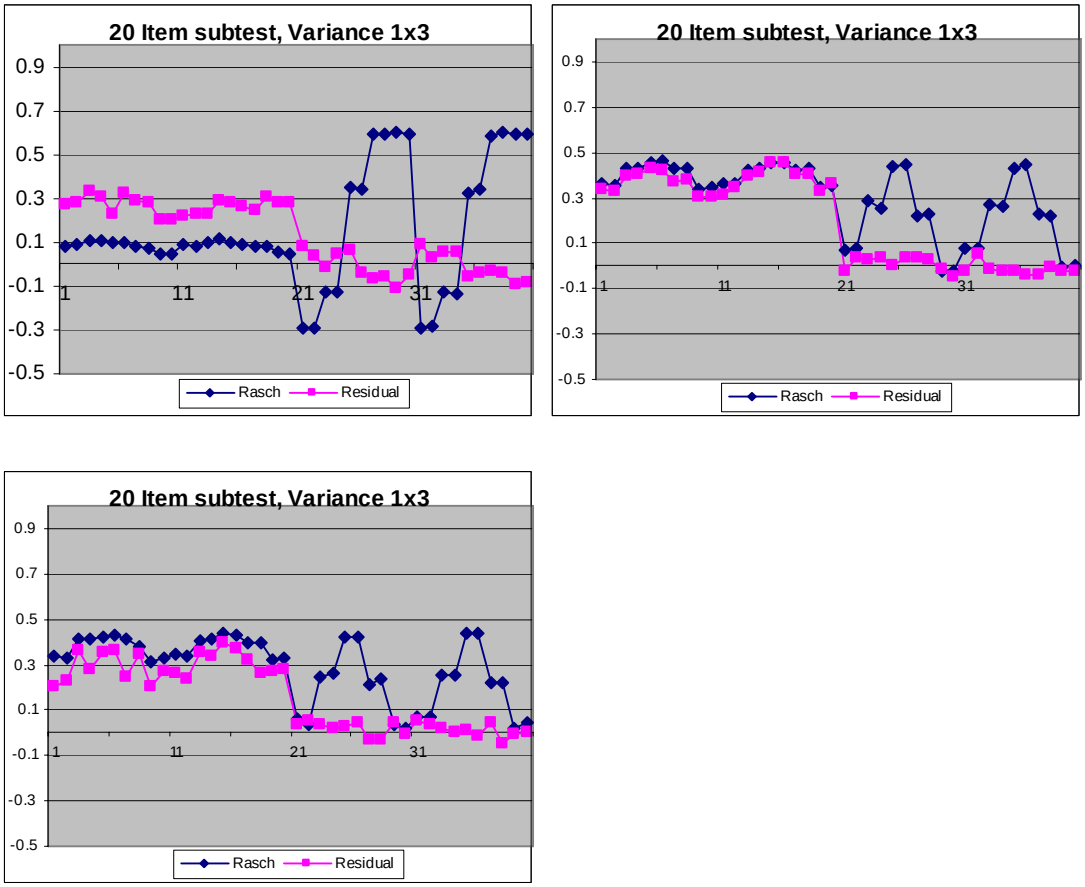
The random contamination residuals met the criteria for detailed exploration in three conditions where: the test range was ±2 for the uncontaminated subtest and ±6 for the contaminated subtest, when the subtest of size was 20 for the 50%, 20% and 5% contamination for the first factor.

Figure 6-33: 50% 20% 5% residual



When an observable factor is indicated in the random contamination condition the interaction occurs between a subtest of twenty items and an increase in the scaling factor of that subtest to 3, As expected from hypothesis 11 the contaminated subtest had smaller pattern values, close to zero, than the all Rasch subtests. Although many other conditions had significant differences between the Rasch and residually weighted data, these conditions had a clear first factor. The other conditions were frequently significant because the Rasch weighted data had Rasch patterns and the random contamination residual had random patterns around zero. The pattern that is occurring in this data is made clear when compared to the Rasch pattern for that condition.

Figure 6-34: 50% 20% 5% comparison



The differences in the 50% contamination between the Rasch and Residually weighted data are obvious in the 20% and 5% contamination conditions. These conditions show that, as expected, Rasch weighted data from the Rasch baseline. The patterns in the residually weighted data show that same Rasch pattern for the Rasch generated subtest, items 1 through 20, and show patterns hovering around zero in the contamination portion of the subtest. In the Residual data this is clearly a non-Rasch pattern for the contaminated subtest and a Rasch pattern for the first twenty items. This pattern of Rasch for one subtest and patterns around zero for the contaminated subtest is the random pattern expected in the residual data when a pattern would exist.

The systematic contamination residuals are explored in the following conditions: The six conditions for 50%, 20% and 5% where the 20 item contaminated subtests had a range of ±6; The two conditions for the 20% and 5% where the 10 item contaminated subtests range was ±2 for the uncontaminated subtest and ±6 for the contaminated subtest; The one condition for the 5% condition the 20 item contaminated subtests the range was ±2 for the uncontaminated subtest and ±6 for the contaminated subtest.

Figure 6-35: Factor 1 Residual patterns over .4% for 50% reversed contamination

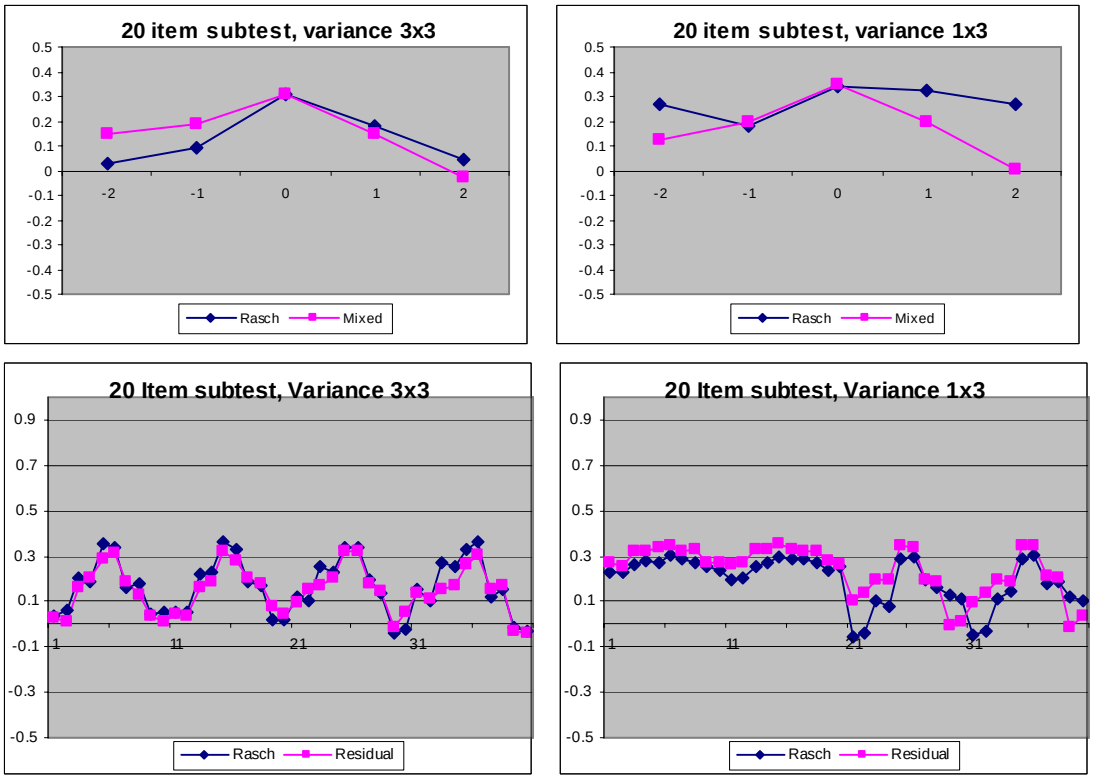
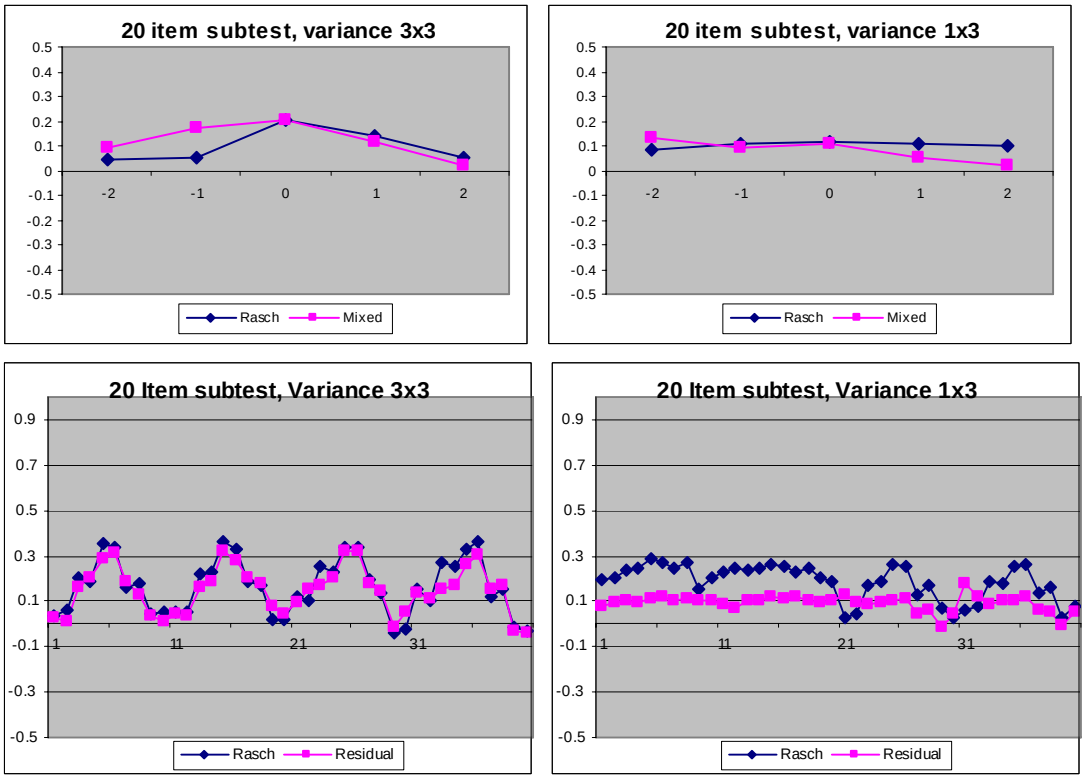
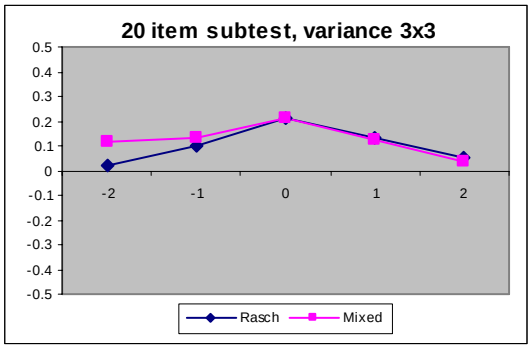


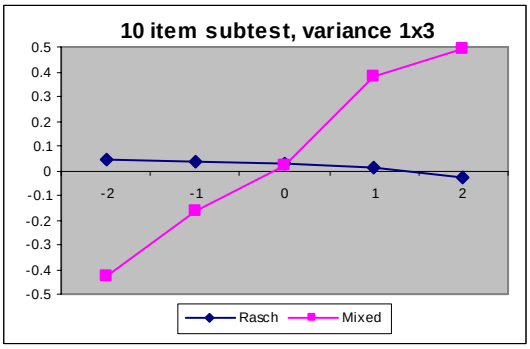
Figure 6-36: Factor 2 Residual patterns over .4% for 50% reversed contamination

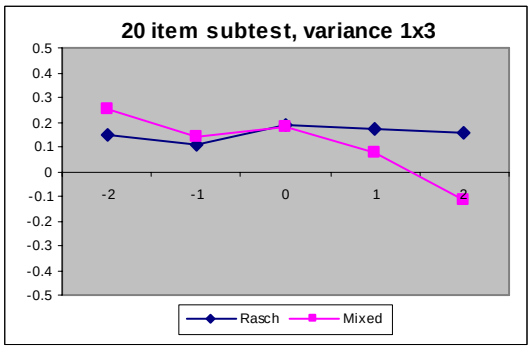


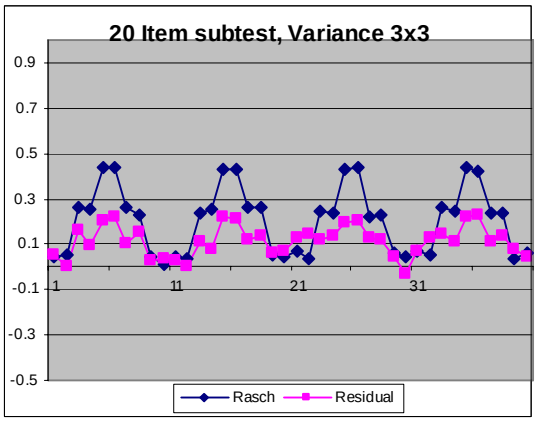
Although these conditions come up as having two factors, they do not look any different than Rasch data. The secondary factor in the 3x3 condition looks like the Rasch data and in the 1x3 condition does not take on any characteristics. The residuals are Rasch-like and the expected reversed Rasch Factor is not clear in the 50% contamination condition.

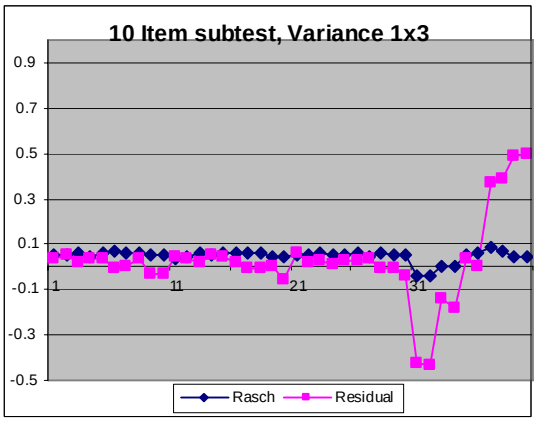
Figure 6-37: Factor 1 Residual patterns over .4% for 20% reversed contamination











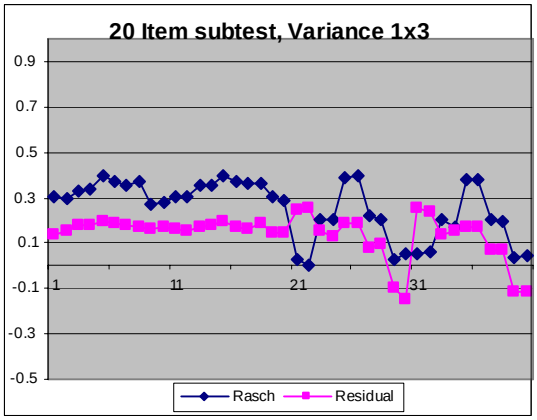


Figure 6-38: Factor 2 Residual patterns over .4% for 20% reversed contamination

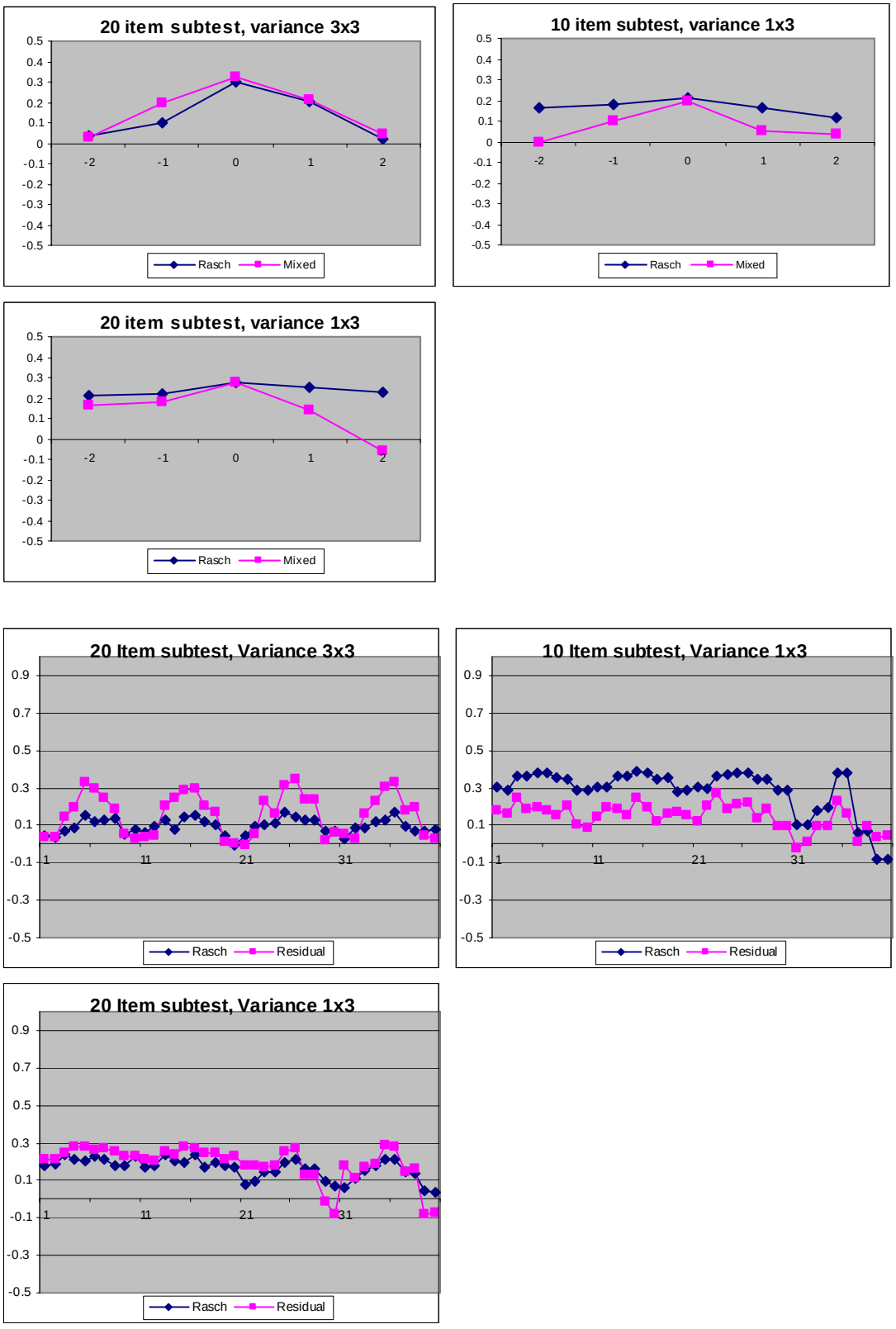
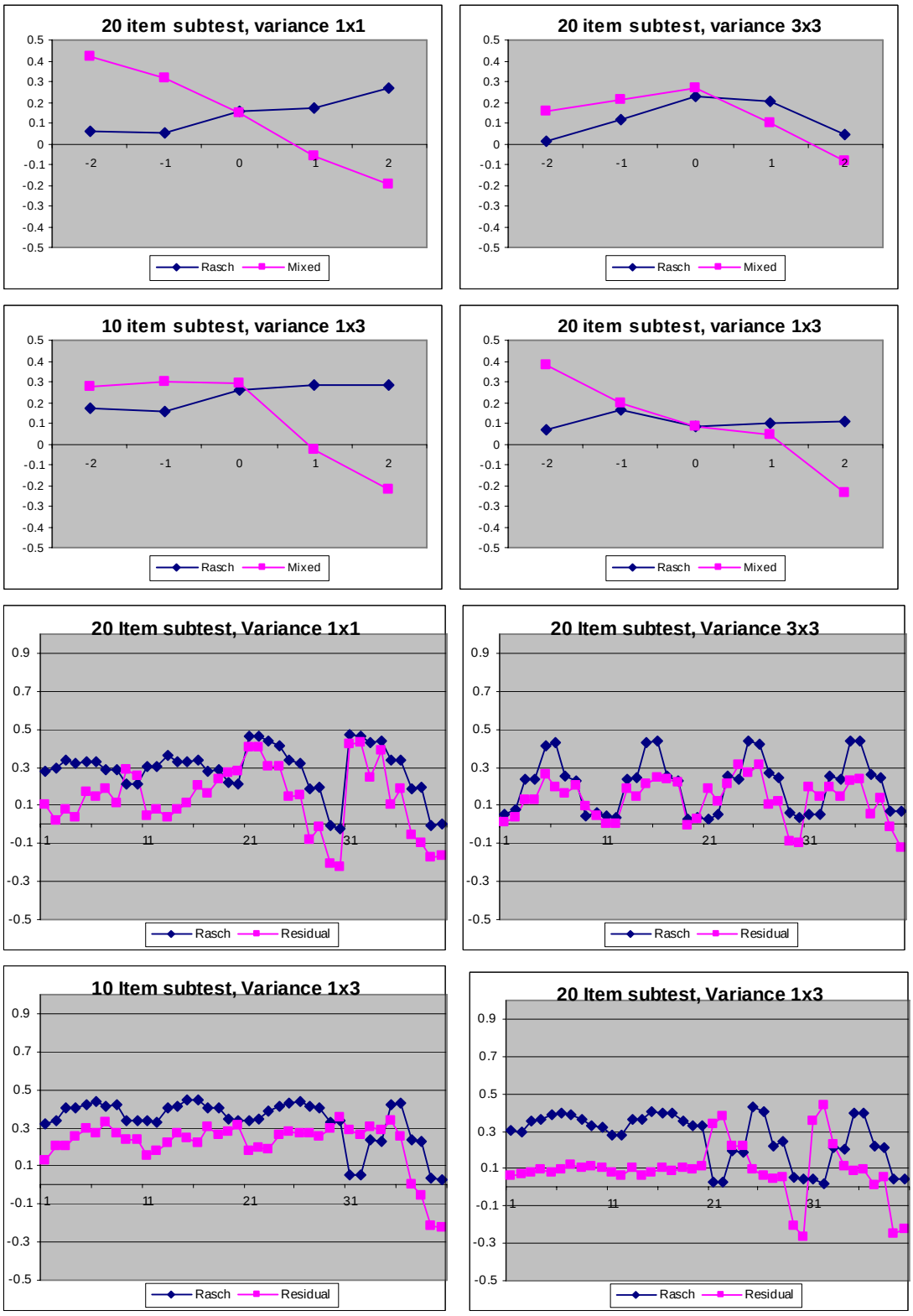
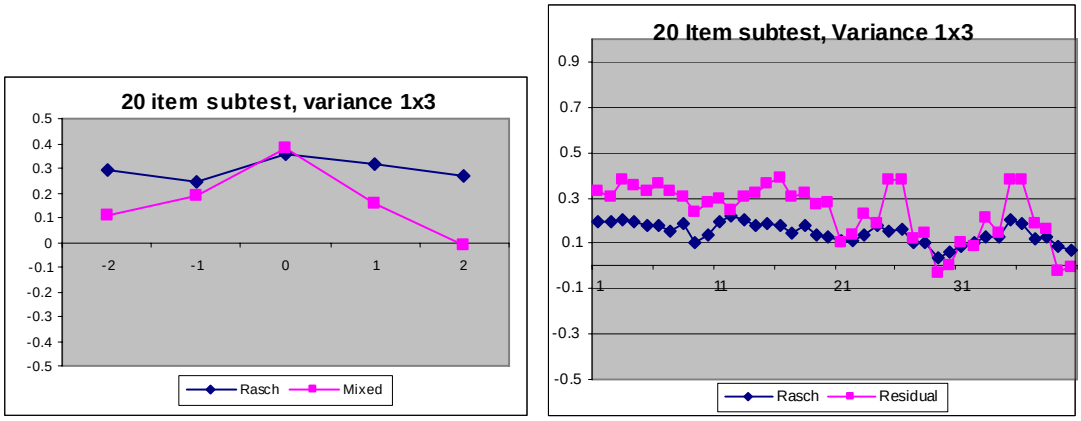


Figure 6-39: Factor 1 residual patterns over .4% for 5% reversed contamination



The two common themes in these patterns for the 5% and 20% conditions are crossing patterns and Rasch patterns. This crossing pattern supports hypothesis 12 that a reversed effect would become more prominent as the strength of effect increased. This pattern shows a reversal effect and a distinct pattern in most cases from the Rasch model. This first factor is a Reversed Rasch factor. In the comparison model the crossing pattern looks a lot like steps. In the 10 1x3 20% condition the first factor is the crossing pattern of Reversed Rasch and indicative of the systematic contamination. Some crossing is apparent in all the 5% conditions, but only one of them has a strong enough second factor that is an expected Rasch pattern. When a pattern is detected as the second factor in the Reversed Rasch contaminated conditions, it is like the Rasch baseline patterns.

Figure 6-40: Factor 2 residual patterns over .4% for 5% reversed contamination



A summary of hypotheses and briefly stated outcomes

1. It is expected that the residual size will:
    a. Proportionately increase as the percent of contamination generated into the data decreases
        i. Proportionate residual size equals 100% = 0.012, 50% = 0.147, 20% =0.239, 5% =0.356.
        ii. Hypothesis Supported
    b. Increase as the contaminated subtest range increases from ±2 to ±6
        i. Residual size for the contaminated subtest range ±2 = 0.717, for ±6 = 6.827.
        ii. Hypothesis supported
    c. Increase as the contaminated subtest size increases from 10 to 20 items
        i. Residual size for the contaminated subtest with 10 items = 0.777, for 20 items = 6.767
        ii. Hypothesis supported
    d. Be larger for the systematic contamination conditions when compared to the random contamination conditions.
        i. Residual size for the systematic contamination = 5.191, for the random contamination = 2.353
        ii. Hypothesis supported
2. In the eight Rasch only generated baseline conditions one factor will be present in the unweighted data.
        i. Two of the eight had small second factors supported by HPA

ii. Hypothesis partially supported

3. In all other 64 systematic or random contamination conditions will have present a second factor in the data for all unweighted datasets.
    i. Hypothesis supported fully in the 50% and 20% conditions with only 2 of 16 exceptions in the 5% condition
    ii. Hypothesis not supported in the 100% conditions with 4 of 16 having two factors.
4. On average, in the Rasch weighted there will be fewer factors extracted from the data when compared to the unweighted data.
    i. Three fewer factors were extracted in the Rasch weighted data in comparison to the unweighted data
    ii. Hypothesis Supported
5. On average it is expected that the systematically contaminated conditions will have more factors than the random contamination conditions when the residual is detectable.
    i. The systematic condition had 8 first and 6 secondary factors, while the random condition had 3 first and no secondary factors.
    ii. Hypothesis Supported
    a. When factors are found for the systematically contaminated conditions there will be two factors: one Rasch and one Reversed Rasch
        i. Hypothesis Supported 6 out of 8 times

b. When factors are found for the random contaminated conditions there will be only one factor in the data which is a suppressed Rasch factor.

ii. Hypothesis Supported 2 out of 2 times

6. Through the use of MANOVA it is expected that when residual misfit is extracted there will be a significant difference between the Rasch weighted patterns and residually weighted patterns in the first and second factors.

i. Hypothesis supported for the first factor

ii. Hypothesis partially supported for the second factor

a. It is expected that differences will be more detectable when the contamination is stronger. Specifically, stronger contamination is measured by: an increase in scaling factor in the contaminated subtest and an increase in the number of items from 10 to 20 items.

i. Number of detected differences for scaling factor 1=9, 3=14.

ii. Number of detected differences for subtest size of 20 item = 15, 10 item = 8

iii. Hypothesis supported, but interacts with selected number of residuals

b. It is expected that as the proportion of contamination increases, fewer residual effects will be significant. The contamination will overwhelm the data in both the residual and Rasch conditions and cancel out differences between the two weighted datasets.

iv. Only the 100% condition second factors for the random contamination selected were not significant.

v. The reversed Rasch contamination were more frequent as the proportion of contamination decreased

vi. Hypothesis partially supported

7. When the Rasch only, baseline data, are randomly split, there should be no visible difference between the two sets of patterns for the first or second factor.

iii. Hypothesis supported by no visible differences in graphs

8. In the random contamination conditions the residually weighted dataset should have suppressed patterns for the first and second factor. The Rasch weighted data should still have strong Rasch patterns. The differenced should be captured with CI differences attributed to residual weighted values close to 0 and Rasch weighted data following Rasch type patterns. These differences should be apparent in graphs. The Wilks' lambda should be significant and larger than the baseline F.

i. Hypothesis supported visually and through analysis

9. In the reversed contamination condition the residually weighted dataset should have strong patterns similar to Rasch weighted data but in a different graphic structure for both the first and second factor. The differences should be captured with CI differences attributed to the differences in patterns particularly in the contaminated subtest. The Wilks' lambda should be significant and larger than the baseline F.

i. Hypothesis supported visually and through analysis

10. It is expected that the subtest in the Rasch baseline conditions will look like the remainder of the exam. Both subtests will have Rasch patterns
    i. The type of Rasch pattern displayed was dependent on the scaling factor.
    ii. Hypothesis partially supported
11. In the random condition it is expected that the contaminated subtest will have significantly smaller pattern values, close to zero, than the all Rasch subtest.
    i. Hypothesis supported visually and through analysis when a factor existed for the random contamination condition.
12. It is expected that patterns in the reverse effect condition should show a reversed pattern in the subtests which becomes more prominent as the strength of the effect increases.
    i. Hypothesis supported visually and through analysis when a factor existed for the reversed effect contamination condition.

Chapter 7: Discussion

The success of this study depended on the ability of the two-class model to parse out a residual class based on an unscalable condition. Next the study focused on the systematic exploration of differences between two weighted datasets, one Rasch and the other residual. The residuals were explored in more detail in the spirit of Tukey to find if there were meaningful relationships left when a residual class was large enough to be explored. Tukey, of course, explores a traditional residual in the context of regression. In the current investigation the residual explored is the proportion assigned to the unscalable class on a case by case basis. In the spirit of Tukey, the graphs provide a tremendous amount of information in the 29 conditions with residuals large enough to explore.

The first goal of the study, to separate out some information from the Rasch class into an unscalable class, was successful. Out of the 64 possible conditions with modeled contamination, the residual was large enough to explore in 29 of those conditions. “Large enough to explore” was operationally defined by a threshold of .4% or the equivalent of at least two simulees. In the 100% contamination condition, the random contamination produced a searchable residual four out of eight times. The reversed contamination condition did not produce anything to explore. These eight conditions should have produced no meaningful residual greater than the true Rasch conditions. The reversed Rasch scaled the difficulties backward for all simulees in the conditions, effectively mirroring the all Rasch condition. The subtest was full Rasch with items in the opposite order. The residuals that are left look nearly identical in

both magnitude and relative pattern as the Rasch baseline condition. Therefore, our success story is altered from 29 out of 64 to 45 out of 72 if one includes the Rasch baseline.

In the current investigation, the model selected was unable to acquire a residual for any condition whose Rasch subtest that had the difficulty parameter scaling factor of 3 and the contaminated condition had a scaling factor of 1. The Rasch subtest with a wide range of ±6 seemed to saturate the model. It is likely that the larger scaling factor overwhelmed the model with the broad range in the Rasch subtest. The contaminated subtest with a range of ±2 was proportionately truncated and seemingly drowned out by the range of the Rasch only subtest. The subtest for random contamination the mixing had its discrepancy at maximum with a ±2 theta being replaced by an unscalable value of .25. When the subtest in the contaminated condition was mixed with Rasch and reverse Rasch, the discrepancy was greatest for ±2, however for ±1 and 0 the Reversed contamination was relatively small compared to the Rasch subtest with a range of ±6.

The opposite overall effect was true for the conditions where the Rasch subtest that had the difficulty parameter scaling factor of 1 and the contamination condition had a scaling factor of 3. In all cases, except for the Rasch only and the 100% reversed Rasch condition, which was effectively like the Rasch only condition, a residual was considered large enough to explore the weighted data. The Rasch data in these conditions for the Rasch subtest was restricted to ±2. This left what could be described as a larger proportional discrepancy amongst the mixing of the Rasch and contamination in the contaminated subtest. Here the truncated Rasch was in the Rasch

only subtest and the contaminated subtest had potential for large discrepancy in its mixing subtest. The ±6 and ±2 difficulties had a relatively large difference in the contaminated condition from the unscalable value. In the Rasch and reverse Rasch conditions some very extreme discrepancies occurred when ±6 and ±2 difficulties were reversed for the same set of items.

Although interesting, the size of residuals along with the number able to meet the threshold value of .4% was merely the first step in a series of investigations. Next, the eigenvalues were explored. Not much was interesting with the unweighted eigenvalues. When a secondary factor was placed into the data and mixed, the factor models mostly detected the second factor. In the 100% contamination condition, the reversed Rasch contamination was effectively the same mathematically as the Rasch condition and showed similar results. The 100% random contamination condition did not add a large factor to the data, which is not surprising considering the data was generated at random for all simulees on those subtests.

When the residual was removed so that the data was weighted to the Rasch class, the first factor strengthened relative to the second factor. The data was better fit to have one factor, although most of the time the secondary factor was still present.

When the data was weighted by the residual class, the reversed condition showed clearer factors than the random condition. In the exploration of residuals with factors the reversed, Rasch contamination often showed up in the residual and a secondary factor was also present in several conditions as a Rasch pattern. In the random contamination most of the time the patterns were scattered around zero.

MANOVAs used in the study as a relative tool more than a tool for testing multivariate significance. The F value in the eight Rasch baseline conditions serves as an expected value. It can effectively be seen as a shift in the non-centrality parameter. When used as a comparison tool of effect size it becomes clear that contaminated models show a difference in factor patterns between the Rasch and residual weighted cases. The effects here are not the meaning of what is happening within the residual but between the residual and the Rasch weighted data. The CI's help support the story of different patterns between the two groups, not within the unscalable class.

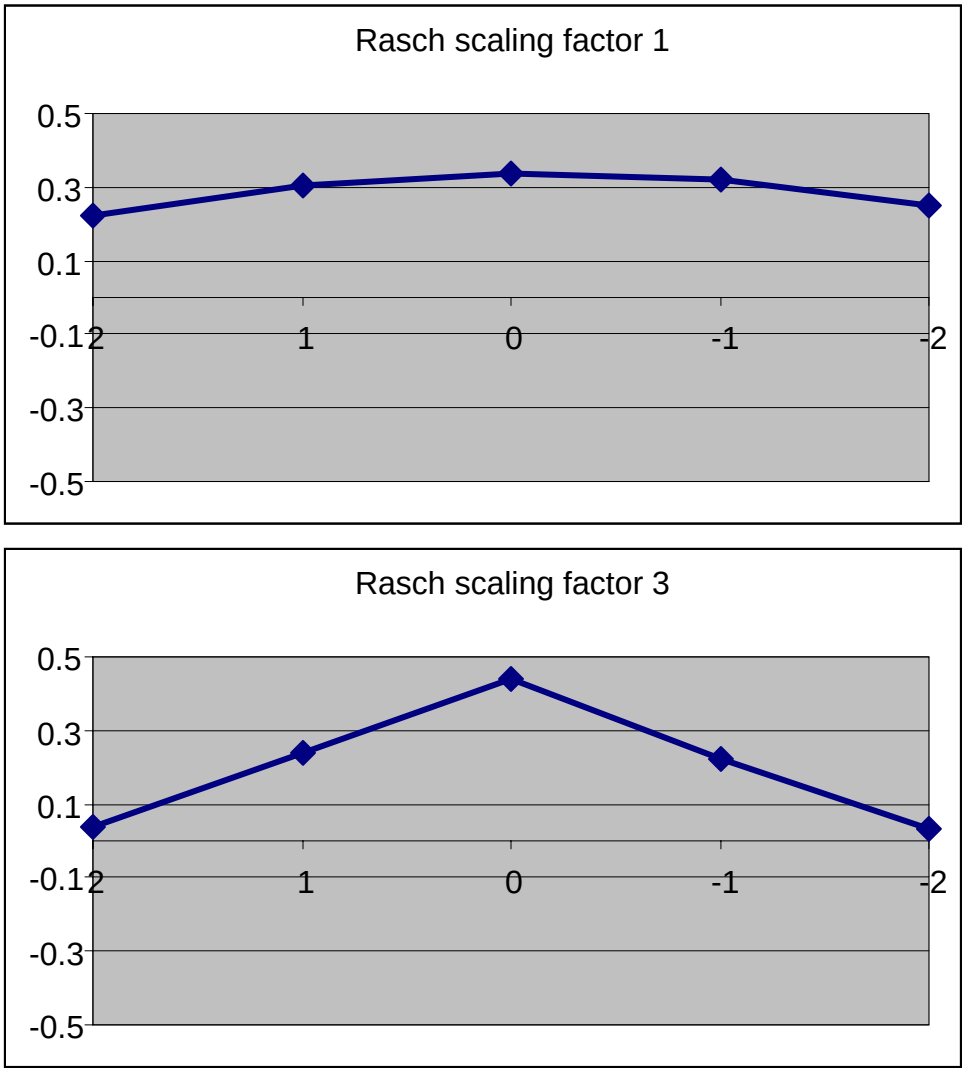
The graphs and the tests of significance show some clear separation between the Rasch weighted data and the residually weighted conditions. The patterns in the randomly weighted conditions often show differences between the Rasch subtest on the two sets of patterns.

It is observed that the differences in the patterns for random and systematic contamination subtest differ from the all Rasch subtest. The Rasch patterns are very clear, taking on a shape that is either a flat with as slight rounding pattern around .3 on average for the subtest with a range from ±2 or an upside down v shape peaking around .3 or .4 for the ±6 subtests. The pattern for the ±2 condition is a less extreme version of the ±6 condition. On the first factor for the conditions with a range of item difficulties of ±6: the extremely ±6 item difficulties have pattern values around zero, the moderate ±3 difficulties had patterns around .3 and the 0 item difficulties had values around .5. As the items approached zero the patterns in the factor were stronger. This was seen in a less extreme condition in the ±2 condition. The ±2 item difficulties have values close to .2 patterns while the difficulties around zero reach a

maximum around .4. This means that items which are closer to the average ability are the strongest patterns. As items diverge from the average theta value of 0 to more extreme values the patterns converge to around zero. The patterns seem to reach a maximum value where maximum information occurs. When little can be discerned in the most extreme item difficulties the factor pattern goes to zero.

The current underlying theory for Rasch only factor patterns is that maximum patterns are displayed at maximum information and the minimum patterns are displayed at minimum information. The secondary factor pattern is just a weak copy of the first pattern. Figure 7-1 shows two examples of the Rasch only maximum information pattern. The scale in figures 7-1 through 7-4 is based on the scaling factor and indicated in the title as either scaling factor 1 (±2) or scaling factor 3 (±6).

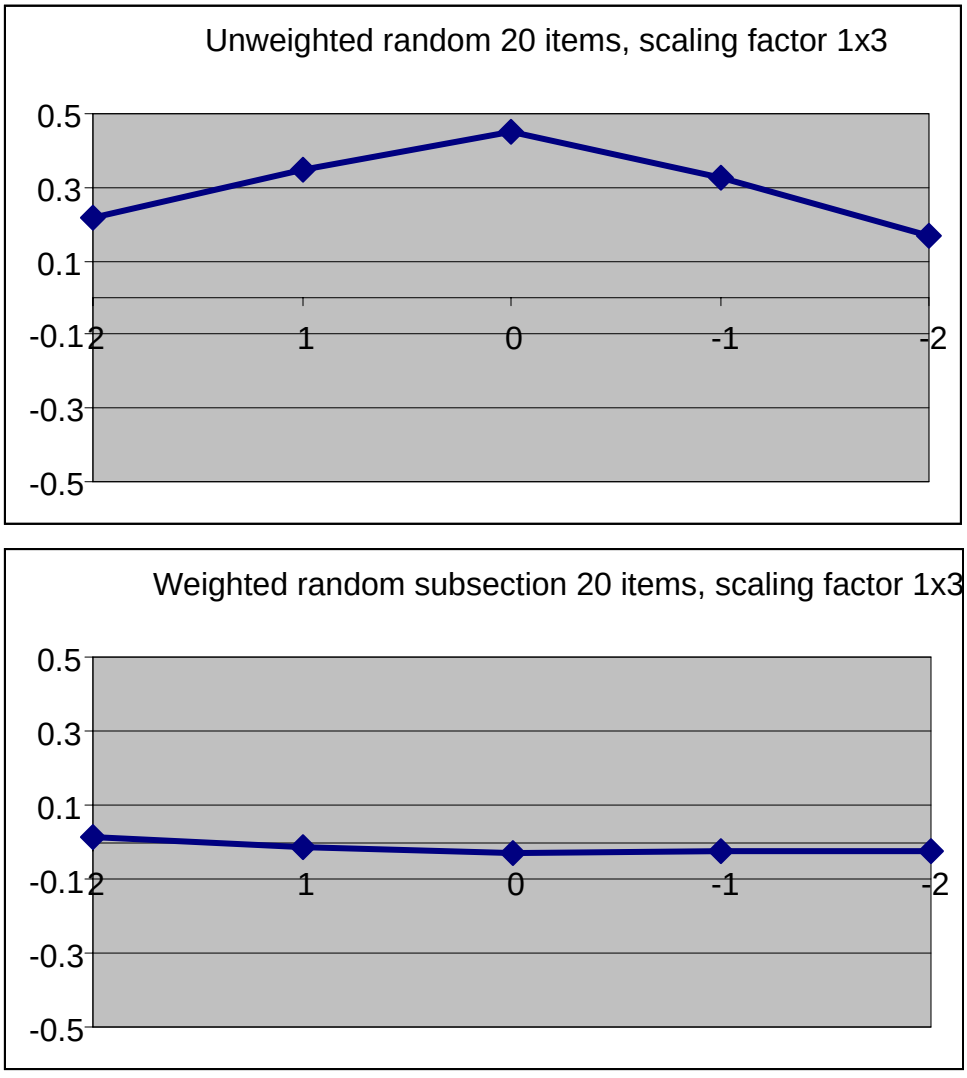
Figure 7-1: Maximum information pattern



In the residual of the random contamination condition, when a factor is determined to exist the obvious pattern that occurs is a separation between the Rasch and random patterns. The random subtest has patterns around zero while the Rasch subtest has patterns that match the Rasch baseline patterns with the largest patterns closest to the mean generating value of zero being largest and those with more extreme difficulties being weaker. Figure 7-2 shows the unweighted and weighted random contamination example. This example shows that random contamination may

still be present and a residual could possibly be extracted in the data even when the model looks like a Rasch patterns.

Figure 7-2: Random contamination pattern



In the reverse conditions the most interesting patterns have a crossing or reversal pattern. The reversal pattern when it is most present in the residual data yields a pattern different from the Rasch maximum information pattern. The reversed pattern has large positive pattern values for large positive item difficulties, small patterns for item difficulties near zero and large negative pattern values for large

negative item difficulties. This reversal contamination pattern is shown in Figures 7-3 and 7-4 with the unweighted example as well.

Figure 7-3: Reversed contamination pattern, example 1

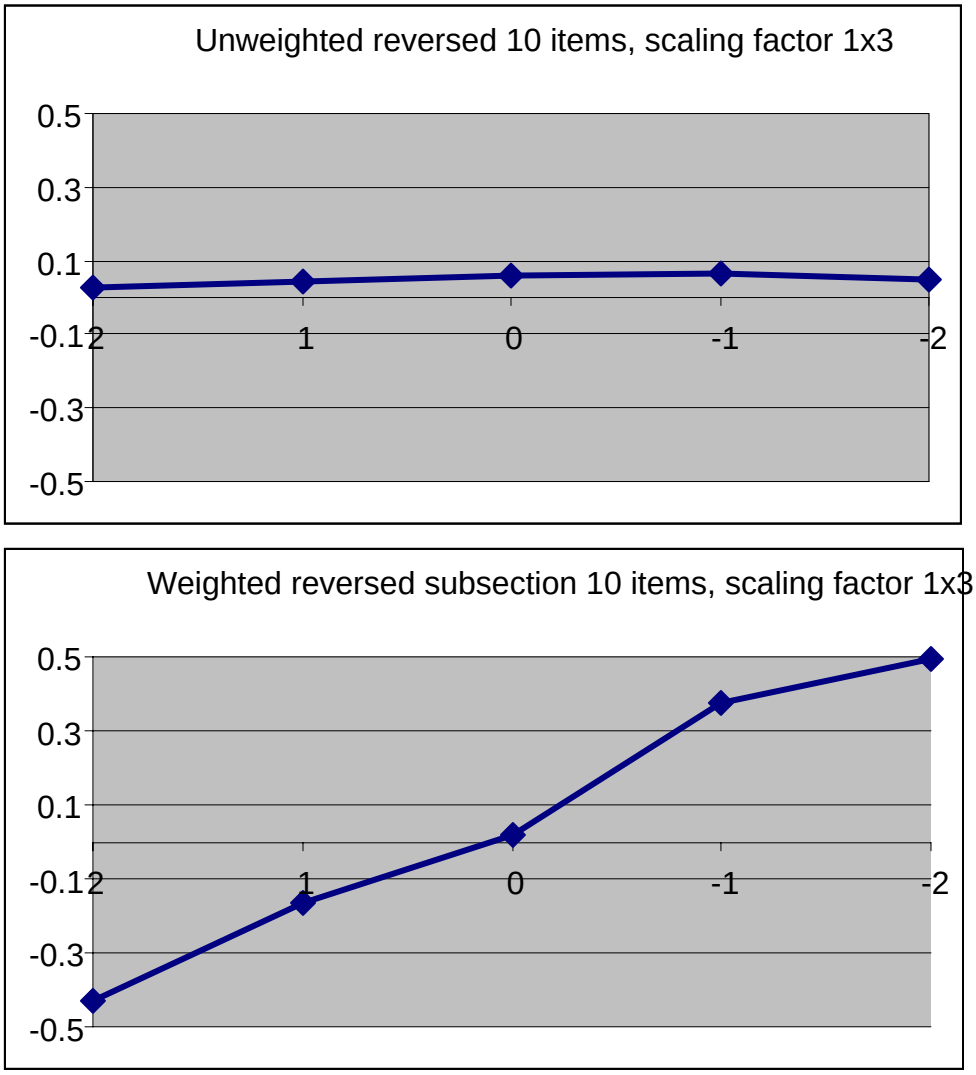
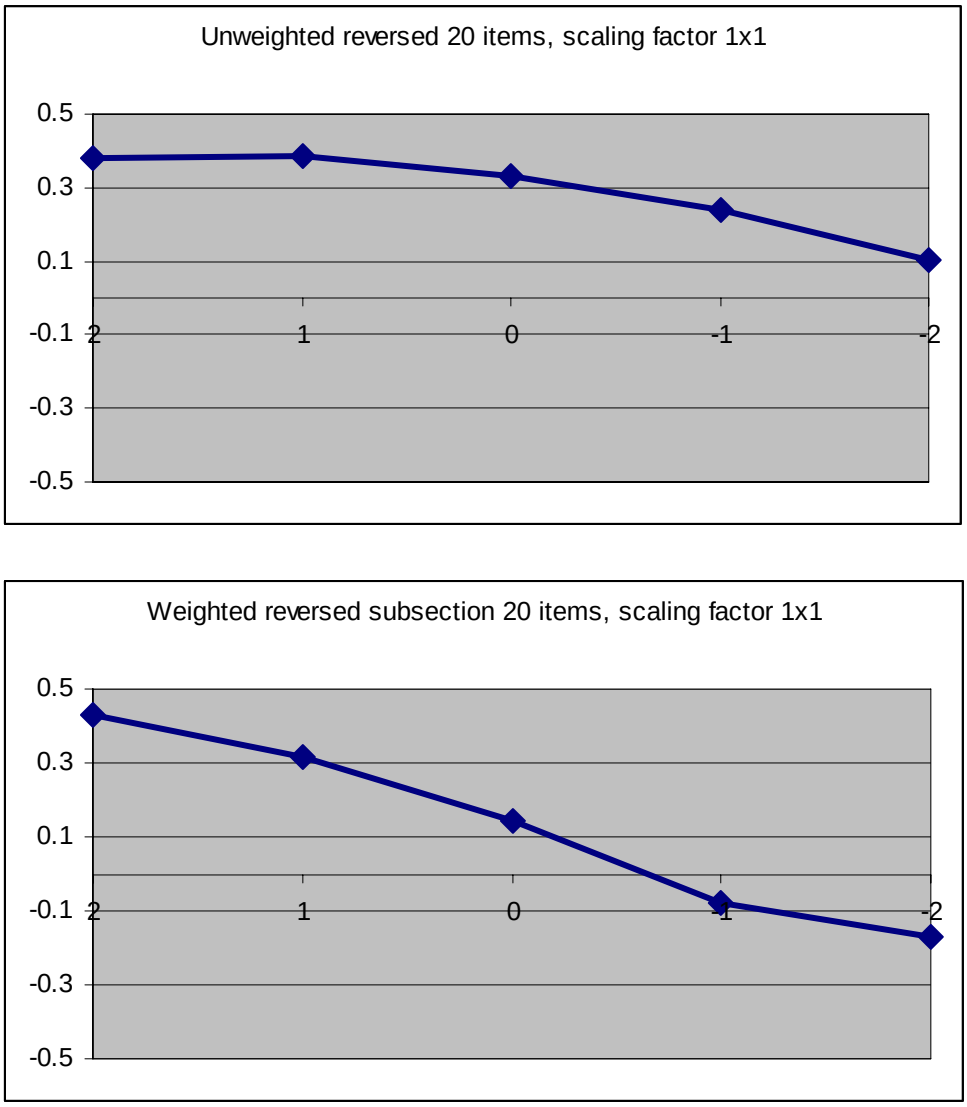


Figure 7-4: Reversed contamination pattern, example 2



In both of the reversed contamination examples, the Rasch model is suspect in the unweighted data. In the first example the Rasch model looks to have suppressed patterns very close to zero, across all levels of difficulty. When the contaminated subtest is explored in the residual weighted data a clear reversed or crossing pattern exists. In the second example the Rasch data takes on some obvious distortion that

could be similar in nature to the reversed pattern. Again when the contaminated subtest is explored in the residual weighted data a reversed pattern exists.

Research Implications

One of the first research implications of the two class models with one class as the residual class is the fact that when the data was truly Rasch under the eight tested conditions, the Rasch model was overwhelmingly classified correctly. This means that as part of a battery of examinations for the testing mechanism, one could use the model with an unscalable class of .25 to determine if anything out of the ordinary can be detected with an unusual size residual. In the 50% reversed contamination condition of the data, twice a residual as small as .5% indicated departure from Rasch in the data. The amount detected did not always indicate the amount truly contaminated but for future research in theoretical and applied work could act as a flag for detection types of competing contamination to the Rasch model.

In cases in which all respondents are equally misinformed or contaminated, the model may not pick up or detect anything unusual, as they may be similar to the reverse Rasch condition or random condition. However, where some students are functioning Rasch and others are functioning in some alternative method, this may be detected by a model as simple as adding an unscalable class to use a residual information. The residual can then be explored.

Just because a residual is detected does not make it meaningful and the opposite was also true. Sometimes the residual was extracted and it looked just like the Rasch weighted data. Other times it looked like random scatter around zero. There were also times when a residual did not indicate a factor but something was clearly

going on. HPA should be used more as a guide than an absolute threshold when it comes to factor extraction, especially when dealing with small residual sizes.

In addition it is conceivable to this researcher that residuals could be misleading. In the 100% random it is possible by chance to place more than would be expected into this class. If a low performing person misses a few easy question and gets two questions right they wouldn't be likely to get correct then they might be a high candidate to be placed in the residual class when they actual are very similar to other individuals. The residual class should be inspected but it may not be that they are truly from the same population.

The factor patterns also show a Rasch pattern which could be explored in more detail to determine if departure from that pattern provides useful information in terms of some alternative to the Rasch model. In the current study it is important to remember that the Rasch baseline is known from examining the baseline models. This permits discussions of departure from this model to be considered adulterated, even in the unweighted factor patterns. If item difficulties and factor patterns generated from those difficulties are know, one would be unlikely to use this technique as a first method as the unweighted factor analysis could simply be compared to the expected unweighted factor model and adulteration of the model could be determined from the raw data. However it is often the case in applied research that the item difficulties and resulting factor patterns are not know before the model is used. This means the baseline model would be unknown and it would be very difficult to know if contamination of the Rasch model was manifest in the data. This technique could serve as a tool to determine, first if there is a residual class of sufficient size to be

considered a flag for departure from the expected model, in this case the Rasch model. Next the residual weighted data can be compared to the Rasch weighted data to determine if differences exist. If differences exist between the two patterns the residually weighted data can be explored in more detail to determine if the factor patterns match some form of know contamination, such as the reversed Rasch or random contamination, or if the contamination is something else yet unidentified.

The most useful immediate outcomes of this model would be in remediation of individuals, evaluation of the exam, and as a guide to inspection of teaching methods. Alternative strategies may help guide an experienced educator to retrain or educate an individual not only to missed content but to a new strategy of thinking. Items on an exam when seen as alternative strategies could be indicators that items may need to be rewritten. Subject matter experts should be used to determine if an alternative pattern discovered on a test could be due to bias. The patterns could also indicate teaching methods need to be employed to incorporate all relevant strategies that may underlie a content area. As in the martial arts comparative strategy mentioned earlier, it is possible that a set of patterns indicates that remediation is not just substantive area but also a strategic one where new methods of teaching may be required.

Direction for future research

The first direction for future research would be to vary some of the existing fixed and manipulated factors. In keeping with the success of this research, it would be interesting to vary the sample size of the test as computing capacity increases. Computing capacity has increased substantially in just the time from the conception

of this dissertation to the current date. It would be extremely useful if one wanted to examine very small residual of sizes in the magnitude of .1% to have much larger datasets. In just a few more years it would be likely that advancement in software and hardware would allow this type of analysis to be done in a fraction of the time. It would also be useful to vary the size of the subtest of contamination and look at some very small contaminated subtests, maybe 10% of items. The range of the subtests could also vary to have a targeted subtest with a standard deviation around zero, say .1.

The model should also be extended to new classes and IRT conditions as well as adding more classes. It would be interesting to use various IRT models as classes. Models could incorporating several different logistic parameter models, such as a 2 and/or 3 parameter logistic model, each as its own class and with its own weighted value to examine . Of particular interest to the research is a model that looks at cumulative and unfolded perspectives of item response theory. A Bayesian model with three classes, one Rasch, one unscalable and the last one a variant of the Hyperbolic Cosine model (HCM) (Andrich & Lou, 1993). This would be particularly useful in extending this research to areas of surveys of agreement. Data can be contaminated using the unfolding approach, as well as the reversed Rasch contamination, should act as a starting point for what a distinct systematic form of contamination would look like.

In future research both the generating model and analytical model can be altered. The examples of alternative models also relate to the generation of data and contamination. Generating data for contamination could also include other models

such as mixing hyperbolic models and alternative IRT models. The analytic models and the generating models could be the manipulated factors of future investigations.It would be useful to explore the baseline Rasch model under more conditions. In particular the interaction between θ and item difficulty should be explored to determine if pattern remain the strongest when item difficulties and θ are close together. This could support the maximum pattern at maximum information theory. If the concept is confirmed it may be useable to detect departure from the Rasch model, both systemic and random.

The current model does not address how to correct the model when alternative strategies are found. Future research can look into more robust models that indicate that a given form of contamination detected better fits a alternative analytical model.

The current model uses a Rasch mixture model to generate the data and then uses a Rasch and unscalable mixture to fit to the resulting data. This is a best case scenario. It would be useful to not only explore models that are generated with contamination but to use models where the data generation process and the estimation process do not match. In these models a review of the estimation model as conducted here in the preliminary analysis would be advisable. A researcher may even have to use strong theory for modeling priors in order to provide useful results.

Conclusion

The concept of a residual to the Rasch condition, although different from that context of regression, still needs exploration of what is left over for there to be meaningful understanding of the residual. In the current investigation, it was found that systematic contamination in smaller percents of the sample size showed some

clear departures from Rasch patterns. In many of the random conditions, the unscalable class was extracting contamination that had no real pattern. A few times for the random contamination an exceptional pattern was discovered; again, this was in the smaller contamination sample size conditions. A clear Rasch pattern has been detected using factor analysis and is identified as a maximum pattern at maximum information.

When residuals are explored and detected with systematic or idiosyncratic patterns, it might benefit the researcher to investigate why this is happening. If a residual exists in a dataset but has no pattern, it just might be guessing, rushing, or other rationale for why a residual looks like no Rasch pattern is present. If the residual has a Rasch pattern, it may indicate something wrong with the dataset that something would be classified as unscalable and then look somewhat like the Rasch class. In the extreme cases when a systemic pattern like the crossing pattern in the reversed Rasch contamination condition are manifest, one might look closely at the underlying curriculum to determine if multiple methods of knowledge or skills are being taught.

Given the test properties, the simple detection of a residual went above and beyond what was expected and was an indicator of a mixture. Lack of a residual over size .4% did not mean contamination was absent, so the current model would only be amongst a battery of exploratory tools to use in application. Further development of a more sophisticated model rather than just an unscalable residual would greatly aid in detection of contamination.

When using real data researchers do not know much about the parameters of items or people and subtests should they exist may be completely lost on a methodologist. In the current investigation the subtest are known to the researcher. In real data the researcher may not know subtest. Substantive experts need relied on either during test construction or after the exam has been developed to determine if areas of an exam can be grouped into subtest. The researcher, although an expert on modeling, may know nothing about the material under investigation. This type of work may even require a team of experts working together as a focus group to evaluate and determine subtests should they exist for an exam.

Overall some patterns exist for the Rasch baseline conditions explored which seem to be driven by the item difficulties proximity to the average θ. In the residual weighted data, patterns different from Rasch patterns are present. The random contamination tends to yield a weak flat pattern. The systematic contamination showed an interesting and opposite effect. In its most apparent form the extreme item difficulties had the largest values in an absolute sense while those closer to zero had relatively small patterns. This contamination is easier to detect in the unweighted data when systematic contamination exists as some adulteration of the data can be found.

Appendix A

Table A-1: Relevant Horn's parallel analysis values

%Contaminated	F1	F2	F3	F4
0	1.678	1.581	1.538	1.476
5	1.685	1.591	1.530	1.478
50	1.987	1.846	1.747	1.687
80	2.658	2.427	2.273	2.119
90	3.495	3.185	2.909	2.642
95	5.275	4.486	3.860	3.274
98	8.648	7.427	6.282	5.478

Appendix B

Table B-1: Factor 1 Wilks' lambda F-values from 100 MANOVAs on Rasch generated data.

%	F-Value	SE	2 positive SE
50%	1.00	0.29	1.59
20%	1.40	0.42	2.24
5%	2.15	0.93	4.02
2%	3.19	1.78	6.75
1%	3.29	1.81	6.91
.4%	3.70	1.84	7.38

Table B-2: Factor 2 Wilks' lambda F-values from 100 MANOVAs on Rasch generated data.

%	F-Value	SE	2 positive SE
50%	1.02	0.29	1.61
20%	1.28	0.38	2.04
5%	2.08	0.56	3.21
2%	3.08	1.19	5.45
1%	3.07	0.89	4.84
.4%	2.88	0.98	4.85

Appendix C

Table C-1: Unweighted first factor patterns for 100% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.36	0.37	0.06	0.08	0.04	0.05	0.34	0.36	0.24	0.24	0.05	0.06	0.04	0.03	0.30	0.33
I2	0.36	0.37	0.09	0.06	0.05	0.09	0.35	0.37	0.23	0.24	0.06	0.03	0.05	0.02	0.27	0.30
I3	0.42	0.44	0.26	0.31	0.29	0.34	0.42	0.43	0.32	0.31	0.25	0.26	0.21	0.18	0.34	0.37
I4	0.42	0.43	0.26	0.30	0.29	0.29	0.42	0.45	0.32	0.31	0.25	0.28	0.23	0.19	0.34	0.38
I5	0.44	0.48	0.47	0.49	0.48	0.49	0.44	0.47	0.34	0.33	0.45	0.43	0.40	0.39	0.33	0.39
I6	0.44	0.47	0.47	0.50	0.48	0.50	0.45	0.47	0.31	0.34	0.43	0.43	0.41	0.39	0.36	0.41
I7	0.40	0.44	0.30	0.30	0.29	0.30	0.42	0.45	0.31	0.30	0.23	0.23	0.21	0.22	0.32	0.37
I8	0.41	0.44	0.29	0.31	0.29	0.32	0.41	0.44	0.30	0.30	0.24	0.23	0.20	0.22	0.30	0.38
I9	0.32	0.35	0.06	0.11	0.05	0.10	0.35	0.37	0.26	0.23	0.05	0.04	0.05	0.07	0.26	0.31
I10	0.31	0.37	0.08	0.04	0.07	0.06	0.35	0.37	0.28	0.23	0.04	0.07	0.04	0.02	0.25	0.30
I11	0.35	0.37	0.05	0.07	0.06	0.07	0.35	0.36	0.23	0.23	0.08	0.06	0.06	0.03	0.26	0.32
I12	0.34	0.36	0.05	0.08	0.07	0.05	0.34	0.38	0.22	0.25	0.07	0.06	0.06	0.03	0.26	0.33
I13	0.41	0.44	0.26	0.30	0.29	0.30	0.41	0.44	0.30	0.33	0.24	0.26	0.25	0.17	0.35	0.37
I14	0.43	0.43	0.27	0.31	0.29	0.28	0.41	0.45	0.30	0.33	0.24	0.27	0.24	0.21	0.32	0.36
I15	0.44	0.46	0.46	0.49	0.50	0.50	0.45	0.46	0.35	0.33	0.43	0.44	0.41	0.40	0.36	0.41
I16	0.42	0.47	0.46	0.49	0.47	0.50	0.45	0.47	0.32	0.33	0.43	0.43	0.43	0.39	0.37	0.42
I17	0.39	0.44	0.31	0.29	0.27	0.29	0.41	0.43	0.32	0.31	0.24	0.23	0.24	0.25	0.33	0.36
I18	0.40	0.44	0.32	0.30	0.27	0.30	0.42	0.43	0.31	0.31	0.25	0.23	0.22	0.24	0.34	0.39
I19	0.33	0.36	0.04	0.06	0.07	0.09	0.34	0.37	0.23	0.25	0.05	0.06	0.04	0.02	0.27	0.29
I20	0.34	0.37	0.07	0.08	0.08	0.08	0.33	0.36	0.25	0.24	0.07	0.07	0.05	0.05	0.26	0.29
I21	0.35	-0.01	0.05	0.02	0.07	0.00	0.34	-0.01	0.23	0.24	0.04	0.05	0.04	0.32	0.29	0.01
I22	0.34	0.00	0.05	-0.01	0.05	0.01	0.34	0.00	0.24	0.23	0.06	0.04	0.06	0.29	0.29	0.04
I23	0.43	0.00	0.25	0.00	0.27	0.00	0.41	0.00	0.32	0.32	0.26	0.23	0.23	0.36	0.35	0.24
I24	0.43	-0.01	0.26	0.01	0.27	0.02	0.42	0.00	0.31	0.31	0.26	0.24	0.21	0.38	0.34	0.20
I25	0.43	0.00	0.45	0.01	0.48	0.01	0.45	0.00	0.31	0.35	0.42	0.43	0.43	0.40	0.37	0.40
I26	0.43	-0.02	0.45	0.01	0.48	-0.03	0.45	0.00	0.34	0.32	0.44	0.42	0.42	0.38	0.35	0.41
I27	0.41	0.00	0.28	0.01	0.29	-0.01	0.42	-0.01	0.31	0.34	0.25	0.27	0.26	0.35	0.32	0.23
I28	0.41	0.00	0.29	-0.02	0.29	0.02	0.41	-0.01	0.32	0.34	0.24	0.29	0.23	0.34	0.31	0.22
I29	0.34	0.00	0.06	-0.01	0.07	0.02	0.34	-0.01	0.27	0.25	0.02	0.06	0.06	0.27	0.25	0.06
I30	0.32	0.01	0.07	0.01	0.07	0.01	0.34	-0.01	0.27	0.25	0.05	0.05	0.03	0.27	0.26	0.06
I31	0.00	0.00	0.00	-0.02	0.00	-0.01	0.00	0.00	0.26	0.23	0.03	0.05	0.34	0.32	0.06	0.02
I32	-0.01	-0.02	0.00	0.00	0.01	0.00	0.01	0.01	0.24	0.24	0.07	0.05	0.33	0.33	0.02	0.06
I33	0.00	0.01	0.00	0.01	0.00	0.00	0.01	-0.01	0.32	0.31	0.23	0.21	0.40	0.36	0.19	0.21
I34	0.01	0.01	0.01	0.00	-0.01	0.00	0.00	-0.01	0.29	0.33	0.25	0.21	0.37	0.37	0.16	0.22
I35	0.01	0.00	0.00	0.01	0.00	-0.02	-0.01	0.00	0.33	0.33	0.42	0.43	0.43	0.40	0.35	0.39
I36	0.01	0.00	0.00	0.00	-0.01	-0.01	0.00	0.00	0.35	0.33	0.44	0.44	0.44	0.39	0.34	0.40
I37	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.28	0.32	0.27	0.24	0.39	0.35	0.19	0.23
I38	-0.01	0.01	0.00	0.02	0.01	0.01	0.00	-0.01	0.30	0.32	0.25	0.26	0.38	0.34	0.20	0.21
I39	0.01	-0.02	0.00	0.03	-0.02	-0.03	0.01	0.01	0.25	0.25	0.06	0.06	0.32	0.29	0.05	0.02
I40	0.00	0.00	0.00	-0.02	-0.01	0.00	0.00	0.00	0.23	0.24	0.05	0.06	0.33	0.27	0.04	0.05

Table C-2: Unweighted first factor patterns for 50% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.34	0.10	0.00	0.00	0.06	-0.01	0.34	-0.01	0.34	0.21	0.01	0.01	0.05	0.00	0.00	0.00
I2	0.35	0.11	-0.01	0.00	0.08	0.02	0.36	0.00	0.34	0.21	0.00	0.01	0.07	0.01	0.00	0.00
I3	0.43	0.13	-0.01	-0.01	0.27	0.04	0.42	-0.01	0.42	0.23	0.01	0.01	0.26	0.00	0.01	0.01
I4	0.42	0.13	0.00	-0.02	0.28	0.04	0.43	0.00	0.43	0.24	0.00	-0.01	0.26	0.01	0.00	0.00
I5	0.43	0.14	0.00	-0.02	0.44	0.09	0.46	0.00	0.45	0.27	0.00	0.01	0.45	0.02	0.00	0.00
I6	0.43	0.16	0.01	-0.01	0.43	0.09	0.46	0.00	0.46	0.27	0.00	-0.01	0.45	0.03	-0.01	0.00
I7	0.42	0.16	0.00	0.00	0.26	0.05	0.43	0.01	0.42	0.24	0.00	0.00	0.25	0.02	0.00	-0.01
I8	0.41	0.15	0.00	0.00	0.27	0.06	0.43	0.00	0.43	0.25	0.00	0.00	0.28	0.00	0.01	0.01
I9	0.33	0.13	-0.02	-0.01	0.06	0.02	0.35	0.00	0.35	0.20	0.00	0.00	0.07	0.00	0.00	-0.01
I10	0.34	0.13	0.00	0.00	0.07	0.03	0.36	0.01	0.35	0.21	0.01	0.00	0.05	0.01	0.00	0.01
I11	0.36	0.08	0.00	0.00	0.04	0.01	0.36	-0.01	0.35	0.19	0.01	0.01	0.06	0.01	0.00	0.00
I12	0.36	0.12	-0.01	0.00	0.05	0.03	0.35	0.00	0.34	0.20	0.01	0.00	0.07	0.00	-0.01	0.00
I13	0.41	0.13	0.00	-0.02	0.27	0.03	0.43	-0.01	0.42	0.24	0.01	0.01	0.27	0.01	-0.01	0.00
I14	0.42	0.12	-0.01	-0.01	0.24	0.03	0.44	-0.01	0.42	0.24	0.00	0.00	0.26	0.03	0.01	0.00
I15	0.43	0.15	0.01	-0.01	0.43	0.09	0.46	0.00	0.45	0.25	0.00	0.00	0.45	0.02	0.00	0.00
I16	0.44	0.14	0.01	-0.02	0.46	0.09	0.46	0.00	0.45	0.27	0.02	0.00	0.46	0.02	0.00	0.00
I17	0.40	0.15	0.00	0.00	0.26	0.07	0.43	0.01	0.41	0.24	0.00	0.00	0.26	0.01	0.00	0.00
I18	0.42	0.14	0.00	-0.02	0.26	0.05	0.42	0.01	0.41	0.24	0.00	0.00	0.25	0.01	-0.01	-0.02
I19	0.34	0.11	0.00	0.00	0.06	0.01	0.33	0.01	0.35	0.21	-0.01	-0.01	0.06	0.01	0.01	0.00
I20	0.34	0.12	0.00	0.00	0.06	0.01	0.34	0.00	0.34	0.20	0.00	0.00	0.06	0.00	0.00	0.00
I21	0.34	-0.02	-0.01	-0.41	0.05	-0.05	0.35	-0.41	0.34	0.19	-0.01	0.12	0.08	-0.01	0.00	0.16
I22	0.35	-0.01	0.01	-0.42	0.08	-0.07	0.36	-0.41	0.35	0.20	0.01	0.12	0.07	-0.01	0.00	0.16
I23	0.42	0.19	0.00	-0.28	0.26	0.17	0.43	-0.27	0.42	0.26	0.00	0.11	0.26	0.01	0.00	0.14
I24	0.41	0.18	-0.01	-0.29	0.26	0.17	0.43	-0.29	0.42	0.25	-0.01	0.11	0.26	0.00	0.00	0.15
I25	0.44	0.37	0.01	0.29	0.44	0.41	0.44	0.31	0.45	0.26	0.01	0.01	0.45	0.03	0.00	0.00
I26	0.44	0.36	0.01	0.29	0.45	0.39	0.46	0.30	0.44	0.25	0.02	-0.01	0.45	0.03	0.00	0.00
I27	0.41	0.50	0.01	0.72	0.26	0.57	0.42	0.72	0.43	0.18	0.00	-0.10	0.25	0.04	0.00	-0.14
I28	0.42	0.49	0.01	0.72	0.25	0.57	0.43	0.72	0.41	0.19	0.01	-0.10	0.25	0.05	0.01	-0.13
I29	0.34	0.54	0.00	0.79	0.07	0.67	0.36	0.79	0.35	0.08	-0.02	-0.12	0.08	0.05	0.00	-0.16
I30	0.34	0.56	0.01	0.78	0.08	0.66	0.35	0.79	0.36	0.08	0.01	-0.12	0.06	0.05	0.01	-0.16
I31	0.20	-0.01	-0.46	-0.41	0.20	-0.07	0.01	-0.41	0.22	0.20	0.16	0.12	0.20	-0.01	0.04	0.16
I32	0.20	-0.02	-0.45	-0.41	0.20	-0.07	0.01	-0.41	0.22	0.20	0.16	0.12	0.20	-0.01	0.04	0.16
I33	0.17	0.18	-0.32	-0.28	0.24	0.17	0.11	-0.27	0.36	0.26	0.15	0.11	0.35	0.00	0.05	0.15
I34	0.19	0.18	-0.29	-0.29	0.23	0.17	0.11	-0.28	0.37	0.26	0.15	0.11	0.35	0.01	0.04	0.15
I35	0.14	0.37	0.34	0.29	0.21	0.40	0.23	0.29	0.45	0.27	0.01	0.00	0.44	0.02	0.00	0.00
I36	0.13	0.35	0.33	0.29	0.20	0.40	0.24	0.30	0.45	0.26	0.01	-0.01	0.45	0.02	0.00	0.00
I37	0.05	0.49	0.75	0.72	0.13	0.57	0.04	0.72	0.37	0.18	-0.14	-0.10	0.38	0.04	-0.03	-0.13
I38	0.04	0.49	0.74	0.72	0.13	0.58	0.04	0.73	0.39	0.18	-0.13	-0.10	0.38	0.05	-0.03	-0.13
I39	-0.03	0.55	0.79	0.79	0.05	0.67	-0.02	0.79	0.23	0.07	-0.16	-0.12	0.24	0.04	-0.04	-0.16
I40	-0.04	0.55	0.79	0.79	0.05	0.66	-0.01	0.79	0.24	0.08	-0.16	-0.12	0.26	0.05	-0.04	-0.16

Table C-3: Unweighted first factor patterns for 20% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.35	0.33	-0.01	-0.02	0.06	0.01	0.34	-0.03	0.35	0.36	0.00	0.00	0.07	0.01	0.00	0.00
I2	0.34	0.34	0.01	-0.02	0.05	0.00	0.35	-0.02	0.33	0.35	0.00	0.00	0.06	0.02	0.01	0.00
I3	0.39	0.40	0.00	-0.02	0.23	0.07	0.43	-0.01	0.42	0.42	0.00	0.01	0.27	0.08	0.00	0.01
I4	0.41	0.40	0.00	-0.02	0.21	0.04	0.42	-0.02	0.41	0.42	-0.01	0.00	0.27	0.08	0.00	0.00
I5	0.42	0.41	-0.01	-0.03	0.39	0.15	0.45	-0.01	0.44	0.45	0.01	0.00	0.46	0.13	0.00	0.00
I6	0.41	0.41	0.01	-0.02	0.38	0.15	0.45	0.00	0.45	0.45	0.00	0.00	0.45	0.14	0.00	0.00
I7	0.36	0.37	0.01	0.00	0.20	0.12	0.43	-0.02	0.42	0.42	0.00	0.00	0.26	0.08	-0.01	0.01
I8	0.38	0.36	0.02	-0.01	0.22	0.11	0.43	-0.02	0.42	0.43	-0.01	0.01	0.25	0.09	0.00	0.00
I9	0.28	0.30	0.00	-0.01	0.04	0.03	0.35	-0.01	0.34	0.35	0.00	0.00	0.06	0.03	0.01	0.00
I10	0.30	0.29	-0.01	0.00	0.05	0.01	0.36	-0.01	0.34	0.35	0.00	0.00	0.09	0.01	0.00	0.00
I11	0.33	0.33	-0.01	0.00	0.05	0.01	0.34	-0.02	0.34	0.36	0.00	-0.01	0.07	0.03	-0.01	0.00
I12	0.36	0.34	0.00	-0.01	0.07	0.01	0.36	-0.03	0.34	0.36	0.00	-0.01	0.06	0.02	-0.01	0.00
I13	0.42	0.39	-0.01	-0.02	0.24	0.06	0.42	0.00	0.41	0.42	0.00	0.00	0.27	0.07	0.00	0.01
I14	0.39	0.41	-0.01	-0.02	0.25	0.05	0.41	-0.02	0.42	0.42	0.00	-0.01	0.26	0.07	0.00	0.01
I15	0.41	0.41	0.02	-0.02	0.38	0.16	0.46	-0.01	0.45	0.45	0.00	0.00	0.45	0.14	0.00	0.00
I16	0.41	0.41	0.00	-0.02	0.39	0.15	0.45	-0.01	0.45	0.46	0.00	0.01	0.44	0.14	0.00	0.00
I17	0.37	0.36	0.01	0.00	0.22	0.10	0.43	0.00	0.41	0.42	0.00	-0.01	0.27	0.08	-0.01	-0.01
I18	0.37	0.37	0.01	0.00	0.22	0.10	0.42	-0.01	0.42	0.42	0.00	0.00	0.27	0.09	0.00	0.00
I19	0.29	0.30	0.02	0.00	0.04	0.02	0.34	0.01	0.34	0.34	0.00	0.00	0.06	0.01	0.01	0.00
I20	0.30	0.30	0.00	0.02	0.05	0.03	0.35	0.00	0.34	0.35	0.00	0.00	0.06	0.02	-0.01	0.00
I21	0.36	0.30	-0.01	-0.50	0.06	-0.02	0.36	-0.49	0.34	0.29	0.01	-0.04	0.08	0.02	0.01	0.04
I22	0.33	0.30	-0.01	-0.49	0.06	-0.03	0.35	-0.49	0.35	0.27	0.01	-0.04	0.06	0.02	0.00	0.04
I23	0.41	0.31	-0.03	-0.27	0.22	0.18	0.42	-0.27	0.41	0.39	0.00	-0.02	0.26	0.08	0.00	0.04
I24	0.40	0.34	-0.03	-0.27	0.23	0.19	0.43	-0.28	0.41	0.40	0.00	-0.02	0.26	0.09	0.00	0.04
I25	0.42	0.27	0.01	0.21	0.39	0.38	0.45	0.20	0.45	0.44	0.00	0.00	0.44	0.13	0.00	0.00
I26	0.41	0.27	0.01	0.21	0.38	0.38	0.45	0.21	0.44	0.45	-0.01	0.01	0.45	0.13	0.00	0.01
I27	0.37	0.17	0.01	0.70	0.22	0.56	0.43	0.71	0.43	0.39	0.01	0.04	0.27	0.16	0.00	-0.02
I28	0.35	0.15	0.00	0.70	0.21	0.54	0.42	0.70	0.41	0.39	0.02	0.04	0.27	0.15	0.01	-0.03
I29	0.30	0.04	0.01	0.83	0.06	0.65	0.36	0.83	0.34	0.24	0.00	0.04	0.06	0.15	-0.01	-0.04
I30	0.30	0.03	0.01	0.84	0.05	0.64	0.36	0.84	0.34	0.24	0.00	0.04	0.06	0.14	0.00	-0.04
I31	0.31	0.30	-0.50	-0.49	0.25	-0.03	0.03	-0.48	0.27	0.28	0.08	-0.04	0.23	0.02	0.00	0.04
I32	0.32	0.30	-0.51	-0.49	0.25	-0.02	0.02	-0.47	0.26	0.27	0.08	-0.04	0.23	0.02	0.00	0.04
I33	0.28	0.33	-0.29	-0.28	0.34	0.18	0.17	-0.26	0.39	0.40	0.08	-0.03	0.37	0.10	0.01	0.04
I34	0.29	0.32	-0.29	-0.29	0.33	0.18	0.17	-0.27	0.39	0.40	0.08	-0.03	0.37	0.08	0.01	0.04
I35	0.20	0.27	0.25	0.20	0.33	0.38	0.37	0.21	0.45	0.44	0.00	0.00	0.45	0.14	0.00	0.00
I36	0.20	0.27	0.25	0.20	0.34	0.38	0.36	0.22	0.45	0.44	0.00	0.00	0.44	0.13	0.01	0.00
I37	0.08	0.16	0.73	0.70	0.30	0.54	0.12	0.69	0.39	0.40	-0.06	0.04	0.40	0.16	0.01	-0.03
I38	0.08	0.17	0.72	0.70	0.30	0.53	0.12	0.71	0.39	0.39	-0.06	0.04	0.39	0.16	0.01	-0.03
I39	-0.07	0.04	0.82	0.83	0.23	0.65	0.01	0.83	0.25	0.25	-0.08	0.04	0.27	0.14	0.00	-0.04
I40	-0.07	0.05	0.83	0.83	0.24	0.64	0.01	0.83	0.26	0.25	-0.08	0.04	0.28	0.14	0.00	-0.04

Table C-4: Unweighted first factor patterns for 5% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.28	0.34	0.07	-0.01	0.04	0.04	0.34	0.02	0.27	0.25	0.00	0.00	0.04	0.04	0.15	-0.01
I2	0.29	0.34	0.07	0.00	0.02	0.04	0.35	0.03	0.26	0.26	-0.01	0.01	0.06	0.05	0.15	0.00
I3	0.30	0.40	0.26	-0.02	0.20	0.15	0.42	0.04	0.32	0.32	0.00	0.01	0.23	0.21	0.17	0.00
I4	0.34	0.37	0.26	-0.01	0.19	0.14	0.42	0.05	0.32	0.31	-0.01	0.00	0.22	0.20	0.16	0.00
I5	0.33	0.37	0.42	0.02	0.41	0.34	0.46	0.05	0.35	0.33	0.00	0.01	0.39	0.36	0.18	0.00
I6	0.35	0.38	0.42	0.01	0.40	0.31	0.46	0.07	0.33	0.33	0.00	0.01	0.41	0.36	0.19	0.01
I7	0.30	0.32	0.23	0.02	0.23	0.21	0.42	0.07	0.32	0.32	0.00	-0.01	0.24	0.23	0.17	0.01
I8	0.31	0.33	0.26	0.03	0.23	0.21	0.42	0.04	0.31	0.33	-0.02	0.00	0.22	0.22	0.18	-0.01
I9	0.22	0.24	0.07	0.01	0.05	0.02	0.36	0.05	0.26	0.27	0.01	0.01	0.05	0.05	0.14	0.01
I10	0.23	0.24	0.05	0.00	0.08	0.04	0.34	0.05	0.26	0.27	0.00	0.00	0.06	0.04	0.14	0.01
I11	0.28	0.35	0.08	0.00	0.05	0.04	0.34	0.03	0.28	0.26	0.00	0.01	0.07	0.04	0.16	0.00
I12	0.26	0.33	0.06	-0.01	0.03	0.04	0.34	0.03	0.27	0.26	0.00	0.01	0.04	0.05	0.14	0.01
I13	0.31	0.37	0.25	-0.01	0.20	0.15	0.41	0.03	0.32	0.32	0.01	0.00	0.22	0.20	0.16	-0.01
I14	0.33	0.37	0.25	-0.02	0.21	0.16	0.42	0.04	0.31	0.33	0.00	0.00	0.23	0.19	0.17	0.00
I15	0.35	0.36	0.42	0.00	0.40	0.32	0.46	0.05	0.35	0.33	0.00	0.00	0.39	0.37	0.18	0.00
I16	0.34	0.36	0.41	0.02	0.39	0.32	0.45	0.05	0.35	0.35	0.00	0.00	0.40	0.37	0.17	0.00
I17	0.30	0.31	0.25	0.02	0.24	0.18	0.42	0.06	0.33	0.32	0.00	-0.01	0.24	0.21	0.17	0.00
I18	0.31	0.32	0.24	0.02	0.25	0.21	0.41	0.05	0.31	0.32	0.00	0.00	0.23	0.22	0.18	-0.01
I19	0.24	0.24	0.07	-0.01	0.03	0.01	0.36	0.05	0.27	0.27	0.02	0.01	0.02	0.05	0.14	0.00
I20	0.23	0.24	0.06	0.00	0.06	0.04	0.36	0.05	0.25	0.26	0.00	0.02	0.04	0.04	0.15	0.00
I21	0.25	0.35	0.06	-0.47	0.04	0.18	0.35	-0.41	0.26	0.18	0.01	-0.15	0.07	0.18	0.15	-0.04
I22	0.28	0.32	0.06	-0.47	0.04	0.17	0.35	-0.43	0.26	0.19	-0.01	-0.15	0.04	0.19	0.15	-0.03
I23	0.31	0.36	0.26	-0.16	0.21	0.31	0.42	-0.12	0.33	0.28	-0.01	-0.09	0.23	0.29	0.18	-0.01
I24	0.35	0.35	0.26	-0.18	0.17	0.29	0.41	-0.13	0.32	0.27	0.01	-0.10	0.23	0.29	0.16	-0.02
I25	0.34	0.30	0.43	0.14	0.40	0.38	0.44	0.16	0.35	0.34	0.00	0.00	0.40	0.36	0.18	0.01
I26	0.35	0.30	0.42	0.13	0.40	0.37	0.45	0.17	0.33	0.33	0.00	0.01	0.39	0.36	0.19	0.00
I27	0.31	0.22	0.25	0.54	0.26	0.40	0.42	0.52	0.32	0.33	0.01	0.11	0.23	0.36	0.18	0.04
I28	0.30	0.19	0.24	0.54	0.24	0.40	0.41	0.51	0.31	0.33	0.00	0.11	0.25	0.35	0.17	0.03
I29	0.24	0.06	0.07	0.82	0.07	0.40	0.36	0.75	0.25	0.28	-0.01	0.16	0.06	0.30	0.14	0.04
I30	0.23	0.06	0.07	0.82	0.07	0.42	0.36	0.74	0.26	0.27	0.01	0.16	0.04	0.30	0.13	0.04
I31	0.26	0.35	-0.03	-0.48	0.27	0.18	0.04	-0.41	0.22	0.19	-0.03	-0.15	0.28	0.20	-0.11	-0.03
I32	0.25	0.35	-0.04	-0.48	0.26	0.19	0.04	-0.42	0.20	0.18	-0.03	-0.15	0.28	0.20	-0.11	-0.03
I33	0.31	0.37	0.21	-0.18	0.37	0.29	0.23	-0.12	0.29	0.29	-0.01	-0.10	0.36	0.30	0.01	-0.02
I34	0.31	0.36	0.20	-0.17	0.37	0.28	0.23	-0.15	0.30	0.29	-0.01	-0.09	0.37	0.28	0.01	-0.01
I35	0.31	0.33	0.41	0.13	0.38	0.38	0.43	0.16	0.34	0.35	0.00	0.01	0.38	0.36	0.18	0.00
I36	0.28	0.30	0.42	0.14	0.42	0.37	0.43	0.17	0.34	0.34	0.00	0.01	0.39	0.36	0.17	0.01
I37	0.26	0.19	0.26	0.56	0.41	0.39	0.19	0.51	0.31	0.33	0.04	0.11	0.34	0.36	0.17	0.03
I38	0.28	0.20	0.25	0.54	0.40	0.41	0.18	0.51	0.29	0.33	0.03	0.11	0.36	0.35	0.17	0.03
I39	0.17	0.06	0.12	0.81	0.35	0.39	0.00	0.75	0.25	0.27	0.05	0.15	0.26	0.30	0.12	0.04
I40	0.15	0.05	0.12	0.83	0.35	0.40	0.00	0.73	0.24	0.29	0.04	0.16	0.25	0.29	0.12	0.04

Table C-5: Unweighted first second patterns for 100% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.05	0.03	0.08	0.06	0.07	0.04	0.07	-0.01	0.25	0.24	0.08	0.03	0.05	0.12	0.20	0.12
I2	0.07	0.01	0.01	0.08	0.07	0.03	0.00	0.00	0.26	0.27	0.05	0.08	0.05	0.09	0.21	0.17
I3	0.06	0.00	0.08	0.05	0.06	0.03	0.06	0.04	0.26	0.27	0.09	0.09	0.14	0.18	0.23	0.17
I4	0.07	0.02	0.10	0.03	0.08	0.05	0.08	0.05	0.27	0.28	0.10	0.05	0.13	0.18	0.22	0.17
I5	0.08	-0.04	0.07	0.06	0.08	0.08	0.06	0.00	0.28	0.29	0.14	0.15	0.18	0.23	0.28	0.19
I6	0.08	0.00	0.12	0.02	0.06	0.04	0.07	0.01	0.30	0.27	0.16	0.13	0.17	0.20	0.26	0.18
I7	0.11	0.01	0.01	0.05	0.07	0.02	0.02	-0.01	0.29	0.29	0.12	0.15	0.17	0.10	0.28	0.18
I8	0.02	-0.02	0.05	0.07	0.02	0.02	0.07	0.05	0.28	0.29	0.10	0.13	0.15	0.12	0.27	0.17
I9	0.08	0.02	0.01	0.08	0.05	0.03	0.03	0.03	0.22	0.26	0.21	0.08	0.08	-0.02	0.22	0.16
I10	0.09	0.03	0.05	0.00	0.04	0.05	0.04	-0.01	0.20	0.26	0.08	0.05	0.11	0.08	0.22	0.18
I11	0.05	0.01	0.07	0.01	0.08	0.07	0.05	0.03	0.24	0.25	0.01	0.01	0.00	0.11	0.22	0.14
I12	0.05	0.00	0.01	0.08	0.03	0.08	0.06	-0.05	0.27	0.23	0.03	0.04	0.09	0.06	0.21	0.11
I13	0.07	0.02	0.09	0.06	0.05	0.09	0.08	-0.02	0.28	0.26	0.12	0.08	0.08	0.20	0.24	0.20
I14	0.03	0.03	0.12	0.02	0.09	0.07	0.05	0.02	0.28	0.26	0.13	0.10	0.10	0.17	0.25	0.20
I15	0.06	0.00	0.11	0.06	0.04	0.04	0.04	0.03	0.28	0.29	0.15	0.17	0.17	0.20	0.25	0.19
I16	0.07	-0.01	0.10	0.08	0.07	0.03	0.06	0.00	0.31	0.30	0.16	0.14	0.16	0.22	0.24	0.15
I17	0.08	0.00	0.04	0.06	0.03	0.05	0.04	0.02	0.28	0.27	0.09	0.14	0.14	0.10	0.24	0.20
I18	0.08	0.00	0.04	0.04	0.02	0.00	0.04	0.01	0.28	0.29	0.11	0.13	0.15	0.10	0.24	0.15
I19	0.02	0.03	0.04	0.03	0.04	0.03	0.04	-0.01	0.26	0.25	0.10	0.08	0.11	0.06	0.24	0.17
I20	0.08	0.00	0.04	0.06	0.00	0.07	0.10	0.04	0.24	0.24	0.05	0.05	0.09	0.09	0.24	0.17
I21	0.02	0.09	0.07	0.06	0.06	0.02	0.08	0.03	0.26	0.27	0.05	0.07	0.01	0.14	0.18	0.09
I22	0.03	0.00	0.11	0.00	0.07	-0.03	0.08	0.04	0.24	0.26	0.02	0.06	0.06	0.17	0.18	0.08
I23	0.03	-0.02	0.12	0.06	0.08	-0.04	0.06	-0.03	0.27	0.27	0.08	0.17	0.10	0.20	0.23	0.10
I24	0.01	0.01	0.11	0.02	0.08	-0.01	0.06	0.01	0.28	0.27	0.07	0.12	0.10	0.19	0.23	0.15
I25	0.09	0.01	0.13	0.02	0.07	0.05	0.05	0.07	0.32	0.28	0.16	0.17	0.15	0.20	0.24	0.19
I26	0.07	0.02	0.10	0.02	0.04	0.04	0.06	0.03	0.30	0.31	0.14	0.16	0.16	0.21	0.28	0.18
I27	0.09	0.03	0.07	0.02	0.05	0.00	0.04	0.03	0.27	0.25	0.10	0.08	0.09	0.21	0.28	0.12
I28	0.07	0.03	0.03	0.04	0.05	0.01	0.01	0.02	0.25	0.25	0.10	0.06	0.14	0.25	0.27	0.12
I29	0.08	0.05	0.04	0.02	0.05	0.04	0.04	0.02	0.21	0.24	0.12	0.13	0.01	0.21	0.24	0.00
I30	0.11	0.04	0.10	0.01	0.06	0.00	0.03	0.05	0.22	0.25	0.06	0.05	0.09	0.21	0.22	0.07
I31	0.03	0.03	0.01	0.01	0.04	0.05	0.04	-0.01	0.21	0.26	0.08	0.06	0.14	0.14	0.06	0.13
I32	0.02	0.09	-0.01	-0.01	-0.01	0.08	0.05	0.02	0.25	0.24	0.07	0.07	0.15	0.14	0.08	0.10
I33	0.02	0.01	0.02	-0.05	0.00	0.04	0.02	0.02	0.26	0.29	0.10	0.17	0.15	0.20	0.17	0.12
I34	0.01	0.04	0.01	0.00	0.00	0.06	0.03	0.04	0.30	0.27	0.14	0.16	0.18	0.19	0.22	0.17
I35	0.05	0.03	-0.01	0.08	0.02	0.02	0.04	0.02	0.27	0.29	0.17	0.17	0.18	0.21	0.27	0.18
I36	0.06	0.05	0.07	-0.01	0.00	0.03	0.04	0.05	0.28	0.30	0.14	0.16	0.14	0.21	0.27	0.18
I37	0.02	-0.04	0.02	0.04	0.05	0.01	0.03	0.03	0.31	0.28	0.07	0.10	0.14	0.24	0.15	0.11
I38	0.01	0.07	0.00	-0.05	0.04	0.03	0.03	0.02	0.29	0.26	0.11	0.09	0.16	0.24	0.14	0.10
I39	-0.01	0.05	0.07	0.02	0.02	0.02	0.01	0.06	0.24	0.25	0.05	0.02	0.14	0.20	0.02	0.10
I40	0.00	0.03	-0.04	0.04	-0.02	-0.03	0.05	0.00	0.27	0.24	0.05	0.07	0.11	0.21	0.06	0.08

Table C-6: Unweighted first second patterns for 50% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.06	0.33	0.08	0.09	0.00	0.07	0.00	0.37	0.01	0.15	0.07	0.07	0.02	0.08	0.35	0.36
I2	0.04	0.30	0.08	0.09	0.01	0.09	-0.01	0.36	0.00	0.14	0.06	0.08	0.03	0.05	0.34	0.36
I3	0.08	0.37	0.29	0.32	0.04	0.30	0.02	0.44	-0.01	0.19	0.29	0.28	0.04	0.27	0.44	0.42
I4	0.08	0.38	0.28	0.31	0.04	0.30	0.00	0.44	0.02	0.18	0.28	0.29	0.04	0.27	0.42	0.44
I5	0.08	0.39	0.48	0.51	0.10	0.45	0.01	0.46	0.00	0.21	0.49	0.49	0.06	0.46	0.45	0.45
I6	0.10	0.38	0.48	0.50	0.11	0.46	0.02	0.46	-0.01	0.20	0.48	0.49	0.07	0.46	0.46	0.47
I7	0.08	0.37	0.28	0.31	0.08	0.27	0.02	0.44	0.00	0.18	0.30	0.30	0.03	0.28	0.42	0.43
I8	0.08	0.36	0.29	0.31	0.07	0.26	0.00	0.43	0.01	0.18	0.27	0.29	0.04	0.29	0.42	0.44
I9	0.09	0.31	0.06	0.08	0.02	0.06	0.01	0.36	0.00	0.16	0.06	0.08	0.01	0.07	0.35	0.36
I10	0.08	0.29	0.08	0.08	0.00	0.06	0.00	0.35	0.01	0.16	0.07	0.08	0.02	0.07	0.35	0.36
I11	0.06	0.31	0.07	0.10	0.03	0.08	0.00	0.37	-0.01	0.16	0.07	0.07	0.02	0.06	0.35	0.35
I12	0.05	0.30	0.08	0.07	0.00	0.11	0.00	0.36	0.01	0.14	0.08	0.07	0.02	0.06	0.34	0.35
I13	0.08	0.37	0.29	0.31	0.05	0.30	-0.02	0.44	-0.01	0.20	0.29	0.29	0.02	0.27	0.42	0.42
I14	0.08	0.38	0.30	0.29	0.07	0.31	0.01	0.44	0.02	0.18	0.27	0.28	0.03	0.26	0.42	0.44
I15	0.08	0.40	0.48	0.50	0.11	0.46	-0.01	0.46	0.01	0.19	0.47	0.49	0.06	0.47	0.45	0.47
I16	0.08	0.39	0.49	0.49	0.10	0.46	0.01	0.46	0.01	0.21	0.48	0.50	0.05	0.47	0.46	0.46
I17	0.09	0.36	0.29	0.30	0.07	0.27	0.00	0.43	0.01	0.19	0.29	0.30	0.02	0.25	0.43	0.43
I18	0.10	0.36	0.29	0.30	0.06	0.27	0.01	0.43	0.01	0.18	0.26	0.29	0.02	0.27	0.43	0.43
I19	0.07	0.31	0.06	0.08	0.01	0.06	0.01	0.37	0.01	0.15	0.09	0.07	0.00	0.05	0.35	0.37
I20	0.08	0.30	0.08	0.06	0.01	0.06	0.00	0.36	0.01	0.15	0.07	0.07	0.01	0.06	0.36	0.36
I21	0.06	0.18	0.07	0.02	0.01	0.27	0.00	0.02	0.00	0.02	0.07	0.01	0.01	0.24	0.35	0.01
I22	0.05	0.19	0.09	0.02	-0.02	0.27	0.00	0.00	0.01	0.02	0.07	0.01	0.01	0.24	0.35	0.01
I23	0.08	0.22	0.30	0.12	0.06	0.27	0.01	0.10	0.01	0.12	0.29	0.11	0.03	0.38	0.42	0.10
I24	0.06	0.21	0.30	0.10	0.08	0.25	-0.02	0.11	-0.02	0.12	0.29	0.11	0.03	0.40	0.43	0.11
I25	0.09	0.19	0.49	0.29	0.11	0.17	0.00	0.25	0.01	0.20	0.48	0.49	0.07	0.47	0.45	0.46
I26	0.08	0.20	0.48	0.29	0.10	0.19	0.00	0.26	0.00	0.20	0.48	0.49	0.08	0.47	0.46	0.46
I27	0.09	0.14	0.30	0.08	0.07	0.05	0.01	0.07	-0.01	0.22	0.28	0.11	0.04	0.38	0.43	0.11
I28	0.09	0.13	0.28	0.08	0.06	0.05	0.01	0.06	0.00	0.22	0.28	0.11	0.04	0.39	0.42	0.11
I29	0.08	0.07	0.07	0.02	0.02	-0.06	0.00	0.01	0.00	0.19	0.06	0.00	0.02	0.23	0.36	0.00
I30	0.08	0.07	0.06	0.02	0.03	-0.06	0.01	0.00	0.00	0.19	0.07	0.00	0.01	0.23	0.35	0.00
I31	-0.08	0.16	0.02	0.01	-0.04	0.27	-0.45	0.00	-0.11	0.02	0.00	0.01	0.05	0.24	0.01	0.01
I32	-0.07	0.17	0.02	0.02	-0.06	0.29	-0.46	0.01	-0.11	0.02	0.00	0.01	0.04	0.25	0.01	0.01
I33	0.19	0.19	0.11	0.12	0.20	0.27	-0.31	0.11	-0.07	0.12	0.10	0.12	0.06	0.39	0.12	0.11
I34	0.21	0.21	0.14	0.13	0.18	0.27	-0.32	0.11	-0.05	0.12	0.11	0.11	0.08	0.40	0.12	0.11
I35	0.44	0.20	0.26	0.29	0.40	0.20	0.32	0.24	0.00	0.20	0.48	0.49	0.06	0.47	0.44	0.46
I36	0.46	0.21	0.25	0.29	0.39	0.19	0.33	0.25	0.00	0.18	0.49	0.49	0.07	0.47	0.45	0.47
I37	0.61	0.14	0.06	0.08	0.55	0.06	0.75	0.06	0.09	0.23	0.12	0.10	0.06	0.38	0.12	0.11
I38	0.62	0.14	0.07	0.09	0.55	0.05	0.75	0.07	0.10	0.24	0.11	0.11	0.05	0.38	0.11	0.11
I39	0.69	0.08	0.00	0.01	0.62	-0.07	0.80	0.00	0.12	0.19	0.01	0.00	0.02	0.23	0.01	0.00
I40	0.69	0.06	0.01	0.01	0.60	-0.06	0.79	0.01	0.13	0.19	0.01	0.00	0.02	0.23	0.00	0.01

Table C-7: Unweighted first second patterns for 20% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.08	0.10	0.06	0.08	0.04	0.08	-0.01	0.37	0.01	0.01	0.07	0.08	0.03	0.06	0.35	0.36
I2	0.10	0.09	0.07	0.08	0.03	0.08	-0.01	0.36	0.01	0.02	0.07	0.06	0.03	0.03	0.34	0.37
I3	0.14	0.13	0.29	0.28	0.12	0.27	-0.02	0.43	0.01	0.01	0.27	0.29	0.06	0.23	0.43	0.43
I4	0.13	0.14	0.27	0.29	0.12	0.30	-0.01	0.43	-0.01	0.02	0.29	0.29	0.07	0.22	0.41	0.43
I5	0.17	0.15	0.48	0.48	0.21	0.42	-0.01	0.46	0.01	0.01	0.48	0.49	0.11	0.39	0.45	0.47
I6	0.15	0.15	0.47	0.48	0.21	0.43	0.00	0.46	0.01	0.02	0.48	0.48	0.12	0.39	0.45	0.46
I7	0.20	0.16	0.29	0.27	0.14	0.22	0.01	0.43	0.01	0.01	0.27	0.31	0.06	0.23	0.42	0.43
I8	0.18	0.16	0.28	0.29	0.14	0.22	0.00	0.43	0.02	0.01	0.29	0.29	0.05	0.22	0.43	0.43
I9	0.17	0.15	0.09	0.07	0.06	0.05	0.01	0.34	0.01	0.00	0.07	0.07	0.01	0.05	0.36	0.36
I10	0.17	0.15	0.06	0.06	0.05	0.05	0.00	0.35	0.03	0.02	0.08	0.06	0.02	0.05	0.35	0.37
I11	0.11	0.10	0.08	0.10	0.04	0.07	-0.03	0.36	0.00	0.01	0.06	0.07	0.03	0.05	0.34	0.37
I12	0.09	0.08	0.05	0.09	0.02	0.07	-0.02	0.36	0.01	0.01	0.08	0.07	0.01	0.05	0.35	0.37
I13	0.14	0.14	0.28	0.29	0.10	0.29	-0.02	0.42	0.00	0.02	0.30	0.29	0.07	0.23	0.43	0.44
I14	0.15	0.12	0.28	0.30	0.09	0.30	-0.01	0.44	0.00	0.00	0.28	0.29	0.05	0.22	0.42	0.43
I15	0.17	0.16	0.47	0.49	0.23	0.43	-0.01	0.46	0.04	0.01	0.49	0.48	0.10	0.39	0.47	0.47
I16	0.18	0.16	0.46	0.50	0.20	0.43	-0.01	0.46	0.00	0.02	0.47	0.48	0.10	0.39	0.46	0.46
I17	0.19	0.16	0.27	0.29	0.12	0.22	0.00	0.42	0.00	0.01	0.29	0.30	0.05	0.22	0.42	0.42
I18	0.19	0.16	0.27	0.27	0.13	0.23	0.00	0.43	0.00	0.03	0.27	0.29	0.05	0.23	0.42	0.43
I19	0.17	0.14	0.06	0.07	0.06	0.05	0.00	0.36	0.01	0.00	0.07	0.06	0.02	0.05	0.35	0.36
I20	0.16	0.12	0.07	0.05	0.03	0.05	0.01	0.36	0.02	0.00	0.07	0.05	0.01	0.06	0.35	0.35
I21	0.08	0.01	0.08	0.01	0.02	0.37	-0.03	0.01	0.03	-0.06	0.08	0.01	0.04	0.23	0.35	0.01
I22	0.09	0.00	0.08	0.01	0.04	0.37	-0.02	0.02	0.01	-0.06	0.07	0.01	0.00	0.23	0.34	0.01
I23	0.13	0.19	0.28	0.22	0.12	0.35	-0.01	0.21	0.01	-0.03	0.28	0.13	0.06	0.35	0.43	0.13
I24	0.14	0.16	0.28	0.19	0.12	0.36	-0.01	0.18	0.01	0.00	0.29	0.13	0.07	0.35	0.42	0.13
I25	0.17	0.36	0.47	0.42	0.22	0.26	-0.02	0.38	0.03	0.01	0.48	0.49	0.10	0.41	0.45	0.45
I26	0.17	0.36	0.46	0.43	0.21	0.25	-0.02	0.40	0.03	0.01	0.49	0.49	0.08	0.40	0.45	0.46
I27	0.18	0.51	0.28	0.15	0.13	0.12	-0.01	0.14	0.03	0.07	0.30	0.13	0.08	0.32	0.42	0.14
I28	0.20	0.52	0.27	0.15	0.15	0.12	0.01	0.15	0.01	0.06	0.28	0.13	0.05	0.33	0.42	0.13
I29	0.16	0.59	0.08	0.05	0.04	-0.04	0.01	0.03	0.01	0.09	0.07	0.00	0.01	0.20	0.35	0.01
I30	0.17	0.60	0.06	0.05	0.03	-0.03	0.00	0.02	0.01	0.09	0.06	0.01	0.04	0.22	0.36	0.01
I31	0.00	0.00	0.01	0.00	0.08	0.38	-0.51	0.01	-0.14	-0.06	0.01	0.01	0.13	0.23	0.01	0.01
I32	-0.01	0.00	0.00	0.02	0.08	0.38	-0.53	0.02	-0.14	-0.07	0.01	0.01	0.14	0.23	0.01	0.01
I33	0.23	0.17	0.19	0.20	0.18	0.36	-0.30	0.19	-0.07	-0.03	0.13	0.13	0.15	0.35	0.13	0.14
I34	0.22	0.19	0.19	0.19	0.18	0.36	-0.29	0.17	-0.06	-0.01	0.13	0.13	0.14	0.34	0.13	0.13
I35	0.44	0.36	0.39	0.43	0.30	0.26	0.24	0.38	0.00	0.03	0.47	0.49	0.11	0.39	0.45	0.46
I36	0.44	0.36	0.39	0.43	0.28	0.29	0.25	0.39	0.01	0.01	0.49	0.49	0.10	0.39	0.45	0.46
I37	0.59	0.51	0.13	0.15	0.34	0.11	0.73	0.14	0.12	0.06	0.13	0.12	0.04	0.33	0.13	0.13
I38	0.59	0.50	0.13	0.15	0.35	0.13	0.74	0.14	0.13	0.08	0.12	0.12	0.06	0.33	0.14	0.13
I39	0.67	0.59	0.03	0.04	0.37	-0.04	0.84	0.02	0.17	0.09	0.01	0.00	-0.02	0.20	0.01	0.01
I40	0.67	0.59	0.02	0.04	0.36	-0.04	0.83	0.03	0.18	0.09	0.01	0.01	-0.03	0.21	0.01	0.01

Table C-8: Unweighted first second patterns for 5% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.21	0.12	0.01	0.05	0.08	0.08	-0.01	0.33	0.22	0.23	0.07	0.07	0.04	0.05	0.19	0.37
I2	0.21	0.12	-0.03	0.06	0.15	0.06	0.01	0.32	0.23	0.22	0.06	0.06	0.02	0.03	0.20	0.38
I3	0.27	0.17	0.00	0.30	0.16	0.22	0.00	0.39	0.27	0.28	0.30	0.28	0.15	0.16	0.25	0.43
I4	0.25	0.19	0.00	0.29	0.17	0.21	0.00	0.39	0.26	0.28	0.29	0.29	0.14	0.17	0.25	0.43
I5	0.29	0.23	0.06	0.48	0.18	0.28	0.00	0.41	0.29	0.29	0.48	0.49	0.22	0.29	0.26	0.45
I6	0.28	0.23	0.06	0.48	0.18	0.29	0.00	0.41	0.29	0.30	0.47	0.48	0.22	0.29	0.26	0.46
I7	0.30	0.25	0.05	0.27	0.12	0.12	0.01	0.39	0.28	0.27	0.29	0.29	0.13	0.15	0.25	0.44
I8	0.26	0.24	0.05	0.28	0.10	0.12	0.01	0.37	0.26	0.26	0.29	0.29	0.16	0.15	0.26	0.42
I9	0.25	0.23	0.02	0.07	0.07	0.08	0.00	0.31	0.22	0.21	0.06	0.07	0.05	0.03	0.20	0.36
I10	0.25	0.24	0.03	0.10	0.07	0.04	0.00	0.32	0.21	0.21	0.07	0.07	0.02	0.02	0.20	0.36
I11	0.20	0.12	0.00	0.08	0.05	0.05	-0.01	0.33	0.22	0.23	0.06	0.07	0.03	0.05	0.21	0.36
I12	0.22	0.13	0.02	0.09	0.08	0.11	-0.02	0.33	0.23	0.22	0.08	0.07	0.06	0.06	0.20	0.37
I13	0.26	0.18	0.00	0.29	0.15	0.21	-0.02	0.39	0.26	0.28	0.28	0.29	0.14	0.17	0.24	0.43
I14	0.25	0.18	0.01	0.29	0.14	0.21	0.00	0.39	0.27	0.26	0.29	0.30	0.11	0.17	0.27	0.42
I15	0.28	0.24	0.05	0.48	0.18	0.30	0.01	0.42	0.27	0.29	0.47	0.48	0.24	0.28	0.26	0.46
I16	0.28	0.24	0.06	0.48	0.20	0.30	0.00	0.41	0.28	0.30	0.48	0.50	0.22	0.27	0.27	0.45
I17	0.28	0.27	0.04	0.27	0.08	0.15	0.02	0.39	0.26	0.25	0.29	0.29	0.11	0.16	0.25	0.43
I18	0.27	0.25	0.05	0.28	0.11	0.14	0.00	0.38	0.28	0.26	0.27	0.28	0.16	0.16	0.24	0.42
I19	0.24	0.24	0.02	0.07	0.09	0.09	0.02	0.31	0.22	0.20	0.08	0.05	0.06	0.05	0.20	0.37
I20	0.23	0.23	0.00	0.09	0.06	0.04	0.00	0.32	0.24	0.22	0.07	0.07	0.03	0.04	0.20	0.37
I21	0.23	0.06	0.00	0.07	0.13	0.28	-0.02	0.01	0.22	0.27	0.08	0.01	0.04	0.27	0.21	0.02
I22	0.21	0.09	-0.03	0.06	0.08	0.29	-0.01	0.00	0.23	0.27	0.07	0.02	0.05	0.29	0.21	0.02
I23	0.27	0.20	0.02	0.27	0.16	0.27	-0.01	0.21	0.26	0.32	0.29	0.20	0.17	0.31	0.24	0.19
I24	0.25	0.21	0.01	0.28	0.16	0.29	-0.01	0.20	0.28	0.31	0.28	0.20	0.14	0.31	0.25	0.19
I25	0.29	0.32	0.04	0.46	0.17	0.26	0.00	0.41	0.28	0.30	0.48	0.49	0.22	0.29	0.25	0.47
I26	0.27	0.34	0.05	0.45	0.18	0.26	0.00	0.40	0.28	0.29	0.48	0.48	0.23	0.27	0.26	0.45
I27	0.26	0.42	0.04	0.19	0.05	0.19	0.02	0.21	0.27	0.23	0.27	0.20	0.16	0.23	0.25	0.18
I28	0.28	0.41	0.04	0.18	0.06	0.20	-0.01	0.22	0.27	0.25	0.29	0.20	0.12	0.21	0.24	0.20
I29	0.23	0.49	0.00	-0.02	-0.01	0.12	0.02	0.09	0.23	0.14	0.06	0.02	0.05	0.12	0.20	0.01
I30	0.25	0.50	0.02	-0.01	0.02	0.10	0.01	0.08	0.22	0.15	0.06	0.02	0.08	0.11	0.22	0.01
I31	0.20	0.06	-0.46	0.07	0.17	0.26	-0.51	0.01	0.23	0.27	0.02	0.01	0.18	0.28	-0.06	0.02
I32	0.21	0.07	-0.42	0.07	0.17	0.27	-0.51	-0.01	0.21	0.27	0.01	0.01	0.17	0.28	-0.06	0.02
I33	0.25	0.21	-0.15	0.29	0.16	0.27	-0.21	0.21	0.30	0.30	0.21	0.20	0.23	0.31	0.07	0.19
I34	0.26	0.20	-0.16	0.27	0.15	0.28	-0.20	0.21	0.29	0.31	0.20	0.20	0.21	0.30	0.07	0.19
I35	0.32	0.31	0.17	0.45	0.20	0.25	0.13	0.41	0.27	0.29	0.47	0.49	0.22	0.27	0.26	0.47
I36	0.34	0.32	0.18	0.46	0.16	0.26	0.14	0.41	0.29	0.30	0.48	0.49	0.23	0.30	0.27	0.45
I37	0.34	0.43	0.54	0.17	0.13	0.20	0.59	0.22	0.27	0.24	0.20	0.20	0.21	0.22	0.17	0.18
I38	0.32	0.42	0.54	0.18	0.15	0.20	0.60	0.21	0.28	0.25	0.20	0.20	0.22	0.22	0.17	0.19
I39	0.31	0.48	0.72	0.00	0.10	0.12	0.82	0.08	0.19	0.14	0.01	0.02	0.17	0.12	0.08	0.01
I40	0.32	0.49	0.72	0.00	0.11	0.13	0.82	0.08	0.18	0.15	0.01	0.02	0.19	0.11	0.09	0.01

Table C-9: Rasch weighted first factor patterns for 100% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.37	0.38	0.06	0.08	0.04	0.06	0.35	0.38	0.25	0.24	0.05	0.06	0.04	0.03	0.30	0.33
I2	0.36	0.38	0.09	0.07	0.05	0.09	0.36	0.38	0.25	0.23	0.06	0.03	0.05	0.02	0.27	0.30
I3	0.42	0.44	0.26	0.31	0.29	0.34	0.42	0.44	0.33	0.32	0.25	0.26	0.21	0.18	0.33	0.37
I4	0.42	0.43	0.26	0.31	0.29	0.30	0.43	0.45	0.32	0.30	0.25	0.28	0.23	0.19	0.34	0.37
I5	0.43	0.46	0.47	0.50	0.48	0.50	0.43	0.46	0.34	0.34	0.45	0.43	0.40	0.39	0.34	0.39
I6	0.44	0.45	0.48	0.51	0.48	0.50	0.44	0.45	0.32	0.35	0.43	0.43	0.41	0.39	0.35	0.41
I7	0.40	0.39	0.30	0.31	0.29	0.30	0.39	0.38	0.30	0.31	0.23	0.23	0.21	0.22	0.32	0.37
I8	0.40	0.39	0.29	0.31	0.29	0.32	0.39	0.38	0.29	0.30	0.24	0.23	0.20	0.22	0.30	0.38
I9	0.29	0.24	0.06	0.10	0.05	0.10	0.31	0.26	0.26	0.23	0.05	0.04	0.05	0.07	0.26	0.32
I10	0.29	0.26	0.07	0.05	0.07	0.04	0.30	0.25	0.26	0.24	0.04	0.07	0.04	0.03	0.25	0.30
I11	0.35	0.38	0.05	0.08	0.06	0.06	0.36	0.38	0.24	0.22	0.08	0.06	0.06	0.03	0.27	0.31
I12	0.34	0.37	0.05	0.08	0.07	0.05	0.35	0.40	0.24	0.26	0.07	0.06	0.06	0.04	0.26	0.33
I13	0.41	0.44	0.26	0.30	0.29	0.30	0.41	0.45	0.31	0.32	0.24	0.26	0.25	0.17	0.35	0.37
I14	0.42	0.43	0.26	0.32	0.29	0.29	0.41	0.45	0.31	0.33	0.24	0.27	0.24	0.21	0.31	0.36
I15	0.44	0.45	0.47	0.50	0.50	0.50	0.44	0.44	0.35	0.33	0.43	0.44	0.41	0.40	0.36	0.41
I16	0.42	0.45	0.46	0.50	0.47	0.50	0.45	0.45	0.32	0.32	0.43	0.43	0.43	0.39	0.36	0.42
I17	0.38	0.39	0.31	0.30	0.27	0.29	0.39	0.37	0.31	0.31	0.24	0.23	0.24	0.25	0.34	0.37
I18	0.40	0.39	0.31	0.29	0.27	0.30	0.39	0.37	0.30	0.30	0.25	0.23	0.22	0.24	0.34	0.39
I19	0.30	0.27	0.05	0.05	0.07	0.08	0.29	0.26	0.22	0.26	0.05	0.06	0.04	0.02	0.26	0.29
I20	0.31	0.27	0.06	0.07	0.08	0.08	0.30	0.26	0.24	0.25	0.07	0.07	0.05	0.05	0.25	0.30
I21	0.36	0.01	0.05	0.02	0.07	0.00	0.35	0.00	0.24	0.24	0.04	0.05	0.04	0.32	0.29	0.03
I22	0.35	0.01	0.05	-0.01	0.05	0.01	0.35	0.01	0.25	0.23	0.06	0.04	0.06	0.29	0.28	0.04
I23	0.43	0.00	0.25	0.01	0.27	-0.01	0.42	0.01	0.32	0.32	0.26	0.23	0.23	0.36	0.36	0.24
I24	0.43	0.00	0.26	0.01	0.27	0.02	0.42	0.01	0.33	0.31	0.26	0.24	0.21	0.38	0.34	0.20
I25	0.43	0.00	0.45	0.01	0.48	0.01	0.45	0.01	0.32	0.36	0.42	0.43	0.43	0.40	0.36	0.40
I26	0.43	0.00	0.46	0.01	0.48	-0.03	0.43	0.02	0.33	0.32	0.44	0.42	0.42	0.38	0.35	0.41
I27	0.40	0.02	0.28	0.01	0.29	-0.01	0.40	-0.01	0.31	0.34	0.25	0.27	0.26	0.35	0.31	0.22
I28	0.40	0.01	0.29	-0.01	0.28	0.02	0.40	0.01	0.30	0.33	0.24	0.29	0.23	0.33	0.31	0.21
I29	0.32	0.01	0.07	-0.02	0.07	0.01	0.30	0.01	0.26	0.24	0.02	0.06	0.06	0.27	0.25	0.06
I30	0.29	0.02	0.08	0.00	0.08	0.01	0.29	-0.01	0.26	0.24	0.05	0.05	0.03	0.27	0.25	0.05
I31	0.00	0.01	-0.01	-0.02	0.00	-0.01	0.01	0.01	0.25	0.22	0.03	0.05	0.34	0.33	0.06	0.03
I32	0.00	-0.01	0.00	0.00	0.01	0.02	0.02	0.02	0.22	0.24	0.07	0.05	0.33	0.33	0.02	0.05
I33	0.01	0.01	0.00	0.01	0.00	0.00	0.02	-0.01	0.32	0.31	0.23	0.21	0.40	0.36	0.18	0.21
I34	0.02	0.02	0.01	0.00	-0.01	0.00	0.00	0.00	0.27	0.33	0.25	0.21	0.37	0.37	0.16	0.22
I35	0.02	0.02	0.00	0.01	0.00	-0.02	0.02	0.01	0.34	0.33	0.42	0.43	0.43	0.40	0.34	0.39
I36	0.02	0.01	0.00	0.00	-0.01	-0.01	0.00	0.01	0.34	0.33	0.44	0.44	0.44	0.39	0.34	0.41
I37	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.28	0.31	0.27	0.24	0.39	0.35	0.19	0.23
I38	0.00	0.02	0.00	0.01	0.01	0.01	0.01	0.01	0.30	0.32	0.25	0.26	0.38	0.34	0.20	0.21
I39	0.02	0.00	0.00	0.03	-0.02	-0.02	0.02	0.01	0.26	0.24	0.06	0.06	0.32	0.28	0.07	0.02
I40	0.01	0.01	0.00	-0.02	-0.01	0.01	0.01	0.02	0.24	0.24	0.05	0.06	0.33	0.26	0.06	0.05

Table C-10: Rasch weighted first factor patterns for 50% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.34	0.10	0.00	0.00	0.06	-0.01	0.35	0.08	0.34	0.17	0.01	0.04	0.05	0.00	0.00	0.23
I2	0.35	0.10	-0.01	0.00	0.08	0.02	0.36	0.09	0.35	0.17	0.00	0.07	0.07	0.01	0.00	0.23
I3	0.43	0.11	-0.01	-0.01	0.27	0.04	0.42	0.10	0.42	0.19	0.00	0.21	0.26	0.00	0.01	0.26
I4	0.42	0.11	0.00	-0.02	0.28	0.04	0.43	0.11	0.43	0.20	0.00	0.19	0.26	0.01	0.00	0.28
I5	0.43	0.12	0.00	-0.02	0.44	0.09	0.46	0.10	0.45	0.22	0.00	0.35	0.45	0.02	0.01	0.27
I6	0.43	0.14	0.01	-0.01	0.43	0.09	0.46	0.10	0.46	0.22	0.00	0.33	0.45	0.03	-0.01	0.30
I7	0.42	0.13	0.00	0.00	0.26	0.05	0.42	0.08	0.42	0.19	0.00	0.16	0.25	0.02	0.00	0.29
I8	0.40	0.12	0.00	0.00	0.27	0.06	0.42	0.07	0.42	0.21	0.00	0.18	0.28	0.00	0.02	0.27
I9	0.32	0.10	-0.02	-0.01	0.06	0.02	0.33	0.04	0.34	0.16	0.00	0.05	0.07	0.00	-0.01	0.26
I10	0.33	0.10	0.00	0.00	0.06	0.02	0.34	0.05	0.34	0.17	0.01	0.05	0.05	0.01	0.00	0.24
I11	0.36	0.07	0.00	0.00	0.05	0.01	0.36	0.09	0.36	0.16	0.00	0.06	0.06	0.01	0.00	0.20
I12	0.36	0.11	-0.01	0.00	0.05	0.03	0.36	0.08	0.34	0.16	0.01	0.05	0.07	0.00	-0.01	0.21
I13	0.41	0.11	0.00	-0.01	0.27	0.03	0.43	0.10	0.42	0.20	0.01	0.22	0.27	0.01	-0.01	0.26
I14	0.42	0.11	-0.01	-0.01	0.24	0.03	0.44	0.11	0.42	0.20	0.00	0.23	0.26	0.03	0.01	0.27
I15	0.43	0.13	0.01	-0.01	0.43	0.09	0.46	0.10	0.45	0.22	0.00	0.36	0.45	0.02	0.00	0.30
I16	0.44	0.12	0.01	-0.02	0.46	0.09	0.46	0.09	0.45	0.22	0.02	0.33	0.46	0.02	0.00	0.29
I17	0.40	0.13	0.00	0.00	0.26	0.07	0.42	0.08	0.41	0.20	0.00	0.19	0.26	0.01	0.00	0.28
I18	0.41	0.12	0.00	-0.02	0.26	0.05	0.42	0.08	0.41	0.19	0.00	0.17	0.25	0.01	-0.01	0.27
I19	0.33	0.08	0.00	0.00	0.05	0.01	0.32	0.05	0.35	0.16	-0.01	0.02	0.06	0.01	0.01	0.23
I20	0.33	0.10	0.00	0.00	0.05	0.01	0.33	0.04	0.34	0.16	0.00	0.02	0.06	0.00	0.00	0.25
I21	0.34	-0.01	-0.01	-0.41	0.05	-0.05	0.35	-0.29	0.34	0.18	-0.01	0.12	0.08	-0.01	0.00	-0.06
I22	0.35	-0.02	0.01	-0.42	0.08	-0.07	0.36	-0.30	0.35	0.19	0.01	0.11	0.07	-0.01	0.00	-0.04
I23	0.42	0.20	0.00	-0.28	0.26	0.17	0.43	-0.13	0.43	0.23	0.00	0.26	0.26	0.01	0.00	0.10
I24	0.41	0.19	-0.01	-0.29	0.26	0.17	0.43	-0.12	0.42	0.22	-0.01	0.23	0.26	0.00	0.00	0.08
I25	0.44	0.38	0.01	0.29	0.44	0.41	0.44	0.35	0.45	0.21	0.01	0.34	0.45	0.03	0.00	0.29
I26	0.44	0.38	0.01	0.30	0.45	0.39	0.45	0.34	0.44	0.21	0.02	0.34	0.45	0.03	0.01	0.30
I27	0.40	0.51	0.01	0.72	0.26	0.57	0.41	0.59	0.43	0.14	0.00	0.19	0.25	0.04	0.00	0.20
I28	0.41	0.51	0.01	0.72	0.25	0.57	0.42	0.59	0.41	0.15	0.01	0.14	0.25	0.05	0.01	0.16
I29	0.33	0.57	0.00	0.79	0.07	0.67	0.35	0.60	0.34	0.04	-0.02	-0.04	0.08	0.05	0.00	0.13
I30	0.33	0.58	0.01	0.78	0.08	0.66	0.34	0.59	0.35	0.04	0.01	-0.03	0.06	0.05	0.01	0.11
I31	0.20	0.00	-0.46	-0.41	0.20	-0.07	0.02	-0.29	0.22	0.20	0.20	0.15	0.20	-0.01	0.02	-0.05
I32	0.21	-0.02	-0.45	-0.41	0.20	-0.07	0.03	-0.29	0.22	0.19	0.20	0.10	0.20	-0.01	0.02	-0.03
I33	0.18	0.18	-0.32	-0.28	0.24	0.17	0.12	-0.12	0.36	0.23	0.19	0.27	0.35	0.00	0.03	0.11
I34	0.19	0.19	-0.29	-0.29	0.23	0.17	0.12	-0.13	0.37	0.24	0.19	0.25	0.35	0.01	0.02	0.14
I35	0.14	0.39	0.34	0.29	0.21	0.40	0.24	0.32	0.44	0.22	0.01	0.33	0.44	0.02	-0.01	0.29
I36	0.13	0.36	0.33	0.29	0.20	0.40	0.25	0.34	0.45	0.21	0.01	0.36	0.45	0.02	0.00	0.30
I37	0.05	0.51	0.75	0.72	0.13	0.57	0.04	0.59	0.37	0.13	-0.17	0.12	0.38	0.04	-0.01	0.18
I38	0.04	0.51	0.74	0.73	0.13	0.58	0.04	0.60	0.39	0.14	-0.16	0.15	0.38	0.05	-0.02	0.19
I39	-0.03	0.58	0.79	0.79	0.05	0.67	-0.02	0.59	0.23	0.03	-0.20	-0.01	0.24	0.04	-0.02	0.12
I40	-0.05	0.58	0.79	0.79	0.05	0.66	-0.01	0.59	0.24	0.04	-0.20	-0.03	0.26	0.05	-0.02	0.10

Table C-11: Rasch weighted first factor patterns for 20% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.35	0.35	-0.01	-0.01	0.06	0.01	0.34	0.36	0.35	0.36	0.00	0.04	0.07	0.01	0.05	0.30
I2	0.34	0.35	0.01	-0.01	0.05	0.00	0.35	0.35	0.33	0.35	0.00	0.05	0.06	0.02	0.05	0.30
I3	0.39	0.41	0.00	0.00	0.22	0.07	0.43	0.43	0.42	0.43	0.00	0.27	0.27	0.07	0.06	0.33
I4	0.41	0.41	0.00	-0.01	0.21	0.04	0.42	0.43	0.41	0.42	-0.01	0.26	0.27	0.08	0.05	0.34
I5	0.42	0.42	-0.01	-0.02	0.38	0.15	0.45	0.46	0.44	0.44	0.01	0.43	0.46	0.12	0.06	0.40
I6	0.41	0.42	0.01	-0.02	0.38	0.15	0.45	0.46	0.45	0.45	0.00	0.44	0.45	0.13	0.07	0.37
I7	0.36	0.38	0.00	-0.02	0.20	0.12	0.42	0.43	0.42	0.41	0.00	0.27	0.26	0.08	0.06	0.35
I8	0.38	0.37	0.02	-0.02	0.22	0.11	0.42	0.43	0.42	0.42	-0.01	0.23	0.25	0.08	0.06	0.37
I9	0.28	0.31	0.00	-0.02	0.04	0.03	0.34	0.34	0.34	0.33	0.00	0.05	0.06	0.02	0.05	0.27
I10	0.30	0.29	-0.02	0.00	0.05	0.01	0.35	0.35	0.34	0.33	0.00	0.01	0.09	0.01	0.05	0.28
I11	0.33	0.35	-0.01	0.01	0.05	0.01	0.35	0.36	0.34	0.36	0.00	0.05	0.07	0.03	0.04	0.31
I12	0.36	0.35	0.00	0.00	0.07	0.01	0.36	0.36	0.35	0.36	0.00	0.04	0.06	0.02	0.04	0.30
I13	0.42	0.40	-0.01	-0.01	0.23	0.06	0.42	0.42	0.41	0.42	0.00	0.24	0.27	0.07	0.06	0.35
I14	0.39	0.42	-0.01	0.00	0.25	0.05	0.41	0.43	0.42	0.43	0.00	0.26	0.26	0.07	0.05	0.36
I15	0.41	0.42	0.02	-0.01	0.38	0.16	0.46	0.46	0.44	0.44	0.00	0.43	0.45	0.13	0.06	0.40
I16	0.42	0.43	0.00	-0.01	0.38	0.15	0.44	0.45	0.45	0.45	0.00	0.43	0.44	0.13	0.06	0.37
I17	0.37	0.37	0.01	-0.02	0.22	0.10	0.42	0.42	0.41	0.41	0.00	0.26	0.27	0.08	0.06	0.37
I18	0.37	0.38	0.01	-0.03	0.22	0.10	0.42	0.43	0.42	0.41	0.00	0.26	0.27	0.08	0.06	0.36
I19	0.29	0.30	0.02	0.00	0.04	0.02	0.33	0.35	0.33	0.32	0.00	0.05	0.06	0.01	0.05	0.31
I20	0.30	0.30	0.00	0.00	0.05	0.03	0.34	0.35	0.34	0.33	0.00	0.05	0.06	0.02	0.04	0.28
I21	0.36	0.32	-0.01	-0.47	0.06	-0.02	0.36	0.07	0.34	0.38	0.01	0.07	0.08	0.03	0.05	0.03
I22	0.33	0.33	-0.01	-0.46	0.06	-0.03	0.35	0.08	0.35	0.36	0.01	0.04	0.06	0.03	0.05	0.00
I23	0.41	0.32	-0.03	-0.25	0.22	0.18	0.42	0.29	0.41	0.43	0.00	0.25	0.26	0.08	0.06	0.21
I24	0.40	0.35	-0.03	-0.25	0.23	0.19	0.43	0.25	0.41	0.44	0.00	0.24	0.26	0.09	0.05	0.21
I25	0.42	0.26	0.01	0.21	0.38	0.38	0.45	0.44	0.45	0.44	0.00	0.43	0.44	0.12	0.05	0.39
I26	0.41	0.26	0.01	0.21	0.37	0.38	0.45	0.45	0.44	0.44	-0.01	0.44	0.45	0.13	0.06	0.40
I27	0.37	0.15	0.01	0.69	0.22	0.56	0.42	0.22	0.42	0.34	0.01	0.22	0.27	0.14	0.04	0.22
I28	0.35	0.14	0.00	0.69	0.21	0.54	0.42	0.23	0.41	0.34	0.02	0.23	0.27	0.14	0.06	0.20
I29	0.30	0.01	0.01	0.83	0.06	0.65	0.35	-0.02	0.34	0.16	0.00	0.07	0.06	0.13	0.05	0.03
I30	0.30	0.01	0.01	0.83	0.05	0.64	0.35	-0.02	0.34	0.16	0.00	0.04	0.06	0.12	0.06	0.05
I31	0.31	0.32	-0.50	-0.46	0.25	-0.03	0.05	0.08	0.27	0.36	0.08	0.07	0.23	0.03	-0.04	0.05
I32	0.33	0.32	-0.51	-0.46	0.25	-0.02	0.05	0.08	0.27	0.36	0.08	0.05	0.23	0.03	-0.04	0.06
I33	0.28	0.33	-0.29	-0.26	0.34	0.18	0.19	0.27	0.39	0.44	0.08	0.26	0.37	0.10	0.01	0.21
I34	0.29	0.32	-0.29	-0.26	0.33	0.18	0.19	0.26	0.39	0.44	0.08	0.25	0.37	0.08	0.00	0.17
I35	0.20	0.26	0.25	0.20	0.33	0.38	0.38	0.43	0.45	0.44	0.00	0.44	0.45	0.13	0.06	0.38
I36	0.20	0.27	0.25	0.21	0.34	0.38	0.36	0.45	0.45	0.44	0.00	0.43	0.44	0.12	0.06	0.38
I37	0.08	0.14	0.73	0.69	0.30	0.54	0.11	0.23	0.38	0.35	-0.06	0.24	0.40	0.15	0.08	0.21
I38	0.08	0.15	0.72	0.70	0.30	0.53	0.11	0.22	0.39	0.34	-0.06	0.24	0.39	0.14	0.07	0.19
I39	-0.07	0.00	0.82	0.83	0.23	0.65	-0.01	0.00	0.24	0.17	-0.08	0.04	0.27	0.12	0.05	0.04
I40	-0.07	0.02	0.83	0.82	0.24	0.64	-0.01	0.01	0.25	0.17	-0.08	0.06	0.28	0.12	0.05	0.04

Table C-12: Rasch weighted first factor patterns for 5% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.28	0.33	0.07	0.04	0.04	0.05	0.34	0.34	0.27	0.28	0.01	0.05	0.04	0.05	0.32	0.30
I2	0.28	0.33	0.07	0.04	0.03	0.04	0.35	0.33	0.26	0.30	-0.01	0.08	0.06	0.05	0.34	0.30
I3	0.31	0.40	0.27	0.20	0.20	0.14	0.42	0.41	0.32	0.34	0.00	0.24	0.23	0.21	0.40	0.36
I4	0.35	0.36	0.26	0.20	0.18	0.14	0.42	0.42	0.32	0.32	-0.01	0.23	0.22	0.20	0.41	0.36
I5	0.33	0.37	0.43	0.32	0.41	0.33	0.46	0.42	0.35	0.33	-0.01	0.42	0.39	0.35	0.42	0.39
I6	0.35	0.38	0.43	0.32	0.40	0.31	0.46	0.43	0.33	0.33	0.00	0.43	0.41	0.36	0.44	0.40
I7	0.29	0.33	0.24	0.18	0.24	0.21	0.42	0.41	0.32	0.29	0.00	0.26	0.24	0.22	0.42	0.39
I8	0.30	0.33	0.27	0.19	0.24	0.21	0.42	0.38	0.31	0.29	-0.01	0.22	0.22	0.22	0.42	0.36
I9	0.22	0.25	0.08	0.04	0.05	0.02	0.36	0.31	0.26	0.21	0.00	0.04	0.05	0.05	0.34	0.33
I10	0.23	0.25	0.05	0.07	0.08	0.05	0.34	0.33	0.27	0.21	0.00	0.06	0.06	0.03	0.34	0.32
I11	0.29	0.34	0.08	0.05	0.05	0.04	0.34	0.35	0.28	0.31	0.00	0.04	0.07	0.04	0.34	0.28
I12	0.26	0.33	0.06	0.06	0.03	0.04	0.34	0.34	0.27	0.30	0.00	0.03	0.04	0.04	0.33	0.28
I13	0.31	0.37	0.26	0.20	0.19	0.14	0.41	0.41	0.32	0.36	0.01	0.24	0.22	0.19	0.41	0.36
I14	0.33	0.37	0.26	0.20	0.21	0.16	0.42	0.41	0.31	0.33	0.00	0.25	0.23	0.19	0.42	0.36
I15	0.35	0.37	0.43	0.31	0.40	0.32	0.46	0.44	0.35	0.33	0.00	0.43	0.39	0.36	0.45	0.41
I16	0.34	0.36	0.42	0.31	0.38	0.31	0.45	0.43	0.36	0.34	0.00	0.44	0.40	0.37	0.45	0.39
I17	0.30	0.31	0.26	0.18	0.24	0.18	0.42	0.40	0.32	0.28	0.00	0.25	0.24	0.21	0.40	0.40
I18	0.31	0.33	0.24	0.19	0.25	0.21	0.41	0.40	0.31	0.29	0.00	0.23	0.23	0.22	0.40	0.35
I19	0.23	0.25	0.08	0.06	0.04	0.02	0.36	0.32	0.27	0.22	0.01	0.03	0.02	0.05	0.35	0.33
I20	0.23	0.25	0.07	0.06	0.06	0.04	0.36	0.33	0.25	0.21	0.01	0.04	0.04	0.05	0.34	0.33
I21	0.26	0.34	0.06	-0.02	0.04	0.17	0.35	0.06	0.26	0.46	0.01	0.03	0.07	0.17	0.34	0.03
I22	0.29	0.31	0.06	-0.02	0.04	0.16	0.35	0.04	0.27	0.46	-0.01	0.05	0.04	0.18	0.35	0.03
I23	0.31	0.36	0.27	0.17	0.21	0.31	0.42	0.25	0.33	0.44	-0.02	0.25	0.23	0.29	0.39	0.20
I24	0.35	0.35	0.27	0.17	0.17	0.29	0.41	0.26	0.32	0.42	0.01	0.24	0.23	0.28	0.41	0.19
I25	0.33	0.30	0.43	0.34	0.40	0.38	0.44	0.42	0.35	0.34	0.00	0.44	0.40	0.36	0.43	0.43
I26	0.35	0.29	0.43	0.33	0.40	0.37	0.45	0.42	0.33	0.33	0.00	0.42	0.39	0.36	0.44	0.41
I27	0.32	0.22	0.26	0.30	0.26	0.40	0.42	0.21	0.32	0.19	0.00	0.27	0.23	0.36	0.41	0.22
I28	0.30	0.19	0.25	0.29	0.24	0.40	0.41	0.24	0.32	0.20	0.00	0.25	0.25	0.35	0.41	0.25
I29	0.23	0.06	0.07	0.24	0.07	0.40	0.36	0.04	0.25	0.00	-0.01	0.06	0.06	0.31	0.33	0.06
I30	0.21	0.07	0.07	0.24	0.08	0.42	0.35	0.02	0.26	-0.02	0.01	0.03	0.04	0.31	0.34	0.05
I31	0.27	0.35	-0.01	-0.03	0.26	0.18	0.07	0.07	0.21	0.47	-0.07	0.05	0.28	0.18	0.05	0.04
I32	0.26	0.34	-0.02	-0.01	0.26	0.18	0.07	0.07	0.19	0.47	-0.07	0.05	0.28	0.19	0.05	0.02
I33	0.30	0.37	0.22	0.18	0.37	0.29	0.24	0.26	0.28	0.43	-0.03	0.25	0.36	0.28	0.23	0.21
I34	0.31	0.36	0.21	0.18	0.37	0.28	0.25	0.25	0.29	0.44	-0.03	0.24	0.37	0.28	0.23	0.21
I35	0.31	0.34	0.41	0.33	0.38	0.38	0.43	0.43	0.34	0.34	0.00	0.43	0.38	0.36	0.42	0.40
I36	0.28	0.30	0.43	0.34	0.42	0.37	0.43	0.43	0.34	0.34	0.00	0.44	0.39	0.36	0.43	0.40
I37	0.27	0.20	0.25	0.29	0.41	0.39	0.18	0.22	0.32	0.19	0.07	0.26	0.34	0.36	0.24	0.22
I38	0.27	0.21	0.25	0.28	0.40	0.41	0.18	0.22	0.29	0.20	0.06	0.25	0.36	0.35	0.23	0.21
I39	0.16	0.07	0.10	0.26	0.35	0.39	-0.03	0.02	0.26	-0.01	0.09	0.07	0.26	0.31	0.03	0.05
I40	0.14	0.05	0.10	0.28	0.35	0.40	-0.02	0.05	0.26	0.00	0.08	0.07	0.25	0.30	0.02	0.05

Table C-13: Rasch weighted second factor patterns for 100% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.03	0.02	0.08	0.04	0.07	0.06	0.04	-0.02	0.25	0.24	0.08	0.03	0.05	0.12	0.20	0.12
I2	0.06	0.06	0.01	0.09	0.07	0.05	0.01	0.01	0.25	0.27	0.05	0.08	0.05	0.10	0.21	0.17
I3	0.04	0.02	0.07	0.07	0.06	0.02	0.08	0.01	0.26	0.27	0.09	0.09	0.14	0.18	0.24	0.17
I4	0.06	0.07	0.09	0.04	0.08	0.04	0.06	0.03	0.27	0.28	0.10	0.05	0.13	0.18	0.22	0.18
I5	0.06	-0.01	0.07	0.03	0.08	0.05	0.06	-0.03	0.28	0.28	0.14	0.15	0.18	0.23	0.27	0.20
I6	0.07	-0.01	0.11	0.01	0.07	0.04	0.04	0.00	0.29	0.27	0.16	0.13	0.17	0.20	0.26	0.18
I7	0.06	0.00	0.02	0.03	0.06	0.01	0.02	0.00	0.30	0.28	0.12	0.15	0.17	0.10	0.28	0.18
I8	0.01	-0.01	0.05	0.05	0.02	0.02	0.06	0.05	0.29	0.29	0.10	0.13	0.15	0.12	0.28	0.17
I9	0.07	0.04	0.01	0.02	0.05	-0.01	0.04	0.02	0.22	0.26	0.21	0.08	0.08	-0.02	0.22	0.15
I10	0.11	0.02	0.05	-0.05	0.04	0.05	0.07	-0.02	0.21	0.24	0.08	0.05	0.11	0.08	0.22	0.18
I11	0.10	0.06	0.07	0.03	0.08	0.07	0.01	0.00	0.23	0.26	0.01	0.01	0.00	0.10	0.21	0.15
I12	0.05	-0.01	0.01	0.07	0.03	0.08	0.08	-0.03	0.25	0.23	0.03	0.04	0.09	0.05	0.22	0.11
I13	0.08	0.06	0.09	0.06	0.05	0.09	0.07	-0.01	0.27	0.26	0.12	0.08	0.08	0.20	0.24	0.20
I14	0.05	0.04	0.11	0.03	0.09	0.05	0.05	0.01	0.27	0.27	0.13	0.10	0.10	0.17	0.26	0.20
I15	0.06	0.02	0.10	0.03	0.04	0.04	0.03	0.07	0.28	0.30	0.15	0.17	0.17	0.20	0.25	0.19
I16	0.06	0.01	0.10	0.06	0.07	0.05	0.06	0.01	0.32	0.30	0.16	0.14	0.16	0.22	0.24	0.15
I17	0.05	-0.02	0.03	0.02	0.03	0.06	0.03	0.00	0.28	0.27	0.09	0.14	0.14	0.09	0.24	0.19
I18	0.01	0.01	0.04	0.01	0.02	0.03	0.07	0.02	0.28	0.30	0.11	0.13	0.15	0.09	0.24	0.15
I19	0.05	0.05	0.04	0.02	0.03	-0.01	0.06	-0.01	0.27	0.24	0.10	0.08	0.11	0.06	0.24	0.17
I20	0.06	0.03	0.04	0.04	-0.01	0.05	0.07	0.02	0.25	0.23	0.05	0.05	0.09	0.09	0.24	0.17
I21	0.03	0.03	0.07	0.09	0.05	0.01	0.05	0.02	0.25	0.26	0.05	0.07	0.01	0.14	0.19	0.07
I22	0.03	0.00	0.11	0.01	0.08	-0.02	0.05	0.05	0.23	0.25	0.02	0.06	0.06	0.17	0.19	0.08
I23	0.01	-0.02	0.11	0.04	0.07	-0.01	0.06	-0.05	0.27	0.27	0.08	0.17	0.10	0.21	0.22	0.09
I24	0.05	0.03	0.11	0.01	0.08	-0.01	0.08	0.00	0.27	0.27	0.07	0.12	0.10	0.19	0.23	0.15
I25	0.09	0.02	0.12	0.03	0.07	0.05	0.03	0.05	0.31	0.27	0.16	0.17	0.15	0.20	0.24	0.19
I26	0.06	0.06	0.09	0.03	0.03	0.04	0.09	0.03	0.30	0.31	0.14	0.16	0.16	0.21	0.28	0.18
I27	0.06	0.01	0.07	0.03	0.06	-0.01	0.08	0.01	0.27	0.25	0.10	0.08	0.09	0.21	0.28	0.13
I28	0.05	0.01	0.02	0.05	0.05	0.00	0.01	0.01	0.27	0.26	0.10	0.06	0.14	0.25	0.27	0.13
I29	0.01	0.06	0.04	0.02	0.05	0.08	0.04	0.03	0.21	0.25	0.12	0.13	0.01	0.21	0.24	-0.01
I30	0.10	0.05	0.08	0.04	0.05	0.00	0.06	0.05	0.22	0.25	0.06	0.05	0.09	0.21	0.22	0.08
I31	0.06	0.04	0.01	0.00	0.03	0.06	0.03	-0.02	0.22	0.26	0.08	0.06	0.14	0.14	0.05	0.12
I32	0.04	0.03	-0.01	0.02	-0.01	0.06	0.05	0.08	0.26	0.24	0.07	0.07	0.15	0.14	0.07	0.11
I33	0.06	0.06	0.03	-0.04	0.00	0.06	0.06	0.00	0.26	0.29	0.10	0.17	0.15	0.20	0.17	0.13
I34	-0.01	0.04	0.01	0.02	0.00	0.08	0.08	0.01	0.31	0.27	0.14	0.16	0.18	0.19	0.21	0.16
I35	0.04	0.02	-0.01	0.08	0.02	0.04	0.02	0.01	0.27	0.28	0.17	0.17	0.18	0.21	0.27	0.18
I36	0.06	0.02	0.08	-0.02	0.00	0.00	0.09	0.01	0.28	0.30	0.14	0.16	0.14	0.21	0.26	0.17
I37	0.02	0.01	0.02	0.02	0.05	0.03	0.05	0.05	0.31	0.28	0.07	0.10	0.14	0.24	0.16	0.11
I38	0.06	0.04	0.00	-0.07	0.04	0.01	0.06	0.02	0.29	0.27	0.11	0.09	0.16	0.24	0.14	0.11
I39	-0.02	0.02	0.07	0.02	0.02	0.05	0.03	0.06	0.24	0.26	0.05	0.02	0.14	0.20	0.02	0.09
I40	0.05	0.02	-0.04	0.02	-0.02	-0.02	0.05	0.06	0.26	0.25	0.05	0.07	0.11	0.21	0.04	0.08

Table C-14: Rasch weighted second factor patterns for 50% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.06	0.34	0.08	0.09	0.00	0.07	0.00	0.28	0.01	0.19	0.07	0.06	0.02	0.08	0.35	0.19
I2	0.04	0.32	0.08	0.09	0.01	0.09	-0.01	0.29	0.00	0.18	0.06	0.01	0.03	0.05	0.35	0.20
I3	0.08	0.39	0.29	0.32	0.04	0.30	0.02	0.34	-0.01	0.23	0.29	0.11	0.04	0.27	0.44	0.24
I4	0.08	0.39	0.28	0.31	0.04	0.30	0.00	0.34	0.02	0.22	0.28	0.18	0.04	0.27	0.42	0.25
I5	0.08	0.39	0.48	0.51	0.10	0.45	0.00	0.34	0.00	0.25	0.49	0.22	0.06	0.46	0.45	0.29
I6	0.10	0.39	0.48	0.50	0.11	0.46	0.01	0.33	-0.01	0.24	0.48	0.24	0.07	0.46	0.46	0.27
I7	0.08	0.37	0.28	0.31	0.08	0.27	0.01	0.32	0.00	0.22	0.29	0.20	0.03	0.28	0.42	0.24
I8	0.08	0.36	0.28	0.31	0.07	0.26	-0.01	0.30	0.01	0.22	0.27	0.17	0.04	0.29	0.42	0.27
I9	0.08	0.30	0.05	0.07	0.02	0.06	-0.01	0.24	0.00	0.19	0.06	0.07	0.01	0.07	0.34	0.16
I10	0.08	0.29	0.08	0.08	0.00	0.06	-0.02	0.24	0.01	0.19	0.06	0.04	0.02	0.07	0.35	0.20
I11	0.06	0.33	0.07	0.10	0.03	0.08	0.00	0.28	-0.01	0.19	0.07	0.03	0.02	0.06	0.36	0.23
I12	0.05	0.32	0.08	0.08	0.00	0.11	0.00	0.29	0.01	0.18	0.09	0.06	0.02	0.06	0.35	0.24
I13	0.08	0.39	0.29	0.31	0.05	0.30	-0.02	0.33	-0.01	0.24	0.29	0.11	0.02	0.27	0.42	0.23
I14	0.08	0.39	0.30	0.29	0.07	0.32	0.00	0.33	0.02	0.22	0.27	0.12	0.03	0.26	0.42	0.24
I15	0.09	0.41	0.48	0.50	0.11	0.46	-0.01	0.34	0.01	0.23	0.47	0.22	0.06	0.47	0.45	0.26
I16	0.08	0.40	0.49	0.49	0.10	0.46	0.01	0.34	0.01	0.24	0.48	0.26	0.05	0.47	0.46	0.25
I17	0.09	0.37	0.29	0.29	0.07	0.27	-0.01	0.31	0.01	0.23	0.28	0.17	0.02	0.25	0.43	0.23
I18	0.10	0.36	0.29	0.30	0.06	0.27	0.00	0.30	0.01	0.23	0.26	0.18	0.02	0.27	0.42	0.25
I19	0.07	0.30	0.06	0.07	0.01	0.06	0.00	0.24	0.01	0.18	0.08	0.08	0.00	0.05	0.34	0.21
I20	0.08	0.29	0.08	0.06	0.01	0.05	-0.02	0.24	0.01	0.18	0.07	0.06	0.01	0.06	0.34	0.18
I21	0.06	0.20	0.07	0.02	0.01	0.27	0.00	0.00	0.00	0.03	0.07	0.06	0.01	0.24	0.35	0.03
I22	0.05	0.21	0.09	0.02	-0.02	0.27	0.00	-0.01	0.01	0.03	0.07	0.08	0.01	0.24	0.36	0.04
I23	0.08	0.23	0.30	0.12	0.06	0.27	0.01	0.11	0.01	0.14	0.29	0.11	0.03	0.38	0.42	0.17
I24	0.07	0.23	0.30	0.11	0.08	0.25	-0.02	0.12	-0.02	0.14	0.29	0.12	0.03	0.40	0.43	0.19
I25	0.09	0.19	0.49	0.29	0.11	0.17	-0.01	0.24	0.01	0.24	0.48	0.24	0.07	0.47	0.45	0.26
I26	0.09	0.19	0.48	0.30	0.10	0.19	-0.01	0.25	0.00	0.24	0.48	0.23	0.08	0.47	0.45	0.26
I27	0.09	0.12	0.30	0.08	0.07	0.05	0.00	0.12	-0.01	0.27	0.28	0.11	0.04	0.38	0.42	0.13
I28	0.09	0.11	0.27	0.08	0.06	0.05	0.00	0.12	0.00	0.26	0.28	0.18	0.04	0.39	0.42	0.17
I29	0.08	0.03	0.06	0.02	0.02	-0.06	-0.01	0.10	0.00	0.23	0.06	0.10	0.02	0.23	0.34	0.07
I30	0.08	0.03	0.06	0.02	0.03	-0.06	0.00	0.11	0.00	0.23	0.07	0.06	0.01	0.23	0.33	0.03
I31	-0.07	0.19	0.02	0.01	-0.04	0.27	-0.45	-0.01	-0.11	0.02	0.00	0.06	0.05	0.24	0.01	0.07
I32	-0.06	0.19	0.02	0.02	-0.06	0.29	-0.45	0.00	-0.11	0.03	0.00	0.05	0.04	0.24	0.01	0.08
I33	0.20	0.21	0.11	0.12	0.20	0.27	-0.31	0.11	-0.07	0.15	0.10	0.09	0.06	0.39	0.12	0.19
I34	0.21	0.23	0.14	0.13	0.18	0.27	-0.31	0.11	-0.05	0.14	0.11	0.12	0.08	0.40	0.12	0.18
I35	0.45	0.19	0.26	0.29	0.40	0.20	0.33	0.24	0.00	0.25	0.48	0.26	0.06	0.47	0.44	0.25
I36	0.46	0.20	0.25	0.29	0.39	0.19	0.33	0.25	0.00	0.23	0.49	0.22	0.07	0.47	0.45	0.26
I37	0.61	0.12	0.06	0.08	0.56	0.06	0.75	0.13	0.09	0.28	0.12	0.19	0.06	0.38	0.12	0.14
I38	0.62	0.11	0.07	0.09	0.55	0.05	0.75	0.13	0.10	0.28	0.11	0.14	0.05	0.39	0.11	0.16
I39	0.70	0.04	0.00	0.01	0.62	-0.07	0.80	0.11	0.12	0.22	0.01	0.06	0.02	0.23	0.01	0.03
I40	0.69	0.03	0.01	0.01	0.60	-0.06	0.79	0.12	0.13	0.23	0.01	0.01	0.02	0.23	0.00	0.08

Table C-15: Rasch weighted second factor patterns for 20% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.08	0.09	0.06	0.08	0.04	0.08	-0.01	-0.02	0.01	0.01	0.07	0.05	0.03	0.06	0.30	0.18
I2	0.10	0.08	0.07	0.08	0.03	0.08	0.00	0.01	0.00	0.03	0.07	0.04	0.03	0.03	0.29	0.19
I3	0.14	0.11	0.29	0.28	0.12	0.27	-0.01	0.00	0.01	0.02	0.27	0.07	0.06	0.24	0.36	0.24
I4	0.13	0.13	0.27	0.29	0.13	0.30	0.00	-0.01	-0.01	0.03	0.29	0.09	0.07	0.23	0.36	0.21
I5	0.17	0.13	0.48	0.48	0.21	0.42	-0.01	0.02	0.01	0.02	0.48	0.15	0.11	0.40	0.38	0.20
I6	0.15	0.13	0.47	0.48	0.21	0.43	-0.01	0.00	0.01	0.03	0.48	0.12	0.12	0.40	0.38	0.23
I7	0.19	0.13	0.29	0.27	0.13	0.22	-0.01	0.02	0.01	0.03	0.27	0.13	0.06	0.23	0.35	0.21
I8	0.18	0.14	0.28	0.29	0.13	0.22	-0.01	0.03	0.02	0.03	0.29	0.14	0.05	0.22	0.35	0.18
I9	0.17	0.12	0.08	0.07	0.05	0.05	-0.02	0.02	0.01	0.00	0.07	0.06	0.01	0.05	0.29	0.18
I10	0.16	0.12	0.06	0.06	0.05	0.06	-0.03	0.05	0.02	0.01	0.08	0.08	0.02	0.05	0.29	0.23
I11	0.11	0.09	0.08	0.10	0.04	0.07	-0.02	0.00	-0.01	0.01	0.06	0.06	0.03	0.05	0.30	0.17
I12	0.09	0.08	0.05	0.09	0.02	0.07	-0.02	0.00	0.01	0.01	0.08	0.10	0.01	0.04	0.31	0.18
I13	0.14	0.13	0.28	0.29	0.10	0.29	-0.02	0.01	0.00	0.04	0.30	0.13	0.07	0.24	0.37	0.24
I14	0.15	0.11	0.28	0.30	0.10	0.30	0.00	0.01	0.00	0.02	0.28	0.08	0.05	0.22	0.37	0.20
I15	0.17	0.14	0.47	0.49	0.23	0.43	-0.01	0.02	0.04	0.04	0.49	0.15	0.10	0.40	0.39	0.20
I16	0.17	0.13	0.46	0.50	0.20	0.43	-0.02	0.03	0.00	0.02	0.47	0.16	0.10	0.40	0.38	0.24
I17	0.18	0.13	0.27	0.28	0.12	0.22	-0.01	0.03	0.00	0.01	0.29	0.12	0.05	0.23	0.34	0.17
I18	0.18	0.14	0.27	0.27	0.12	0.23	-0.02	0.04	0.00	0.04	0.27	0.10	0.05	0.24	0.35	0.19
I19	0.16	0.11	0.06	0.07	0.06	0.05	-0.02	0.04	0.01	0.01	0.07	0.04	0.02	0.05	0.28	0.18
I20	0.15	0.09	0.07	0.04	0.03	0.05	-0.01	0.03	0.02	0.00	0.07	0.00	0.01	0.06	0.29	0.17
I21	0.08	0.02	0.08	0.03	0.03	0.37	-0.02	-0.09	0.03	-0.41	0.08	0.04	0.04	0.23	0.31	0.08
I22	0.09	0.00	0.08	0.03	0.04	0.37	-0.01	-0.20	0.00	-0.41	0.07	0.10	0.00	0.23	0.29	0.10
I23	0.13	0.20	0.28	0.23	0.12	0.35	-0.01	-0.04	0.01	-0.22	0.28	0.11	0.06	0.35	0.37	0.15
I24	0.14	0.17	0.28	0.20	0.12	0.36	-0.01	-0.04	0.01	-0.21	0.29	0.11	0.07	0.35	0.37	0.14
I25	0.17	0.36	0.47	0.42	0.22	0.26	-0.02	0.12	0.03	0.02	0.48	0.17	0.10	0.42	0.38	0.20
I26	0.17	0.37	0.46	0.43	0.21	0.25	-0.02	0.09	0.03	0.04	0.49	0.14	0.08	0.41	0.38	0.21
I27	0.18	0.52	0.28	0.14	0.13	0.12	-0.03	0.29	0.03	0.29	0.30	0.13	0.08	0.34	0.35	0.16
I28	0.19	0.53	0.27	0.15	0.15	0.12	-0.01	0.28	0.01	0.29	0.28	0.13	0.05	0.35	0.35	0.16
I29	0.15	0.61	0.08	0.03	0.04	-0.03	-0.02	0.53	0.01	0.47	0.07	0.07	0.01	0.21	0.28	0.09
I30	0.16	0.61	0.06	0.03	0.02	-0.03	-0.02	0.45	0.01	0.48	0.06	0.07	0.04	0.23	0.29	0.07
I31	0.00	0.01	0.01	0.02	0.09	0.38	-0.49	-0.12	-0.16	-0.42	0.01	0.03	0.13	0.23	0.10	0.06
I32	0.00	0.00	0.00	0.03	0.08	0.38	-0.50	-0.14	-0.16	-0.42	0.01	0.09	0.14	0.22	0.10	0.11
I33	0.24	0.18	0.19	0.22	0.18	0.36	-0.27	-0.05	-0.07	-0.21	0.13	0.09	0.15	0.35	0.18	0.16
I34	0.22	0.19	0.19	0.21	0.19	0.36	-0.26	-0.06	-0.08	-0.21	0.13	0.12	0.14	0.34	0.19	0.18
I35	0.44	0.37	0.39	0.44	0.30	0.26	0.25	0.11	0.00	0.04	0.47	0.12	0.11	0.40	0.38	0.21
I36	0.44	0.36	0.39	0.43	0.28	0.29	0.26	0.10	0.01	0.03	0.49	0.17	0.10	0.39	0.38	0.22
I37	0.59	0.52	0.13	0.14	0.34	0.11	0.73	0.26	0.13	0.28	0.13	0.10	0.04	0.34	0.06	0.14
I38	0.59	0.51	0.13	0.14	0.35	0.14	0.73	0.29	0.14	0.30	0.12	0.07	0.06	0.35	0.07	0.14
I39	0.67	0.61	0.03	0.02	0.36	-0.04	0.83	0.46	0.19	0.47	0.01	0.07	-0.02	0.22	-0.08	0.05
I40	0.67	0.60	0.02	0.03	0.36	-0.04	0.83	0.46	0.19	0.47	0.01	0.08	-0.03	0.23	-0.08	0.04

Table C-16: Rasch weighted second factor patterns for 5% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.20	0.13	0.01	0.01	0.07	0.08	0.00	0.08	0.22	0.20	0.07	0.06	0.04	0.04	0.04	0.19
I2	0.21	0.12	-0.03	0.03	0.15	0.06	0.02	0.06	0.23	0.18	0.06	0.01	0.02	0.03	0.05	0.19
I3	0.26	0.17	0.00	0.11	0.16	0.22	0.00	0.07	0.27	0.25	0.30	0.10	0.15	0.16	0.06	0.20
I4	0.25	0.19	0.00	0.11	0.17	0.21	0.00	0.08	0.26	0.27	0.29	0.12	0.14	0.17	0.02	0.20
I5	0.29	0.23	0.05	0.20	0.18	0.28	0.00	0.11	0.28	0.28	0.48	0.16	0.22	0.29	0.05	0.18
I6	0.28	0.23	0.05	0.19	0.18	0.30	-0.01	0.11	0.29	0.30	0.47	0.13	0.22	0.29	0.04	0.18
I7	0.31	0.24	0.04	0.11	0.11	0.12	0.00	0.11	0.28	0.30	0.28	0.09	0.13	0.16	0.03	0.16
I8	0.27	0.23	0.04	0.11	0.10	0.13	0.00	0.08	0.26	0.32	0.29	0.14	0.16	0.15	0.03	0.19
I9	0.25	0.22	0.01	0.03	0.07	0.08	-0.01	0.10	0.22	0.27	0.06	0.05	0.05	0.03	0.03	0.10
I10	0.25	0.22	0.01	0.04	0.07	0.02	-0.02	0.08	0.20	0.27	0.07	0.05	0.02	0.03	0.02	0.13
I11	0.19	0.12	0.00	0.04	0.05	0.05	0.00	0.07	0.22	0.18	0.06	0.04	0.03	0.05	0.04	0.19
I12	0.22	0.14	0.02	0.03	0.08	0.10	-0.01	0.08	0.23	0.18	0.08	0.07	0.06	0.07	0.05	0.22
I13	0.27	0.18	0.00	0.10	0.16	0.23	-0.01	0.08	0.26	0.23	0.28	0.11	0.14	0.17	0.03	0.20
I14	0.25	0.18	0.01	0.10	0.15	0.21	0.00	0.08	0.27	0.27	0.29	0.15	0.11	0.17	0.04	0.18
I15	0.27	0.23	0.04	0.19	0.18	0.30	0.00	0.10	0.27	0.29	0.47	0.14	0.24	0.29	0.02	0.19
I16	0.28	0.24	0.05	0.19	0.21	0.30	0.00	0.10	0.27	0.30	0.48	0.16	0.22	0.28	0.03	0.18
I17	0.28	0.26	0.03	0.11	0.08	0.14	0.01	0.11	0.27	0.30	0.29	0.10	0.11	0.17	0.03	0.15
I18	0.27	0.24	0.04	0.11	0.11	0.13	-0.01	0.09	0.28	0.28	0.27	0.15	0.16	0.16	0.04	0.17
I19	0.25	0.22	0.01	0.01	0.09	0.08	0.00	0.07	0.22	0.25	0.08	0.08	0.06	0.05	0.01	0.13
I20	0.23	0.22	0.00	0.04	0.06	0.04	-0.03	0.06	0.24	0.27	0.08	0.09	0.03	0.04	0.03	0.13
I21	0.22	0.07	0.00	-0.09	0.13	0.29	-0.01	0.01	0.22	0.00	0.08	0.08	0.04	0.29	0.02	0.11
I22	0.20	0.11	-0.03	-0.09	0.08	0.30	0.00	0.05	0.23	0.01	0.07	0.09	0.05	0.31	0.02	0.11
I23	0.27	0.21	0.01	0.07	0.16	0.27	0.00	0.07	0.26	0.17	0.29	0.09	0.17	0.32	0.04	0.14
I24	0.24	0.21	0.00	0.04	0.16	0.29	0.00	0.00	0.28	0.17	0.28	0.13	0.14	0.32	0.03	0.18
I25	0.29	0.32	0.03	0.22	0.17	0.27	-0.01	0.11	0.28	0.30	0.48	0.15	0.22	0.29	0.03	0.15
I26	0.27	0.34	0.04	0.22	0.18	0.26	0.00	0.10	0.28	0.30	0.48	0.17	0.23	0.28	0.04	0.16
I27	0.26	0.41	0.03	0.28	0.05	0.20	0.00	0.13	0.26	0.38	0.27	0.06	0.16	0.22	0.03	0.10
I28	0.28	0.41	0.03	0.29	0.06	0.19	-0.01	0.13	0.26	0.40	0.29	0.13	0.12	0.21	0.03	0.11
I29	0.23	0.48	0.00	0.41	-0.01	0.12	0.00	0.20	0.23	0.44	0.06	0.08	0.05	0.11	0.03	0.04
I30	0.26	0.48	0.02	0.40	0.01	0.10	0.00	0.23	0.22	0.46	0.05	0.10	0.08	0.10	0.02	0.06
I31	0.19	0.08	-0.47	-0.09	0.18	0.27	-0.40	0.03	0.25	0.00	0.03	0.04	0.18	0.30	0.08	0.09
I32	0.20	0.09	-0.43	-0.09	0.17	0.28	-0.41	-0.02	0.23	0.00	0.02	0.04	0.17	0.30	0.10	0.10
I33	0.26	0.21	-0.16	0.06	0.16	0.28	-0.16	0.06	0.31	0.16	0.21	0.09	0.23	0.32	0.08	0.13
I34	0.26	0.20	-0.17	0.05	0.15	0.29	-0.13	0.06	0.30	0.16	0.20	0.11	0.21	0.31	0.06	0.13
I35	0.32	0.30	0.16	0.23	0.20	0.26	0.15	0.09	0.28	0.30	0.47	0.15	0.22	0.28	0.03	0.20
I36	0.34	0.32	0.17	0.23	0.16	0.26	0.15	0.11	0.29	0.30	0.48	0.14	0.23	0.30	0.02	0.19
I37	0.34	0.42	0.55	0.30	0.13	0.20	0.54	0.14	0.26	0.39	0.20	0.12	0.21	0.22	-0.02	0.12
I38	0.33	0.41	0.55	0.30	0.15	0.20	0.57	0.11	0.27	0.40	0.20	0.10	0.22	0.22	-0.02	0.13
I39	0.32	0.47	0.73	0.42	0.09	0.12	0.80	0.20	0.18	0.44	0.01	0.03	0.17	0.12	-0.04	0.08
I40	0.33	0.48	0.74	0.39	0.10	0.12	0.80	0.22	0.17	0.45	0.00	0.07	0.19	0.10	-0.01	0.07

Table C-17: Residual weighted first factor patterns for 100% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.1	0.04	NA	NA	NA	NA	0.17	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I2	0.11	0.1	NA	NA	NA	NA	0.1	0.1	NA	NA	NA	NA	NA	NA	NA	NA
I3	0.09	0.06	NA	NA	NA	NA	0.1	0.05	NA	NA	NA	NA	NA	NA	NA	NA
I4	0.09	0.08	NA	NA	NA	NA	0.13	0.13	NA	NA	NA	NA	NA	NA	NA	NA
I5	0.04	0.06	NA	NA	NA	NA	0.12	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I6	0.03	0.07	NA	NA	NA	NA	0.08	0.12	NA	NA	NA	NA	NA	NA	NA	NA
I7	0.12	0.02	NA	NA	NA	NA	0.03	0.09	NA	NA	NA	NA	NA	NA	NA	NA
I8	0.05	0.06	NA	NA	NA	NA	0.12	0.12	NA	NA	NA	NA	NA	NA	NA	NA
I9	-0	-0	NA	NA	NA	NA	0.1	0.06	NA	NA	NA	NA	NA	NA	NA	NA
I10	0.02	0.03	NA	NA	NA	NA	-0	0.02	NA	NA	NA	NA	NA	NA	NA	NA
I11	0.09	0.04	NA	NA	NA	NA	0.11	0.15	NA	NA	NA	NA	NA	NA	NA	NA
I12	0.05	0.07	NA	NA	NA	NA	0.1	0.11	NA	NA	NA	NA	NA	NA	NA	NA
I13	0.06	-0	NA	NA	NA	NA	-0	0.04	NA	NA	NA	NA	NA	NA	NA	NA
I14	0.09	0.06	NA	NA	NA	NA	0.1	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I15	0.07	0.06	NA	NA	NA	NA	0.1	0.11	NA	NA	NA	NA	NA	NA	NA	NA
I16	0.07	0.02	NA	NA	NA	NA	0	0.05	NA	NA	NA	NA	NA	NA	NA	NA
I17	0.09	0.01	NA	NA	NA	NA	0.07	-0	NA	NA	NA	NA	NA	NA	NA	NA
I18	0.07	0.08	NA	NA	NA	NA	0.04	0.03	NA	NA	NA	NA	NA	NA	NA	NA
I19	0.01	-0	NA	NA	NA	NA	0.12	-0	NA	NA	NA	NA	NA	NA	NA	NA
I20	0.11	0	NA	NA	NA	NA	0.11	0.04	NA	NA	NA	NA	NA	NA	NA	NA
I21	-0	0.04	NA	NA	NA	NA	0.08	-0	NA	NA	NA	NA	NA	NA	NA	NA
I22	0.06	0.01	NA	NA	NA	NA	0.05	0	NA	NA	NA	NA	NA	NA	NA	NA
I23	0.1	0.12	NA	NA	NA	NA	0.15	0.02	NA	NA	NA	NA	NA	NA	NA	NA
I24	0.12	-0	NA	NA	NA	NA	0.08	0.02	NA	NA	NA	NA	NA	NA	NA	NA
I25	0.02	0.04	NA	NA	NA	NA	0.09	0.01	NA	NA	NA	NA	NA	NA	NA	NA
I26	0.09	0.09	NA	NA	NA	NA	0.07	0.06	NA	NA	NA	NA	NA	NA	NA	NA
I27	0.11	0.01	NA	NA	NA	NA	0.09	-0	NA	NA	NA	NA	NA	NA	NA	NA
I28	0.15	0.02	NA	NA	NA	NA	0.05	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I29	-0	-0	NA	NA	NA	NA	0	-0.1	NA	NA	NA	NA	NA	NA	NA	NA
I30	0.02	-0	NA	NA	NA	NA	0.13	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I31	0.04	-0	NA	NA	NA	NA	0.03	0.04	NA	NA	NA	NA	NA	NA	NA	NA
I32	-0	0.1	NA	NA	NA	NA	0.01	0.09	NA	NA	NA	NA	NA	NA	NA	NA
I33	0.15	-0.1	NA	NA	NA	NA	0.01	-0	NA	NA	NA	NA	NA	NA	NA	NA
I34	0.1	0.01	NA	NA	NA	NA	-0	0.02	NA	NA	NA	NA	NA	NA	NA	NA
I35	0.02	0.09	NA	NA	NA	NA	0.06	-0	NA	NA	NA	NA	NA	NA	NA	NA
I36	0.01	-0	NA	NA	NA	NA	0.01	0.01	NA	NA	NA	NA	NA	NA	NA	NA
I37	-0	0.01	NA	NA	NA	NA	0.03	0.09	NA	NA	NA	NA	NA	NA	NA	NA
I38	-0	-0	NA	NA	NA	NA	-0	-0	NA	NA	NA	NA	NA	NA	NA	NA
I39	0.06	0.04	NA	NA	NA	NA	0.06	-0	NA	NA	NA	NA	NA	NA	NA	NA
I40	-0.1	0.06	NA	NA	NA	NA	0.01	0.07	NA	NA	NA	NA	NA	NA	NA	NA

Table C-18: Residual weighted first factor patterns for 50% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.08	0.1	NA	NA	NA	NA	0.14	0.27	NA	-0	NA	0.03	NA	NA	0.02	0.27
I2	0.16	0.17	NA	NA	NA	NA	0.13	0.28	NA	0.12	NA	0.01	NA	NA	-0	0.26
I3	0.08	0.15	NA	NA	NA	NA	0.17	0.33	NA	0.06	NA	0.16	NA	NA	0.06	0.32
I4	0.07	0.06	NA	NA	NA	NA	0.17	0.3	NA	0.06	NA	0.21	NA	NA	0.01	0.32
I5	0.18	0.18	NA	NA	NA	NA	0.17	0.23	NA	0.03	NA	0.29	NA	NA	-0	0.33
I6	0.17	0.2	NA	NA	NA	NA	0.1	0.33	NA	-0	NA	0.31	NA	NA	0.09	0.35
I7	0.21	0.17	NA	NA	NA	NA	0.05	0.29	NA	0.07	NA	0.19	NA	NA	0.06	0.32
I8	0.05	0.15	NA	NA	NA	NA	0.23	0.28	NA	-0	NA	0.13	NA	NA	0.06	0.33
I9	0.13	0.01	NA	NA	NA	NA	0.16	0.21	NA	0.05	NA	0.04	NA	NA	0.09	0.27
I10	-0	0.06	NA	NA	NA	NA	0.08	0.21	NA	0.03	NA	0.01	NA	NA	0.02	0.27
I11	0.15	0.12	NA	NA	NA	NA	0.02	0.22	NA	0.04	NA	0.04	NA	NA	0.05	0.26
I12	0.05	0.11	NA	NA	NA	NA	0.13	0.23	NA	0.12	NA	0.04	NA	NA	0.05	0.27
I13	0.07	0.14	NA	NA	NA	NA	0.1	0.23	NA	-0	NA	0.17	NA	NA	0.03	0.33
I14	0.08	0.07	NA	NA	NA	NA	0.07	0.28	NA	0.09	NA	0.18	NA	NA	0.03	0.33
I15	0.2	0.12	NA	NA	NA	NA	0.14	0.28	NA	0.05	NA	0.32	NA	NA	0.11	0.36
I16	0.05	0.09	NA	NA	NA	NA	0.17	0.26	NA	0.01	NA	0.28	NA	NA	0.05	0.33
I17	0.17	0.18	NA	NA	NA	NA	0.11	0.24	NA	0.03	NA	0.21	NA	NA	0.09	0.32
I18	0.15	0.18	NA	NA	NA	NA	0.08	0.31	NA	0.02	NA	0.18	NA	NA	0.02	0.32
I19	0.15	0.08	NA	NA	NA	NA	0.06	0.28	NA	0.09	NA	0.08	NA	NA	0.07	0.28
I20	0.04	0.04	NA	NA	NA	NA	-0	0.28	NA	0.07	NA	0.04	NA	NA	0.1	0.26
I21	0.11	-0	NA	NA	NA	NA	0.11	0.08	NA	0.18	NA	0.1	NA	NA	0.04	0.1
I22	0.04	0.04	NA	NA	NA	NA	0.14	0.04	NA	0.18	NA	0.16	NA	NA	0.09	0.14
I23	0.09	0.03	NA	NA	NA	NA	0.12	-0	NA	0.08	NA	0.17	NA	NA	-0	0.2
I24	0.1	-0	NA	NA	NA	NA	0.18	0.05	NA	0.12	NA	0.21	NA	NA	0.1	0.2
I25	0.09	0.05	NA	NA	NA	NA	0.17	0.07	NA	0.01	NA	0.32	NA	NA	0.02	0.35
I26	0.11	-0	NA	NA	NA	NA	0.12	-0	NA	0.1	NA	0.32	NA	NA	0	0.34
I27	0.1	-0.1	NA	NA	NA	NA	0.17	-0.1	NA	0.02	NA	0.17	NA	NA	-0	0.19
I28	0.22	-0	NA	NA	NA	NA	0.08	-0.1	NA	0.03	NA	0.14	NA	NA	0.09	0.19
I29	0	-0	NA	NA	NA	NA	0.01	-0.1	NA	-0.1	NA	-0	NA	NA	-0	-0
I30	0.12	-0.1	NA	NA	NA	NA	0.07	-0.1	NA	-0.1	NA	0.05	NA	NA	0.04	0.01
I31	-0.1	0.01	NA	NA	NA	NA	-0.1	0.09	NA	0.22	NA	0.13	NA	NA	-0	0.1
I32	-0	0.1	NA	NA	NA	NA	0.05	0.03	NA	0.11	NA	0.11	NA	NA	-0	0.14
I33	-0	0.03	NA	NA	NA	NA	-0	0.05	NA	0.27	NA	0.16	NA	NA	0.03	0.19
I34	0	0.1	NA	NA	NA	NA	0.02	0.05	NA	0.11	NA	0.17	NA	NA	0.03	0.18
I35	0.02	0.07	NA	NA	NA	NA	-0	-0.1	NA	0.07	NA	0.26	NA	NA	0.05	0.35
I36	-0	0.02	NA	NA	NA	NA	0.04	-0	NA	0	NA	0.3	NA	NA	0.02	0.35
I37	0.09	-0	NA	NA	NA	NA	0.03	-0	NA	-0	NA	0.15	NA	NA	0.05	0.21
I38	0.11	-0	NA	NA	NA	NA	-0	-0	NA	0	NA	0.17	NA	NA	0.06	0.2
I39	-0	0.04	NA	NA	NA	NA	-0	-0.1	NA	-0.1	NA	-0	NA	NA	0.04	-0
I40	-0.1	-0.1	NA	NA	NA	NA	-0.1	-0.1	NA	-0.1	NA	-0	NA	NA	0.01	0.03

Table C-19: Residual weighted first factor patterns for 20% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	0.19	NA	0.07	NA	NA	0.14	0.34	0.07	0.11	NA	0.05	NA	NA	0.04	0.14
I2	NA	0.13	NA	0.06	NA	NA	0.1	0.33	0.17	0.11	NA	0.01	NA	NA	0.05	0.15
I3	NA	0.23	NA	0.18	NA	NA	0.2	0.39	-0	0.05	NA	0.16	NA	NA	0.02	0.18
I4	NA	0.2	NA	0.19	NA	NA	0.02	0.41	0.06	0.05	NA	0.1	NA	NA	0.04	0.18
I5	NA	0.13	NA	0.22	NA	NA	0.19	0.43	0	0.01	NA	0.2	NA	NA	0.03	0.2
I6	NA	0.16	NA	0.16	NA	NA	0.17	0.42	0.06	0.12	NA	0.22	NA	NA	-0	0.19
I7	NA	0.2	NA	0.15	NA	NA	0.09	0.37	0.06	0.12	NA	0.1	NA	NA	0.01	0.18
I8	NA	0.17	NA	0.03	NA	NA	0.07	0.38	0.08	0.09	NA	0.15	NA	NA	0.04	0.17
I9	NA	0.16	NA	0.04	NA	NA	0.15	0.31	0.07	0.15	NA	0.03	NA	NA	-0	0.16
I10	NA	0.13	NA	0.03	NA	NA	0.14	0.3	0.19	0.14	NA	0.03	NA	NA	-0	0.17
I11	NA	0.14	NA	0.04	NA	NA	0.09	0.32	0.06	0.02	NA	0.03	NA	NA	0.04	0.16
I12	NA	0.21	NA	0.03	NA	NA	0.14	0.35	0.1	0.07	NA	-0	NA	NA	0.03	0.15
I13	NA	0.2	NA	0.11	NA	NA	0.19	0.4	0.1	0.06	NA	0.11	NA	NA	0.02	0.17
I14	NA	0.15	NA	0.08	NA	NA	0.04	0.42	0.19	0.01	NA	0.08	NA	NA	0.05	0.18
I15	NA	0.3	NA	0.19	NA	NA	0.11	0.45	0.06	0.09	NA	0.22	NA	NA	0.04	0.2
I16	NA	0.19	NA	0.13	NA	NA	0.16	0.46	0	0.04	NA	0.21	NA	NA	0.02	0.17
I17	NA	0.16	NA	0.11	NA	NA	0.15	0.41	-0	0.13	NA	0.12	NA	NA	-0	0.16
I18	NA	0.19	NA	0.09	NA	NA	0.13	0.41	0.16	0.06	NA	0.14	NA	NA	-0	0.19
I19	NA	0.08	NA	0.03	NA	NA	0.15	0.33	-0	0.14	NA	0.06	NA	NA	0	0.15
I20	NA	0.12	NA	0.01	NA	NA	0.16	0.37	0.22	0.19	NA	0.07	NA	NA	-0.1	0.15
I21	NA	0.02	NA	-0.1	NA	NA	0.12	-0	-0	0.23	NA	0.13	NA	NA	0.06	0.25
I22	NA	0.01	NA	-0.1	NA	NA	0.08	0.04	0.11	0.32	NA	0.14	NA	NA	0.02	0.26
I23	NA	0.01	NA	0.27	NA	NA	0.14	0.03	0.12	0.25	NA	0.12	NA	NA	0.03	0.15
I24	NA	0.07	NA	-0	NA	NA	0.1	0.04	0.06	0.23	NA	0.14	NA	NA	0.01	0.13
I25	NA	-0.1	NA	-0	NA	NA	0.14	0	0.12	0.06	NA	0.2	NA	NA	0.03	0.19
I26	NA	-0	NA	0.06	NA	NA	0.14	0.04	-0	0.06	NA	0.21	NA	NA	0.03	0.19
I27	NA	-0.1	NA	-0	NA	NA	0.2	0.03	0.09	-0.1	NA	0.13	NA	NA	0.03	0.08
I28	NA	-0.1	NA	-0	NA	NA	0.03	0.03	0.08	-0.1	NA	0.12	NA	NA	-0	0.1
I29	NA	-0	NA	0.08	NA	NA	0.19	-0	-0	-0.2	NA	0.05	NA	NA	-0	-0.1
I30	NA	-0	NA	0.04	NA	NA	0.11	-0	0.18	-0.2	NA	-0	NA	NA	-0	-0.2
I31	NA	0.06	NA	-0.1	NA	NA	0.05	-0	0.15	0.31	NA	0.07	NA	NA	-0.4	0.26
I32	NA	-0	NA	0.07	NA	NA	0.03	0.05	0.11	0.29	NA	0.13	NA	NA	-0.4	0.24
I33	NA	-0.1	NA	0.01	NA	NA	0.09	-0	0.14	0.29	NA	0.14	NA	NA	-0.1	0.13
I34	NA	0.02	NA	0.04	NA	NA	0	-0	0.22	0.18	NA	0.11	NA	NA	-0.2	0.15
I35	NA	-0.1	NA	0.09	NA	NA	-0.1	-0	0.16	0.12	NA	0.22	NA	NA	0.04	0.17
I36	NA	0.13	NA	0.12	NA	NA	0.07	-0	-0.1	0.12	NA	0.23	NA	NA	0	0.17
I37	NA	0.03	NA	0.03	NA	NA	-0.1	-0	0.02	-0.1	NA	0.11	NA	NA	0.37	0.07
I38	NA	-0.1	NA	0.05	NA	NA	-0	-0	-0.1	-0	NA	0.14	NA	NA	0.39	0.07
I39	NA	-0	NA	0.02	NA	NA	-0.1	-0	-0	-0.2	NA	0.08	NA	NA	0.49	-0.1
I40	NA	-0	NA	0.06	NA	NA	-0	-0	0.08	-0.2	NA	0.04	NA	NA	0.5	-0.1

Table C-20: Residual weighted first factor patterns for 5% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	0.26	NA	0	NA	NA	0.16	0.21	0.06	0.1	NA	0.01	NA	NA	0.13	0.06
I2	NA	0.18	NA	0.01	NA	NA	0.14	0.23	-0	0.02	NA	0.03	NA	NA	0.21	0.07
I3	NA	0.24	NA	0.14	NA	NA	0.19	0.36	0.12	0.08	NA	0.13	NA	NA	0.2	0.08
I4	NA	0.12	NA	0.08	NA	NA	0.09	0.28	0.03	0.04	NA	0.13	NA	NA	0.25	0.09
I5	NA	0.13	NA	0.19	NA	NA	0.35	0.36	0.13	0.17	NA	0.26	NA	NA	0.29	0.08
I6	NA	0.18	NA	0.22	NA	NA	0.26	0.36	0.17	0.15	NA	0.2	NA	NA	0.27	0.09
I7	NA	0.11	NA	0.03	NA	NA	0.13	0.24	0.11	0.19	NA	0.16	NA	NA	0.33	0.12
I8	NA	0.28	NA	0.18	NA	NA	0.24	0.35	0.08	0.11	NA	0.21	NA	NA	0.27	0.11
I9	NA	0.17	NA	0.03	NA	NA	0.23	0.21	0.19	0.29	NA	0.09	NA	NA	0.24	0.11
I10	NA	0.11	NA	0	NA	NA	0.2	0.27	0.27	0.25	NA	0.04	NA	NA	0.23	0.11
I11	NA	0.14	NA	-0	NA	NA	0.1	0.26	0.07	0.04	NA	0	NA	NA	0.15	0.08
I12	NA	0.21	NA	0.03	NA	NA	0.15	0.24	0.04	0.08	NA	0	NA	NA	0.18	0.06
I13	NA	0.18	NA	0.13	NA	NA	0.15	0.35	0.01	0.04	NA	0.18	NA	NA	0.22	0.1
I14	NA	0.08	NA	0.14	NA	NA	0.14	0.34	0.08	0.08	NA	0.14	NA	NA	0.27	0.06
I15	NA	0.18	NA	0.09	NA	NA	0.26	0.4	-0	0.11	NA	0.21	NA	NA	0.25	0.08
I16	NA	0.32	NA	0.14	NA	NA	0.21	0.37	-0.1	0.2	NA	0.25	NA	NA	0.22	0.1
I17	NA	0.1	NA	0.13	NA	NA	0.23	0.32	0.22	0.16	NA	0.23	NA	NA	0.31	0.09
I18	NA	0.1	NA	0.06	NA	NA	0.12	0.26	0.01	0.24	NA	0.22	NA	NA	0.26	0.1
I19	NA	0.17	NA	0.01	NA	NA	0.27	0.27	0.17	0.27	NA	-0	NA	NA	0.28	0.1
I20	NA	0.2	NA	0.01	NA	NA	0.21	0.28	0.01	0.28	NA	0.03	NA	NA	0.31	0.11
I21	NA	0.1	NA	0.02	NA	NA	0.13	0.03	0.13	0.41	NA	0.18	NA	NA	0.18	0.34
I22	NA	0.1	NA	0.06	NA	NA	0.17	0.06	0.02	0.4	NA	0.12	NA	NA	0.19	0.38
I23	NA	0.05	NA	-0	NA	NA	0.19	0.04	0.08	0.31	NA	0.21	NA	NA	0.19	0.22
I24	NA	0.06	NA	0.05	NA	NA	0.22	0.02	0.07	0.31	NA	0.31	NA	NA	0.27	0.22
I25	NA	-0	NA	-0	NA	NA	0.05	0.02	0.06	0.14	NA	0.27	NA	NA	0.28	0.09
I26	NA	0.14	NA	0.08	NA	NA	0.13	0.05	0.12	0.15	NA	0.32	NA	NA	0.27	0.07
I27	NA	-0	NA	0.03	NA	NA	0.21	-0	0.22	-0.1	NA	0.1	NA	NA	0.27	0.05
I28	NA	-0	NA	0	NA	NA	0.04	-0	0.24	-0	NA	0.12	NA	NA	0.25	0.05
I29	NA	-0	NA	0.02	NA	NA	0.21	0.05	0.2	-0.2	NA	-0.1	NA	NA	0.3	-0.2
I30	NA	-0.1	NA	0.06	NA	NA	0.19	-0	0.16	-0.2	NA	-0.1	NA	NA	0.35	-0.3
I31	NA	-0.1	NA	0.05	NA	NA	0.1	0.05	0.23	0.42	NA	0.19	NA	NA	0.29	0.35
I32	NA	-0	NA	0.06	NA	NA	0.07	0.03	0.3	0.43	NA	0.14	NA	NA	0.27	0.44
I33	NA	0.1	NA	0.11	NA	NA	0.1	0.02	0.33	0.24	NA	0.19	NA	NA	0.3	0.23
I34	NA	0.08	NA	0.06	NA	NA	-0	0	0.21	0.39	NA	0.15	NA	NA	0.29	0.11
I35	NA	0.01	NA	0.05	NA	NA	0.02	0.01	-0	0.1	NA	0.23	NA	NA	0.34	0.09
I36	NA	-0	NA	-0	NA	NA	0.03	-0	0.1	0.19	NA	0.24	NA	NA	0.25	0.1
I37	NA	0.02	NA	-0	NA	NA	-0	0.04	-0	-0.1	NA	0.05	NA	NA	0	0.01
I38	NA	-0	NA	0.09	NA	NA	-0.1	-0	-0.1	-0.1	NA	0.14	NA	NA	-0.1	0.06
I39	NA	-0	NA	0.08	NA	NA	-0.1	-0	-0.1	-0.2	NA	-0	NA	NA	-0.2	-0.3
I40	NA	-0	NA	0.02	NA	NA	-0.2	-0	-0.1	-0.2	NA	-0.1	NA	NA	-0.2	-0.2

Table C-21: Residual weighted second factor patterns for 100% contamination

	Random								Reversed							
RANGE	1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3	
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	-0	0.03	NA	NA	NA	NA	0.05	0.07	NA	NA	NA	NA	NA	NA	NA	NA
I2	0.05	0	NA	NA	NA	NA	0.1	0.04	NA	NA	NA	NA	NA	NA	NA	NA
I3	0.09	0.02	NA	NA	NA	NA	0.03	0.14	NA	NA	NA	NA	NA	NA	NA	NA
I4	0.04	0.08	NA	NA	NA	NA	0.04	0.09	NA	NA	NA	NA	NA	NA	NA	NA
I5	0.12	0.02	NA	NA	NA	NA	0.07	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I6	0.13	0.03	NA	NA	NA	NA	-0	0.11	NA	NA	NA	NA	NA	NA	NA	NA
I7	0.08	0.03	NA	NA	NA	NA	0.07	0.05	NA	NA	NA	NA	NA	NA	NA	NA
I8	0.02	0.05	NA	NA	NA	NA	-0	0.09	NA	NA	NA	NA	NA	NA	NA	NA
I9	-0	-0	NA	NA	NA	NA	0.04	0.03	NA	NA	NA	NA	NA	NA	NA	NA
I10	0.13	0.1	NA	NA	NA	NA	0.04	0.03	NA	NA	NA	NA	NA	NA	NA	NA
I11	0.09	0.08	NA	NA	NA	NA	0.09	0.07	NA	NA	NA	NA	NA	NA	NA	NA
I12	0.09	0.07	NA	NA	NA	NA	0.02	0.09	NA	NA	NA	NA	NA	NA	NA	NA
I13	0.05	0.12	NA	NA	NA	NA	0.1	0.04	NA	NA	NA	NA	NA	NA	NA	NA
I14	0.05	0.08	NA	NA	NA	NA	0.01	0.01	NA	NA	NA	NA	NA	NA	NA	NA
I15	0.1	0.07	NA	NA	NA	NA	0.08	0	NA	NA	NA	NA	NA	NA	NA	NA
I16	0.11	0.06	NA	NA	NA	NA	0.11	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I17	0.07	0.04	NA	NA	NA	NA	0.07	0.1	NA	NA	NA	NA	NA	NA	NA	NA
I18	0.06	0.02	NA	NA	NA	NA	0.04	0.03	NA	NA	NA	NA	NA	NA	NA	NA
I19	0.03	0.09	NA	NA	NA	NA	0.01	0.06	NA	NA	NA	NA	NA	NA	NA	NA
I20	0.02	0.02	NA	NA	NA	NA	0.02	0.04	NA	NA	NA	NA	NA	NA	NA	NA
I21	0.06	0.04	NA	NA	NA	NA	0.05	-0	NA	NA	NA	NA	NA	NA	NA	NA
I22	0.13	-0	NA	NA	NA	NA	0.05	-0	NA	NA	NA	NA	NA	NA	NA	NA
I23	0.11	0.04	NA	NA	NA	NA	0.06	-0	NA	NA	NA	NA	NA	NA	NA	NA
I24	0.1	0.11	NA	NA	NA	NA	0.16	0.01	NA	NA	NA	NA	NA	NA	NA	NA
I25	0.07	-0	NA	NA	NA	NA	0.05	0.03	NA	NA	NA	NA	NA	NA	NA	NA
I26	0.1	0.03	NA	NA	NA	NA	0.08	0.03	NA	NA	NA	NA	NA	NA	NA	NA
I27	0.02	0.03	NA	NA	NA	NA	0.13	0.02	NA	NA	NA	NA	NA	NA	NA	NA
I28	-0	-0	NA	NA	NA	NA	0.11	0.07	NA	NA	NA	NA	NA	NA	NA	NA
I29	0.08	0.07	NA	NA	NA	NA	0.09	-0	NA	NA	NA	NA	NA	NA	NA	NA
I30	0.12	0.02	NA	NA	NA	NA	0.09	-0	NA	NA	NA	NA	NA	NA	NA	NA
I31	0	0.04	NA	NA	NA	NA	0.05	0.07	NA	NA	NA	NA	NA	NA	NA	NA
I32	-0	0.04	NA	NA	NA	NA	-0.1	0	NA	NA	NA	NA	NA	NA	NA	NA
I33	-0	0.11	NA	NA	NA	NA	-0	0.06	NA	NA	NA	NA	NA	NA	NA	NA
I34	0.05	-0	NA	NA	NA	NA	-0	0.02	NA	NA	NA	NA	NA	NA	NA	NA
I35	-0	0.05	NA	NA	NA	NA	0.01	-0	NA	NA	NA	NA	NA	NA	NA	NA
I36	-0	0.01	NA	NA	NA	NA	0.07	0.08	NA	NA	NA	NA	NA	NA	NA	NA
I37	0.02	-0	NA	NA	NA	NA	0.02	0.01	NA	NA	NA	NA	NA	NA	NA	NA
I38	0.02	0.03	NA	NA	NA	NA	0.01	-0	NA	NA	NA	NA	NA	NA	NA	NA
I39	0.09	-0	NA	NA	NA	NA	-0	0.07	NA	NA	NA	NA	NA	NA	NA	NA
I40	-0.1	0.01	NA	NA	NA	NA	0.01	0.06	NA	NA	NA	NA	NA	NA	NA	NA

Table C-22: Residual weighted second factor patterns for 50% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	0.14	0.12	NA	NA	NA	NA	0.04	0.07	NA	0.08	NA	0.08	NA	NA	0.08	0.06
I2	0.06	0.13	NA	NA	NA	NA	0.14	0.09	NA	0.05	NA	0.06	NA	NA	0.1	-0
I3	0.08	0.11	NA	NA	NA	NA	0.08	0.09	NA	0	NA	0.13	NA	NA	0.11	0.17
I4	-0	0.1	NA	NA	NA	NA	0.09	0.06	NA	0.02	NA	0.1	NA	NA	0.1	0.14
I5	0.13	0.1	NA	NA	NA	NA	0.14	0.11	NA	0.04	NA	0.21	NA	NA	0.11	0.15
I6	0.06	0.09	NA	NA	NA	NA	0	0.1	NA	0.11	NA	0.21	NA	NA	0.12	0.09
I7	0.08	0.04	NA	NA	NA	NA	0.14	0.06	NA	-0	NA	0.11	NA	NA	0.11	0.11
I8	0.15	0.09	NA	NA	NA	NA	0.14	0.08	NA	0.08	NA	0.16	NA	NA	0.11	0.12
I9	0.02	0.11	NA	NA	NA	NA	0.07	0.05	NA	0.03	NA	0.03	NA	NA	0.1	0.06
I10	0.1	0.11	NA	NA	NA	NA	0.09	0.11	NA	0.11	NA	0.11	NA	NA	0.1	0.06
I11	0.04	0.17	NA	NA	NA	NA	0.11	0.08	NA	0.05	NA	0.02	NA	NA	0.09	0.01
I12	0.12	0.14	NA	NA	NA	NA	0.17	0.09	NA	0.09	NA	0.02	NA	NA	0.07	0.05
I13	0.06	0.06	NA	NA	NA	NA	0.01	0.12	NA	0.13	NA	0.13	NA	NA	0.1	0.12
I14	0.06	0.08	NA	NA	NA	NA	0.1	0.07	NA	0.05	NA	0.09	NA	NA	0.11	0.03
I15	0.08	0.1	NA	NA	NA	NA	0.1	0.11	NA	0.05	NA	0.2	NA	NA	0.12	0.1
I16	0.05	0.1	NA	NA	NA	NA	0.07	0.13	NA	-0	NA	0.2	NA	NA	0.11	0.17
I17	0.06	0.11	NA	NA	NA	NA	0.03	0.12	NA	0.08	NA	0.18	NA	NA	0.12	0.06
I18	0.04	0.08	NA	NA	NA	NA	0.06	0.09	NA	0.05	NA	0.11	NA	NA	0.1	0.12
I19	0.07	0.06	NA	NA	NA	NA	0.19	0.09	NA	0.09	NA	0.05	NA	NA	0.1	0.07
I20	0.14	0.19	NA	NA	NA	NA	0.11	0.08	NA	0.03	NA	0.02	NA	NA	0.1	0.11
I21	-0	0.03	NA	NA	NA	NA	0.11	-0	NA	0.04	NA	0.09	NA	NA	0.13	0.08
I22	0.08	0.03	NA	NA	NA	NA	0.04	0.03	NA	-0.1	NA	0.13	NA	NA	0.1	0.01
I23	0.16	0.04	NA	NA	NA	NA	0.08	0.09	NA	-0.1	NA	0.15	NA	NA	0.08	0.03
I24	0.09	0.07	NA	NA	NA	NA	-0	0.08	NA	0.01	NA	0.18	NA	NA	0.09	0.17
I25	0.08	-0	NA	NA	NA	NA	0.12	-0	NA	0.11	NA	0.19	NA	NA	0.11	0.07
I26	0.06	-0	NA	NA	NA	NA	-0	0.06	NA	0.06	NA	0.21	NA	NA	0.11	0.07
I27	0.08	-0.1	NA	NA	NA	NA	0.05	-0	NA	0.15	NA	0.11	NA	NA	0.05	0.11
I28	0.16	0	NA	NA	NA	NA	0.26	-0	NA	0.07	NA	0.13	NA	NA	0.06	0.01
I29	0.16	-0	NA	NA	NA	NA	0.14	0.01	NA	0.14	NA	0.04	NA	NA	-0	0.13
I30	0.09	0.04	NA	NA	NA	NA	0.17	-0	NA	0.11	NA	-0	NA	NA	0.04	0.05
I31	0.07	-0	NA	NA	NA	NA	-0	0.02	NA	-0	NA	0.09	NA	NA	0.18	0.02
I32	0.04	0.08	NA	NA	NA	NA	0.05	0.07	NA	-0	NA	0.06	NA	NA	0.12	0.12
I33	0	0.02	NA	NA	NA	NA	0.07	0.05	NA	0.03	NA	0.19	NA	NA	0.09	-0.1
I34	0.05	0.03	NA	NA	NA	NA	0.08	0.08	NA	0.1	NA	0.16	NA	NA	0.1	0.05
I35	0	0.1	NA	NA	NA	NA	0.07	0.02	NA	0.06	NA	0.19	NA	NA	0.11	0.06
I36	0.07	-0	NA	NA	NA	NA	-0	-0.1	NA	0.08	NA	0.22	NA	NA	0.12	-0
I37	0.06	0	NA	NA	NA	NA	-0.1	0.02	NA	0.15	NA	0.12	NA	NA	0.06	0.08
I38	-0.1	0.01	NA	NA	NA	NA	0.03	0.02	NA	0.05	NA	0.11	NA	NA	0.05	0.04
I39	-0.1	0.05	NA	NA	NA	NA	-0	-0.1	NA	0.16	NA	0	NA	NA	-0	0
I40	0	-0	NA	NA	NA	NA	0.07	-0	NA	0.08	NA	0.05	NA	NA	0.05	0.03

Table C-23: Residual weighted second factor patterns for 20% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	0.07	NA	0	NA	NA	0.09	0.06	0.15	0.02	NA	0	NA	NA	0.18	0.21
I2	NA	0	NA	0	NA	NA	0.08	0.1	0.05	0.01	NA	0	NA	NA	0.16	0.22
I3	NA	0.04	NA	0	NA	NA	0.12	0.12	0.18	0.03	NA	0	NA	NA	0.25	0.25
I4	NA	-0	NA	0.25	NA	NA	0.03	0.08	0.1	0.03	NA	0.01	NA	NA	0.18	0.28
I5	NA	0.27	NA	-0.1	NA	NA	0.09	0.12	0.09	0.11	NA	0	NA	NA	0.2	0.28
I6	NA	0.2	NA	0	NA	NA	0.1	0.09	0.13	-0	NA	0.41	NA	NA	0.18	0.26
I7	NA	0.07	NA	0.17	NA	NA	0.1	0.12	0.08	0.24	NA	0.01	NA	NA	0.16	0.27
I8	NA	0.19	NA	0.74	NA	NA	0.08	0.08	0.11	0.03	NA	0.99	NA	NA	0.2	0.26
I9	NA	0	NA	0	NA	NA	0.07	0.13	0.15	0	NA	0.24	NA	NA	0.1	0.23
I10	NA	0.03	NA	0	NA	NA	0.03	0.08	0.04	-0.1	NA	0	NA	NA	0.09	0.23
I11	NA	0	NA	0	NA	NA	0.06	0.09	0.04	0	NA	0	NA	NA	0.15	0.21
I12	NA	0.07	NA	0	NA	NA	0.07	0.08	0.08	0	NA	0	NA	NA	0.19	0.2
I13	NA	-0	NA	0	NA	NA	0.08	0.1	0.11	0.01	NA	0	NA	NA	0.19	0.25
I14	NA	0.01	NA	0.25	NA	NA	0.07	0.1	0.11	0.04	NA	0	NA	NA	0.15	0.24
I15	NA	0.1	NA	0.21	NA	NA	0.06	-0	0.08	0.18	NA	0.99	NA	NA	0.24	0.28
I16	NA	0.15	NA	0.25	NA	NA	0.07	0.06	0.17	0.11	NA	0.01	NA	NA	0.2	0.27
I17	NA	0.16	NA	-0.1	NA	NA	0.06	0.01	0.11	0.22	NA	0.24	NA	NA	0.12	0.25
I18	NA	0.22	NA	0.21	NA	NA	0.09	0.06	0.05	-0	NA	-1	NA	NA	0.16	0.25
I19	NA	-0	NA	0.5	NA	NA	0.04	0.18	0.13	-0.1	NA	0	NA	NA	0.17	0.21
I20	NA	-0	NA	0	NA	NA	0.03	0.07	0.1	0.01	NA	0.41	NA	NA	0.15	0.23
I21	NA	0	NA	0.25	NA	NA	0.07	0.01	0.13	-0	NA	0	NA	NA	0.12	0.18
I22	NA	-0	NA	0	NA	NA	0.03	0.03	0.13	0	NA	0	NA	NA	0.21	0.18
I23	NA	0.1	NA	0.25	NA	NA	0.09	0.09	0.12	0.01	NA	0.01	NA	NA	0.27	0.17
I24	NA	0.1	NA	0	NA	NA	0.03	0.01	0.03	0.01	NA	0	NA	NA	0.18	0.18
I25	NA	0.04	NA	0.23	NA	NA	-0	0.07	0.04	0.12	NA	-0.4	NA	NA	0.22	0.26
I26	NA	0.14	NA	0.21	NA	NA	0.03	0.07	0.15	0.23	NA	0.2	NA	NA	0.22	0.27
I27	NA	0.25	NA	0	NA	NA	0.05	0.14	0.03	-0	NA	0.01	NA	NA	0.14	0.13
I28	NA	0.23	NA	-0	NA	NA	0.05	0.1	0.05	0.14	NA	0.75	NA	NA	0.19	0.13
I29	NA	-0	NA	0	NA	NA	-0	0.08	-0	0.04	NA	0	NA	NA	0.1	-0
I30	NA	-0	NA	0.18	NA	NA	0.01	0.05	-0.1	0.01	NA	0	NA	NA	0.1	-0.1
I31	NA	0.08	NA	0.3	NA	NA	0.07	0.06	0.11	0.07	NA	0.75	NA	NA	-0	0.18
I32	NA	0.07	NA	0.21	NA	NA	-0	0.03	0.1	0.06	NA	0.2	NA	NA	0.01	0.11
I33	NA	0.11	NA	0.21	NA	NA	0.03	0.02	0.1	0.2	NA	0.99	NA	NA	0.1	0.17
I34	NA	0.15	NA	-0.3	NA	NA	0.02	0.08	0.12	0.17	NA	0.99	NA	NA	0.09	0.19
I35	NA	0.07	NA	0.11	NA	NA	0.08	0.04	0.06	0.13	NA	0	NA	NA	0.23	0.29
I36	NA	-0.1	NA	0	NA	NA	0.03	0.16	0.1	0.18	NA	0.99	NA	NA	0.16	0.28
I37	NA	-0	NA	0.03	NA	NA	0.06	0.1	-0.1	-0	NA	0	NA	NA	0.01	0.15
I38	NA	0.12	NA	0.11	NA	NA	-0	-0	0.04	-0	NA	-0.4	NA	NA	0.09	0.16
I39	NA	-0	NA	-0.2	NA	NA	-0.1	0.06	-0	-0.1	NA	-0.7	NA	NA	0.03	-0.1
I40	NA	0.16	NA	-0.3	NA	NA	0.07	0.02	-0	-0.1	NA	-0.7	NA	NA	0.04	-0.1

Table C-24: Residual weighted second factor patterns for 5% contamination

	Random								Reversed							
RANGE		1x1		3x3		3x1		1x3		1x1		3x3		3x1		1x3
SUBTEST	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20
I1	NA	0.16	NA	0.03	NA	NA	0.09	0.13	0.06	0.1	NA	0.02	NA	NA	0.09	0.33
I2	NA	0.19	NA	0.02	NA	NA	0.08	0.17	0.01	0.12	NA	0	NA	NA	0.06	0.3
I3	NA	0.19	NA	0.08	NA	NA	0.14	0.1	0.01	0.06	NA	0.14	NA	NA	0.14	0.38
I4	NA	0.03	NA	0.12	NA	NA	0.06	0.13	0.04	0.09	NA	0.12	NA	NA	0.12	0.36
I5	NA	0.07	NA	0.16	NA	NA	0.1	0.14	0.07	0.16	NA	0.25	NA	NA	0.17	0.33
I6	NA	0.19	NA	0.09	NA	NA	0.11	0.19	0.12	0.06	NA	0.22	NA	NA	0.16	0.36
I7	NA	0.08	NA	0.12	NA	NA	0.15	0.15	0.07	0.1	NA	0.13	NA	NA	0.14	0.33
I8	NA	0.13	NA	0.18	NA	NA	0.08	0.17	0.18	0.02	NA	0.16	NA	NA	0.16	0.31
I9	NA	0.21	NA	0.01	NA	NA	0.01	0.1	0.06	0.07	NA	0.01	NA	NA	0.1	0.24
I10	NA	0.09	NA	0.01	NA	NA	0.17	0.13	0.2	0.04	NA	0	NA	NA	0.19	0.28
I11	NA	0.06	NA	0.02	NA	NA	0.15	0.05	0.08	0.14	NA	0.03	NA	NA	0.1	0.3
I12	NA	0.09	NA	0	NA	NA	0.09	0.1	0.01	0.12	NA	0.03	NA	NA	0.09	0.24
I13	NA	-0	NA	0.12	NA	NA	0.05	0.05	-0	0.08	NA	0.08	NA	NA	0.09	0.3
I14	NA	0.15	NA	0.09	NA	NA	0.02	0.15	0.16	0.05	NA	0.13	NA	NA	0.14	0.32
I15	NA	0.06	NA	0.26	NA	NA	0.2	0.15	0.02	0.12	NA	0.26	NA	NA	0.15	0.36
I16	NA	0.1	NA	0.15	NA	NA	0.2	0.13	0.13	0.08	NA	0.25	NA	NA	0.13	0.39
I17	NA	0.15	NA	0.11	NA	NA	0.1	0.08	0.03	0.04	NA	0.08	NA	NA	0.13	0.31
I18	NA	0.21	NA	0.07	NA	NA	0.25	0.12	0.02	0.13	NA	0.05	NA	NA	0.17	0.32
I19	NA	0.12	NA	-0	NA	NA	0.02	0.04	0.13	0.14	NA	0.02	NA	NA	0.14	0.27
I20	NA	0.1	NA	0.02	NA	NA	0.06	0.11	0.09	0.08	NA	0.02	NA	NA	0.16	0.28
I21	NA	0.04	NA	-0	NA	NA	0.15	-0	0.11	0.17	NA	0.03	NA	NA	0.11	0.11
I22	NA	0.05	NA	0.03	NA	NA	0.05	0.03	-0	-0	NA	0.06	NA	NA	0.1	0.14
I23	NA	0.01	NA	0.08	NA	NA	-0	-0	-0.1	0.09	NA	0.1	NA	NA	0.18	0.23
I24	NA	-0.1	NA	0.04	NA	NA	0.1	0.06	0.13	0.14	NA	0.13	NA	NA	0.1	0.19
I25	NA	0.05	NA	0.07	NA	NA	0.09	-0.1	0.1	0.18	NA	0.2	NA	NA	0.17	0.38
I26	NA	-0.1	NA	0.06	NA	NA	0.17	-0	0.09	0.06	NA	0.26	NA	NA	0.18	0.38
I27	NA	0	NA	0.1	NA	NA	0.17	-0.1	-0	0.09	NA	0.16	NA	NA	0.15	0.12
I28	NA	-0.1	NA	0.07	NA	NA	0.04	0.06	0.06	0.07	NA	0.11	NA	NA	0.17	0.15
I29	NA	0.02	NA	0.02	NA	NA	0.12	0	0.12	-0	NA	0.01	NA	NA	0.11	-0
I30	NA	0.06	NA	0	NA	NA	0.18	-0.1	0.11	-0	NA	-0	NA	NA	0.15	0
I31	NA	0.05	NA	0.03	NA	NA	0.03	0.04	0.17	0.19	NA	0.03	NA	NA	0.2	0.11
I32	NA	0.13	NA	0.12	NA	NA	0.04	-0	0.14	0.2	NA	0.02	NA	NA	0.21	0.08
I33	NA	0.05	NA	-0.1	NA	NA	0.09	0.03	0.06	0.14	NA	0.11	NA	NA	0.14	0.21
I34	NA	-0.1	NA	0.01	NA	NA	0.03	-0	0.22	0.11	NA	0.1	NA	NA	0.18	0.14
I35	NA	-0	NA	0.06	NA	NA	0.06	0.07	0.05	0.07	NA	0.25	NA	NA	0.14	0.38
I36	NA	0.04	NA	0.02	NA	NA	-0	0.06	0.04	0.05	NA	0.25	NA	NA	0.15	0.38
I37	NA	-0	NA	0	NA	NA	-0	-0	0.06	-0	NA	0.22	NA	NA	0.03	0.19
I38	NA	-0.1	NA	0	NA	NA	-0.1	0.06	-0	-0	NA	0.12	NA	NA	-0	0.16
I39	NA	0.02	NA	0	NA	NA	0.02	0.1	-0	0.01	NA	0	NA	NA	-0.2	-0
I40	NA	-0	NA	0.02	NA	NA	0.04	-0	0.01	-0.1	NA	0.02	NA	NA	-0.2	-0

Appendix D

Table D-1: average residual values within subtests for factor 1 of the all Rasch baseline condition

RANGE	1X1		3X3		3X1		1X3	
SUBTEST	10	20	10	20	10	20	10	20
Rasch1								
-2	0.24	0.20	0.01	0.01	0.01	0.02	0.23	0.28
-1	0.33	0.33	0.15	0.14	0.16	0.13	0.32	0.34
0	0.36	0.33	0.36	0.30	0.39	0.37	0.37	0.34
1	0.34	0.31	0.14	0.14	0.14	0.11	0.30	0.31
2	0.21	0.33	0.00	0.01	0.00	0.01	0.25	0.25
Rasch2								
-2	0.27	0.24	0.02	0.01	0.27	0.28	0.01	0.00
-1	0.32	0.34	0.17	0.14	0.38	0.29	0.11	0.13
0	0.30	0.29	0.34	0.35	0.48	0.31	0.31	0.33
1	0.29	0.34	0.18	0.11	0.29	0.27	0.14	0.13
2	0.27	0.21	0.01	0.01	0.27	0.19	0.01	0.02

Table D-2: average residual values within subtests for factor 1 of the 100% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	0.0625	0.0631	NA	NA	NA	NA	0.1025	0.1086
	-1	0.1012	0.0732	NA	NA	NA	NA	0.0860	0.0700
	0	0.0532	0.0560	NA	NA	NA	NA	0.0770	0.0911
	1	0.0974	0.0426	NA	NA	NA	NA	0.0665	0.0524
	2	0.0210	-0.0013	NA	NA	NA	NA	0.0726	0.0324
Mixture	-2	0.0014	0.0287	NA	NA	NA	NA	0.0214	0.0308
	-1	0.1272	0.0128	NA	NA	NA	NA	0.0004	0.0110
	0	0.0156	0.0469	NA	NA	NA	NA	0.0347	0.0191
	1	-0.0479	0.0051	NA	NA	NA	NA	-0.0009	0.0301
	2	0.0041	0.0097	NA	NA	NA	NA	0.0345	0.0120
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	NA	NA	NA	NA	NA	NA	NA
	-1	NA	NA	NA	NA	NA	NA	NA	NA
	0	NA	NA	NA	NA	NA	NA	NA	NA
	1	NA	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA	NA
Mixture	-2	NA	NA	NA	NA	NA	NA	NA	NA
	-1	NA	NA	NA	NA	NA	NA	NA	NA
	0	NA	NA	NA	NA	NA	NA	NA	NA
	1	NA	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA	NA

Table D-3: average residual values within subtests for factor 1 of the 50% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	0.0973	0.1237	NA	NA	NA	NA	0.1124	0.2505
	-1	0.0791	0.0918	NA	NA	NA	NA	0.1544	0.2014
	0	0.1321	0.1460	NA	NA	NA	NA	0.1440	0.2734
	1	0.1495	0.1692	NA	NA	NA	NA	0.1187	0.2801
	2	0.0667	0.0461	NA	NA	NA	NA	0.0593	0.2426
Mixture	-2	-0.0356	0.0369	NA	NA	NA	NA	-0.0144	0.0596
	-1	0.0017	0.0342	NA	NA	NA	NA	0.0008	0.0346
	0	-0.0006	0.0293	NA	NA	NA	NA	0.0013	-0.0171
	1	0.0962	-0.0436	NA	NA	NA	NA	-0.0106	-0.0505
	2	-0.0490	-0.0242	NA	NA	NA	NA	-0.0480	-0.0845
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.0628	NA	0.0307	NA	NA	0.0348	0.2651
	-1	NA	0.0479	NA	0.0944	NA	NA	0.0299	0.1765
	0	NA	0.0186	NA	0.3105	NA	NA	0.0425	0.3434
	1	NA	0.0220	NA	0.1790	NA	NA	0.0506	0.3242
	2	NA	0.0598	NA	0.0450	NA	NA	0.0526	0.2709
Mixture	-2	NA	0.1742	NA	0.1463	NA	NA	-0.0135	0.1203
	-1	NA	0.1462	NA	0.1906	NA	NA	0.0298	0.1924
	0	NA	0.0442	NA	0.3099	NA	NA	0.0360	0.3457
	1	NA	0.0050	NA	0.1479	NA	NA	0.0527	0.1997
	2	NA	-0.0762	NA	-0.0302	NA	NA	0.0265	0.0062

Table D-4: average residual values within subtests for factor 1 of the 20% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.1673	NA	0.0490	NA	NA	0.1110	0.3323
	-1	NA	0.1698	NA	0.1645	NA	NA	0.1368	0.2560
	0	NA	0.1953	NA	0.1754	NA	NA	0.1533	0.4405
	1	NA	0.1788	NA	0.0962	NA	NA	0.1110	0.3932
	2	NA	0.1217	NA	0.0278	NA	NA	0.1515	0.3267
Mixture	-2	NA	0.0224	NA	-0.0358	NA	NA	0.0422	0.0102
	-1	NA	0.0031	NA	0.0723	NA	NA	0.0481	0.0096
	0	NA	-0.0109	NA	0.0657	NA	NA	-0.0188	-0.0059
	1	NA	-0.0464	NA	0.0020	NA	NA	-0.0652	0.0026
	2	NA	-0.0356	NA	0.0506	NA	NA	-0.0932	-0.0264
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.0765	NA	0.0213	NA	NA	0.0420	0.1506
	-1	NA	0.0971	NA	0.0970	NA	NA	0.0389	0.1066
	0	NA	0.0650	NA	0.2130	NA	NA	0.0255	0.1886
	1	NA	0.1017	NA	0.1280	NA	NA	0.0090	0.1747
	2	NA	0.1563	NA	0.0495	NA	NA	-0.0256	0.1562
Mixture	-2	NA	0.2876	NA	0.1170	NA	NA	-0.4300	0.2508
	-1	NA	0.2369	NA	0.1282	NA	NA	-0.1630	0.1417
	0	NA	0.0897	NA	0.2153	NA	NA	0.0199	0.1796
	1	NA	-0.0634	NA	0.1249	NA	NA	0.3780	0.0783
	2	NA	-0.2044	NA	0.0332	NA	NA	0.4946	-0.1182

Table D-5: average residual values within subtests for factor 1 of the 5% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.1960	NA	0.0074	NA	NA	0.1397	0.2339
	-1	NA	0.1264	NA	0.0848	NA	NA	0.1397	0.2340
	0	NA	0.2031	NA	0.1596	NA	NA	0.2109	0.3713
	1	NA	0.1464	NA	0.1014	NA	NA	0.1616	0.2928
	2	NA	0.1633	NA	0.0100	NA	NA	0.2204	0.2562
Mixture	-2	NA	0.0247	NA	0.0478	NA	NA	0.0873	0.0434
	-1	NA	0.0715	NA	0.0529	NA	NA	0.0474	0.0179
	0	NA	0.0256	NA	0.0251	NA	NA	0.0247	0.0166
	1	NA	-0.0167	NA	0.0256	NA	NA	-0.0733	-0.0163
	2	NA	-0.0370	NA	0.0436	NA	NA	-0.1187	0.0101
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	0.0518	0.0603	NA	0.0132	NA	NA	0.1741	0.0662
	-1	0.1654	0.0504	NA	0.1189	NA	NA	0.1539	0.1641
	0	0.0659	0.1577	NA	0.2305	NA	NA	0.2620	0.0869
	1	0.1475	0.1758	NA	0.2052	NA	NA	0.2805	0.1025
	2	0.1665	0.2711	NA	0.0400	NA	NA	0.2862	0.1065
Mixture	-2	0.2643	0.4166	NA	0.1591	NA	NA	0.2759	0.3772
	-1	0.2708	0.3122	NA	0.2154	NA	NA	0.2968	0.1957
	0	0.0335	0.1461	NA	0.2644	NA	NA	0.2959	0.0847
	1	-0.0410	-0.0618	NA	0.1023	NA	NA	-0.0252	0.0409
	2	-0.1209	-0.1924	NA	-0.0801	NA	NA	-0.2214	-0.2355

Table D-6: average residual values within subtests for factor 2 of the 100% contamination condition.

RANGE	1X1		3X3		3X1		1X3	
SUBTEST	10	20	10	20	10	20	10	20
Rasch1								
-2	0.15	0.17	0.00	0.00	0.01	0.01	0.14	0.13
-1	0.15	0.18	0.06	0.09	0.12	0.07	0.22	0.12
0	0.16	0.22	0.21	0.20	0.18	0.18	0.20	0.18
1	0.21	0.20	0.09	0.10	0.07	0.07	0.20	0.17
2	0.14	0.16	0.02	0.00	0.01	0.01	0.14	0.11
Rasch2								
-2	0.12	0.13	0.02	0.01	0.07	0.14	0.02	0.01
-1	0.11	0.17	0.12	0.09	0.09	0.20	0.08	0.04
0	0.21	0.22	0.28	0.17	0.18	0.22	0.15	0.17
1	0.13	0.20	0.12	0.07	0.16	0.15	0.05	0.11
2	0.20	0.17	0.01	0.00	0.09	0.14	0.01	0.01

Table D-7: average residual values within subtests for factor 2 of the 100% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	0.0673	0.0452	NA	NA	NA	NA	0.0617	0.0668
	-1	0.0851	0.0497	NA	NA	NA	NA	0.0689	0.0876
	0	0.1057	0.0421	NA	NA	NA	NA	0.0627	0.0697
	1	0.0410	0.0351	NA	NA	NA	NA	0.0688	0.0703
	2	0.0646	0.0489	NA	NA	NA	NA	0.0482	0.0413
Mixture	-2	-0.0126	0.0175	NA	NA	NA	NA	-0.0129	0.0104
	-1	0.0176	0.0594	NA	NA	NA	NA	-0.0091	0.0227
	0	-0.0168	0.0195	NA	NA	NA	NA	0.0437	0.0295
	1	0.0218	0.0042	NA	NA	NA	NA	0.0172	0.0185
	2	0.0173	0.0237	NA	NA	NA	NA	-0.0109	0.0262
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	NA	NA	NA	NA	NA	NA	NA
	-1	NA	NA	NA	NA	NA	NA	NA	NA
	0	NA	NA	NA	NA	NA	NA	NA	NA
	1	NA	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA	NA
Mixture	-2	NA	NA	NA	NA	NA	NA	NA	NA
	-1	NA	NA	NA	NA	NA	NA	NA	NA
	0	NA	NA	NA	NA	NA	NA	NA	NA
	1	NA	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA	NA

Table D-8: average residual values within subtests for factor 2 of the 50% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	0.0679	0.1399	NA	NA	NA	NA	0.0989	0.0836
	-1	0.0781	0.0841	NA	NA	NA	NA	0.0738	0.0630
	0	0.0759	0.0985	NA	NA	NA	NA	0.0664	0.1112
	1	0.0953	0.0799	NA	NA	NA	NA	0.1144	0.0881
	2	0.0942	0.1154	NA	NA	NA	NA	0.1286	0.0824
Mixture	-2	0.0583	0.0279	NA	NA	NA	NA	0.0036	0.0239
	-1	0.0251	0.0379	NA	NA	NA	NA	0.0758	0.0769
	0	0.0378	0.0126	NA	NA	NA	NA	0.0271	-0.0036
	1	-0.0014	-0.0153	NA	NA	NA	NA	-0.0266	-0.0033
	2	-0.0368	0.0081	NA	NA	NA	NA	0.0325	-0.0386
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.0672	NA	0.0464	NA	NA	0.1567	0.0829
	-1	NA	0.0303	NA	0.0558	NA	NA	0.0800	0.1051
	0	NA	0.0489	NA	0.2024	NA	NA	0.1522	0.1175
	1	NA	0.0450	NA	0.1394	NA	NA	0.1160	0.1118
	2	NA	0.0637	NA	0.0531	NA	NA	0.1597	0.1001
Mixture	-2	NA	-0.0264	NA	0.0913	NA	NA	0.0224	0.1306
	-1	NA	0.0077	NA	0.1694	NA	NA	0.0252	0.0912
	0	NA	0.0753	NA	0.2013	NA	NA	0.1541	0.1110
	1	NA	0.1070	NA	0.1194	NA	NA	0.0442	0.0547
	2	NA	0.1200	NA	0.0204	NA	NA	-0.0096	0.0187

Table D-9: average residual values within subtests for factor 2 of the 20% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.1200	NA	0.0201	NA	NA	0.1362	0.0756
	-1	NA	0.1095	NA	0.0475	NA	NA	0.0980	0.0876
	0	NA	0.0949	NA	0.0852	NA	NA	0.1823	0.0795
	1	NA	0.0920	NA	0.1143	NA	NA	0.1300	0.0832
	2	NA	0.0729	NA	0.0083	NA	NA	0.1430	0.0458
Mixture	-2	NA	0.0043	NA	0.0637	NA	NA	0.0609	0.0368
	-1	NA	0.0156	NA	0.0043	NA	NA	-0.0242	0.0420
	0	NA	-0.0070	NA	0.0225	NA	NA	-0.0068	0.0333
	1	NA	0.0030	NA	0.0390	NA	NA	-0.0598	0.0388
	2	NA	-0.0085	NA	0.0397	NA	NA	-0.0776	0.0069
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.0805	NA	0.0358	NA	NA	0.1670	0.2095
	-1	NA	0.1183	NA	0.0992	NA	NA	0.1812	0.2179
	0	NA	0.1177	NA	0.3031	NA	NA	0.2093	0.2721
	1	NA	0.0879	NA	0.2032	NA	NA	0.1613	0.2553
	2	NA	0.1032	NA	0.0236	NA	NA	0.1172	0.2250
Mixture	-2	NA	0.1163	NA	0.0312	NA	NA	-0.0058	0.1640
	-1	NA	0.0932	NA	0.1945	NA	NA	0.0961	0.1777
	0	NA	0.0871	NA	0.3225	NA	NA	0.1937	0.2723
	1	NA	0.0151	NA	0.2109	NA	NA	0.0500	0.1416
	2	NA	-0.0421	NA	0.0401	NA	NA	0.0380	-0.0607

Table D-10: average residual values within subtests for factor 2 of the 5% contamination condition.

	TYPE	RANDOM							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	NA	0.1262	NA	0.0182	NA	NA	0.1004	0.1115
	-1	NA	0.0820	NA	0.0562	NA	NA	0.0582	0.0747
	0	NA	0.1037	NA	0.1673	NA	NA	0.1447	0.1494
	1	NA	0.1428	NA	0.1188	NA	NA	0.1302	0.1306
	2	NA	0.1303	NA	0.0089	NA	NA	0.0906	0.0951
Mixture	-2	NA	0.0663	NA	0.0340	NA	NA	0.0380	0.0036
	-1	NA	-0.0205	NA	0.0175	NA	NA	0.0589	0.0174
	0	NA	0.0003	NA	0.0520	NA	NA	0.0312	-0.0011
	1	NA	-0.0617	NA	0.0446	NA	NA	-0.0629	0.0066
	2	NA	0.0199	NA	0.0101	NA	NA	0.0299	0.0093
	TYPE	REVERSE							
	RANGE	1X1		3X3		3X1		1X3	
	SUBTEST	10	20	10	20	10	20	10	20
Rasch	-2	0.0422	0.1214	NA	0.0206	NA	NA	0.0911	0.2936
	-1	0.0537	0.0781	NA	0.1169	NA	NA	0.1027	0.2420
	0	0.0885	0.1046	NA	0.2445	NA	NA	0.1594	0.3598
	1	0.0615	0.0749	NA	0.1044	NA	NA	0.1528	0.3163
	2	0.1184	0.0833	NA	0.0151	NA	NA	0.1416	0.2653
Mixture	-2	0.1550	0.1373	NA	0.0388	NA	NA	0.2020	0.1080
	-1	0.1414	0.1195	NA	0.1103	NA	NA	0.1601	0.1920
	0	0.0464	0.0903	NA	0.2398	NA	NA	0.1461	0.3781
	1	0.0146	0.0373	NA	0.1548	NA	NA	-0.0024	0.1542
	2	-0.0139	-0.0280	NA	0.0038	NA	NA	-0.1862	-0.0118

Appendix E

The following is an annotated version of the Macro BAY_MIRT used to generate data.

```
/*****************************************************************
The SAS Macro entitled BAY_MIRT generates data submits the data to
WINBUGS calls back pertinent information into SAS and begins the
analysis process for factor analysis. Definitions of pre Macro
structuring of the file.
location    = is a macro variable for the ROOT folded location for
files in the Macro.
WINB        = Location of WINbugs on the computer.
DOS         =Easy fix to work with DOS commands, no spaces
ITEMNUM           =    number of items
REPS        =   number of replications within each replication
FRSTREP     =   first replication, typically 1.
SAMPLESZ    =   number of cases per replication
SAMPLEC     =   used to add commas to WINbugs code
SAMPLET1    =   used for output of WINbugs commas
SAMPLET2    =   for output of WINbugs data lines
libname is location of MIXIRT on current computer
*****************************************************************
****/

*OPTIONS NONOTES;
%LET LOCATION           =     C:\NEW2OUTPUT;
%LET WINB               = C:\WinBUGS14;
%LET DOS                = C:\SASWINBUGS;
%LET    ITEMNUM         =   40;
%LET    REPS            =   50;
%LET  FRSTREP           =     1;
%LET    SAMPLESZ  =   500;
%LET    SAMPLEC         =   %EVAL(&SAMPLESZ+1);
%LET    SAMPLET1  =   %EVAL(&SAMPLEC+27);
%LET    SAMPLET2  =   %EVAL(&SAMPLEt1+1);
%Let  EFFECT            =     0;
Libname MIXIRT "&LOCATION";

/*****************************************************************
/*GENERATES THE BASIC TEST TO BE MANIPULATED IN THE SIMULATION*/
/*these values are changed based on the range in the code*/

DATA MIXIRT.ITEM_DIFF;
    INPUT   ITEM1-ITEM&ITEMNUM;
DATALINES;
2 2 1 1 0 0 -1 -1 -2 -2 2 2 1 1 0 0 -1 -1 -2 -2 2 2 1 1 0 0 -1 -1 -2
-2 2 2 1 1 0 0 -1 -1 -2 -2
;
RUN;

/* generates WINBbugs CODE. MATCH SAMPLE SIZE if it is changed!*/
```

```
DATA MIXIRT.IRTDATA;
    INPUT LD $80.;
        ARRAY ITEM[&ITEMNUM]; /*FORMAT DATASET FOR BASE*/
            DO I = 1 TO &ITEMNUM;
            END;
CARDS;
list(N=500, I=40, G=2, alpha=c(2,2),
class=c(
1,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA),
r=structure(.Data=c(
RUN;

DATA MIXIRT.DIM;
LD = "), .Dim = c(&SAMPLESZ, &ITEMNUM))";
RUN;

DATA MIXIRT.SET1;
      INPUT RANDOM RASCH;
CARDS;
0 1
;
RUN;

/*******************************************************************/
/*generates the Rasch code to run in WINbugs*/
DATA model;
INPUT model $80.;
CARDS;
model
{

   for( j in 1 : N ) {
      for( k in 1 : I ) {
```

```
p1[j,k] <- exp(theta[j]-b[k,class[1]])/(1+exp(theta[j]-
b[k,class[1]]))
p2[j,k] <- 0.25
p[j,k] <- p2[j,k]*prop1[j]+p1[j,k]*(1-prop1[j])
r[j,k] ~ dbern(p[j , k])
      }
   }
   for( k in 1 : I ) {
      for( c in 1 : 1 ) {
         b[k , c] ~ dnorm( 0.0,0.25)
      }
   }
   for( j in 1 : N ) {
      theta[j] ~ dnorm( 0.0,tau)
      class[j] ~ dcat(pi[1:G])
      prop1[j] <- class[j] - 1
   }
   pi[1:G] ~ ddirch(alpha[])
   tau ~ dgamma( 0.5,1)
}

;
run;

/*OUTPUTS IRT TO TXT FILE FOR WINBUGS*/
DATA _NULL_;
SET model;
FILE "&LOCATION..\sas to winbugs\IRTMODEL.txt";
PUT model;
RUN;

/*MERGE WITH TERTA FOR EASY */
DATA MIXIRT.TMERGE;
    DO REPS = 1 TO &ITEMNUM;
        DO P = 1 TO &ITEMNUM;
            OUTPUT;
        END;
    END;
RUN;

/**********BEGIN BAY_MIRT MACRO******************/
MACRO BAY_MIRT;
/* BC changes the proportion of the contamination in the
contaminated subtest, 1 = no contamination, 0 = all contamination */
      %DO BC= 2 %TO 2;
            %If &BC=1 %THEN %Do;
                  %LET        BRASCH      = 0;
                  %LET        RAS   = 1;
            %END;
            %If &BC=2 %THEN %Do;
                  %LET        BRASCH      = .5;
                  %LET        RAS   = .5;
            %END;
            %If &BC=3 %THEN %Do;
                  %LET        BRASCH      = .8;
                  %LET        RAS   = .2;
            %END;
```

```
            %If &BC=4 %THEN %Do;
                  %LET        BRASCH      = .95;
                  %LET        RAS   = .05;
            %END;
            %If &BC=5 %THEN %Do;
                  %LET        BRASCH      = 1;
                  %LET        RAS   = 0;
            %END;
/* MIX = Type of contamination 1 random, 2 reversed difficulties */
      %DO MIX=1 %TO 1;
            %If &MIX=1 %THEN %Do; %Let    MIX     = 1;    %END;
            %If &MIX=2 %THEN %Do; %Let    MIX     = 2;    %END;
/* TIR  changes the scaling factor for the test range. */
      %DO TIR=3 %TO 3;
            %If &TIR=1 %THEN %Do;
                        %Let    IRANGE          =     1;
                        %Let  SUBRANGE    =     1;
            %END;
            %If &TIR=2 %THEN %Do;
                        %Let    IRANGE          =     3;
                        %Let  SUBRANGE    =     3;
            %END;
            %If &TIR=3 %THEN %Do;
                        %Let    IRANGE          =     3;
                        %Let  SUBRANGE    =     1;
            %END;
            %If &TIR=4 %THEN %Do;
                        %Let    IRANGE          =     1;
                        %Let  SUBRANGE    =     3;
            %END;
/* WC changes the size of the contaminated subtest*/
      %DO WC=1 %TO 1;
            %If &WC=1 %THEN %Do;
                  %LET    WRS       = 30;
                  %LET  WNRB  = 31;
            %END;
                %If &WC=2 %THEN %Do;

                  %LET  WRS   = 20;
                  %LET  WNRB  = 21;
            %END
/*NRP can be used to have fully contaminated replicantes(not used)*/
      %DO NRP=1 %TO 1;
            %If &NRP=1 %THEN %Do;   %Let   PROP      = 1;   %END;
/* Can be used to change the mean of theta*/
      %DO T=4 %TO 4;
            %If &T=4 %THEN %Do; %Let    THETA = 0;  %END;
/*NUMBER OF REPLICATIONS*/
      %DO R = &FRSTREP %TO &REPS;
/*****************************************************************
Macro statements used for do loops
BPR   = THE PROPORTION OF BETWEEN DATA THAT IS ALL RASCH DATA. USED
TO END FIRST DO LOOP
BPNR  = THE PROPORTION OF BETWEEN DATA THAT HAS NON-RASCH DATA. USED
TO BEGIN DO LOOP
BNRPROP     = THE PROPORTION OF DATA IN NON-RASCH SECTION THAT IS
ALL NON RASCH, ENDS THIS LOOP
```

```
BNRPROP     = THE PROPORTION OF DATA IN NON-RASCH SECTION THAT IS
SPLIT RASCH/NON RASCH, BEGINS FINAL LOOP
*****************************************************************

%LET BPR          = INT(&BRASCH*&SAMPLESZ);
%LET BPNR         = &BPR+1;
%LET BNRPROP      =
INT(&BRASCH*&SAMPLESZ)+INT(&RAS*&SAMPLESZ*&PROP);
%LET BNRPROPB     = &BNRPROP+1;

/*BEGIN DATA GENERATION*/
/*Rescales the test by the scaling factor to generate ranges for
subtests*/
DATA ITEM_DIFF_TEST;
    SET MIXIRT.ITEM_DIFF;
        ARRAY ITEM[&ITEMNUM];
            DO I = 1 TO &WRS;
                ITEM[I]=ITEM[I]*&IRANGE;
                  END;
            DO I = &WNRB TO &ITEMNUM;
                ITEM[I]=ITEM[I]*&SUBRANGE;
            END;
RUN;
/*ENDS TRANDFORMATION OF TEST*/

DATA MIXIRT.THETA&BC&MIX&TIR&WC&NRP&T&R;
    SET ITEM_DIFF_TEST;
            DO RESPONSET1 = 1 TO &BPR;
                  DIST = RANNOR(0);
               ARRAY ITEM[&ITEMNUM];
               ARRAY PROB[&ITEMNUM];
                  DO P = 1 TO &ITEMNUM;
                  PROB[P] =   (EXP((&THETA+DIST)-
ITEM[P]))/(1+(EXP(((&THETA+DIST)-ITEM[P]))));
                END;
                    OUTPUT;
            END;
    IF &MIX = 1 THEN DO;
        DO RESPONSET1 =  &BPNR TO &BNRPROP;
            DIST = RANNOR(0);
                DO P = 1 TO &WRS;
                    PROB[P] = (EXP((&THETA+DIST)-
ITEM[P]))/(1+(EXP(((&THETA+DIST)-ITEM[P]))));
                END;
                DO P = &WNRB TO &ITEMNUM;
                    PROB[P] = 0.25;
                END;
                    OUTPUT;
        END;
            DO RESPONSET1 =  &BNRPROPB TO &SAMPLESZ;
                DO P = 1 TO &ITEMNUM;
                    PROB[P] = 0.25;
                END;
                    OUTPUT;
        END;
    END;
    IF &MIX = 2 THEN DO;
```

```
           DO RESPONSET1 =  &BPNR TO &BNRPROP;
                  DIST = RANNOR(0);
                DO P = 1 TO &WRS;
                    PROB[P] = (EXP((&THETA+DIST)-
ITEM[P]))/(1+(EXP(((&THETA+DIST)-ITEM[P]))));
                END;
                DO P = &WNRB TO &ITEMNUM;
                    PROB[P] = (EXP(((&THETA+DIST)-
&EFFECT)+ITEM[P]))/(1+(EXP((((&THETA+DIST)-&EFFECT)+ITEM[P]))));
                END;
             OUTPUT;
        END;
            DO RESPONSET1 =  &BNRPROPB TO &SAMPLESZ;
                  DIST = RANNOR(0);
                       DO P = 1 TO &ITEMNUM;
                    PROB[P] = (EXP(((&THETA+DIST)-
&EFFECT)+ITEM[P]))/(1+(EXP((((&THETA+DIST)-&EFFECT)+ITEM[P]))));
                END;
                    OUTPUT;
        END;
    END;
RUN;

/*FROM THE PROBABLITIES, A RESPONSE SET FOR 500 PEOPLE OVER THE
TWENTY ITEM TEST IS GENERATED*/
DATA MIXIRT.THETA&BC&MIX&TIR&WC&NRP&T&R;
    SET MIXIRT.THETA&BC&MIX&TIR&WC&NRP&T&R;
        ARRAY ITEM[&ITEMNUM];
        ARRAY PROB[&ITEMNUM];
                DO T = 1 TO &ITEMNUM;
                    ITEM[T]     = (RANBIN(0,1,PROB[T]));
                END;
            DROP  T I P RESPONSET1 DIST;
RUN;

/********CODE TO GENERATE TXT DATA FOR WINBUGS***/
DATA IRTDATA&BC&MIX&TIR&WC&NRP&T&R;
    SET MIXIRT.IRTDATA;
RUN;
PROC APPEND
    BASE=IRTDATA&BC&MIX&TIR&WC&NRP&T&R
DATA=MIXIRT.THETA&BC&MIX&TIR&WC&NRP&T&R FORCE;
RUN;
DATA IRTDATA&BC&MIX&TIR&WC&NRP&T&R;
    SET IRTDATA&BC&MIX&TIR&WC&NRP&T&R;
COM=",";
IF (_N_ > 1 AND _N_ < &SAMPLET1) THEN COM2=",";
RUN;

PROC APPEND
    BASE=IRTDATA&BC&MIX&TIR&WC&NRP&T&R DATA=MIXIRT.DIM FORCE;
RUN;

/******************CODE TO format data for WINbugs **************/
DATA _NULL_;
    SET IRTDATA&BC&MIX&TIR&WC&NRP&T&R;
FILE "&LOCATION.\sas to winbugs\IRTDATA&BC&MIX&TIR&WC&NRP&T&R..txt";
```

```
PUT  LD;
IF (_N_ > 28 AND _N_ <&SAMPLET2) THEN PUT @1
(ITEM1 COM ITEM2 COM ITEM3 COM ITEM4 COM ITEM5 COM ITEM6 COM ITEM7
COM ITEM8 COM ITEM9 COM ITEM10 COM
ITEM11 COM ITEM12 COM ITEM13 COM ITEM14 COM ITEM15 COM ITEM16 COM
ITEM17 COM ITEM18 COM ITEM19 COM ITEM20 COM
ITEM21 COM ITEM22 COM ITEM23 COM ITEM24 COM ITEM25 COM ITEM26 COM
ITEM27 COM ITEM28 COM ITEM29 COM ITEM30 COM
ITEM31 COM ITEM32 COM ITEM33 COM ITEM34 COM ITEM35 COM ITEM36 COM
ITEM37 COM ITEM38 COM ITEM39 COM ITEM40 COM2) (&ITEMNUM*1.);
RUN;

/******************CODE TO run winbugs ****/
DATA _NULL_;
/*File location to house this .txt file called batchirt*/
/*INITS ARE USED IF INITS ARE SET*/
/*DIC NOT USED WITH MIXTURE*/
FILE "&WINB.\BatchIRT.txt";
PUT // @@
#1 "display('log')"
#2 "check('&LOCATION.\sas to winbugs/IRTMODEL.txt')"
#3 "data ('&LOCATION.\sas to
winbugs/IRTDATA&BC&MIX&TIR&WC&NRP&T&R..txt')"
#4 "compile(1)"
/*
#5 "inits(1,'&LOCATION.\sas to winbugs/INIT1.txt')"
#6 "inits(2,'&LOCATION.\sas to winbugs/INIT2.txt')"
#7 "inits(3,'&LOCATION.\sas to winbugs/INIT3.txt')"
*/
#8 "gen.inits()"
#9 "update(2000)"
#10 "set(theta)"
#11 "set (b)"
/*#12 "set(p)"*/
/*#13 "set (class)"*/
#14 "set (prop1)"
#15 "set (pi)"
/*#15 "dic.set()"*/
#16 "update(5000)"
#18 "stats(*)"
/*#19 "dic.stats()"*/
#20 "gr()"
/*#21 "coda(*,'&LOCATION.\sas to winbugs/codairt.txt')"*/
#22 "save('&LOCATION.\sas to
winbugs/bugslog&BC&MIX&TIR&WC&NRP&T&R..txt')"
#23 "quit()"
;
RUN;

DATA _NULL_;
FILE "&DOS.\runIRT.bat";
PUT "CD &WINB";
PUT "WinBUGS14.exe /PAR BatchIRT.txt";
PUT "exit";
RUN;

DATA _NULL_;
```

```
X "&DOS.\runIRT.bat";
RUN;
QUIT;

/*end not generate*/

DATA MIXIRT.BUGSLOG&BC&MIX&TIR&WC&NRP&T&R;
INFILE "&LOCATION.\sas to
winbugs\bugslog&BC&MIX&TIR&WC&NRP&T&R..txt" FIRSTOBS=20
DELIMITER="09"x OBS =11570;
INPUT  node $ mean   sd  MCerror    TWONHALF    median
NINETYSEVENNHALF    start   sample;
RUN;

DATA PROP;
    SET MIXIRT.BUGSLOG&BC&MIX&TIR&WC&NRP&T&R;
    PR = index(NODE, "prop1");
    IF PR > 0;
    RANDOM  =   MEAN;
    RASCH   =   1-MEAN;
    KEEP RANDOM RASCH;
RUN;

DATA PROP;
      SET MIXIRT.SET1 PROP;
RUN;

DATA MIXIRT.FINAL&BC&MIX&TIR&WC&NRP&T&R;
    MERGE MIXIRT.THETA&BC&MIX&TIR&WC&NRP&T&R PROP;
      ITEMSUM = SUM(of ITEM1-ITEM&ITEMNUM)/&ITEMNUM;
RUN;
                                            %END;
                                        %END;
                                    %END;
                                %END;
                            %END;
                        %END;
    %END;


%MEND BAY_MIRT;
```

References

Andreassen, S., Woldbye, M., Falck, B., & Andersen, S.K. (1987). MUNIN: A causal probabilistic network for interpretation of electromyographic findings. *Proceedings of the 10th International Joint Conference on Artificial Intelligence* (pp. 366-372). Milan: Kaufmann.

Bohm, D. (1980). *Wholeness and the implicate order.* London, England: Routledge.

Brooks, S. P., & Gelman, A. (1998). Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, **7**, 434-455.

Capra, F. (1982) *The turning point.* New York, NY: Bantam Books.

Chen, T., & Davison, M. L. (1996) A multidimensional scaling, paired comparisons approach to assessing unidimensionality in the Rasch model. In Wilson, M. & Engelhard, G. (Eds.), *Objective measurement: theory into practice* (Vol.3). Stanford, CT: Ablex.

Chung, H., Loken, E., & Schafer, J. (2004) Difficulties in drawing inferences with finite-mixture models: A simple example with a simple solution. *The American Statistician*, 58, 2, 152-158.

Chyn, S., Tang, K. L., & Way, W. D. (1995) Investigation of IRT-based assembly of TOEFL test. (ETS Research Report TR-9.) Princeton, New Jersey: Educational Testing Service.

Davier, M., & Carstensen, C. (2007) *Multivariate and mixture distribution Rasch models: extensions and applications.* New York, NY: Springer

Dayton, C. M. (1998). *Latent class scaling analysis.* Quantitative Applications in the

Social Sciences Series No. 126. Thousand Oaks, CA: Sage Publications.

Dayton, C. M., & Macready, G. B. (1980) Scaling model with response errors and intrinsically unscalable respondents. *Psychometrika*, *45*, 3, 343-356

Dayton, C. M., & Macready, G. B. (2007) "Latent Class Analysis in Psychometrics." In Rao, C. R. & Sinharay, S. (Eds) *Handbook of Statistics, 26,* 421-446, Amsterdam, Netherlands: Elsevier Science Pub Co.

Einstein, A. (1961). *Relativity. The special and the general theory* (Lawson, R. W. trans.). New York, NY: Crown Trade Paperbacks (1916).

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004*). Bayesian data analysis* (2nd Ed.) Boca Raton, FL: Champion and Hall/CRC.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99-129). Hillsdale, NJ: Lawrence Erlbaum Associates.

.Goodman, L.A., (1975). A new model for scaling response patterns: an application of the Quasi independence concept. *Journal of the American Statistical Association*, 70, 755-768.

Gorsuch, R. L. (1983). *Factor analysis* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Hoaglin, D. C., Mosteller, F. & Tukey, K. W., eds. (1985b*). Understanding Robust and Exploratory Data Analysis.* New York: Wiley.

Hoaglin, D. C. (2003). John W. Tukey and data analysis. *Statistical Science*, 18, 3, 311-318

Kant, I. (1961). *Critique of pure reason.* (Smith, N. K., trans.). New York, NY: St Martin's Press (Original work published 1781).

Kelderman, H. (1984). Loglinear Rasch model tests. *Psychometrica*,49,2 223-245.

Kelderman, H., & Macready, G. (1990) The use of loglinear models for assessing differential item functioning across manifest and latent examinee groups. *Journal of Educational Measurement*. 27,4, 307-327.

Kuhn, T. S. (1996). *The structural of scientific revolution* (3rd ed.). Chicago, IL: The University of Chicago Press.

Kolen, M. J., & Brennan, R. L. (1995). *Test equating: Methods and practices*. NewYork: Springer-Verlag

Lunn, D.J., Thomas, A., Best, N., & Spiegelhalter, D. (2000) WinBUGS -- A Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10:325--337.

Mead, R. (1976). Fit of data to the Rasch model through the analysis of residuals. Doctoral dissertation. University of Chicago.

Minium, E. W., King, B. M., & Bear, G. (1993) *Statistical reasoning in psychology and education* (3nd ed.) New York, NY: John Wiley & Sons, Inc.

Mislevy, R.J. (1986). Bayes model estimates in item response models. *Psychometrika*, 51, 177-195.

Mislevy, R.J., & Chang, H. (2000). Does adaptive testing violate local independence? *Psychometrika*, 65, 2 149-156.

Mislevy, R. J., & Verhelst, N. (1990). Modeling item responses when different subjects employ different solution strategies. *Psychometrika*, 55, 195-215.

Mislevy, R.J., & Wilson, M.R. (1996). Marginal maximum likelihood estimation for a psychometric model of discontinuous development. *Psychometrika*, 61, 41-71.

Mosteller, F., & Tukey, J. W. (1977), *Data Analysis and Regression: A Second Course in Statistics.* Addison-Wesley. Reading, MA.

Muthén, B. (2008). Latent variable hybrids: Overview of old and new models. In Hancock, G. R., & Samuelsen, K. M. (Eds.), *Advances in latent variable mixture models,* 1-24. Charlotte, NC: Information Age Publishing, Inc.

Rupp, A. (2003) Item Response Modeling With BILOG-MG and MULTILOG for Windows. *International Journal of Testing,* 3, 4, 365-384.

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement*, 14, 271-282.

Smith, R. M. (1986) Person fit analysis with the Rasch model. *Educational and Psychological Measurement.* June, 46, 359-372

Smith, R. M. (1988) The distributional properties of Rasch standardized residuals. *Educational and psychological measurement, 48*, 657-667

Smith, R. M. (1991) The distributional properties of Rasch item fit statistics. *Educational and Psychological Measurement* 51, 541-565

Spiegelhalter, D. J., Thomas, A., Best, N. G., & Lunn, D. (2003). WinBUGS version 1.4: user manual. Cambridge Medical Research Council Biostatistics Unit. Retrieved from http://www.mrc-bsu.cam.ac.uk/bugs/

Tabachnick B., & Fidell L. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper and Row.

Wainer, H., & Wright, B. (1980) Robust estimation of ability in the Rasch model *Psychometrika*, *45*, 373-391

Wilson, M. (1989). Saltus: A psychometric model of discontinuity in cognitive development. *Psychological Bulletin*, 105(2), 276-289.

Wright, B.D. (1995) Diagnosing person misfit. *Rasch Measurement Transactions*, 9(2), p430-1

Wright B. D., & Tennant A. (1996) Sample size again. *Rasch Measurement Transactions 9*:4, 468

Yamamoto, K. (1989) HYBRID model of IRT and latent class models. (ETS Research Report RR-89-41.) Princeton, New Jersey: Educational Testing Services.

Yamamoto, K. (1995) Estimating the effects of Test Length and Test Time on Parameter Estimation Using the HYBRID model (ETS Research Report RR-95-2.) Princeton, New Jersey: Educational Testing Service.

Zimowski, M., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). BILOG-MG 3: Item analysis and test scoring with binary logistic models. Chicago, IL: Scientific Software.