Measurement and Research Department Reports 2003-3

The Componential Nedelsky Model: A First Exploration

Timo Bechger
Gunter Maris

Citogroep

Measurement and Research Department Reports 2003-3

THE COMPONENTIAL NEDELSKY MODEL: A FIRST EXPLORATION

Timo Bechger Gunter Maris

CITO, NATIONAL INSTITUTE FOR EDUCATIONAL MEASUREMENT

ARNHEM

Citogroep
Arnhem, March 3, 2003

Abstract

We describe a variant of the Nedelsky model for multiple choice items. The Nedelsky model is based upon the simple idea that subjects reject some of the alternative answers and then draw their response from the remaining ones. The same idea provides the foundation for the componential Nedelsky model. Here, a multiple choice item is considered as a stimulus with known characteristics that the respondent is expected to recognize. The options give interpretations of the stimulus.

1. Introduction

The Nedelsky model (NM) model is based upon the idea that a test-taker responds to a multiple choice (MC) question by first eliminating the answers he recognizes as wrong and then guesses at random from the remaining answers. In this paper, we continue with the development of the NM by making more explicit assumptions about the answer process. We assume that a MC item consists of: a stimulus, an instruction to respond, and a number of alternative answers that each give an interpretation of the stimulus. The stimulus can take many forms; e.g., a picture, a speech fragment, a newspaper article, etc. but any stimulus may be characterized by a number of key properties or "content elements" (CEs) that respondents must recognize in order interpret the stimulus correctly. Each alternative answer provides an interpretation of the stimulus. The correct alternative has all relevant properties of the stimulus while item writers have been careful to ensure that the incorrect alternatives lack one or more. A respondent who has recognized all CEs is certain to choose the correct answer.

Consider, for example, an item showing a picture of a car on a traffic circle with his right blinker on. Respondents are instructed to choose the alternative that they believe to give a correct interpretation of the situation. Response alternatives may include, for instance, a) the driver may turn off the traffic circle, b) the driver should give priority to the cyclist, etc. Items of this kind are common in traffic examinations. A similar situation arises in examinations for wine tasters, where the stimulus would be say a wine produced in the valley of the Rio Duero and examinees are supposed to recognized its particular taste.

In this paper we describe the componential Nedelsky model and variants of the model including models suitable for opinion questions. The model was inspired by section 3.2 in the thesis of Javier Revuela (2000) who in turn was inspired by the

Nedelsky model.

2. A Componential Nedelsky Model

The componential Nedelsky model (CNM) will be developed as a variant of the NM. We will first give the basic ingredients of the NM. Consider a MC item i with $J_i + 1$ options arbitrarily indexed $0, 1, ..., J_i$. For convenience, 0 indexes the correct alternative. Let the random variable S_{ij} indicate whether alternative j is recognized as wrong, and define $\mathbf{S}_i$ by the vector $(S_{i0}, S_{i1}, ..., S_{iJ_i})$. We refer to $\mathbf{S}_i$ as *a latent subset.* The random variable $S_i^+ \equiv \sum_{j=1}^{J_i} S_{ij}$ denotes the number of distractors that are exposed.

In the NM, the solution process is assumed to consist of two stages. In the first stage, a respondent eliminates the answers he recognizes to be wrong. Formally, this means that he draws a latent subset. Once a latent subset is chosen, a respondent guesses at random from the remaining answers. Thus, the conditional probability of responding with option j to item i, given latent subset $\mathbf{s}_i$, is given by:

$$\Pr(X_i = j | \mathbf{S}_i = \mathbf{s}_i) = \frac{1 - s_{ij}}{v(s_i^+)}, \tag{1}$$

where $X_i = j$ denotes the event that the respondent chooses alternative j, and $v(s_i^+) \equiv \sum_{h=0}^{J_i}(1 - s_{ih})$ the number of alternatives to choose from. Combining the two stages of the answer process, we find that the conditional probability of choosing option j with item i is equal to

$$\Pr(X_i = j | \theta) = \sum_{\mathbf{s}_i} \frac{1 - s_{ij}}{v(s_i^+)} \Pr(\mathbf{S}_i = \mathbf{s}_i | \theta) \tag{2}$$

We now relate the content structure of the items to the latent subset. The content structure of an item can be represented by means of a *structure matrix* $\mathbf{T}_i$. The rows of the structure matrix represent the alternatives and the relevant content elements are represented by its columns. One is free to re-order the rows and columns of the structure matrix. For convenience, we number the *rows*, $\mathbf{t}_{ij}$, of the structure matrix from 0 to J_i so that each row corresponds to the alternative with the same index. For any $j \in \{0, ..., J_i\}$ and $f \in \{0, ..., c_i\}$, the (j, f) entry of $\mathbf{T}_i$, represented by t_{ijf}, equals 1 if CE f is missing in alternative j, and 0 otherwise. Assume, for instance, that there are five alternatives, that the stimulus has four relevant CEs, and that:

$$\mathbf{T}_i = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \tag{3}$$

The first row represents the correct alternative. It is seen that no CEs are missing in the correct alternative while the distractors systematically lack certain CEs; the last incorrect alternative is seen to be unrelated to the problem. For later reference we refer to an item with this particular structure as a *chain item.* This particular chain item will serve as an example throughout this section. Remember that the structure matrix can take any form. It could, for instance, be assumed that the structure matrix equals $\mathbf{T}_i = (0, 1, \cdots, 1)^t$. This item could be called a *know/don't know item.*

Let $\mathbf{E}_i = (E_{i0}, ..., E_{ic_i})$ represent a latent profile; $E_{ih} = 1$ if a respondent recognizes CE ih, and zero otherwise. Realizations of $\mathbf{E}_i$ are denoted by $\mathbf{e}_i = (e_{i0}, e_{i2}, ..., e_{ic_i})$. The ordering of the CEs is assumed to equal that of the columns of the structure matrix; i.e., both E_{if} and t_{ijf} refer to the CE indexed by if. We assume that a distractor is excluded ($s_{ij} = 1$) if the respondent recognizes any CE that is missing in the alternative ($t_{ijf} = 1$, and $e_{if} = 1$ for some f). Otherwise, the alternative is included. This implies that

$$s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij}) = 1 - \prod_f (1 - e_{if})^{t_{ijf}} \quad . \tag{4}$$

Following Maris (1995), this function is called ***a condensation rule (CR).*** Consider the previous example. A respondent with latent profile $(0,0,1,0)$ excludes the last two alternatives and his latent subset will be $(0,0,0,1,1)$. Note that the same subset corresponds to profile $(0,0,1,1)$. In general, the mapping from profiles to subsets is many-to-one; each profile corresponds to only one subset but different profiles may give rise to the same subset.

It is assumed that

$$\Pr(E_{if} = e_{if}|\theta) = \frac{\exp(\theta - \zeta_{if})^{e_{if}}}{1 + \exp(\theta - \zeta_{if})}. \tag{5}$$

The person parameter θ may be interpreted as the ability to recognize the ith stimulus and similar signals, and the parameter ζ_{if} as the difficulty to recognize the CE indexed by if. It is seen that E_{if} is modelled as a latent Rasch item. Assuming independence among the CEs given θ, the probability that a subject with ability θ chooses any latent profile $\mathbf{e}_i$ is given by the likelihood of c_i independent Rasch items, that is,

$$\Pr(\mathbf{E}_i = \mathbf{e}_i|\theta) = \prod_f \Pr(E_{if} = 1|\theta)^{e_{if}} \left[\Pr(E_{if} = 0|\theta)\right]^{1-e_{if}} \tag{6a}$$

$$= \frac{\exp\left[\theta e_i^+ - \sum_f e_{if}\zeta_{if}\right]}{\prod_f \left[1 + \exp(\theta - \zeta_{if})\right]}, \tag{6b}$$

where $e_i^+ = \sum_f e_{if}$ denotes the number of CEs recognized and $\sum_f e_{if}\zeta_{if}$ may be considered a location parameter for profile $\mathbf{e}_i$. It can be shown that

$$\Pr(S_{ij} = 1|\theta) = 1 - \prod_f \Pr(E_{if} = 0|\theta)^{t_{ijf}} \tag{7}$$

which equals the CR with e_{if} replaced by $\Pr(E_{if} = 1|\theta)$. It follows that $\Pr(S_{i0} = 0|\theta) = 1$; the correct alternative is always in the subset. While this assumption appeared somewhat arbitrarily in the NM, it follows from substantive arguments in the CNM. Note further that a respondent who recognizes all CEs will exclude all distractors and choose the correct alternative with probability 1. In general, item writers should consider whether it is reasonable that a respondent who recognizes all CEs should always choose the correct answer.

Note that the relation between the response mapping and the probability $\Pr(S_{ij} = 1|\theta)$ is not coincidental. It will hold for most, if not all, condensation rules. The reason is that $\Pr(S_{ij} = 1|\theta)$ can be written as $E\left[f(\mathbf{E}_i, \mathbf{t}_{ij})|\theta\right]$; that is, the expectation of a function of the random variables $E_{i1}, .., E_{ic_i}$ with the row of the structure matrix acting as parameters. In general,

$$E\left[f(\mathbf{E}_i, \mathbf{t}_{ij})|\theta\right] = f(E\left[\mathbf{E}_i\right])$$

if the function f is linear in each of its arguments. Most condensation rules can be written in this way. With $c_i = 2$ and all $t_{ijf} = 1$, for instance, the present condensation rule can be written as

$$s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij}) = 1 - \left[1 - e_{i1} - e_{i2} + e_{i1}e_{i2}\right].$$

Since the E_{if}s are independent given θ,

$$E[s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij})|\theta] = 1 - \left[1 - E[E_{i1}|\theta] - E[E_{i2}|\theta] + E[E_{i1}|\theta]E[E_{i2}|\theta]\right].$$

It should be noted that the elements of the subset need not be independent in the CNM as in the NM. Specifically,

$$\Pr(\mathbf{S}_i = \mathbf{s}_i|\theta) = \sum_{\mathbf{e}_i} \Pr(\mathbf{S}_i = \mathbf{s}_i|\mathbf{E}_i = \mathbf{e}_i) \Pr(\mathbf{E}_i = \mathbf{e}_i|\theta), \tag{8}$$

where

$$\Pr(\mathbf{S}_i = \mathbf{s}_i|\mathbf{E}_i = \mathbf{e}_i) = \prod_{j=0}^{J_i} \left\{s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij})^{s_{ij}} \left[1 - s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij})\right]^{1-s_{ij}}\right\} \tag{9}$$

may be called *the product CR*; it is one if the subset corresponds to the profile and zero otherwise.

Finally,

$$\Pr(X_i = j|\theta; \mathbf{T}_i) = \sum_{\mathbf{e}_i} \Pr(X_i = j|\mathbf{E}_i = \mathbf{e}_i; \mathbf{T}_i) \Pr(\mathbf{E}_i = \mathbf{e}_i|\theta) \tag{10}$$

$$= \sum_{\mathbf{e}_i} \frac{1 - s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij})}{\sum_{a=0}^{J_i} (1 - s_{ia}(\mathbf{e}_i, \mathbf{t}_{ia}))} \Pr(\mathbf{E}_i = \mathbf{e}_i|\theta), \tag{11}$$

where $\Pr(X_i = j|\mathbf{E}_i = \mathbf{e}_i; \mathbf{T}_i)$ is the probability of drawing j from the subset corresponding to profile $\mathbf{e}_i$. It is seen that $\left(\sum_{a=0}^{J_i} (1 - s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij}))\right)^{-1}$ must always be positive, that is, respondents must find at least one of the alternatives suitable for inclusion in their latent subset. This is ensured when the structure matrix has a row with unit entries corresponding to the correct alternative. Note that $\Pr(X_i = j|\theta; \mathbf{T}_i)$ has the same form as the corresponding probability in the NM. As remarked in the previous section, the CNM reduces to the NM if $\Pr(S_{ij} = 1|\theta) = \Pr(E_{if} = 1|\theta)$ as, for instance, in a know/don't know item.

Consider again the chain item, with structure matrix (3). The following table contains the conditional probabilities $\Pr(X_i = j|\mathbf{E}_i = \mathbf{e}_i; \mathbf{T}_i)$ for all possible profiles:

Alternatives

Profiles:	0	1	2	3	4
$(1, \cdot, \cdot, \cdot)$	1	0	0	0	0
$(0, 1, \cdot, \cdot)$	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$(0, 0, 1, \cdot)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0
$(0, 0, 0, 1)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	0
$(0, 0, 0, 0)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

It is seen that each possible profile corresponds to one and only one latent subset and that there are only five subsets with positive probability. Thus, compared to the NM, the CNM reduces the number of possible latent subsets.

3. Irrelevant Content Elements

We believe that the structure matrix may be useful for item writers in the sense that it provides a "vocabulary for items"; that is, a systematic means to classify items. We have already used this to distinguish don't know items from chain items. In addition, it could serve as a tool to investigate the properties of items before we administer them to subjects and guide the construction of items. Here is an example of what we have in mind.

Suppose that all entries in the gth column of $\mathbf{T}_i$ are 0 which means that all alternative answers share the corresponding CE. Let $\mathbf{E}_i^{(g)}$ denote the vector $\mathbf{E}_i$ without the g-th element. The shared CE is *irrelevant* in the sense that the respondents choice is independent of recognition of this CE. Specifically

$$\begin{aligned}\Pr(X_i = j|\mathbf{E}_i = \mathbf{e}_i; \mathbf{T}_i) &= \Pr(X_i = j|\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}, E_{ig} = 1; \mathbf{T}_i) \\ &= \Pr(X_i = j|\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}, E_{ig} = 0; \mathbf{T}_i) \quad . \end{aligned} \tag{12}$$

It follows that

$$\Pr(X_i = j|\theta; \mathbf{T}_i) = \sum_{\mathbf{e}_i} \Pr(X_i = j|\mathbf{E}_i = \mathbf{e}_i; \mathbf{T}_i) \Pr(\mathbf{E}_i = \mathbf{e}_i|\theta) \tag{13}$$

$$= \sum_{\mathbf{e}_i^{(g)}} \Pr(X_i = j|\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}; \mathbf{T}_i) \left(\Pr(\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}, E_{ig} = 1|\theta) + \Pr(\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}, E_{ig} = 0|\theta)\right)$$

$$= \sum_{\mathbf{e}_i^{(g)}} \Pr(X_i = j|\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}; \mathbf{T}_i) \Pr(\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}|\theta) \Pr(E_{ig} = 1|\theta) + \Pr(E_{ig} = 0|\theta)$$

$$= \sum_{\mathbf{e}_i^{(g)}} \Pr(X_i = j|\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}; \mathbf{T}_i) \Pr(\mathbf{E}_i^{(g)} = \mathbf{e}_i^{(g)}|\theta).$$

This means that we may delete the gth column of the structure matrix without consequences. In further research we hope to learn more about the role of the structure matrix. We know for instance, that the information about θ is related to the structure matrix but we do not know how.

4. The Interpretation of Profile Elements

The interpretation of the profiles is a bit more intricate than one might think at first sight. This is illustrated by means of the following know/don't know item: $\mathbf{T} = (0, 1)^t$. If E_1 equals 1 we have that S_0 equals zero whereas S_1 equals 1. That is, the respondent knows which of the two alternatives is correct. If E_1 equals zero we have that both S_0 and S_1 equal zero. That is, the respondent doesn't know which of the two alternatives is correct. We see that E_1 equals 1 means that the respondent recognizes the presence of the content element whereas E_1 equal to zero does not mean that the respondent recognizes the absence of the content element but only that he or she doesn't know. This reveals a strong substantive hypothesis that is implied by the CNM. It is not possible for a respondent to be misled to think that a distractor is correct.

5. Variants of the CNM Obtained by Limiting the Number of Possible Profiles

5.1. A Sequential Chain Nedelsky Model

Assume that the item has a chain structure. Assume further that the CEs are sequentially ordered in the sense that a CE is recognized only if the previous CE is recognized. Consider, for example, a chain item with structure matrix (3). Then, the sequential ordering hypothesis implies the following probabilities:

$$\Pr(\mathbf{E}_i = (1,1,1,1)|\theta) = \prod_{f=0}^{3} \Pr(E_{if} = 1|\theta) \tag{14}$$

$$\Pr(\mathbf{E}_i = (0,1,1,1)|\theta) = \Pr(E_{i0} = 0|\theta)) \prod_{f=1}^{3} \Pr(E_{if} = 1|\theta)$$

$$\Pr(\mathbf{E}_i = (0,0,1,1)|\theta) = \Pr(E_{i1} = 0|\theta) \prod_{f=2}^{3} \Pr(E_{if} = 1|\theta)$$

$$\Pr(\mathbf{E}_i = (0,0,0,1)|\theta) = \Pr(E_{i2} = 0|\theta) \Pr(E_{i3} = 1|\theta)$$

$$\Pr(\mathbf{E}_i = (0,0,0,0)|\theta) = \Pr(E_{i3} = 0|\theta)).$$

Note that the only profiles with positive probability are those that correspond to rows in the structure matrix and there is one and only one latent subset that corresponds to each profile. Specifically, the probabilities of latent subsets are now:

$$\Pr(\mathbf{S}_i = (0, t_{ij1}, ..., t_{ijc_i})|\theta) = \Pr(\mathbf{E}_i = \mathbf{1} - \mathbf{t}_{ij}|\theta). \tag{15}$$

In contrast to the model where CEs are independent and all possible profiles have positive probability, the sequential model implies that one needs to know all relevant CEs in order to be certain about the correct answer. It seems natural to expect an ordering of the CE location parameters; the fourth CE is learned first, then, the third, etc.

The sequential hypothesis may also be applied to items that have a different structure. Consider, for instance, the following structure matrix

$$\mathbf{T}_i = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \tag{16}$$

The corresponding item could be called an *incomplete chain item.* Now, the possible profiles and the corresponding latent subsets are:

$$\mathbf{E}_i = (1,1,1) \Leftrightarrow \mathbf{S}_i = (0,1,1) \tag{17}$$

$$\mathbf{E}_i = (0,1,1) \Leftrightarrow \mathbf{S}_i = (0,0,1)$$

$$\mathbf{E}_i = (0,0,1) \Leftrightarrow \mathbf{S}_i = (0,0,1)$$

$$\mathbf{E}_i = (0,0,0) \Leftrightarrow \mathbf{S}_i = (0,0,0)$$

Figure (1) shows the manifest probabilities with the CNM and the sequential model using this incomplete chain item.Figure (1) reveals that the stimulus appears slightly more difficult to recognize when a sequential model is used with the same parameters. The figure also suggest that, in practice, both models cannot be distinguished. We could, however, use the sequential model to have more interpretable parameters.

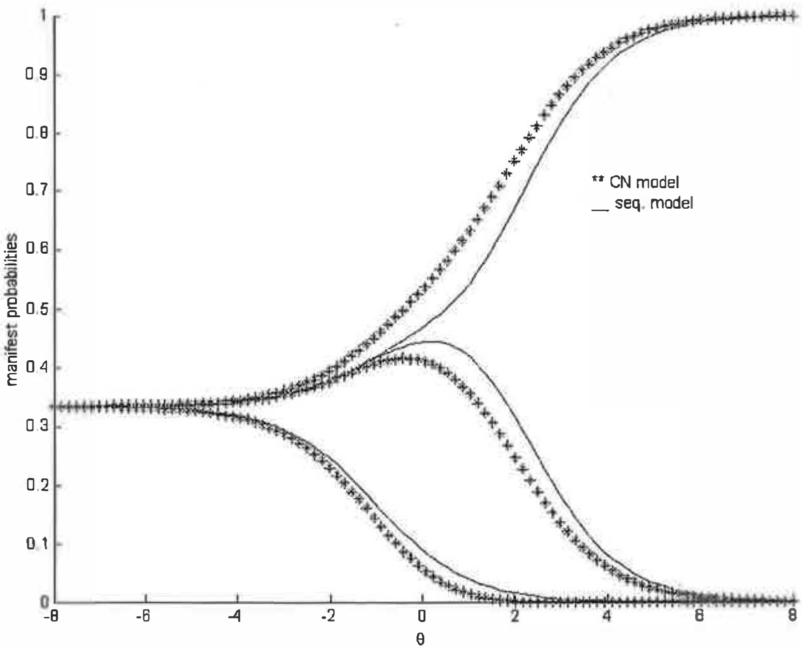

FIGURE 1.

Manifest probabilities $\Pr(X_i = j|\theta)$ under the CNM model with unrestricted probabilities and the sequential model using $\zeta = (2, 1, -1)^t$.

5.2. Bock's Nominal Response Model

Consider an item where the structure matrix is a square matrix of the form

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \tag{18}$$

and each alternative constitutes a unique CE. Hence, there is no correct alternative which reflects the nominal character of the options, and $c_i = J_i + 1$. Further, we condition upon the event that a respondent recognizes one and only one CE. We assume that respondents with other profiles will refuse to answer because none of the alternatives will look suitable to them. It follows from Equation (4) that $s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij}) = 0$ if the jth CE is recognized and 1 otherwise. This means that the probability to choose alternative j corresponds to the probability that a respondent recognizes CE j and none of the others. It will be shown that, conditional upon a response, this peculiar CNM is equivalent to *Bock's (1972) nominal response model (NRM)* if an option-specific discrimination parameter is introduced. Note that $\Pr(S_{ij} = 0|\theta) = \prod_{f \neq j} \Pr(E_{if} = 0|\theta)$ and the model is not formally equivalent to the NM.

First, under the present assumptions:

$$\Pr(X_i = k|\theta) = \frac{\Pr(E_{ik} = 1|\theta) \prod_{f \neq k} \Pr(E_{if} = 0|\theta)}{\sum_{h=0}^{c_i} \Pr(E_{ih} = 1|\theta) \prod_{f \neq h} \Pr(E_{if} = 0|\theta)}. \tag{19}$$

Assume further that the model is extended with option-specific discrimination parameters. That is, $\Pr(E_{if}=0|\theta)=(1+\exp(a_{if}\theta-\zeta_{if}))^{-1}$ so that

$$\begin{aligned}\Pr(E_{ik}=1|\theta)\prod_{f\neq k}\Pr(E_{if}=0|\theta) \qquad\qquad\qquad (20)\\ =\frac{\exp(a_{ik}\theta-\zeta_{ik})}{1+\exp(a_{if}\theta-\zeta_{if})}\prod_{f\neq k}\frac{1}{1+\exp(a_{if}\theta-\zeta_{if})}\\ =\frac{\exp(a_{ik}\theta-\zeta_{ik})}{\prod_f\left[1+\exp(-a_{if}\theta+\zeta_{if})\right]}.\end{aligned}$$

It follows that

$$\begin{aligned}\Pr(X_i=k|\theta)=\frac{\Pr(E_{ik}=1|\theta)\prod_{f\neq k}\Pr(E_{if}=0|\theta)}{\sum_{h=0}^{c_i}\Pr(E_{ih}=1|\theta)\prod_{f\neq h}\Pr(E_{if}=0|\theta)} \qquad (21)\\ =\frac{\frac{\exp(a_{ik}\theta-\zeta_{ik})}{\prod_f\left[1+\exp(-a_{if}\theta+\zeta_{if})\right]}}{\sum_{h=0}^{c_i}\frac{\exp(a_{ih}\theta-\zeta_{ih})}{\prod_f\left[1+\exp(-a_{if}\theta+\zeta_{if})\right]}}\\ =\frac{\exp(a_{ik}\theta-\zeta_{ik})}{\sum_{h=0}^{c_i}\exp(a_{ih}\theta-\zeta_{ih})}\end{aligned}$$

which is the NRM. It is seen that, from the viewpoint of the CNM, the NRM implies strong assumptions on the solution process which might be reasonable for items that measure the respondent's opinion on a particular subject but not for items that measure ability or achievement. Of course, this is not an opportune way to derive the NRM, which can be seen to derive directly from Luce's choice axiom (Luce, 1959) with $\exp(a_{ik}\theta-\zeta_{ik})$ being the attractiveness of the alternative indexed k for a respondent with latent trait value θ. An alternative model CNM for opinion questions will be developed in the next section.

6. A Nominal CNM To Measure Opinions

Consider an item asking the subject to choose among a number of options; for example, Spanish mandarins (Clemenvilla, Clemenules, Satsuma, etc.[1]). If the

[1]Clemenvilla and Clemenules are varieties of the clementina cultivated in Villa-Real and Nules, respectively. Such varieties are obtained by grafting. Satsumas are for export only. The name

respondent has no opinion he may choose the last option stating "none of the above ". Consider an item with four options and structure matrix:

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

The condensation rule is adapted for this specific situation. That is,

$$s_{ij}(\mathbf{e}_i, \mathbf{t}_{ij}) = 1 - \prod_f [1 - e_{if}(1 - t_{ijf})] + \prod_f t_{ijf}(1 - e_{if}). \tag{22}$$

For all options, except for the none-of-the-above option, $\prod_f t_{ijf}(1 - e_{if}) = 0$ and $s_{ij} = 1 - e_j$ which implies that alternative j is taken into consideration if $e_j = 1$ which we interpret as that the respondent agrees with the option. For the last option, $s_{iJ_i} = \prod_f(1 - e_{if})$ and it is chosen if and only if all e_{if} are zero, that is, when the respondent holds no opinion. Both functions are linear in the elements of the profile so that

$$\Pr(S_{ij} = 1|\theta) = 1 - \prod_f [1 - \Pr(E_{ij} = 1|\theta)(1 - t_{ijf})] + \prod_f t_{ijf} \Pr(E_{ij} = 0|\theta)$$

$$= \begin{cases} \Pr(E_{ij} = 0|\theta) \text{ for } j < J_i \\ \prod_f \Pr(E_{if} = 0|\theta) \text{ for } j = J_i \end{cases}$$

The model implies that respondents with multiple opinions will guess among the options they agree with. If they hold no opinion on the subject they will choose the non-of-the above category.

In the case of mandarins, the latent trait would be interpreted as the attitude

mandarins is chosen to remind us of their Chinese origin. The orange apple was important from china around 1900 and the man who important the first tree is given a statue that can be visited in Villa-Real.

towards mandarins. The situation is more complex when the attitude would be bipolar, as in an item asking for preference for political candidates. This can be solved by including an option-specific discrimination parameter. In the case of an item asking for the respondents preference for political candidates the discrimination may be negative if the candidate is known as right-wing and positive if it is a left-wing candidate.

7. Discussion

An analysis with the CNM requires that items have been carefully constructed. We have demonstrated that the content structure of the alternative answers may be expressed schematically using the structure matrix. The structure matrix is a very useful device. It serves to classify items, for example, as know/don't know items, or Nedelsky items, and it serves to see in what ways the correct answer may be found. In short, it helps us to talk about the items. Furthermore, under the CNM the fact that distractor j is chosen provides information about ability above and beyond the information provided by the fact that a distractor is chosen.

Define a constructed response item as a stimulus with an instruction to respond. The CNM is suitable for constructed response items provided the answers can be used to tell which CEs the respondents have recognized, that is, the content structure of the stimulus dictates how the responses should be scored and it is assumed that this can be done without error. If this is the case, we observed the latent profiles. If the number of CE's recognized is used to score the answers the Rasch model can be seen to apply (see Equation (6a)). If it is unlikely that the response can be interpreted without error, the model should be extended to include a "profile recognition step". One might also consider using two independent "judges" to interpret the response and use only the consistent responses to estimate the parameters. We then have an incomplete dataset where each respondent has its own pattern of missing data. This kind of data may subsequently be analyzed using the Rasch model.

A further topic for future research is to consider models with a parametric condensation rule, that is, models where the condensation rule is estimated from the data.

References

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

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New-York: Wiley.

Revuelta, J. (2000). *A psychometric model for multiple choice items.* Doctoral Dissertation Universidad Autonoma de Madrid.