

Conceptual Notes on Models for Discrete Polytomous Item Responses

Gideon J. Mellenbergh

University of Amsterdam

The following types of discrete item responses are distinguished: nominal-dichotomous, ordinal-dichotomous, nominal-polytomous, and ordinal-polytomous. Bock (1972) presented a model for nominal-polytomous item responses that, when applied to dichotomous responses, yields Birnbaum's (1968) two-parameter logistic model. Applying Bock's model to ordinal-polytomous items leads to a conceptual problem. The ordinal nature of the response variable must be preserved; this can be achieved using three different methods. A number of existing models are derived using these three methods. The structure of these models is similar, but they differ in the interpretation and qualities of their parameters. Informa-

tion, parameter invariance, log-odds differences invariance, and model violation also are discussed. Information and parameter invariance of dichotomous item response theory (IRT) also apply to polytomous IRT. Specific objectivity of the Rasch model for dichotomous items is a special case of log-odds differences invariance of polytomous items. Differential item functioning of dichotomous IRT is a special case of measurement model violation of polytomous IRT. *Index terms: adjacent categories, continuation ratios, cumulative probabilities, differential item functioning, log-odds differences invariance, measurement model violation, parameter invariance, polytomous IRT models.*

Items requiring discrete responses frequently are used in psychological and educational measurement. These items can use nominal or ordinal response categories, and can have either two response categories (dichotomous) or more than two response categories (polytomous). Combining these properties yields four types of discrete item responses: nominal-dichotomous (ND), ordinal-dichotomous (OD), nominal-polytomous (NP), and ordinal-polytomous (OP). An example of a discrete ND item is having a respondent select one of two political parties; an example of an OD item is an examinee's answer to an achievement test item, which is evaluated as "correct" or "incorrect;" an example of an NP item is having a respondent answer a question on a psychiatric interview, which is classified according to personality type; and an example of an OP item is an examinee's essay score, which is graded A, B, C, or D.

Models for polytomous items are presented here. The discussion is restricted to models for items that measure one continuous latent variable (i.e., unidimensional latent trait models). Only models with parametric item response functions are considered, which excludes nonparametric models such as Molenaar's (1982) generalization of Mokken's (1970) model. Models with a guessing parameter, such as Birnbaum's (1968) three-parameter logistic model and Thissen & Steinberg's (1984) model, are excluded. The discussion is on a conceptual level and addresses differences and similarities between the different item response models.

Nominal-Polytomous Item Responses

In an NP item, a person's answer is classified into exactly one of K nominal categories. It is assumed that the distribution of the person's responses over the K categories is multinomial. Person j 's probability of responding to the k th category of the i th item is denoted by τ_{ijk} .

Bock (1972) presented the following model for these probabilities:

$$\tau_{ijk} = \exp(b_{ik} + a_{ik}\theta_j) / \sum_{k=1}^K \exp(b_{ik} + a_{ik}\theta_j), \tag{1}$$

where θ_j is person j 's value on a continuous latent trait, and a_{ik} and b_{ik} are parameters for the k th category of the i th item. The model in Equation 1 must be constrained. Therefore, Bock (1972) set the sums of the within-item parameters equal to 0. For convenience here, the parameters of the first category of each item are set equal to 0; that is,

$$a_{i1} = b_{i1} = 0, \quad i = 1, 2, \dots, n, \tag{2}$$

where n is the number of items.

Equation 1 is transformed by conceptually splitting the nominal response variable with K categories into a series of $(K - 1)$ dichotomous variables. Because the scale type is nominal, one category can be selected arbitrarily for reference. For convenience, the first category is selected and the k th and the first category are combined into one dichotomous variable. In this way the K -category nominal response variable is split into $(K - 1)$ dichotomous variables. Figure 1 shows a four-category ($K = 4$) nominal response variable split into three dichotomous variables.

Figure 1
A Four-Category Nominal Variable Split Into Three Dichotomous Variables

Four-Category Variable	Three Dichotomous Variables		
	Category 2 versus 1	Category 3 versus 1	Category 4 versus 1
τ_{ij1}	τ_{ij1}	τ_{ij1}	τ_{ij1}
τ_{ij2}	τ_{ij2}		
τ_{ij3}		τ_{ij3}	
τ_{ij4}			τ_{ij4}

The odds of selecting the k th category versus the first category are defined as the ratio of their probabilities— (τ_{ijk}/τ_{ij1}) . The log odds is defined as the natural logarithm of the odds— $\ln(\tau_{ijk}/\tau_{ij1})$. If the log odds equals 0, the person does not prefer one category over the other. If the log odds is larger than 0, the person tends to prefer the k th category over the first; and if the log odds is smaller than 0, the person tends to prefer the first category over the k th.

Bock's model (Equations 1 and 2) can be reformulated in terms of $(K - 1)$ log odds,

$$\ln(\tau_{ijk}/\tau_{ij1}) = b_{ik} - b_{i1} + (a_{ik} - a_{i1})\theta_j = b_{ik} + a_{ik}\theta_j, \quad k = 2, 3, \dots, K, \tag{3}$$

where the last term follows from the constraint in Equation 2. The model in Equation 3 consists of $(K - 1)$ simultaneous linear equations and is equivalent to the model in Equations 1 and 2.

The graphical representation of Equation 3 consists of $(K - 1)$ straight lines. b_{ik} is the intercept (i.e., the value of the log odds at $\theta = 0$) and a_{ik} is the slope of the line (i.e., the rate of increase of the log odds). a_{ik} can be interpreted as the discrimination between the k th and the first category. At the point where the line intersects the θ axis,

$$\ln(\tau_{ijk}/\tau_{ij1}) = b_{ik} + a_{ik}\theta_j = 0. \tag{4}$$

Equation 4 shows that the line intersects the θ axis at the point $\theta = -b_{ik}/a_{ik}$. At the point of intersection, the log odds is 0 (i.e., the person does not prefer the k th category over the first). For θ values larger than the point of intersection, the person tends to prefer the k th category over the first; for θ values smaller than this point the person tends to prefer the first category. Therefore, the point of intersection can be interpreted as the boundary between the k th and the first category.

Ordinal-Polytomous Item Responses

Bock's (1972) model for nominal item responses also can be used to construct models for ordinal item responses. However, the ordinal nature of the response categories must be preserved. For OD responses, the application is straightforward. The highest category (e.g., the correct answer) is the second category ($k = 2$) and the lowest category (e.g., the incorrect answer) is the first ($k = 1$) category. The result is:

$$\ln(\tau_{ij2}/\tau_{ij1}) = b_{i2} + a_{i2}\theta_j = a_i(\theta_j - b_i), \tag{5}$$

where $a_i = a_{i2}$ and $b_i = -b_{i2}/a_i$. Equation 5 is equivalent to Birnbaum's (1968) two-parameter logistic model (2PLM).

For dichotomous items, Bock's (1972) model (Equations 1 and 2) yields the same result for nominal and ordinal item responses. To use Bock's model for polytomous items, the ordinal nature of the response categories must be preserved.

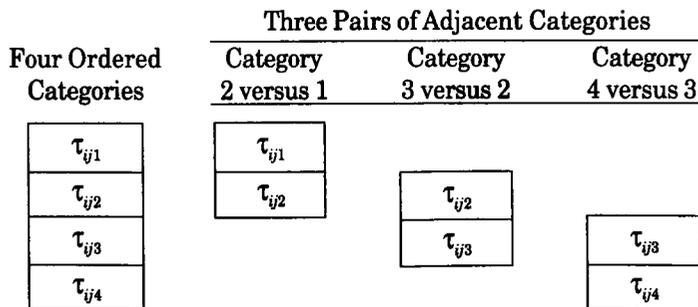
The order of the response categories is preserved by using contiguous categories or groups of categories. This is achieved by using one of three types of methods to split the OP response variable into dichotomies (Agresti, 1990, sec. 9.3), which yields the following types of models: (1) adjacent-category models, (2) cumulative probability models, and (3) continuation-ratio models. The forms of the models are comparable, but the interpretation of the parameters is very different (Molenaar, 1983). For OD items, the differences between the three types of models disappear and the interpretational complexity of the parameters does not exist.

Adjacent-Category Models

In adjacent-category models, the K categories of the ordinal response variable are split using $(K - 1)$ adjacent categories. Figure 2 shows a four-category ordinal variable split into three adjacent category pairs. The three dichotomies are formed using contiguous categories; thus, the ordinal nature of the four categories is preserved.

A general, adjacent-category model is obtained by applying Bock's (1972) model (Equations 1 and 2) to the log odds of the adjacent k th and $(k + 1)$ th categories,

Figure 2
 A Four-Category Ordinal Variable Split Into Three Adjacent Pairs



$$\ln(\tau_{ijk+1}/\tau_{ijk}) = b_{ik+1} - b_{ik} + (a_{ik+1} - a_{ik})\theta_j = b'_{ik} + a'_{ik}\theta_j, \quad k=1, 2, \dots, K-1, \quad (6)$$

where the prime is used to indicate that the parameters have different values and interpretations than those of Equation 3, and where

$$b'_{ik} = b_{ik+1} - b_{ik} \quad (7)$$

and

$$a'_{ik} = a_{ik+1} - a_{ik}. \quad (8)$$

Equation 6 contains the intercept b'_{ik} , which is specific for the i th item and the k th category. The intercept can be restricted by decomposing it into separate item and category parameters; that is,

$$b'_{ik} = d_i + e_k, \quad \sum_{k=1}^{K-1} e_k = 0, \quad (9)$$

where d_i and e_k are parameters for the i th item and k th category, respectively.

Some special cases of Equations 6 and 9 are well known. For example, assuming that within an item the slope parameters are equal yields Muraki's (1992) generalized partial credit model:

$$\ln(\tau_{ijk+1}/\tau_{ijk}) = b'_{ik} + a'_i\theta_j, \quad k=1, 2, \dots, K-1. \quad (10)$$

Combining Equations 9 and 10 yields Muraki's (1992) rating scale generalized partial credit model. Assuming that the slope parameters within and between items are equal yields Masters' (1982) partial credit model:

$$\ln(\tau_{ijk+1}/\tau_{ijk}) = b'_{ik} + a'\theta_j, \quad k=1, 2, \dots, K-1. \quad (11)$$

Combining Equations 9 and 11 yields Andrich's (1978) rating scale partial credit model.

Cumulative Probability Models

For cumulative probability models, the K categories of the ordinal variable are split into $(K-1)$ cumulative probabilities. Figure 3 shows a four-category ordered variable split into three cumulative probabilities. The ordinal nature of the variable is preserved by using contiguous groups of categories.

The upward cumulative probabilities are defined as:

$$\tau_{ijk}^* = \tau_{ij1} + \dots + \tau_{ijk}, \quad k=1, 2, \dots, K-1. \quad (12)$$

Figure 3
A Four-Category Ordinal Variable Split Into Three Cumulative Probabilities
Three Cumulative Probabilities

Four Ordered Categories	Categories 2, 3 and 4 versus 1	Categories 3 and 4 versus 1 and 2	Category 4 versus 1, 2 and 3
τ_{ij1}	$\tau_{ij1}^* = \tau_{ij1}$	$\tau_{ij2}^* =$	$\tau_{ij3}^* =$
τ_{ij2}	$\tau_{ij2} +$	$\tau_{ij1} + \tau_{ij2}$	$\tau_{ij1} +$
τ_{ij3}	$\tau_{ij3} +$	$\tau_{ij3} +$	$\tau_{ij2} + \tau_{ij3}$
τ_{ij4}	τ_{ij4}	τ_{ij4}	τ_{ij4}

Applying Bock's (1972) model (Equations 1 and 2) to the log odds of the downward and upward cumulative probabilities yields

$$\ln\left[\frac{1 - \tau_{ijk}^*}{\tau_{ijk}^*}\right] = b''_{ik} + a''_{ik}\theta_j, \quad k = 1, 2, \dots, K - 1, \tag{13}$$

where the double prime is used to indicate that the parameters have other values and interpretations than those of Equation 3 and of the adjacent-category model (Equation 6).

The slope parameters of Equation 13 must be constrained. A property of the upward cumulative probabilities (Equation 12) is that τ_{ijk+1}^* is larger than or equal to τ_{ijk}^* for all values of θ . Therefore, the straight lines of the model in Equation 13 cannot intersect for any value of θ ; that is, the slope parameters of the categories within an item must be equal. This necessary restriction yields

$$\ln\left[\frac{1 - \tau_{ijk}^*}{\tau_{ijk}^*}\right] = b''_{ik} + a_i''\theta, \quad k = 1, 2, \dots, K - 1. \tag{14}$$

Equation 14 is equivalent to the homogeneous case of Samejima's (1969) graded response model. She also described a heterogeneous case (Samejima, 1972), which is not a cumulative probability model. In the heterogeneous case of Samejima's graded response model, the category slope parameters need not be equal within an item. Combining Equations 9 and 14 yields Muraki's (1990) rating scale version of Samejima's model.

Continuation-Ratio Models

The third way of preserving the ordinal nature of the response variable is splitting the K categories into $(K - 1)$ continuation ratios. The three continuation ratios for a four-category ordered variable are given in Figure 4. The ordinal nature of the variable is preserved by using one contiguous category and a group of categories.

Figure 4
A Four-Category Ordinal Variable Split Into Three Continuation Ratios

Four Ordered Categories	Three Continuation Ratios		
	Categories 2, 3 and 4 versus 1	Categories 3 and 4 versus 2	Category 4 versus 3
τ_{ij1}	τ_{ij1}		
τ_{ij2}	$\tau_{ij2} +$	τ_{ij2}	
τ_{ij3}	$\tau_{ij3} +$	$\tau_{ij3} +$	τ_{ij3}
τ_{ij4}	τ_{ij4}	τ_{ij4}	τ_{ij4}

Applying Bock's (1972) model (Equations 1 and 2) to the log odds of continuation ratios yields

$$\ln\left[\frac{1 - \tau_{ijk}^*}{\tau_{ijk}^*}\right] = b'''_{ik} + a'''_{ik}\theta_j, \quad k = 1, 2, \dots, K - 1. \tag{15}$$

where the triple prime is used to indicate that the parameters have other values and interpretations than those of Equation 3, the adjacent-category models, and the cumulative probability models.

Assuming that the slope parameters within and between items are equal yields Tutz' (1990) sequential model,

$$\ln\left[\frac{1 - \tau_{ijk}^*}{\tau_{ijk}^*}\right] = b'''_{ik} + a'''\theta_j, \quad k = 1, 2, \dots, K - 1. \tag{16}$$

Combining Equations 9 and 16 yields the rating scale version of Tutz' model.

Discussion

The parameters of the three types of models have different values and interpretations. The b parameter is the intercept and the a parameter is the slope of the linear model for the log odds; however, the odds are defined differently for the three types of models. In adjacent-category models the odds of the $(k + 1)$ th and the k th category are used; in cumulative probability models the last $(K - k)$ categories and the first k categories are collapsed, and the odds of the combined last $(K - k)$ categories and the combined first k categories are used; and in continuation-ratio models the last $(K - k)$ categories are collapsed and the odds of the combined last $(K - k)$ categories and the k th category are used.

The log odds is 0 for the three type of models at the point at which the straight lines intersect the θ axis. Therefore, the intersection can be interpreted as a boundary, but the interpretation differs between the three types of models. In adjacent-category models the intersection point is the value of θ at which the probability of responding to the $(k + 1)$ th category equals the probability of responding to the k th category; thus, the intersection point is the boundary between the given $(k + 1)$ th and k th category. In cumulative probability models, the intersection is the value of θ at which the probability of responding to the $(K - k)$ highest categories equals the probability of responding to the k lowest categories. It is the boundary between the collapsed $(K - k)$ highest categories and the collapsed k lowest categories. In continuation-ratio models, the intersection is the value of θ at which the probability of responding to the $(K - k)$ highest categories equals the probability of responding to the k th category. It is the boundary between the given collapsed $(K - k)$ highest categories and the k th category.

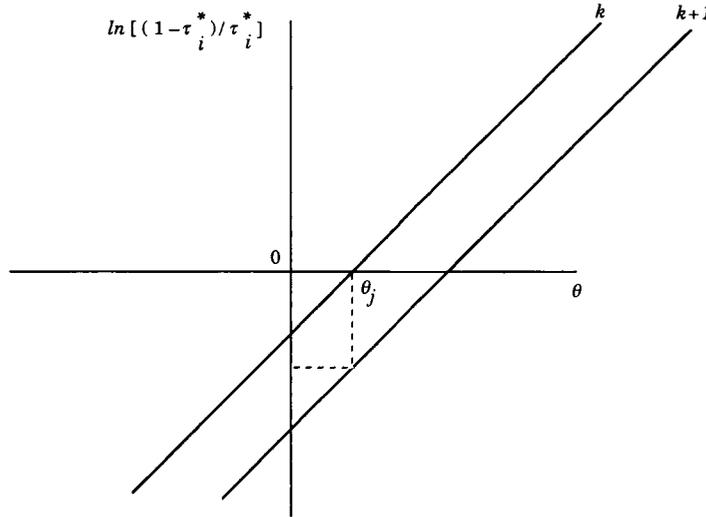
The cumulative probability models differ in two respects from the adjacent-category and continuation-ratio models. First, the use of cumulative probabilities implies that slope parameters of different categories within an item must be constrained to be equal (see Equation 14). Although it may be of no practical use, theoretically it is not necessary to constrain within-item slope parameters in the other two types of models.

A second difference between cumulative probability models and the other two types of models concerns the boundaries. Suppose that person j 's θ is at the boundary where the k th line intersects the θ axis. For the cumulative probability model, using Equation 14 it follows that $1 - \tau_{ijk}^* = \tau_{ijk}^*$, and using Equation 12 it follows that $(1 - \tau_{ijk+1}^*) \leq \tau_{ijk+1}^*$. Thus, for the j th person and $(k + 1)$ th category, Equation 14 is equal to or smaller than 0. Therefore, the $(k + 1)$ th straight line coincides or is to the right of the k th line (see Figure 5). Therefore, in cumulative probability models the boundaries are in the same order as the item categories. Using the adjacent-category model (Equation 6), it follows that at the boundary $\tau_{ijk+1} = \tau_{ijk}$, which does not imply anything about the relationship between τ_{ijk+2} and τ_{ijk+1} . At the boundary, the $(k + 1)$ th line can be to the left or to the right of the k th line. Using the continuation-ratio model (Equation 15), it follows that at the boundary $1 - \tau_{ijk}^* = \tau_{ijk}^*$, which does not imply anything about the relationship between $1 - \tau_{ijk+1}^*$ and τ_{ijk+1} . At the boundary, the $(k + 1)$ th line can be to the left or to the right of the k th line. Therefore, in adjacent-category and continuation-ratio models the boundaries are not necessarily ordered in the same way as the item categories. All three types of models preserve the ordinal nature of the response variable, but only cumulative probability models yield ordered boundaries.

The adjacent-category models differ in the following respect from the other two types of models. In adjacent-category models, Bock's (1972) model (Equations 1 and 2) is applied directly to two adjacent categories. In the other two types of models, the categories must be collapsed. Therefore, Bock's model is applied to newly defined probabilities.

For dichotomous items, where $K = 2$, the differences between the three types of models disappear. The models in Equations 6, 14, and 15 reduce to Birnbaum's (1968) 2PLM (Equation 5).

Figure 5
 Graphical Representation of the k th and $(k+1)$ th Line of the Model in Equation 14 Demonstrating That $(1 - \tau_{jk+1}^*) < \tau_{jk+1}^*$ Implies That the $(k+1)$ th Line Is to the Right of the k th Line



Generalizations of IRT Measurement Concepts

In dichotomous IRT models the concept of information is used to assess measurement precision. Information is defined as the inverse of the variance of the maximum likelihood estimate, $\hat{\theta}$, at θ (see, e.g., Hambleton & Swaminathan, 1985, sec. 5.4). This definition of information also can be applied to the polytomous item response models given in Equations 3, 6, 14, 15, and their special cases.

In Birnbaum’s (1968) 2PLM (Equation 5) for dichotomous items, the a and b parameters are specific to the item and do not depend on the person; in this sense the item parameters are person invariant. The θ parameter is specific for the person and does not depend on the item; in this sense the person parameter is item invariant. These properties also apply to the parameters of polytomous item response models (Equations 3, 6, 14, 15, and their special cases). The a and b parameters are specific for the item and are person invariant, and the θ parameter is specific for the person and is item invariant.

The Rasch (1960) model for dichotomous item responses is the special case of Birnbaum’s (1968) 2PLM (Equation 5), where the discrimination parameters of the items are set equal to each other (i.e., $a_i = a$). The Rasch model has the property of “specific objectivity.” A measurement is specifically objective if (1) the comparison of two persons does not depend on the item, and (2) the comparison of two items does not depend on the person (Fischer, 1987, p. 567). This concept can be generalized for nominal item responses, but also can be similarly generalized for adjacent-category, cumulative probability, and continuation-ratio models. If, in Equation 3, the a parameters are constrained to be equal between items (i.e., $a_k = a_{k'}$), the log-odds differences of two different persons, j and j' , are

$$\ln(\tau_{ijk} / \tau_{ij1}) - \ln(\tau_{ij'k} / \tau_{ij'1}) = a_k(\theta_j - \theta_{j'}), \quad k = 2, 3, \dots, K, \tag{17}$$

which do not depend on the item. The log-odds differences of two different items, i and i' , are

$$\ln(\tau_{ijk} / \tau_{ij1}) - \ln(\tau_{i'jk} / \tau_{i'j1}) = b_{ik} - b_{i'k}, \quad k = 2, 3, \dots, K, \tag{18}$$

which do not depend on the person. Therefore, under the constraint $a_{ik} = a_k$ the log-odds differences of two different persons are item invariant and the log-odds differences of two different items are person invariant. In other words, in using log-odds differences the comparison of two persons does not depend on the item and the comparison of two items does not depend on the person. Masters' partial credit model (Equation 11; $a'_{ik} = a'$) and Tutz' sequential model (Equation 16; $a'''_{ik} = a'''$) have this property.

Measurement Violations

The concept of differential item functioning (DIF) is used in IRT for dichotomous items. The usual definition of DIF is that the a and/or b parameters differ between groups (see, e.g., Hambleton & Swaminathan, 1985, sec. 13.2). If both the a and b parameters differ between groups, DIF is nonuniform; if only the b parameter differs between the groups, DIF is uniform. In this definition, group membership is an observed nominal variable, such as gender or ethnicity. Kelderman (1989) noted that group membership also can be a latent nominal variable, such as the latent class variable of curriculum mastery or nonmastery.

The concept of DIF can be generalized to the concept of measurement violation. Birnbaum's (1968) 2PLM (Equation 5) is a measurement model. If the model is misspecified and does not fit the data, it can be extended by one or more other variables called violators (Oort, 1992). The violators can be observed variables or latent variables that are either nominal or quantitative. Examples of observed and latent nominal violators are gender and curriculum mastery/nonmastery, respectively. Examples of observed and latent quantitative violators are age and the latent trait of social desirable answering, respectively.

The concept of violation also can be applied to polytomous item responses. If the models in Equations 3, 6, 14, and 15 are misspecified, they can be extended with one or more violators. The violator model is discussed for one violator, denoted by V , and the nominal item response model in Equation 3. The corresponding violator models for Equations 6, 14, 15, and their special cases are constructed in the same way and the extension of the model to more than one violator is straightforward.

The general one-violator model using Equation 3 is obtained by extending the model with the violator and the product of the violator and the latent trait,

$$\ln(\tau_{ijk}/\tau_{ij1}) = b_{ik} + a_{ik}\theta_j + c_{ik}v_j + d_{ik}\theta_j v_j, \quad k = 2, 3, \dots, K, \quad (19)$$

where c_{ik} and d_{ik} are parameters specific for the i th item and k th category, and v_j is person j 's value on the violator. In the case of a latent continuous variable V , the violator must be standardized with, for example, mean 0 and variance 1 to identify the model. If $c_{ik} \neq 0$ and $d_{ik} \neq 0$, the model is the nonuniform one-violator model; if $c_{ik} \neq 0$ and $d_{ik} = 0$, the model is the uniform one-violator model for NP item responses. For dichotomous items ($K = 2$) and an observed nominal violator, such as gender, where $V_j = 0$ for males and $V_j = 1$ for females, the violator model reduces to

$$\ln(\tau_{ij2}/\tau_{ij1}) = b_{i2} + a_{i2}\theta_j \quad (20)$$

for males and

$$\ln(\tau_{ij2}/\tau_{ij1}) = b_{i2} + a_{i2}\theta_j + c_{i2} + d_{i2}\theta_j = b_{i2}^* + a_{i2}^*\theta_j \quad (21)$$

for females, where $b_{i2}^* = b_{i2} + c_{i2}$ and $a_{i2}^* = a_{i2} + d_{i2}$. Equations 20 and 21 are equivalent to Birnbaum's (1968) 2PLM with different parameters for males and females, which is the conventional DIF model.

The one-violator model in Equations 20 and 21 for dichotomous items permits a large number of specifications. The violation can be uniform or nonuniform, the violator can be an observed or latent variable, and the violator can be nominal or quantitative. The one-violator model in Equation 19 for polytomous items yields an extra number of possible specifications. The ordinal response variable can be split into

order-preserving dichotomies in three ways (adjacent categories, cumulative probabilities, and continuation ratios); and the $(K - 1)$ c parameters and the $(K - 1)$ d parameters can be free parameters or can be restricted (e.g., $c_{ik} = c_i$ or $c_{ik} = c$ and $d_{ik} = d_i$ or $d_{ik} = d$).

Conclusions

Bock's (1972) model is a one-dimensional latent trait model for NP item responses. Models for OP item responses can be derived from it. The ordinal nature of the item response variable can be preserved by using adjacent-category, cumulative-probability, or continuation-ratio splittings. The corresponding models have a similar structure, but the interpretation of the parameters is different. It seems plausible that the type of model used to preserve the ordinal nature should be determined by features of the item and by the cognitive processes involved in answering the item; however, the literature does not provide much support for preferring one model over another. For dichotomous items, Bock's model for NP items and the three models for OP items all reduce to the same model—Birnbaum's (1968) 2PLM. Therefore, the difference in parameter interpretation does not exist for dichotomous items.

The concepts of person invariance of item parameters, item invariance of the person parameter, and information of dichotomous IRT also apply to polytomous IRT. Under the constraint that the a parameters of polytomous item response models are equal between items, the log-odds differences of two different items are person invariant and the log-odds differences of two different persons are item invariant. For the Rasch model for dichotomous items these properties are equivalent to the property of specific objectivity.

The concept of DIF of dichotomous IRT can be generalized. DIF implies that the original measurement model is misspecified and that the model must be extended with one or more violators. In DIF the violators are usually observed or latent nominal variables, but violators also can be observed or latent quantitative variables. If polytomous item response models are misspecified they also can be extended with a violator of a different type (observed or latent, and nominal or quantitative). For dichotomous items and an observed-nominal violator, the one-violator model for polytomous items reduces to the conventional concept of DIF. Finally, the general one-violator model for polytomous items is complex because it permits a large number of different specifications.

References

- Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561–573.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (pp. 397–424). Reading MA: Addison-Wesley.
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, *37*, 29–51.
- Fischer, G. H. (1987). Applying the principles of specific objectivity and of generalizability to the measurement of change. *Psychometrika*, *52*, 565–587.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer-Nijhoff.
- Kelderman, H. (1989). Item bias detection using log linear IRT. *Psychometrika*, *54*, 681–697.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149–174.
- Mokken, R. J. (1970). *A theory and procedure of scale analysis*. The Hague, The Netherlands: Mouton.
- Molenaar, I. W. (1982). Mokken scaling revisited. *Kwantitatieve Methoden*, *3*, 145–164.
- Molenaar, I. W. (1983). *Item steps* (Heymans Bulletin HB-83-630-EX). Groningen, The Netherlands: University of Groningen, Vakgroep Statistiek en Meettheorie FSW.
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, *14*, 59–71.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, *16*, 159–176.
- Oort, F. J. (1992). Using restricted factor analysis to detect item bias. *Methodika*, *6*, 150–166.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish

- Institute for Educational Research. Expanded edition, University of Chicago Press, 1980.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*, No. 17.
- Samejima, F. (1972). A general model for free-response data. *Psychometric Monograph*, No. 18.
- Thissen, D., & Steinberg, L. (1984). A response model for multiple-choice items. *Psychometrika*, 49, 501–519.
- Tutz, G. (1990). Sequential item response models with an ordered response. *British Journal of Mathemat-*

cal and Statistical Psychology, 43, 39–55.

Acknowledgments

The author thanks Bert van Engelenburg, Henk Kelderman, Frans Oort, and Ferry Reith for their valuable comments.

Author's Address

Send requests for reprints or further information to Gideon J. Mellenbergh, Faculty of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. Internet: ml_dorman@macmail.psy.uva.nl.