A Simple and Effective Method to Test the Dimensionality Axiom of the Rasch Model

Arnold L. van den Wollenberg

A simple method is introduced to test the dimensionality axiom in the Rasch model. The method consists of a suitable partitioning of the data set. This partitioning is based on the responses to some of the items being investigated. The items used for this purpose are called the external criterion items. Simulation data are presented indicating the effectiveness of the method in instances where the traditional test procedures fail.

It was recently pointed out by several authors (Gustafsson, 1980; Gustafsson & Lindblad, 1978; Stelzl, 1979; van den Wollenberg, 1979, 1980a) that the traditional global test statistics for the Rasch model, such as the conditional likelihood ratio test (Anderson, 1973), the test of Martin-Löf (1973), the test of Wright and Panchapakesan (1969), and the Fischer-Scheiblechner test (Fischer & Scheiblechner, 1970), are essentially insensitive to violation of the unidimensionality axiom. These tests are standardly applied to a raw score partitioning of the data set. The insensitivity of the tests is associated with this partitioning (van den Wollenberg, 1979). When these tests are applied to other types of partitioning, violation of the unidimensionality axiom may become manifest, although this is in no way guaranteed to occur.

Martin-Löf (1973) presented a method to test whether two sets of items measure the same latent trait. This test, however, requires an a priori hypothesis about which items belong to one trait and which to the other. This type of knowledge is as a rule not available, so the method of Martin-Löf has only limited applicability.

Van den Wollenberg (1979, 1980a) introduced two new test statistics, Q_1 and Q_2. The Q_1 statistic is obtained as follows. For each item-level combination the following statistic is defined:

$$q_{ri} = \frac{n_{ri} - E(n_{ri})}{E(n_{ri})} + \frac{n_{r\bar{i}} - E(n_{r\bar{i}})}{E(n_{r\bar{i}})} \quad . \qquad [1]$$

Here n_{ri} is the number of subjects in level-group r positively responding to item i and $n_{r\bar{i}}$ is the number of subjects in level-group r negatively responding to this item. The expectation is obtained as

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 6, No. 1, Winter 1982, pp. 83–91
© Copyright 1982 Applied Psychological Measurement Inc.
0146-6216/82/010083-09$1.45

$$E(n_{ri}) = n_r \times \pi_{ri} = n_r \left(\frac{\varepsilon_i \; \gamma_{r-1}^{(i)}}{\gamma_r} \right) , \qquad [2]$$

where γ_r is the basic symmetric function of order r and $\gamma_{r-1}^{(i)}$ is its partial derivative with respect to the logarithmic item parameter ε_i; n_r is the total number of subjects in level-group r.

The overall statistic Q_1 is obtained as

$$Q_1 = \frac{k-1}{k} \sum_r \sum_i q_{ri} \qquad \begin{array}{l} (r=1, \ldots, k-1) \\ (i=1, \ldots, k \quad) \end{array} \qquad [3]$$

Equation 3 is a sum of terms that are dependent to some degree; the addition theorem of χ^2 variates does not apply. This problem can be resolved by incorporating the complete variance-covariance matrix of the deviates $n_{ri} - E(n_{ri})$, as is done in the construction of the Martin-Löf statistic. When this is done, however, a matrix inversion is involved for each level group.

Van den Wollenberg (1980a) showed that in the case of equal item parameters, the correction factor $(k-1)/k$ can analytically be substituted for the variance-covariance matrix. For this case, the Martin-Löf statistic and Q_1 are analytically equivalent. When the item parameters are not equal, the correction factor is only an approximate substitute for the variance-covariance matrix. In simulation studies, van den Wollenberg (1980a) showed that the approximation is very good. He obtained correlations between Q_1 and the Martin-Löf statistic that were .99 or higher.

Van den Wollenberg (1980b) argued that Q_1 also closely resembles the Wright-Panchapakesan statistic. This last statistic has two important deficiencies that cause it to be heavily biased. When the deficiencies are corrected, the Wright-Panchapakesan statistic becomes equivalent to Q_1 (van den Wollenberg, 1980b). Q_1 is sensitive to the same effects as the other global test statistics and therefore is essentially insensitive to violation of the unidimensionality axiom.

The statistic Q_1 has some practical advantages. It is easily computed, it gives contributions to the overall statistic for each item-level combination, which is very valuable for the study of fit, and its construction is analogous to Q_2, which is especially sensitive to violation of the unidimensionality axiom.

The statistic Q_2 is also a sum of individual statistics. For each level group the second-order observed frequencies are compared with their expectations. For each item pair the following statistic is obtained:

$$q_{rij} = \frac{d^2}{E(n_{rij})} + \frac{d^2}{E(n_{ri\bar{j}})} + \frac{d^2}{E(n_{ri\bar{j}})} + \frac{d^2}{E(n_{r\bar{i}\bar{j}})} . \qquad [4]$$

The statistic is obtained by comparing the cells of a two by two observed contingency table with the corresponding expectations. The quantity d^2 is the difference between the observed and the expected frequency and is, of course, equal for all cells of the table.

The second-order expected frequency $E(n_{rij})$ is obtained as

$$E(n_{rij}) = n_r \left(\frac{\varepsilon_i \varepsilon_j \; \gamma_{r-2}^{(i,j)}}{\gamma_r} \right) , \qquad [5]$$

where $\gamma_{r-2}^{(i,j)}$ is the second-order partial derivative of the basic symmetric function with respect to the item parameters ε_i and ε_j. For each level group the following overall statistic is obtained:

$$Q_{2(r)} = \frac{k-3}{k-1} \sum_{i} \sum_{j} q_{rij} \qquad \begin{matrix}(i=1,\ \ldots,\ k-1)\\(j=i+1,\ \ldots,\ k)\end{matrix} \quad . \qquad [6]$$

Again, a correction factor is involved to compensate for the negligence of the covariance between the statistics given by Equation 4. (For a more detailed discussion, see van den Wollenberg, 1980a.)

There still exist several problems in the application of the Q_2 statistic (van den Wollenberg, 1980a):

1. Whereas only the overall item parameters are needed for Q_1, the item parameters of every subsample have to be estimated in order to obtain Q_2. This not only requires considerably more computing time, but it can also prove impossible to estimate the item parameters in some of the subsamples because one or more items have been passed or failed by all subjects in the subsample.
2. Because the data set has to be partitioned into $k - 1$ subsamples (where k is the number of items), the number of subjects in the subsamples can become rather small, which makes the occurrence of small expected frequencies rather likely. This may cause Q_2 to be a fairly unstable statistic in some practical applications.
3. It is as yet not clear whether and how subsamples can be combined into larger subsamples in order to obtain better parameter estimates and more stable statistics.

In the following a simple procedure is described, which amounts to a suitable partitioning of the data set, and which makes violation of the dimensionality axiom very likely to manifest itself. For this purpose the traditional test statistics can be used. The Q_1 statistic will be used in the present study because of its practical advantages; the correspondence between Q_1 and the other global statistics is so high that the results obtained with Q_1 may be considered to hold for the other statistics to the same degree.

A Simple Method and Its Rationale

The traditional test statistics and the Q_1 statistic amount to a more or less direct inspection of the equality of item parameter estimates over subsamples. These subsamples are, as a rule, obtained by splitting the sample according to raw score. When the raw score partitioning of the data set is used, inspecting equality of the item parameters is equivalent to inspection of parallelism of the item characteristic curves; the differences between item parameter estimates should be equal across subsamples.

When more than one, say two, latent traits underly the data, the raw score is dependent upon both latent traits, and the raw score estimates an "observed trait," which is a combination of the two latent traits. The traditional tests only inspect the item characteristic curves on parallelism on the "observed trait." When parallelism holds, multidimensionality is not detected.

The situation described above is quite feasible when two Rasch homogeneous tests are amalgamated into one. The raw score is the sum of the raw scores of the subtests. When, in a given sample, the subtests have (approximately) equal raw score distributions, the observed total score is (approximately) dependent upon both subtests to the same extent. Now it becomes quite likely that the item characteristic curves on the two latent traits project as a set of parallel curves on the observed trait. Van den Wollenberg (1979, 1980a) gave a set of sufficient conditions for a two-dimensional data set to behave as a perfectly Rasch homogeneous data set, in the sense that all item parameter estimates are, within chance limits, equal over all subsamples of the raw score partitioning.

Even when the above conditions are only met to a certain degree, the sensitivity of the test statistics for violation of the unidimensionality axiom may be decisively undermined. It must, therefore, be concluded that the available global test procedures based on a raw score partitioning of the data set may fail to detect multidimensionality.

One alternative to the intended testing procedures is Q_2, in which local stochastic independence is tested by means of the association between item pairs. When unidimensionality is violated and only one (combination) trait is obtained in the analysis, then there will still be inter-item association left and, as a consequence, local stochastic independence will be violated.

The second possibility is to use a partitioning criterion that is differentially related to the underlying traits. However, a priori knowledge about the underlying traits will, as a rule, not be available; when it is known that more than one latent trait underlies the data, the dichotomous Rasch model should not be used anyhow. Thus, a method is called for that is not dependent upon substantial knowledge about the test being analyzed.

Consider the following case. A test consists of items, some of which appeal to trait A (e.g., items 1, 2, 3 and 4), whereas the other items appeal to trait B (e.g., items 5 through 8). The mean subject parameters on both traits are equal, i. e., $\bar{\xi}_A = \bar{\xi}_B$. One of the items is removed from the test. Assume, without loss of generality, that this is an A-type item, say item 1. Now consider two groups of subjects, those responding positively to item 1 and those responding negatively to this item. These groups of subjects will be designated as the A^+ and the A^- sample, respectively.

On the average, the A^+ sample consists of subjects with relatively high parameters for trait A, whereas the A^- sample consists of subjects with relatively low A parameters. When, again without loss of generality, it is assumed that the traits are uncorrelated, the average trait parameter for trait B will be equal for both subsamples:

	Sample		
Trait	A^+		A^-
A	$\bar{\xi}_A^+$	$>$	$\bar{\xi}_A^-$
B	$\bar{\xi}_B^+$	$=$	$\bar{\xi}_B^-$

The state of affairs can also be formulated in another way: For the A^+ sample the A items are relatively easy as compared to the B items, which are actually of intermediate difficulty. For the A^- subsample the situation is reversed, the A items being relatively difficult in comparision with the B items.

The differences in relative difficulty between the subsamples must reflect in the item parameter estimates of the two subsamples; in subsample A^+ the A items are the more easy ones, whereas in subsample A^- the B items are easier. The traditional tests, inspecting the equality of item parameter estimates over subsamples, will be successful in detecting violation of the unidimensionality axiom. Of course, the item that is used as a partitioning criterion, item 1 in this case, should be excluded from the test to be analyzed. This item is, in an artifactual way, extremely easy in the A^+ subsample and extremely difficult in the A^- subsample. It may be obvious that the correlation between the latent traits and the equality of the mean subject parameters of the latent traits influence the *extent* to which the violation becomes manifest but are of no importance for the effect as such.

It is also possible to use more than one item as an external partitioning criterion. When the sum score of the external criterion items is used for this purpose, the danger exists that the partitioning criterion suffers the same evil that it is supposed to be a remedy for. That is, when the items constituting the external criterion stem from both latent traits, the partitioning criterion is no longer differentially related to the underlying latent traits, which is necessary in order to detect violation of the unidimensionality axiom.

There is yet another way in which more items can be used to give a partitioning of the data set, that is, by using the response pattern of the criterion items instead of the sum score. With two items there are four subsamples: (1) + +, (2) + −, (3) − +, and (4) − −.

When the items stem from the same dimension, say trait A, the following situation holds:

	Criterion Pattern						
Trait	−−		−+		+−		++
A	$\bar{\xi}_A^{--}$	$<$	$\bar{\xi}_A^{-+}$	$=$	$\bar{\xi}_A^{+-}$	$<$	$\bar{\xi}_A^{++}$
B	$\bar{\xi}_B^{--}$	$=$	$\bar{\xi}_B^{-+}$	$=$	$\bar{\xi}_B^{+-}$	$=$	$\bar{\xi}_B^{++}$

It is obvious that the violation of unidimensionality now is detected in quite the same way as in the one-criterion item case. When the two items are from different subtests, the pattern partitioning gives the following situation with respect to the mean subject parameters:

	Criterion Pattern									
Trait	−−		−+		+−		++			
A	$\bar{\xi}_A^{--}$	$=$	$\bar{\xi}_A^{-+}$	$<$	$\bar{\xi}_A^{+-}$	$=$	$\bar{\xi}_A^{++}$	$(\bar{\xi}_B^{--}$	$=$	$\bar{\xi}_B^{+-})$
B	$\bar{\xi}_B^{--}$	$<$	$\bar{\xi}_B^{-+}$	$>$	$\bar{\xi}_B^{+-}$	$<$	$\bar{\xi}_B^{++}$	$(\bar{\xi}_B^{-+}$	$=$	$\bar{\xi}_B^{++})$

It is seen that the effects of violation of unidimensionality will especially become manifest in the +− and the −+ subsamples. In the +− subsample the subjects will have relatively high parameter values on trait A and simultaneously low values on trait B; for the −+ subsample the situation is reversed. The differences between parameter estimates of the items of the two subtests will diverge on the basis of two effects on the mean subject-parameters.

The subsamples ++ and −− are not differentiated with respect to the latent traits A and B and therefore are not expected to contribute to the significance of the overall test statistic. They are of minor importance in the course of testing the unidimensionality axiom.

A Simulation Study

The sensitivity of the present test procedure with respect to violation of the dimensionality axiom was investigated by means of simulated data. For the generation of the data sets, two different methods were used, both of which are described by van den Wollenberg (1979, 1980a). A short description of both methods will be presented here.

Method

Method A generates data such that chance fluctuations are present; the statistics should behave as χ^2 with the appropriate number of degrees of freedom. The steps involved in method A are as follows:

1. For the items fixed parameter values are chosen.
2. A subject parameter ξ_v is sampled from the standard normal distribution, although any other distribution would do.
3. Given the item parameters and the subject parameter, the response probabilities are obtained by means of the basic formula of the Rasch model:

$$\pi_{vi} = \frac{\exp(\xi_v - \sigma_i)}{1 + \exp(\xi_v - \sigma_i)} . \qquad [7]$$

4. Corresponding to the response probabilities under Number 3, k (number of items) random numbers are sampled from the uniform distribution with range (0,1).
5. The response probabilities under Number 3 are compared with the random numbers under Number 4; a simulee is said to give a positive response to a given item when the response probability exceeds the corresponding random number.
6. For every simulee steps 2 through 5 are repeated.

The method described above gives a data set that is in accordance with the Rasch model; the test statistics have χ^2 properties. Multidimensionality of the data set is introduced by sampling a vector-valued subject parameter under Number 3 and by letting some of the items appeal to one element of the vector-valued parameter and other items to another element.

Method B generates the expected number of each response vector, given the item and subject parameters. When the model holds, this method leads to statistics that are equal to zero, except for rounding errors. This method is especially appropriate for studying the effects of model violations, as the value of the statistics is totally due to violation; no chance fluctuations are involved. Steps 1 through 3 are the same as in method A. The succeeding steps are the following:

4. For a given simulee the probabilities of all response patterns are obtained.
5. The probability of each response pattern is summated over all simulees. This summation is the expected frequency for each response pattern.
6. Response vectors are generated up to the expected number rounded to the nearest integer.

Data sets that were obtained by means of both of the above methods are reported upon. The data sets shared the following characteristics: A total of 4,000 subjects was sampled, using eight items. All item parameters were equal ($\sigma_i = 0.0$; $i = 1, ..., k$). The subject parameters were sampled from the bivariate standard normal distribution. The first four items appealed to the first subject parameter, the others to the second subject parameter; the correlation between the subject parameters was zero. For the data set generated according to method A, one hundred replications were obtained, which was, of course, not necessary for method B.

Results

In Table 1 a comparison is offered of the Q_1 and Q_2 statistics based on a high-low partitioning of the data set on the one hand and the Q_1 statistic in connection with an external criterion item on the other.[1] As can be observed, the Q_1 statistic associated with a high-low-partitioning of the data set is insensitive to multidimensionality. The Q_2 statistic reacts very clearly to the model violation, as it should. Multidimensionality is also clearly detected by the Q_1 statistic in connection with an external criterion item. It must be noted that one degree of freedom is lost in the last testing procedure because one item is removed from the test and used for the partitioning of the data set. Indeed, the partitioning of the data set by means of an item of the test is effective in detecting violation of the dimensionality axiom; the Q_2 statistic does the job even better.

In Table 2 results are presented pertaining to model tests making use of two external criterion items. In the first part of the table, the results are given for the situation in which both criterion items appeal to the same latent trait, whereas in the second part the items appeal to different latent traits. The remaining number of items in the test is six. When both items appeal to the same dimension, the partitioning criterion is differentially related to the items in the test and the test of the model should

[1] Q_2 should preferably be used with a complete raw score partitioning (see van den Wollenberg, 1980a). However, with other types of partitioning it can be used heuristically, which has been done here for the sake of comparison.

Table 1
Comparison of the Statistics Q_1 and Q_2 Based on a High-Low Partitioning and the Statistic Q_1 Based on a Partitioning by Means of an External Criterion Item, for Data Generation Methods A and B

Partitioning Method	Data Generation Method B (1 Replication)	 A (100 Replications)
High-Low	Q_1 = .03	$\bar{Q}_1$ = 6.72 (df = 7)
	Q_2 = 755.46	$\bar{Q}_2$ = 906.97 (df = 40)
External Criterion Item	Q_1 = 163.22	$\bar{Q}_1$ = 191.93 (df = 6)

show the violation of the model, as indeed it does. The raw score partitioning gives three subsamples, whereas the response pattern partitioning leads to four subsamples; this explains the difference in the number of degrees of freedom.

It is a remarkable fact that for the fluctuation-free-data, both types of partitioning lead to the same value for the Q_1 statistic; this point will be discussed below. When the items appeal to different dimensions, the raw score partitioning criterion loses its effectiveness, as may be apparent from Table 2. In this case the response pattern partitioning remains effective and even becomes more sensitive to the model violation.

The use of more than one item as partitioning criterion is thus quite possible. However, when the raw score of these items is used, the danger exists that the criterion loses its effectiveness. This does not happen when the response pattern partitioning is used.

In Table 3 the results of Table 2 are inspected more closely. To this end only the data sets generated by means of method B were used. Here the results for the analysis with the criterion items ap-

Table 2
Testing the Dimensionality Axiom by Means of Two External Criterion Items Using Both a Raw Score and a Response Pattern Partitioning

Partitioning Method	Data Generation Method B	 A
Both Criterion Items Appealing to the Same Dimension		
Raw Score	Q_1 = 219.36	$\bar{Q}_1$ = 259.52 (df = 10)
Response Pattern	Q_1 = 219.36	$\bar{Q}_1$ = 260.30 (df = 15)
Criterion Items Appealing to Different Dimensions		
Raw Score	Q_1 = 0.0	$\bar{Q}_1$ = 8.61 (df = 10)
Response Pattern	Q_1 = 280.62	$\bar{Q}_1$ = 333.16 (df = 15)

Table 3
Contributions of the Subsamples to the Overall Statistic for Model Tests With Two External Criterion Items Using Data Generation Method B

Raw Score	Contribution	Pattern	Contribution
Both Criterion Items Appealing to the Same Dimension			
0	106.76	00	106.76
1	0.01	01	0.01
		10	0.01
2	112.59	11	112.59
Q_1	219.36		219.36
Criterion Items Appealing to Different Dimensions			
0	0.00	00	0.00
1	0.00	01	140.31
		10	140.31
2	0.00	11	0.00
Q_1	0.00		280.62

pealing to the same dimension are reported, just like the results for the two-dimensional criterion. In Table 3, however, the contributions of the subsamples are given.

When the raw score partitioning is used in connection with an external criterion appealing to just one of the dimensions, then score group 1 does not contribute to the value of the test statistic. The explanation is obvious. This group consists of subjects for which the items are of intermediate difficulty, which also holds for the items appealing to trait B, and there is no differentiation between items of trait A and trait B; consequently, multidimensionality is not detected. Looking at the response pattern partitioning, the same can be said for the groups with response patterns 01 and 10. This does also explain the fact that the overall statistics are equal for these two types of partitioning of the data set.

When the items appeal to different dimensions, the raw score partitioning totally fails, as was indicated earlier. When the results for the response pattern partitioning are inspected, subsamples can again be observed that do not contribute to the overall test statistic. Again, the explanation can be found in the fact that in these groups there is no differential relation to the underlying traits. For the group with response pattern 00, both subject parameters on trait A and trait B are expected to be below average; for the group with response pattern 11 both the A parameters and the B parameters are relatively high. In both instances there is no difference with respect to the underlying dimensions, and it is this differential relation that is necessary to detect multidimensionality.

The violation of the dimensionality axiom not only becomes manifest in the value of the Q_1 statistic, when test items are used as external criterion, but also the pattern of contributions over the subsamples is predictable. This can give useful added information when the model is tested.

Discussion

In the present study a method was presented to test the dimensionality axiom of the Rasch model. This method is to be looked upon as a useful alternative for Q_2 (van den Wollenberg, 1980a) in cases where the latter methods meets practical objections.

Whenever possible, Q_2 is to be preferred, as it is more sensitive to violations and because the comparison of observed and expected second-order frequencies gives useful information with respect to the sources of model violation. However, the practical disadvantages of Q_2, such as computing time, the risk that for some subsamples the parameters may be inestimable, and the instability of the statistic when there are small expected frequencies, are not to be underestimated in practical applications.

In the present analyses the tests to be analyzed in fact consisted of two Rasch homogeneous subtests; each item appealed to only one latent trait. However, it is also quite feasible that an item appeals to some linear combination of the underlying traits. In this case, the differential relation of the item to the underlying traits can be at issue; now several analyses with different external criterion items must be performed in order to get a good check on the dimensionality assumption.

In the present study a two-dimensional structure was studied. The present results can easily be generalized to more than two dimensions.

As was argued, the pattern of contributions over subsamples can also be used in the evaluation of model fit. However, it should be kept in mind that the specific results obtained here are also dependent upon the specific raw score distributions of the different subtests and the value of the item parameters. Thus, the present results with regard to the pattern of contributions cannot be generalized blindly. Nevertheless, it can be stated that a two-dimensional structure will tend to show a pattern of contributions quite similar to the one reported here.

References

Andersen, E. B. A goodness of fit test for the Rasch model. *Psychometrika,* 1973, *38,* 123-140.

Fischer, G. H., & Scheiblechner, H. H. Algorithmen und Programmen für das probabilistische Testmodel von Rasch. *Psychologische Beiträge,* 1970, *12,* 23-51.

Gustafsson, J. E. Testing and obtaining fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology,* 1980, *33,* 205-233.

Gustafsson, J. E., & Lindblad, T. *The Rasch model for dichotomous items: A solution of the conditional estimation problem for long tests and some thoughts about item screening procedures* (Report No. 67). Mölndal, Sweden: University of Göteberg, Institute of Education, June 1978. (Paper presented at the European Conference on Psychometrics and Mathematical Psychology, Uppsala, Sweden, June 1978.)

Martin-Löf, P. *Statistika Modeller, Anteckningar från seminarier Lasåret 1969-1970 utarbetade av Rolf Sunberg obetydligt ändrat nytryk.* 2:a uppl. Stockholm: Institutet för Försäkringsmatematik och Matematisk Statistik vid Stockholms Universitet, October 1973.

Stelzl, I. Ist der Modelltest des Rasch-Modells geeignet Homogenitäts Hypothesen zu prüfen? Ein Berich über Simulation Studien mit inhomogene Daten. *Zeitschrift für experimentalle und angewandte Psychologie,* 1979, *26,* 652-672.

Wollenberg, A. L. van den. *The Rasch model and time limit tests.* Unpublished doctoral dissertation, Nijmegen, 1979.

Wollenberg, A. L. van den. *Two new test statistics for the Rasch model* (Internal Report 80-MA-01). Nijmegen, The Netherlands: Katholieke Universiteit Nijmegen, Vakgroep Mathematical Psychologie, Psychologisch Laboratorium, August 1980. (a) (In press, *Psychometrika)*

Wollenberg, A. L. van den. *On the Wright-Panchapakesan goodness of fit test for the Rasch model* (Internal Report 80-MA-02). Nijmegen, The Netherlands: Katholieke Universiteit Nijmegen, Vakgroep Mathematische Psychologie, Psychologisch Laboratorium, 1980. (b)

Wright, B. D., & Panchapakesan, N. A procedure for sample-free item analysis. *Educational and Psychological Measurement,* 1969, *29,* 23-48.

Author's Address

Send requests for reprints or further information to A. L. van den Wollenberg, Vakgroep Mathematische Psychologie, Psychologisch Laboratorium, Postbus 9104, 6500 HE NIJMEGEN, The Netherlands.