

## **Loss of Information in Estimating Item Parameters in Incomplete Designs**

**Theo J.H.M. Eggen**  
**Norman D. Verhelst**



**Loss of Information in Estimating Item Parameters in  
Incomplete Designs**

Theo J.H.M. Eggen  
Norman D. Verhelst

Citogroep  
Arnhem, 2004

**Cito groep**  
Postbus 1034 6801 MG Arnhem  
Kenniscentrum

8501 005 6334



This manuscript has been submitted for publication. No part of this manuscript may be copied or reproduced without permission.

## **Abstract**

In this paper, the efficiency of conditional maximum likelihood (CML) and marginal maximum likelihood (MML) estimation of the item parameters of the Rasch model in incomplete designs is studied. The use of the concept of F-information (Eggen, 2000) is generalized to incomplete testing designs. The standardized determinant of the F-information matrix is used for a scalar measure of information in a set of item parameters. In this paper, the relation between the normalization of the Rasch model and this determinant is clarified. It is shown that comparing estimation methods with the defined information efficiency is independent of the chosen normalization.

In examples, information comparisons are conducted. It is found that for both CML and MML some information is lost in all incomplete designs compared to complete designs. A general trend is that with increasing test booklet length the efficiency of an incomplete to a complete design and also the efficiency of CML compared to MML is increasing. The main differences between CML and MML is seen in relation to the length of the test booklet. It will be demonstrated that with very small booklets, there is a substantial loss in information (about 35%) with CML estimation, while this loss is only about 10% in MML estimation. However, with increasing test length, the differences between CML and MML quickly disappear.



## 1. Introduction

In the Rasch model, two methods are commonly used to consistently estimate the item parameters: conditional maximum likelihood (CML) and marginal maximum likelihood (MML). In Eggen (2000), the loss of information resulting from the use of CML estimation was studied. The concept of F-information was shown to be useful in quantifying this loss. It was shown that CML estimation normally involves some loss of information with respect to the item parameters compared to using the full likelihood (JML), but the loss is small. Also a small loss was found in comparing the information in CML with MML estimation. The reported efficiencies of CML compared to JML and to MML are larger than 92%.

The results in Eggen (2000) only concern complete testing designs. In the present study the loss of information in incomplete designs will be treated. First, a short review of the concept of F-information and the expressions for CML and MML in the Rasch model will be given. For comparing information matrices and computing the information efficiency, a determinant criterion is used. It will be shown that this criterion has some unique properties. Next, the generalization of the use of the concepts in incomplete designs in general and for the efficiency comparison in particular will be given. Some examples will be used to illustrate the loss of information and the efficiency of CML versus MML in incomplete designs with the Rasch model. A distinction will be made between the possible loss due only to the incompleteness of the design, and the loss due to the design and the estimation method combined.

## 2. F-information and the Rasch model

In Eggen (2000), the F-information concept was used to study the loss of information in models with two (vector) parameters where there is interest in only one of the two parameters and in which, by conditioning or marginalizing, one parameter is in a sense ignored in the inference. The F-information is defined as generalized Fisher information and it expresses the information in a two-parameter distribution with respect to one of the two parameters. If  $p(x; \omega)$  is the two-parameter distribution with  $\omega^T = (\psi^T, \tau^T)$  and  $S_{p; \omega}(X) = \partial \ln p(X; \omega) / \partial \omega$  is the efficient score statistic, then the Fisher information matrix is given by

$$I_p(\omega) = \mathcal{E}[S_{p; \omega} \cdot S_{p; \omega}^T] = \mathcal{E} \begin{bmatrix} S_{p; \psi} & S_{p; \psi}^T & S_{p; \psi} & S_{p; \tau}^T \\ S_{p; \tau} & S_{p; \tau}^T & S_{p; \tau} & S_{p; \psi}^T \end{bmatrix} = \begin{bmatrix} I_p^{\psi\psi^T} & I_p^{\psi\tau^T} \\ I_p^{\tau\psi^T} & I_p^{\tau\tau^T} \end{bmatrix}. \quad (1)$$

The F-information in  $p$  with respect to  $\psi$  is then given by

$$I_p(\psi; \omega) = \mathcal{E} S_{p;\psi} S_{p;\psi}^T - \mathcal{E} S_{p;\psi} S_{p;\tau}^T \cdot \mathcal{E} (S_{p;\tau} S_{p;\tau}^T)^{-1} \cdot \mathcal{E} S_{p;\tau} S_{p;\psi}^T. \quad (2)$$

Next consider the (vector) statistic  $T = T(X)$  with distribution  $g(t; \omega)$ , and write  $p(x; \omega)$  as the product of  $g(t; \omega)$  and the conditional distribution of  $x$  given  $t : f(x|t; \omega)$ . In Eggen (2000), it was shown that in two-parameter models which can be decomposed as  $p(x; \omega) = f(x|t; \psi) \cdot g(t; \omega)$ , the F-information with respect to  $\psi$  in  $p(x; \omega)$  is the sum of the F-information in the conditional model  $f(x|t; \psi)$ , which only depends on  $\psi$ , and the F-information in the model  $g(t; \omega)$ :  $I_p(\psi; \omega) = I_f(\psi; \omega) + I_g(\psi; \omega)$ . It is readily understood that if one uses only the conditional model for inference on  $\psi$  instead of the full model, there is no loss of information if  $I_g(\psi; \omega) = 0$ . Bhapkar (1989) has shown that a sufficient condition for this is that the distribution of  $T$  is at least weakly ancillary with respect to  $\psi$ . In the event the distribution is completely independent of  $\psi$ , the ancillary case, of course also no information is lost.

In the Rasch model for dichotomously scored items, the probability of getting the answer pattern  $x$ , with responses  $X_{vi} = x_{vi}$  (0 or 1) of persons  $v = 1, \dots, n$  on items  $i = 1, \dots, k$ , is given by

$$p(x; \beta, \theta) = \frac{\exp\{-\sum_i (\sum_v x_{vi}) \beta_i + \sum_v (\sum_i x_{vi}) \theta_v\}}{\prod_i \prod_v \{1 + \exp(\theta_v - \beta_i)\}}. \quad (3)$$

In (3),  $\beta^T = (\beta_1, \dots, \beta_k)$  is the item parameter and  $\theta^T = (\theta_1, \dots, \theta_n)$  the person or ability parameter.  $T_v = \sum_i X_{vi}$ , the sum score of a person is sufficient for  $\theta_v$ , for  $v = 1, \dots, n$  and the distribution of  $T$  is given by

$$g(t; \beta, \theta) = \prod_{v=1}^n g(t_v; \beta, \theta_v) = \prod_{v=1}^n \frac{\exp(\theta_v t_v) \cdot \gamma_{t_v}(\beta)}{\prod_{i=1}^k \{1 + \exp(\theta_v - \beta_i)\}}, \quad (4)$$

in which

$$\gamma_t(\beta) = \sum_{\sum y_i = t} \exp(-\sum_{i=1}^k \beta_i y_i) \quad (5)$$



are the so-called basic symmetric functions of order  $t$ . In (5), the summation runs across all answer patterns  $(y_1, \dots, y_k), y_i \in \{0, 1\}$  for which  $\sum_{i=1}^k y_i = t$ . It is easily checked that the conditional distribution of the answer patterns  $x$  given the scores  $t$  is only dependent on the item parameter:  $p(x; \beta, \theta) = f(x|t; \beta) g(t; \beta, \theta)$ ; in CML estimation of the item parameters only the conditional distribution is used.

Weak ancillarity is the key condition for losing no information in estimating the item parameters with CML instead of using the full model. When the model belongs to the exponential family, which the Rasch model (3) clearly does, the fulfillment of this condition is readily checked. (See Bhapkar (1989) and Eggen (2000)).

$T$  is weakly ancillary for  $\beta$  if and only if there exist functions of  $\beta$  only and independent of the data,  $w_j(\beta)$ ,  $j = 1, \dots, k$ , and  $v(\beta)$ , such that:

$$\frac{\partial \ln \gamma_t(\beta)}{\partial \beta_j} = w_j(\beta) \cdot t + v(\beta), \text{ for } j = 1, \dots, k \quad (6)$$

for all  $t$ .

In the Rasch model, this partial derivative is given by (for  $j = 1, \dots, k$ ):

$$\frac{\partial \ln \gamma_t(\beta)}{\partial \beta_j} = - \frac{e^{-\beta_j} \cdot \gamma_{t-1}^{(j)}}{\gamma_t(\beta)}, \quad (7)$$

in which

$$\gamma_{t-1}^{(j)} := \partial \gamma_t(\beta) / \partial e^{-\beta_j}, \text{ for } j = 1, \dots, k. \quad (8)$$

It is well known (Molenaar, 1995) that expression (7) is equal to the conditional probability of answering item  $j$  correctly given the score  $t$ :  $P(X_j = 1 | t)$ . In general if  $t \neq 0$ , rewrite (7) as

$$\frac{\partial \ln \gamma_t(\beta)}{\partial \beta_j} = - \frac{e^{-\beta_j} \cdot \gamma_{t-1}^{(j)}}{\gamma_t(\beta)} \cdot \frac{t}{t} = - \frac{e^{-\beta_j} \cdot \gamma_{t-1}^{(j)}}{\sum_{i=1}^k e^{-\beta_i} \cdot \gamma_{t-1}^{(i)}} \cdot t, \quad (9)$$

in which the recursive formula  $\gamma_t(\beta) \cdot t = \sum_{i=1}^k e^{-\beta_i} \cdot \gamma_{t-1}^{(i)}$  (Fischer, 1974) is used.

Then in expression (6),  $v(\beta) = 0$ , and  $w_j(\beta) = -e^{-\beta_j} \cdot \gamma_{t-1}^{(j)} / \sum_i e^{-\beta_i} \cdot \gamma_{t-1}^{(i)}$ , which is not only a function of the item parameters  $\beta$ , but also dependent on the data,  $t$ . Therefore,  $T$  is in general not weakly ancillary for  $\beta$ .

There is, however, an interesting case which yields weak ancillarity. This is the case if all item parameters are equal:  $\beta_j = \beta, j = 1, \dots, k$ . Then (9) simplifies to:

$$\frac{\partial \ln \gamma_t(\beta)}{\partial \beta_j} = \frac{e^{-\beta} \cdot \gamma_{t-1}^{(j)}}{k \cdot e^{-\beta} \cdot \gamma_{t-1}^{(j)}} \cdot t = \frac{1}{k} \cdot t. \quad (10)$$

This means that, in this special case, there is no loss of information in estimating the item parameters if CML is used instead of the full model.

For the Rasch model, the expressions for the F-information with respect to the item parameter  $\beta$  are given in Eggen (2000). For instance, in the full Rasch model (3), evaluating expression (2) gives the F-information. The expressions in the cases of MML and CML estimation of the item parameters are given in full detail in Eggen (2000).

If MML estimation of the item parameters in the Rasch model is used, we have a two-parameter distribution  $p_m(x; \beta, \xi)$ , the first parameter being the item parameter  $\beta$  and the second,  $\xi$ , the parameters of the ability distribution. In this case,  $\omega^T = (\beta^T, \xi^T)$  and the F-information with respect to  $\beta$  is formally given by

$$I_{p_m}(\beta; \omega) = I_{p_m}^{\beta\beta^T} - I_{p_m}^{\beta\xi^T} [I_{p_m}^{\xi\xi^T}]^{-1} I_{p_m}^{\xi\beta^T}. \quad (11)$$

The F-information is expressed in terms of the Fisher information matrix (1). In this case,  $\psi = \beta$  and  $\tau^T = \xi^T$ . So,  $I_{p_m}^{\beta\beta^T}$  is then the item parameter submatrix of the Fisher information matrix.  $I_{p_m}^{\xi\xi^T}$  is the part of the ability distribution parameters. If we use a normal distribution, this matrix has four elements ( $\xi^T = (\mu, \sigma^2)$ ). And, finally,  $I_{p_m}^{\beta\xi^T}$  is the part in which each element includes a partial derivative of the likelihood with respect to both an item parameter and an ability distribution parameter.

In the case of CML estimation of the item parameters, the conditional distribution of the answer pattern given the sufficient statistic,  $f(x|t; \beta)$ , is only dependent on  $\beta$  and the F-information in this case is equal to the expected conditional Fisher information:

$$I_f(\beta; \omega) = \mathcal{E}[S_{f,\beta} \cdot S_{f,\beta}^T] = \mathcal{E} \left( \frac{\partial \ln f(x|t; \beta)}{\partial \beta} \cdot \frac{\partial \ln f(x|t; \beta)}{\partial \beta}^T \right) = \mathcal{E}_T I_f(\beta | T). \quad (12)$$

In which the latter expectation is taken with respect to the distribution of  $T$ , which is usually derived from the distribution of the ability  $\theta$ . If a normal ability distribution with mean  $\mu$  and variance  $\sigma^2$  is used, the exact expressions for computing the F-information for CML and MML estimation in the Rasch model are given in Eggen (2000).

It should be understood that the expressions for the F-information matrices in the Rasch model, given in (11) and (12), are only useful, if the model is properly normalized. It is well known that in the Rasch model not all  $k$  but only  $(k-1)$  item parameters are free. A commonly used normalization is to fix the value of one parameter to zero: e.g.  $\beta_1 = 0$ . In Eggen (2004, p.13-60) this normalization was also used and a scalar quantification of the F-information matrices was then given by their determinants. And, following Pukelsheim (1993), the information efficiency was used to compare the F-information in two models. For comparing MML and CML estimation of  $k$  items in the Rasch model, this is computed as

$$\text{INFEFF}(\beta; f: p_m) = \left( \frac{\det(I_f^{*(i)}(\beta; \omega))}{\det(I_{p_m}^{*(i)}(\beta; \omega))} \right)^{1/(k-1)}, \quad (13)$$

in which  $I^{*(i)}$  denotes a F-information matrix after normalization on item  $i$ . In the next section, the relation between the normalization of the model and the information matrices will be treated in detail.

### 3. Normalization, information, and the determinant

In Eggen (2000), it was seen that the F-information matrices of CML and MML estimation of the item parameters in the Rasch model are double centered and therefore not of full rank. This is a consequence of the well-known indeterminacy in the model, which can be solved by imposing one linear restriction on the parameters (Molenaar, 1995). This normalization of the parameters always has an influence on the information matrix. In order to compare different estimation methods on the information, it seems to be reasonable to use the same normalization. In the sequel it will become clear that for comparing this condition can be relaxed somewhat.

In MML, a common normalization is to set the population mean  $\mu$  equal to some constant. But when CML is used, the population mean does not need to be defined and, consequently this normalization cannot be used for comparing CML and MML. Therefore, only normalizations on the item parameters will be considered.

Proper normalizations of the parameter space in the Rasch model can be characterized by the following equation:

$$d_0 + \sum_{i=1}^k d_i \beta_i = 0, \text{ with the restriction that } \sum_{i=1}^k d_i \neq 0. \quad (14)$$

Without loss of generality, it is assumed in the sequel that  $d_0 = 0$ .

The most commonly used normalizations are twofold. The first is that one of the item parameters is fixed to an arbitrary constant,

$$\beta_i = 0. \quad (15)$$

The other is that a linear restriction is put on the mean or the sum of the item parameters, for instance,  $\sum_i \beta_i = 0$ .

It is easily checked that both normalizations are special cases of (14) (with  $d_0 = 0$ ).

The normalization in (14) is equivalent with saying that there will be at least one  $i$  with  $d_i \neq 0$ , such that  $\beta_i$  can be written as

$$\beta_i = \sum_{j \neq i} c_j \beta_j, \quad (16)$$

with the restriction that  $\sum_{j \neq i}^k c_j \neq 1$  and with  $c_j = -d_j/d_i$ . Henceforth in this paper, (16) will be used for the normalization.

### 3.1. The influence of the normalization on the information matrices

The F-information matrices of the item parameters in the Rasch model given in (11) for MML and in (12) for CML estimation are matrices in which the normalization is discarded. These  $(k \times k)$ -matrices are determined as if all  $k$  item parameters are free. These double-centered matrices will be called the non-normalized F-information matrices. If the model is properly normalized, it has only  $k-1$  free parameters, and the normalized F-information matrices will be  $(k-1) \times (k-1)$ . Next, we will express the normalized in the non-normalized F-information.

Although normalization by fixing one item parameter (15) is a special case of normalization (16), this case will be treated first because, in this case, the results are easy and straightforward. Furthermore, it will be shown that for the computation of the information efficiency in the more general case, the results of this special case suffice.

#### Normalization by fixing one item

A normalized F-information matrix is deduced from the non-normalized F-information by simply deleting the row and the column corresponding to the fixed item. For the determinant of such a matrix, the following lemma holds.

#### **Lemma 1**

Let  $A$  be a  $k \times k$  matrix which is double centered, that is,

$$A\mathbf{1} = \mathbf{0} \text{ and } \mathbf{1}^T A = \mathbf{0}^T \quad (17)$$

and  $A^{(i)}$ , a principal submatrix of  $A$ , which results if the  $i^{th}$  row and the  $i^{th}$  column of  $A$  are deleted. Then the determinants of all principal submatrices, or all the principal minors, are equal:

$$\det A^{(i)} = \det A^{(j)}, \quad i, j = 1, \dots, k. \quad (18)$$

#### **Proof Lemma 1**

Because the principal minors do not change if rows and columns of the matrix  $A$  are permuted in the same way, it is sufficient to prove that  $\det A^{(1)} = \det A^{(2)}$ .

Define two  $(k-1) \times (k-1)$ -matrices  $P$  and  $D$  by:

$$P = \begin{pmatrix} 1 & \mathbf{1}_{(k-2)}^T \\ \mathbf{0}_{(k-2)} & I_{k-2} \end{pmatrix} \text{ and } D = \text{diag}(-1, \mathbf{1}_{(k-2)}^T).$$

Then it is easily checked that  $\det P = 1$  and  $\det D = -1$ .

Furthermore, it yields that, because  $A$  is double centered (17)

$$A^{(1)} = DPA^{(2)}P^T D^T.$$

And the result follows directly.  $\square$

If an F-information matrix after a normalization by fixing item  $i$  is denoted by  $\Gamma^{*(i)}(\beta; \omega)$ , then it yields  $\det(\Gamma^{*(i)}(\beta; \omega)) = \det(\Gamma^{*(j)}(\beta; \omega))$ , for  $i, j = 1, \dots, k$ . This is true for the F-information matrices in the MML (11) and in the CML (12) case. Consequently, not only the determinant of the F-information, but also the information efficiency (13), used in comparing MML and CML, is independent of the item which is fixed in the normalization:

$$\left( \frac{\det(\Gamma_f^{*(i)}(\beta; \omega))}{\det(\Gamma_{p_m}^{*(j)}(\beta; \omega))} \right)^{1/(k-1)} = \left( \frac{\det(\Gamma_f^{*(l)}(\beta; \omega))}{\det(\Gamma_{p_m}^{*(m)}(\beta; \omega))} \right)^{1/(k-1)}, \text{ for } i, j, l, m = 1, \dots, k. \quad (19)$$

#### Normalization by a general linear restriction

If the normalization is established by putting a linear restriction on the item parameters as in (16), the non-normalized F-information matrices in (11) and (12) can also be expressed in the normalized F-information matrices. Denote by  $S_{p; \omega}^{(i)}$  the efficient score  $(k-1)$ -vector of distribution  $p$  with respect to parameter  $\omega$  if the normalization is on item  $i$  as in (16). (In the MML case  $\omega^T = (\beta^T, \xi^T)$ ). The elements of this score vector have a simple relation with the elements of the non-normalized score vector  $S_{p; \omega}$ .

The partial derivatives, under the restriction (16), with respect to an item parameter  $\beta_j^*$  ( $j \neq i$ ) of the log likelihood (in both the MML and CML case) are expressed in the non-restricted derivatives as

$$\frac{\partial \ln L}{\partial \beta_j^*} = \sum_{l=1}^k \frac{\partial \ln L}{\partial \beta_l} \frac{\partial \beta_l}{\partial \beta_j^*} = \frac{\partial \ln L}{\partial \beta_j} + c_j \frac{\partial \ln L}{\partial \beta_i}, \text{ for } j = 1, 2, \dots, i-1, i+1, \dots, k. \quad (20)$$

Under the restriction (16) there are only  $(k-1)$  free item parameters  $\beta_j^*$  instead of  $k$ . In the sequel we will drop the  $*$  from the notation of the item parameters, but it should be understood that in all normalized information matrices only  $(k-1)$  item parameters are considered.

Without loss of generality it is assumed in the sequel that  $i = 1$  in (16) and (20).

If we define a  $k \times (k-1)$  matrix  $K$  by

$$K = \begin{bmatrix} c^T \\ I_{k-1} \end{bmatrix}, \text{ with } c^T = (c_2, \dots, c_k), \quad (21)$$

then it is seen that, in the CML case,

$$S_{f;\beta}^{(i)T} = S_{f;\beta}^T K . \quad (22)$$

We see that the normalized F-information matrix, denoted by  $I_f^{(i)}(\beta;\omega)$ , using (12) and (22), is related to the non-normalized F-information matrix as

$$I_f^{(i)}(\beta;\omega) = K^T I_f(\beta;\omega) K . \quad (23)$$

In the MML case with the normalization (16), the population parameters  $\xi$  are free. Therefore we consider the  $(k+2) \times (k+1)$ -matrix  $K^*$  :

$$K^* = \begin{bmatrix} K & \mathbf{0}_{(k \times 2)} \\ \mathbf{0}_{(2 \times k-1)}^T & I_2 \end{bmatrix} \quad (24)$$

and it is easily checked that

$$S_{p_m;\omega}^{(i)T} = S_{p_m;\omega}^T K^* . \quad (25)$$

With (25), it follows that, in the MML case, the relation between the normalized Fisher information matrix, to denote by  $I_{p_m}^{(i)}(\omega)$ , and the non-normalized is

$$I_{p_m}^{(i)}(\omega) = K^{*T} I_{p_m}(\omega) K^* . \quad (26)$$

And by using (11), it can be shown that the relation between the F-information matrix in the normalized case, denoted by  $I_{p_m}^{(i)}(\beta;\omega)$ , and the non-normalized is given by

$$I_{p_m}^{(i)}(\beta;\omega) = K^T I_{p_m}(\beta;\omega) K . \quad (27)$$

A second lemma on the determinant of double-centered matrices with a special structure is given below.

**Lemma 2**

Let  $A^*$  be a symmetric  $k \times k$  matrix partitioned as follows:

$$A^* = \begin{bmatrix} \alpha & a^T \\ a & A \end{bmatrix} \quad (28)$$

$$\text{with } a = -A \mathbf{1} \text{ and } \alpha = -\mathbf{1}^T a = \mathbf{1}^T A \mathbf{1} . \quad (29)$$

So  $A$  is a  $(k-1) \times (k-1)$  matrix.

Next, consider a  $(k-1)$ -vector  $c^T = (c_2, \dots, c_k)$  and a  $k \times (k-1)$ -matrix  $K$  with the following structure:

$$K = \begin{bmatrix} c^T \\ I_{k-1} \end{bmatrix} . \quad (30)$$

Then, for any double-centered matrix  $A^*$  with the structure given in (28) and (29) and for any matrix  $K$  as given in (30), it holds that

$$\det(K^T A^* K) = g(c) \det A , \quad (31)$$

with  $g(c) = (\det(I - c \mathbf{1}^T))^2$  a function of  $c$ , which is independent of  $A^*$ .

**Proof Lemma 2**

Define  $B^* = K^T A^* K = \begin{bmatrix} b & B \end{bmatrix}$

where  $b = \alpha c + a$

and  $B = c a^T + A$ .

Because  $a = -A \mathbf{1}$ , it follows that

$$B = -c \mathbf{1}^T A + A = A(I - c \mathbf{1}^T), \text{ from which it follows that}$$

$$\det B = \det A \cdot \det(I - c \mathbf{1}^T) .$$

Because  $a = -A \mathbf{1}$  and  $\alpha = -\mathbf{1}^T a = \mathbf{1}^T A \mathbf{1}$ , it follows that

$$B \mathbf{1} = -b ,$$

which means that the matrix  $B^*$  is row-centered, i.e the elements of each row sum to zero.



Now

$$K^T A^* K = B^* K = (K^T B^{*T})^T \text{ where } B^{*T} = \begin{bmatrix} \mathbf{b}^T \\ B^T \end{bmatrix} \text{ is column-centered.}$$

By that it follows that

$$\det(K^T A^* K) = \det B^T \cdot \det(I - c \mathbf{1}^T) = \det A \cdot (\det(I - c \mathbf{1}^T))^2. \square$$

It will be clear that the normalized F-information matrices for CML (23) and for MML (27) have exactly the structure of the matrix presented in lemma 2 and thus yields (31). For the determinants of the F-information matrices, we then have that, for CML,

$$\det I_f^{(i)}(\beta; \omega) = \det(K^T I_f(\beta; \omega) K) = g(c) \cdot \det I_f^{*(i)}(\beta; \omega), \quad (32)$$

in which  $I_f^{*(i)}(\beta; \omega)$  is constructed as before from the non-normalized CML F-information matrix by dropping the row and the column corresponding to the item  $i$  which is used for normalization.

For MML we have similarly:

$$\det I_{p_m}^{(i)}(\beta; \omega) = \det(K^T I_{p_m}(\beta; \omega) K) = g(c) \cdot \det I_{p_m}^{*(i)}(\beta; \omega). \quad (33)$$

Note that both (32) and (33) yield for any  $i = 1, \dots, k$  used in the normalization. It should be understood that the F-information matrices change when the normalization is on a different item, but, in the determinants, these differences will only be reflected in different  $g(c)$ . This is because (see lemma 1) the principal minors of the non-normalized F-information matrices are independent of the item  $i$ .

If we compare CML with MML estimation, using the same normalization, this has the consequence that the computed efficiency is completely independent of the normalization chosen. Thus, for  $i = 1, \dots, k$ :

$$\left( \frac{\det(I_f^{(i)}(\beta; \omega))}{\det(I_{p_m}^{(i)}(\beta; \omega))} \right)^{1/(k-1)} = \left( \frac{g(c) \cdot \det(I_f^{*(i)}(\beta; \omega))}{g(c) \cdot \det(I_{p_m}^{*(i)}(\beta; \omega))} \right)^{1/(k-1)} = \left( \frac{\det(I_f^{*(i)}(\beta; \omega))}{\det(I_{p_m}^{*(i)}(\beta; \omega))} \right)^{1/(k-1)}. \quad (34)$$

It can be noted, that the strict condition to use the same normalization in computing the efficiency can be relaxed somewhat. It suffices to demand that the normalization should result in the same value of  $g(c)$ .

This result, combined with (19), implies that any efficiency comparison on the estimation of the item parameters with the determinants of the F-information matrices is independent of the normalization of the Rasch model.

In Eggen (2000), the trace function was used to compare the F-information matrices. The trace is very easily computed, but is not as useful as the determinant for comparing. If we look at the relation of the normalized and the non-normalized F-information, using the general notation of Lemma 2, then it is easily shown that

$$(K^T A K)_{ii} = A_{ii} + 2c_i a_i + c_i^2 \alpha \quad (35)$$

with  $a_i$  and  $c_i$  the  $i^{th}$  element of  $\mathbf{a}$  and  $\mathbf{c}$  respectively. Because of (35)

$$\text{tr}(K^T A K) = \text{tr}(A) + 2\mathbf{c}^T \mathbf{a} + \alpha \mathbf{c}^T \mathbf{c} . \quad (36)$$

It can be seen that there is no simple relation between the traces of the matrices. Consequently the comparison of F-information matrices with the trace function will always be dependent on the normalization chosen.

#### 4. F-information in incomplete designs

In incomplete testing designs, not all items are administered to all persons. Define the design variable  $d_{vj} = 1$  if item  $j$ ,  $j = 1, \dots, k$ , was administered to person  $v$ ,  $v = 1, \dots, n$  and, otherwise. The design vector of a person is denoted by  $\mathbf{d}_v^T = (d_{v1}, \dots, d_{vk})$  and the complete design vector by  $\mathbf{d}^T = (\mathbf{d}_1^T, \dots, \mathbf{d}_n^T)$ . Furthermore, assume that the response variable  $X_{vj}$  takes an arbitrary constant value  $c$  if an item  $j$  is not answered by person  $v$ :  $X_{vj} = c$  if  $d_{vj} = 0$ ; if  $d_{vj} = 1$  the values of the response variable are, as in the complete data case, equal to 1 or 0.

In the Rasch model, the probability of getting answer pattern  $\mathbf{x}$  given the design vector  $\mathbf{d}$  is given by

$$p(x|d; \beta, \theta) = \prod_{v=1}^n \prod_{i=1}^k p(x_{vi} | d_{vi}; \beta_i, \theta_v) = \prod_{v=1}^n \prod_{i=1}^k \frac{\exp[(\theta_v - \beta_i)x_{vi}d_{vi}]}{\{1 + \exp(\theta_v - \beta_i)\}^{d_{vi}}}. \quad (37)$$

In general, one could consider the design variable to be random with distribution  $p(d; \phi)$ . Throughout this paper it will be assumed that this distribution is independent of the parameters  $\beta$  and  $\theta$ . So, for inference on the parameters  $\beta$  and  $\theta$ , it is justified to consider only the conditional distribution of the answer pattern given the design, as specified in (37) instead of the simultaneous distribution of  $(x, d)$ . This can also be understood in terms of the loss of F-information. The following decomposition is true:

$$p(x, d; \beta, \theta, \phi) = p(x | d; \beta, \theta) p(d; \phi). \quad (38)$$

The simultaneous distribution of  $(x, d)$  is a two-parameter distribution: the first parameter is  $(\beta, \theta)$  and the second is  $\phi$ . Because the design distribution is assumed to be independent of the first parameter, the ancillary case, no information is lost if only the conditional distribution,  $p(x | d; \beta, \theta)$ , is considered in the inference on  $(\beta, \theta)$ . For more details on stochastic designs, see Rubin (1976) and Eggen (2004, p.97-134).

The generalization to incomplete designs of CML and MML estimation of the item parameters is well known (Molenaar, 1995). In CML, the conditional distribution of the answer pattern given the scores,  $T_v = \sum_i d_{vi} X_{vi}$  for  $v = 1, \dots, n$ , and the design used is:

$$f(x | t, d; \beta) = \prod_{v=1}^n f(x_v | t_v, d_v; \beta) = \prod_{v=1}^n \frac{\exp[-\sum_{i=1}^k d_{vi} x_{vi} \beta_i]}{\gamma_{t_v}(\beta \odot d_v)}, \quad (39)$$

in which

$$\gamma(\beta \odot d_v) = \sum_{\sum_i d_{vi} y_i = t_v} \exp\left(-\sum_{i=1}^k y_i d_{vi} \beta_i\right) \quad (40)$$

are the basic symmetric functions as in complete designs (5). However, in (40), they are defined for the item parameter vector multiplied directly with the design vector:  $\beta \odot d_v = (\beta_1 d_{v1}, \dots, \beta_i d_{vi}, \dots, \beta_k d_{vk})^T$ , which means that only those items will be taken into account which are administered to a person. Furthermore, the summation in (40) runs across all answer patterns  $(y_1, \dots, y_k)$ ,  $y_i d_{vi} \in \{0, 1\}$  for which  $\sum_{i=1}^k y_i d_{vi} = t_v$ .

If the ability distribution is denoted by  $h(\theta; \xi)$  then the marginal distribution, used in MML estimation is given by:

$$\begin{aligned}
 p_m(x|d; \beta, \xi) &= \prod_{v=1}^n \int_{-\infty}^{\infty} \prod_{i=1}^k p(x_{vi} | \theta, d_{vi}; \beta_i) h(\theta; \xi) d\theta \\
 &= \prod_{v=1}^n \int_{-\infty}^{\infty} \prod_{i=1}^k \frac{\exp[(\theta - \beta_i)x_{vi}d_{vi}]}{\{1 + \exp(\theta - \beta_i)\}^{d_{vi}}} h(\theta; \xi) d\theta
 \end{aligned} \tag{41}$$

The F-information with respect to the item parameters in CML as well as in MML estimation is defined as expected information. In efficiency comparisons in complete testing designs, it suffices therefore to derive this information for one randomly selected person from the ability distribution. There is no need to consider more persons because each has an independent equal expected contribution to the F-information. This is not true in incomplete designs because there is no contribution to the information on  $\beta_i$  if  $d_{vi} = 0$ . Therefore, the total information in the whole sample of persons  $v = 1, \dots, n$ , will be considered, i.e., the design must be taken in account. An elegant and easy way to avoid complicated formulae which contain the design indicators  $d_{vi}$ , is to consider the contribution of every person  $v$  with design vector  $d_v$  separately, as displayed in Figure 1.

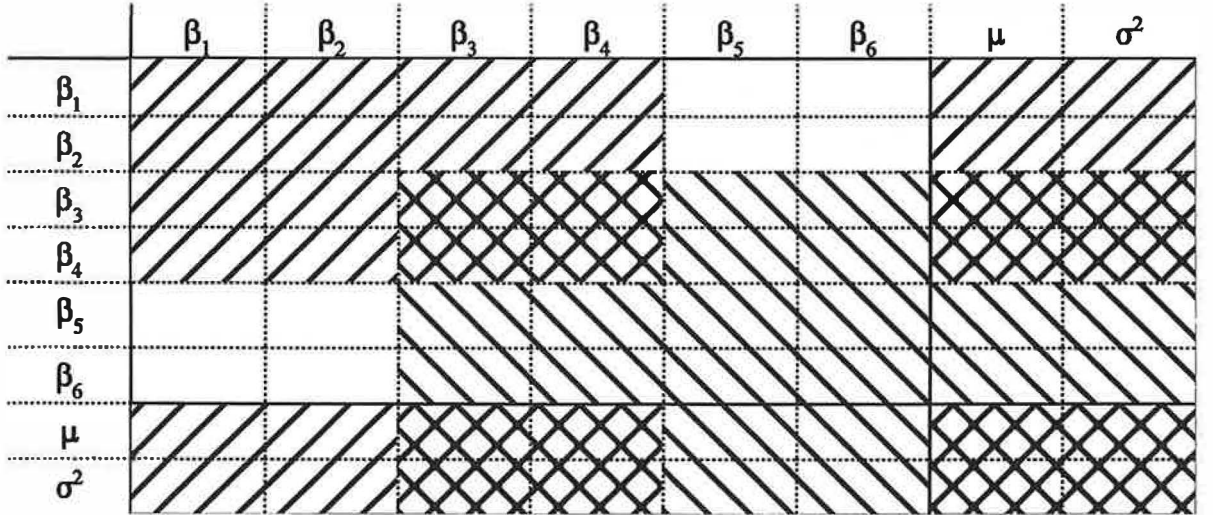




Figure 1. Schematic representation of the elements of the information matrices in incomplete designs by the example  $d_1^T = (1, 1, 1, 1, 0, 0)$  , and  $d_2^T = (0, 0, 1, 1, 1, 1)$  

With design vector  $d_1$ , only the items 1 to 4 are administered, and for such an administration the information matrix can be computed as in a complete design with these four items. It will be clear that with such a test no information is collected with respect to the items 5 and 6. Using design vector  $d_2$  (independently to another person randomly drawn from the same ability distribution), and applying the same procedure we now collect information on the parameters of the items 3 to 6 and on the two population parameters. Because of the independence, the information from both observations can be added, leading to the situation as depicted in Figure 1: for some cells (doubly shaded) the two observations have a contribution, for others (singly shaded) only one observation contributes, and for some of the cells, there is no contribution at all. For example, there is no information available on the items 1 and 5 jointly, because this pair has not been observed jointly.

Notice that the contribution of several design vectors to the same cell is not equal. Consider, for example, the cell on the main diagonal with respect to the parameter  $\mu$ : the amount of information on this parameter depends on the number of item parameters and their specific value for each design vector  $d_v$ .

## 5. The information comparison in incomplete designs

In complete testing designs, only one randomly selected person from the ability distribution was needed to compare the F-information matrices for two models, and the comparison of two models was the only sensible comparison to make. With incomplete designs, however, efficiency comparisons can be made along several lines: one could compare the efficiency of two models, given that the data will be collected in a given design, but one could also compare the efficiency of two different designs under the same model.

A good efficiency measure for the first problem is (13), where the information matrix  $I^{*(i)}$  can represent the total F-information from a sample of size  $n$  with data collected in some (incomplete) test design. As long as  $n$  and the design are the same under the two models, the results will be unaffected by the sample size. This can easily be understood from a simple example. Suppose we collect data using a complete design and the sample size  $n=2$ . The elements of the total information matrix under either model will be twice the value of the elements in the matrix for a single person and consequently the determinants of both total F-information matrices will be  $2^{(k-1)}$  times the value of the determinants of the single

observation, but the factor  $2^{(k-1)}$  will cancel in numerator and denominator of the right hand side of (13). Of course, the efficiency of CML compared to MML may vary according to the design.

Before presenting a more general efficiency measure, we first discuss the practical implication of the exponent  $1/(k-1)$  in the efficiency measure (13). The  $(k-1)$ -th root of the determinant of the normalized F-information matrix under some model, where the information matrix is computed using a single person, can be understood as the total amount of information per person. So we could define

$$m_*(I_{1,f}) = [\det(I_f(\beta; \omega))]^{1/(k-1)} \quad (42)$$

where  $f$  refers to the model,  $k$  is the number of items,  $*$  refers to the normalization used and the '1' in subscript refers to the number of observations used in deriving the information matrix. From now, in the notation of the information matrices the indices referring to the normalization are dropped. Then it is easy to show that

$$m_*(I_{t,f}) = t \cdot m_*(I_{1,f}) . \quad (43)$$

Since the information matrix is additive in the number of persons, the entries in the non-normalized as well as the normalized information matrices based on  $t$  independent observations are  $t$  times the corresponding values with a single person. The normalized information matrix has  $(k-1)$  rows and columns and therefore the determinant of the normalized information matrix based on  $t$  persons will be  $t^{(k-1)}$  times the determinant of the matrix based on a single person. By taking the  $(k-1)$ -th root, the assertion follows.

If we want to compare two designs under the same model, there is no simple and unequivocal measure of efficiency, as can be seen from the following simple example. Suppose we want to determine the efficiency of the design used in Figure 1 compared to a complete design. To apply the incomplete design, one needs at least a sample size of 2, while the complete design can be used with a single observation. If one uses these numbers, the incomplete design may appear as the most efficient one, but one could argue that the comparison is not fair because of unequal sample sizes. If on the other hand, one uses equal sample sizes, the complete design will appear the most efficient because in the incomplete

design either of the two persons has answered only to a subset of the  $k$  items. So, to make a fair comparison, it is not sufficient to consider only the differences in sample sizes.

It seems reasonable, therefore, to develop an efficiency measure which takes in some respect the cost of the design implementation into account, by expressing the information measure relative to the cost of implementation of the design. The measure of efficiency is then the ratio of these two relative information measures. So the measure proposed in (13) can then be generalized to

$$\text{INFEFF}(\beta; f(D_1):p_m(D_2), C) = \frac{C(D_2)}{C(D_1)} \left( \frac{\det(I_{f(D_1)}(\beta; \omega))}{\det(I_{p_m(D_2)}(\beta; \omega))} \right)^{1/(k-1)}. \quad (44)$$

Where  $C(D)$  represents the cost of the data collection according to the design  $D$  and  $I_{f(D)}(\beta; \omega)$  represents the F-information matrix under model  $f$  using design  $D$ , under a proper normalization which is common to both models.

Formula (13) is easily seen as a special case, where the two models differ but the designs are the same, so that their cost ratio equals one.

In the next section, examples will be studied for another special case of (44), where the designs differ, but the models are the same, either CML or MML, so that we obtain a measure of the relative efficiency of two designs. An interesting question in this respect is whether and to what extent this measure will vary across different models.

An important issue is the definition of the cost function  $C(D)$ , which can range from a simple model with a fixed unit cost per item answer, to more realistic functions with different costs per test taker and per answer within a test taker up to complicated functions which take the risk of errors in the implementation of the design into account.

Within the present paper we will confine ourselves to two cost functions: a cost function with a unit cost per observed answer, and another one with a unit cost per test taker. We illustrate both cases using a simple example, where the design of Figure 1 is compared to the complete design.

Consider the first cost function ( $C_1$ ) with a single response of a person to an item as unit cost. Suppose the information measure in the complete design is based on  $n_c$  test takers, and the information in the incomplete design on  $n_i$  test takers. We will assume for simplicity that  $n_i$  is an integer multiple of the number of different design vectors used in the incomplete design. Then it will be clear that (see also Figure 1)  $C_1(D_c) = 6 \cdot n_c$ , while  $C_1(D_i) = 4 \cdot n_i$ .

Application of (44) yields a comparison of both designs in terms of amount of information per single response. Since both cost functions are multiples of the sample sizes, and in view of (43) it will be clear that the efficiency measure (44) in combination with cost function  $C_1$  is invariant under the choice of any (positive) sample sizes, either in the complete as in the incomplete design. Although such a measure may not be directly useful in planning data collections, it gives a kind of intrinsic comparison between designs. In fact, it can give a sensible answer to the question whether a single response contributes the same amount of information to the estimation of the item parameters in an incomplete design compared to another incomplete design or to a complete design.

The second cost function  $C_2$ , which is equal or proportional to the sample sizes, yields a more practical result. It is useful for comparing the efficiency of two designs with the same model. The efficiency measure compares the amount of information per test taker in two different designs. This means that the cost functions of the design in Figure 1 and the complete design are respectively:  $C_2(D_i) = n_i$  and  $C_2(D_c) = n_c$ . Suppose that for the example of Figure 1 (and a given set of parameter values) we find  $\text{INFEFF}(\beta; f(D_i); f(D_c), C_2) = 0.8$ . This means that the amount of information delivered per test taker in the incomplete design is 80% of the information per test taker in the complete design, or, conversely, if we would plan a data collection with 1000 test takers in the complete design, we would need  $1000/0.8 = 1250$  test takers in the incomplete design to collect the same amount of information with the incomplete design.

As a final note, it should be stressed that equating the total amount of information in two different designs by manipulating the sample sizes, does not mean that these two designs are therefore equivalent in all possible respects. This can easily be seen from Figure 1: whatever the sample size in the incomplete design used there, there will be always eight zero cells (the cells without any shading). It is beyond the scope of the present paper, however, to predict in a detailed way all possible implications of such and similar patterns in the information matrix.

## 6. Examples of comparing the efficiency in incomplete designs

The efficiency comparisons are conducted with two item banks, each with 18 items. The item bank size of 18 was chosen because at that point the difference in efficiency of CML compared to MML in complete designs (Eggen, 2000) is negligibly. In all examples, a normal ability



distribution with mean 0 and variance 1 was used. Two different sets of item parameters were considered:

1. all items equal to 0:  $\beta_i = 0, i = 1, \dots, 18$ .
2. symmetric around 0, ranging from -2.25 to 2.25 in steps of 0.25, omitting 0 itself. This will be denoted by  $\beta_i = -2.25 (0.25) 2.25, i = 1, \dots, 18$ .

With the first set of items, there is, as was seen before, no loss of information in using CML compared to MML, and the sole loss due to the incompleteness can be studied. In the second situation, the extra loss resulting from the use of CML can be considered.

### 6.1. The designs

The incomplete designs considered are in common use and are all so-called anchor-item designs. In order to be able to estimate the item parameters from the resulting data, there is an overlap between the tests or booklets administered. A booklet is a number of items which are administered together to a person. All the designs considered share the properties that every booklet is the same length and that every item has the same number of observations. Furthermore, it is assumed that the booklets are randomly assigned to the students. Three different types of designs are distinguished:

1. Item-interlaced anchoring design (Vale, 1986). In these designs there are as many booklets as there are items. The first booklet begins with the first item and contains sequentially all items until it has its established test length. The second booklet starts with the second item. The final booklet contains the last item and one or more of the first items. These designs can have tests or booklets with of 2, 3, or  $k-1$  items. The limiting case with  $k$  items gives a complete design. In the sequel, the following notation will be used for these designs:  $\Pi; \#i; \#b$ , in which  $\Pi$  is the abbreviation of the type of design,  $\#i$  is the length of a booklet in the design and  $\#b$  is the number of booklets in the design.

In Figure 2 two examples are given in case the total number of items is  $k=18$ .

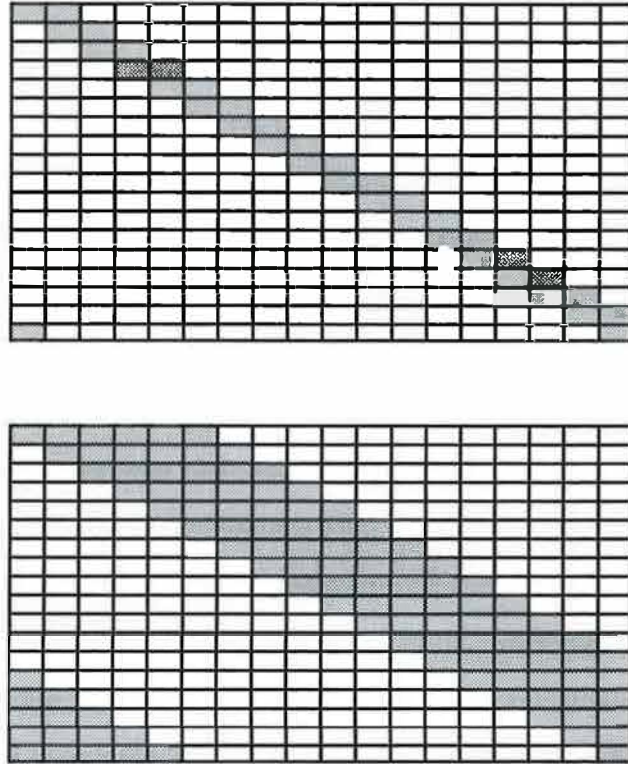


Figure 2. Item-interlaced anchoring designs  $II;2;18$  and  $II;6;18$

In design  $II;2;18$ : the lengths of the booklets are 2: the first booklet contains item number 1 and 2, the second item number 2 and 3, etcetera, until the eighteenth booklet with item number 18 and 1. In  $II;6;18$  the test lengths are 6: booklet 1 has the items 1, 2, 3, 4, 5, and 6; booklet 2 the items 2, 3, 4, 5, 6, and 7, and so on.

2. Block-interlaced anchoring design. In these designs, a booklet contains two blocks of items. These two blocks are interlaced as in an item-interlaced anchoring design with booklets with two blocks. Each block can contain more than one item, but the number of items per block are equal. The general notation is  $BI; \#i; \#b$ , with  $\#i$  the total number of items in the two blocks.

In the case of 18 items, 4 of these designs can be established. The block size is 1, 2, 3, or 6 items. These designs then have respectively 18 booklets with 2 items, 9 booklets with 4 items, 6 booklets with 6 items, and 3 booklets with 12 items. The block-interlaced design with 6 booklets, each having two blocks of 3 items is given in Figure 3.

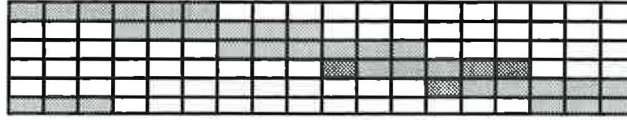


Figure 3. Block-interlaced anchoring design BI;6;6

Remark that the item interlaced design II;2;18 in Figure 2 is the same as the block interlaced design BI;2;18 with blocksize 1.

3. Balanced block designs. In balanced block designs, not only all the blocks of items have an equal number of observations. Also, the pairs of blocks of items have an equal number of observations. The notation we will use for that is  $BB;(\#bl.bls),\#b$ , in which  $\#bl.bls$  is the number of blocks in a booklet multiplied by the size of the blocks. (This is of course the number of items in a booklet). With 18 items in total, there are several balanced block designs possible (Cochran & Cox, 1957). Only designs with no more booklets than items are considered. Two examples are given in Figure 4.

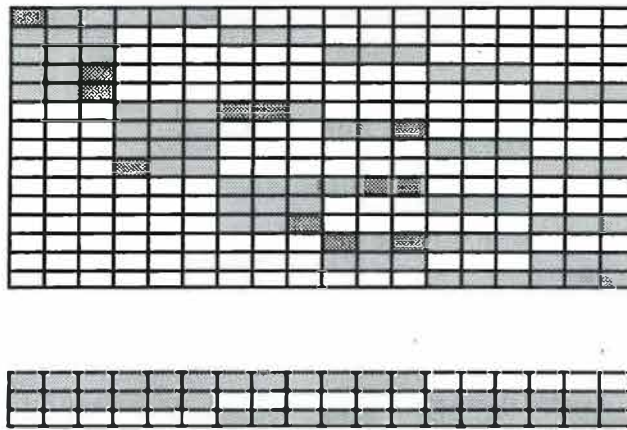


Figure 4. Balanced block designs  $BB;(2.3);15$  and  $BB;(2.6);3$

It can be seen that the BB;(2.6);3 is the same as the block interlaced design BI;12;3. The balanced block designs with 18 items and no more than 18 booklets considered next are given in Table 1.

*Table 1. Balanced block designs with 18 items and no more than 18 booklets*

code	total number of blocks	block size	number of blocks per booklet	length of booklet	number of booklets
BB;(17.1);18	18	1	17	17	18
BB;(8.2);9	9	2	8	16	9
BB;(5.3);6	6	3	5	15	6
BB;(2.6);3	3	6	2	12	3
BB;(5.2);18	9	2	5	10	18
BB;(3.3);10	6	3	3	9	10
BB;(4.2);18	9	2	4	8	18
BB;(2.3);15	6	3	2	6	15

## 6.2. Results for an observed response as unit of cost

Using the cost function  $C_1$ , see section 5, comparisons between different estimation methods and different designs are possible. It is noted that with this cost function all the results are given per observed response. In the Tables 2 to 7, the information per observed response and the information efficiency will be given. The entries in the columns CML-det and MML-det are  $\text{CML-det} = 10^3 \cdot [\det(I_{f(D)})]^{1/(k-1)} / C_1(D)$  and  $\text{MML-det} = 10^3 \cdot [\det(I_{p_m(D)})]^{1/(k-1)} / C_1(D)$  respectively. In computing the determinants and the efficiencies (44) the normalization was established by fixing the first item.

### Results for item-interlaced anchoring designs

For 18 items with  $\beta_i = 0$  in the Rasch model the standardized determinants of the F-information matrices per observation for CML and MML are given, followed by the information efficiency of CML compared to MML. For reference, the results for a complete design are also reported in Table 2. At each entry in Table 2 also the efficiency of using the

design compared to the complete design, using the same estimation method, is given between brackets.

*Table 2. Information comparison for item-interlaced anchoring designs;  $\beta_i = 0, i=1, \dots, 18$*

design	CML-det (% of complete)	MML-det (% of complete)	INFEFF( $\beta; f(D); p_m(D), C_1$ )
complete	9.68	9.68	1
II;2;18	3.40 (0.351)	8.28 (0.855)	0.411
II;3;18	5.49 (0.567)	8.49 (0.877)	0.646
II;4;18	6.75 (0.697)	8.67 (0.896)	0.778
II;5;18	7.56 (0.781)	8.82 (0.911)	0.857
II;6;18	8.11 (0.838)	8.96 (0.926)	0.906
II;7;18	8.51 (0.879)	9.07 (0.937)	0.937
II;8;18	8.79 (0.908)	9.18 (0.948)	0.958
II;9;18	9.00 (0.930)	9.27 (0.958)	0.971
II;10;18	9.17 (0.947)	9.35 (0.966)	0.980
II;11;18	9.29 (0.960)	9.42 (0.973)	0.986
II;12;18	9.36 (0.967)	9.46 (0.977)	0.989
II;13;18	9.46 (0.977)	9.52 (0.983)	0.993
II;14;18	9.52 (0.983)	9.56 (0.988)	0.995
II;15;18	9.57 (0.989)	9.60 (0.992)	0.997
II;16;18	9.61 (0.993)	9.63 (0.995)	0.998
II;17;18	9.65 (0.997)	9.66 (0.998)	0.999

It was shown, for this set of item parameters, that there is no difference in the information between CML and MML estimation in complete designs. And the first interesting observation from Table 2 is that in any incomplete design, there is, irrespective of the estimation method, at least some loss of information compared to a complete testing design. For both estimation methods, CML and MML, the information increases with the test booklet length. It can be seen that the loss of information due to the design with CML estimation is quite large if we have a small number of items in the booklets. This is not true in MML estimation, where the loss due to the design with a few items is also rather small. The efficiency compared to a complete

design is already at 0.855 with test length of two items. However, the difference in the efficiency between CML and MML vanishes rather quickly with increasing test booklet length. The efficiency is already above 0.95 with an 8-item test booklet. From test lengths of 12 items or more there is hardly any difference in the loss due to the design with CML and MML estimation.

The comparisons for the item bank with items with different difficulty parameters are given in Table 3.

*Table 3. Information comparison for item interlaced anchoring designs:*

$$\beta_i = -2.25 \text{ (0.25)} 2.25, i = 1, \dots, 18$$

design	CML-det (% of complete)	MML-det (% of complete)	INFEFF( $\beta; f(D); p_m(D), C_1$ )
complete	7.15	7.16	0.999
II;2;18	2.32 (0.324)	6.18 (0.863)	0.377
II;3;18	3.80 (0.531)	6.35 (0.887)	0.598
II;4;18	4.73 (0.662)	6.48 (0.905)	0.730
II;5;18	5.35 (0.748)	6.58 (0.919)	0.813
II;6;18	5.78 (0.808)	6.67 (0.932)	0.868
II;7;18	6.10 (0.853)	6.74 (0.941)	0.905
II;8;18	6.34 (0.887)	6.80 (0.950)	0.932
II;9;18	6.52 (0.912)	6.86 (0.958)	0.951
II;10;18	6.66 (0.931)	6.91 (0.965)	0.964
II;11;18	6.77 (0.947)	6.96 (0.972)	0.974
II;12;18	6.86 (0.959)	6.99 (0.976)	0.981
II;13;18	6.93 (0.969)	7.03 (0.982)	0.986
II;14;18	6.99 (0.978)	7.06 (0.986)	0.990
II;15;18	7.04 (0.985)	7.09 (0.990)	0.993
II;16;18	7.08 (0.990)	7.11 (0.993)	0.996
II;17;18	7.12 (0.996)	7.13 (0.996)	0.997

Note that, in this situation, there is a small loss of information if CML instead of MML estimation is used in a complete testing design. Compared to the results in Table 2, where all item parameters are equal, it can be seen that, for all designs, the information efficiency of CML estimation compared to MML estimation is a little bit lower. It is true that in CML the efficiency compared to a complete design is always a bit larger if the item parameters are equal than if there are different item parameters. For MML, the differences between the two item sets are even smaller. However for MML, for booklets longer than 10 items the same trend is seen as with CML, while for shorter booklets the differences are in the opposite direction. For example, in the II;4;18 design, CML compared to a complete design gives 0.838 for equal parameters and for different parameters we have 0.808. For MML, we have in that same design 0.926 for equal parameters and 0.932 for unequal parameters. However, in the II;12;18 design MML shows 0.978 for all  $\beta = 0$  and 0.976 and for differing  $\beta$  for the efficiency compared to a complete design.

However, the general picture is the same: with very few items MML is more efficient, but this advantage disappears quickly with longer test booklets.

#### Results for block-interlaced anchoring designs

Table 4 and 5 give the results for 18 items with  $\beta_i = 0$  and for 18 items with  $\beta_i = -2.25 (0.25) 2.25, i=1, \dots, 18$  respectively. The efficiency of an incomplete design compared to a complete design is reported in parentheses in Tables 4 and 5. Two designs in these tables are the same as an item-interlaced design and a balanced block design respectively. This is indicated between brackets in the first column of the tables.

*Table 4. Information comparison for block-interlaced designs; 18 items with  $\beta_i = 0$*

design	CML-det (% of complete)	MML-det (% of complete)	INFEFF( $\beta; f(D); p_m(D), C_1$ )
complete	9.68	9.68	1
BI;2;18 (II)	3.40 (0.351)	8.28 (0.855)	0.411
BI;4;9	6.53 (0.675)	8.65 (0.894)	0.755
BI;6;6	7.95 (0.821)	8.93 (0.923)	0.891
BI;12;3 (BB)	9.36 (0.967)	9.46 (0.977)	0.989

Table 5. Information comparison for block-interlaced anchoring designs;

$$\beta_i = -2.25 (0.25) 2.25, i = 1, \dots, 18$$

design	CML-det (% of complete)	MML-det (% of complete)	INFEFF( $\beta; f(D); p_m(D), C_1$ )
complete	7.15	7.16	0.999
BI;2;18 (II)	2.32 (0.324)	6.18 (0.863)	0.377
BI;4;9	4.61 (0.645)	6.47 (0.904)	0.713
BI;6;6	5.71 (0.799)	6.66 (0.930)	0.858
BI;12;3 (BB)	6.86 (0.959)	6.99 (0.976)	0.981

We see the same trends in the results for the block-interlaced designs as in the item-interlaced designs. It is seen in Table 4 and 5 that in all cases there is loss of information due to any incomplete design. Again, this loss is substantial for CML but not for MML, for very short test lengths. But for longer tests (larger than 6), the differences between CML and MML are relatively small. It can again be seen that in CML the efficiency compared to a complete design is always a bit larger in the case of equal item parameters than in the case of different item parameters. For MML, this is again only true for the designs with the longer booklets.

We see no large differences if we compare the results of item-interlaced and block-interlaced designs. However, if we compare the results for designs with equal booklet length in Table 2 and 4 (and in Table 3 and 5), the comparison always goes in the same direction. For the design with the short booklets, the information in CML estimation and the information in MML estimation as well as in the CML versus MML efficiency is always a bit larger in item interlaced than in block interlaced designs. With 12 items in a booklet no differences are left.

#### Results for balanced block designs

The results of the information comparison in balanced block designs are given in Table 6 and 7. Two designs in these tables are the same as an item-interlaced design and a block-interlaced design respectively. This is indicated between brackets in the first columns of the tables.



Table 6. Information comparison for balanced block designs;  $\beta_i = 0, i = 1, \dots, 18$

design	CML-det (% of complete)	MML-det (% of complete)	INFEFF( $\beta; f(D); p_m(D), C_1$ )
complete	9.68	9.68	1
BB;(17.1);18 <sub>(II)</sub>	9.65 (0.997)	9.66 (0.998)	0.999
BB;(8.2);9	9.61 (0.993)	9.63 (0.995)	0.998
BB;(5.3);6	9.57 (0.989)	9.60 (0.992)	0.997
BB;(2.6);3 <sub>(BI)</sub>	9.36 (0.967)	9.46 (0.977)	0.989
BB;(5.2);18	9.22 (0.952)	9.37 (0.968)	0.983
BB;(3.3);10	9.07 (0.937)	9.29 (0.960)	0.976
BB;(4.2);18	8.94 (0.924)	9.23 (0.953)	0.969
BB;(3.2);15	8.33 (0.860)	8.99 (0.929)	0.927

All the results in Table 6 and 7 confirm the earlier findings. Although all the differences are very small, the differences and trends reported with the II and the IB designs are found again. The general result is that both estimation methods, CML and MML, are quite efficient in balanced block designs. Because of the ordering in the tables, it is easily checked that with increasing test length the information in CML, in MML, and also in the efficiency of CML compared to MML is increasing.

Table 7. Information comparison balanced block designs;

$$\beta_i = -2.25 \text{ (0.25) } 2.25, i = 1, \dots, 18$$

design	CML-det (% of complete)	MML-det (% of complete)	INFEFF( $\beta; f(D); p_m(D), C_1$ )
complete	7.15	7.16	0.999
BB;(17.1);18 <sub>(m)</sub>	7.12 (0.996)	7.13 (0.996)	0.997
BB;(8.2);9	7.08 (0.990)	7.11 (0.993)	0.996
BB;(5.3);6	7.04 (0.985)	7.09 (0.990)	0.993
BB;(2.6);3 <sub>(BI)</sub>	6.86 (0.959)	6.99 (0.976)	0.981
BB;(5.2);18	6.70 (0.937)	6.91 (0.965)	0.969
BB;(3.3);10	6.56 (0.917)	6.86 (0.958)	0.956
BB;(4.2);18	6.42 (0.898)	6.81 (0.951)	0.944
BB;(3.2);15	5.87 (0.821)	6.66 (0.930)	0.881

Finally, we can compare the results of the balanced block designs with the item- interlaced designs. The differences between the designs are very small. Only for designs with booklets with 10 items or less, it is seen that with booklets of the same length the CML and the MML information and the CML-MML information efficiency is larger for balanced block designs than for item-interlaced designs. And because the item-interlaced designs did a bit better than the block-interlaced designs, in general the balanced block designs can be preferred.

### 6.3. Results for a test taker as unit of cost

If the cost function  $C_2$ , defined in section 5, is used in the efficiency comparisons, we have a unit cost per test taker. The results can be used for comparing different designs with the same estimation method. Using this cost function, the results are easily translated to the number of persons needed in applying the design in order to get the same information as compared to a number of persons in a standard design. In Table 8, for the item bank with  $\beta_i = -2.25 \text{ (0.25) } 2.25, i = 1, \dots, 18$ , for both CML and MML the results are given. In both situations the standard of the information gathered in a complete design with a sample of 1000 students is taken.

Table 8. Efficiency of designs with cost function  $C_2$  and the item bank with

$\beta_i = -2.25 (0.25) 2.25, i = 1, \dots, 18$

Design	CML sample size	MML sample size
complete	1000	1000
II;2;18	27737	10427
II;3;18	11289	6765
II;4;18	6802	4972
II;5;18	4811	3917
II;6;18	3711	3220
II;7;18	4014	2731
II;8;18	2537	2369
II;9;18	2193	2087
II;10;18	1932	1865
II;11;18	1728	1683
II;12;18	1563	1536
II;13;18	1428	1410
II;14;18	1315	1303
II;15;18	1218	1211
II;16;18	1136	1132
II;17;18	1063	1063
BI;2;18 (II)	27737	10427
BI;4;9	6979	4979
BI;6;6	3756	3225
BI;12;3 (BB)	1563	1536
BB;(17.1);18(II)	1063	1063
BB;(8.2);9	1136	1132
BB;(5.3);6	1218	1211
BB;(2.6);3 (BI)	1563	1536
BB;(5.2);18	1920	1865
BB;(3.3);10	2179	2087
BB;(4.2);18	2505	2365
BB;(3.2);15	3654	3225

The general trends in the results in comparing the designs with the cost function  $C_2$  are the same as reported in section 6.2 for the cost function  $C_1$ . For both CML and MML, the complete design is most efficient and furthermore it is clear that with a constant cost per test taker it is better to have more items per test taker. Interesting in the results, is the direct translation of the efficiency in the number of observations which can be very useful in planning designs in practice. Although, the results in Table 8 are quite obvious, one example will be given for illustration. If we are planning to have booklets with 6 items, in CML we need 3711 persons in an item-interlaced design, 3756 in an block interlaced design and 3654 in a balanced block design to get the same information as in a complete design with 1000 persons. This indicates a clear preference for a balanced block design in the CML case. For MML, the needed numbers are 3220, 3225 and 3225 respectively, which indicates hardly any difference between the designs.

## 7. Conclusion

In this study the comparison of the efficiency of CML and MML estimation of the item parameters of the Rasch model in incomplete designs was studied. The use of the concept of F-information (Eggen, 2000) was generalized to incomplete testing designs. It was shown that the only case in which CML and MML are equally efficient is when all item parameters are equal.

The standardized determinant of the F-information matrix is used for a scalar measure of information with respect to set of item parameters. In this paper, the relation between the normalization of the Rasch model and this determinant was clarified. It was shown that the value of the determinants depends on a specific normalization. But comparing estimation methods with the defined information efficiency was shown to be independent of the chosen normalization.

Information comparisons were carried out with two item banks, two different cost functions of data collection and three different types of incomplete designs: item-interlaced, block-interlaced, and balanced block designs. The first general observation was that for both CML and MML some information is lost in all incomplete designs compared to complete designs. Also the efficiency of CML compared to MML is in an incomplete design always smaller than in a complete design. The general trend is that with increasing test booklet length the

information in CML, in MML, and also in the efficiency of CML compared to MML is increasing. The main differences between CML and MML were seen in relation to the length of the test booklet. It was demonstrated that with very small booklets (4 or less items), there is a substantial loss in information (more than 35%) with CML estimation, while this loss is only about 10% in MML estimation. However, with increasing test length, the differences between CML and MML quickly disappear. With test booklets with 12 items, the lowest reported efficiency of CML compared to MML was 0.979.

Between the results of the two considered sets of items no large differences were reported. The CML-MML efficiencies are somewhat higher with equal item parameters. Also the CML efficiency compared to a complete design is higher in case of equal parameters, while this is not found for all designs for the MML efficiency compared to a complete design.

No large differences were found between the design types studied. Their main results were the same. A slight advantage was seen for balanced block designs compared to item-interlaced designs, while these designs are expected to perform a bit better than block-interlaced designs. The differences between the designs were a bit larger for CML than for MML.



## 8. References

- Bhappkar, V.P. (1989). Conditioning on ancillary statistics and loss of information in the presence of nuisance parameters. *Journal of Statistical Planning and Inference*, 21, 139-160.
- Cochran, W.G. & Cox, G.M. (1957). *Experimental designs. Second edition*. New York: Wiley.
- Eggen, T.J.H.M. (2000). On the loss of information in Conditional maximum likelihood estimation of item parameters. *Psychometrika*, 65, 337-362.
- Eggen, T.J.H.M. (2004). *Contributions to the theory and practice of computerized adaptive testing*. Arnhem: Citogroep.
- Fischer, G.H. (1974). *Einführung in die Theorie psychologischer Tests*. [Introduction to mental test theory.] Bern: Huber.
- Molenaar, I.W. (1995). Estimation of item parameters. In: Fischer, G.H. & Molenaar, I.W. (Eds.). *Rasch Models* (p.39-51). New York: Springer.
- Pukelsheim, F. (1993). *Optimal design of experiments*. New York: Wiley.
- Rubin, D.B. (1976). Inference and missing data. *Biometrika*, 63, 581-592.
- Vale, C.D. (1986). Linking item parameters onto a common scale. *Applied Psychological Measurement*, 10, 333-344.







1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that this is crucial for ensuring transparency and accountability in the organization's operations.

2. The second part outlines the various methods and tools used to collect and analyze data. This includes both traditional manual methods and modern digital technologies, highlighting the benefits of each approach.

3. The third part focuses on the role of human resources in the data collection process. It discusses how training and support for staff can improve the quality and reliability of the data collected.

4. The fourth part addresses the challenges and limitations of data collection. It identifies common pitfalls and provides strategies to overcome them, ensuring that the data remains valid and useful for decision-making.

5. The fifth part concludes by summarizing the key findings and recommendations. It stresses the need for continuous improvement and regular updates to the data collection process to keep it relevant and effective.