# Computation of the Molenaar Sijtsma Statistic

L. Andries van der Ark[1]

Department of Methodology and Statistics, Tilburg University, P. O. Box 90153, 5000 LE, Tilburg, The Netherlands `a.vdark@uvt.nl`

**Summary.** The Molenaar Sijtsma statistic is an estimate of the reliability of a test score. In some special cases, computation of the Molenaar Sijtsma statistic requires provisional measures. These provisional measures have not been fully described in the literature, and we show that they have not been implemented in the software. We describe the required provisional measures as to allow the computation of the Molenaar Sijtsma statistic for all data sets.

**Key words:** Molenaar Sijtsma Statistic, Reliability, Psychological Test Construction.

## 1 Introduction

Psychological and educational tests are often used for the classification of respondents. For example, a clinical psychologist may decide that one patient needs special treatment and another patient does not, based on their scores on a psychological test; and the decision of an admission committee of a university may strongly depend on the student's score on an educational test. A valid classification requires that the test scores are reliable, which can be investigated by a reliability statistic. Most well known reliability statistics (e.g., Cronbach's alpha [1], lambda-2 [2], the greatest lower bounds [3]) are lower bounds to the reliability. The Molenaar Sijtsma statistic (MS) [5, 6, 8, 9]. gives a direct estimate of the reliability of a test score. Simulation studies showed that MS was almost unbiased and had less bias and smaller variance than other reliability statistics [9, 11]. Therefore, MS gives a more accurate estimate of the reliability than other well known reliability statistics. MS is implemented in the software package MSP5.0 [7].

In some special cases MS cannot be computed straightforwardly and provisional measures are required [6] but these have never been discussed in detail and, as we show in Sect. 4, have not been implemented in the software package MSP5.0. Therefore, the researcher is left in the dark what to do in these special cases. This paper discusses all details of the computation of MS, so as to allow the computation of MS in all cases. For reasons of space we do not discuss details of the rationale of MS and its background theory. We refer the interested reader to [5, 6, 8, 9].

Assume that a test consists of $J$ items, indexed by $i$ and $j$. Each item has $m+1$ ordered answer categories $0, \ldots, m$; indexed by $g$ and $h$. The items scores are denoted by $X_1, \ldots, X_J$. Assume that $N$ respondents have responded to the $J$ items and there are no missing values. For each respondent the test score $X = \sum X_i$ is used for classification. In classical test theory the expected value of a respondent's test score over independent replications is called the *true score* and is denoted by $T$ [4]. $T$ is unobservable. Let $\sigma^2(X)$ and $\sigma^2(T)$ denote the population variance of the test score and the true score, respectively. Under the assumptions of the classical test theory, the reliability of $X$ is defined as $\rho_{XX'} = \sigma^2(T)/\sigma^2(X)$ [4]. Since $\sigma^2(T)$ is unobservable, the reliability cannot be computed directly and must be estimated.

Let $\pi_{g(i)} = P(X_i \geq g)$ denote the marginal cumulative probability of obtaining a score of at least $g$ on item $i$, and let $\pi_{g(i),h(j)} = P(X_i \geq g, X_j \geq h)$ denote the joint cumulative probability of obtaining a score of at least $g$ on item $i$ and at least $h$ on item $j$. Molenaar and Sijtsma [6] showed that $\sigma^2(T) = \sum_{i=1}^{J} \sum_{g=1}^{m} \sum_{j=1}^{J} \sum_{h=1}^{m} [\pi_{g(i),h(j)} - \pi_{g(i)} \times \pi_{h(j)}]$, and, therefore, the reliability of $X$ can be expressed as

$$\rho_{XX'} = \sum_{i=1}^{J} \sum_{g=1}^{m} \sum_{j=1}^{J} \sum_{h=1}^{m} \left[ \frac{\pi_{g(i),h(j)} - \pi_{g(i)} \times \pi_{h(j)}}{\sigma^2(X)} \right]. \tag{1}$$

MS estimates the reliability of $X$ by plugging in estimates for each term in Equation 1. The following estimates are straightforward because they only depend on observable item scores.

- The population variance of the test score, $\sigma^2(X)$, is estimated by the (biased) sample variance

$$S^2(X) = \frac{1}{N} \sum_{n=1}^{N} (X_n - \overline{X})^2.$$

- The marginal cumulative probabilities $\pi_{g(i)}$ and $\pi_{h(j)}$ are estimated by the corresponding marginal cumulative proportions in the sample, denoted $P_{g(i)}$ and $P_{h(j)}$, respectively.
- If $i \neq j$, the joint cumulative probabilities $\pi_{g(i),h(j)}$ are estimated by the corresponding observable joint cumulative proportions in the sample, denoted $P_{g(i),h(i)}$.

If $i = j$, $\pi_{g(i),h(i)}$ is the joint probability of obtaining at least score $g$ and at least score $h$ on item $i$ in two independent replications. Estimation is not straightforward because the corresponding joint cumulative proportions in the sample are unobservable in a single test administration. Two cases are distinguished. In Case I, there are no marginal cumulative proportions with exactly the same values. Case I, which requires no provisional measures, is discussed in Sect. 2. In Case II, one or more marginal cumulative proportions have exactly the same value. Case II, which requires provisional measures, is discussed in Sect. 3. In Sect. 4, we show that MSP5.0 can produce an incorrect MS.

## 2 Case I: The Computation of MS When No Provisional Measures Are Needed

Case I is explained using the first numerical example, which consists of four items, each with three ordered categories. Table 1 shows the marginal cumulative proportions. Marginal cumulative proportions $P_{0(i)}$ $(i = 1, \ldots, J)$ equal 1 by definition and are not informative.

**Table 1.** Marginal cumulative proportions of the first numerical example

|  | $i = 1$ | $i = 2$ | $i = 3$ | $i = 4$ |
|---|---|---|---|---|
| $P_{0(i)}$ | 1.00 | 1.00 | 1.00 | 1.00 |
| $P_{1(i)}$ | .90 | .80 | .70 | .60 |
| $P_{2(i)}$ | .50 | .40 | .30 | .20 |

The first step in estimating the unobservable joint cumulative probabilities $\pi_{g(i),h(i)}$ is to rank all the informative marginal cumulative proportions from small to large. For the first numerical example, Table 1 shows that this rank order is

$$P_{2(4)} < P_{2(3)} < P_{2(2)} < P_{2(1)} < P_{1(4)} < P_{1(3)} < P_{1(2)} < P_{1(1)}. \qquad (2)$$

The second step in estimating the joint cumulative probabilities $\pi_{g(i),h(i)}$ is to create a matrix of joint cumulative proportions in which the rows and columns are ordered by the size of the corresponding marginal cumulative proportions (cf. Equation 2 in the first step). Table 2 shows this matrix of joint cumulative proportions for the first numerical example. NA indicates that a joint cumulative proportion is unobservable and must estimated. Assume that joint cumulative proportion $P_{g(i),h(i)}$ is in the cell with row $r$ and column $c$. For convenience, $P_{g(i),h(i)}$ is denoted $P_{r,c}$, and the corresponding marginal cumulative probabilities $P_r$ and $P_c$, respectively. For example, $P_{2(4),1(4)}$ is in row 1 and column 5 of Table 2 and is, therefore, denoted $P_{1,5}$.

The third step in estimating the unobservable joint cumulative probability $\pi_{g(i),h(i)}$ is to define

1. the *lower neighboring joint cumulative proportion*: $P_{lo} = P_{r+1,c}$,

**Table 2.** Marginal cumulative proportions (boldface) and joint cumulative proportions of the first numerical example

| | | $P_{2(4)}$ | $P_{2(3)}$ | $P_{2(2)}$ | $P_{2(1)}$ | $P_{1(4)}$ | $P_{1(3)}$ | $P_{1(2)}$ | $P_{1(1)}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | **.20** | **.30** | **.40** | **.50** | **.60** | **.70** | **.80** | **.90** |
| $P_{2(4)}$ | **.20** | NA | .20 | .20 | .20 | NA | .20 | .20 | .20 |
| $P_{2(3)}$ | **.30** | .20 | NA | .30 | .30 | .30 | NA | .30 | .30 |
| $P_{2(2)}$ | **.40** | .20 | .30 | NA | .40 | .40 | .40 | NA | .40 |
| $P_{2(1)}$ | **.50** | .20 | .30 | .40 | NA | .50 | .50 | .50 | NA |
| $P_{1(4)}$ | **.60** | NA | .30 | .40 | .50 | NA | .60 | .60 | .60 |
| $P_{1(3)}$ | **.70** | .20 | NA | .40 | .50 | .60 | NA | .70 | .70 |
| $P_{1(2)}$ | **.80** | .20 | .30 | NA | .50 | .60 | .70 | NA | .80 |
| $P_{1(1)}$ | **.90** | .20 | .30 | .40 | NA | .60 | .70 | .80 | NA |

2. the *right-hand neighboring joint cumulative proportion*: $P_{ri} = P_{r,c+1}$,
3. the *upper neighboring joint cumulative proportion*: $P_{up} = P_{r-1,c}$, and
4. the *left-hand neighboring joint cumulative proportion*: $P_{le} = P_{r,c-1}$.

It may be noted that not all four neighboring joint cumulative proportions need exist. For example, for $P_{1,5}$, $P_{up}$ does not exist, $P_{lo} = .30$, $P_{le} = .20$, and $P_{ri} = .20$.

The fourth step is to estimate the unobservable joint cumulative probability $\pi_{g(i),h(i)}$ eight times using the following eight different estimates (see, [6], for the derivation).

$$P^{(1)}_{r,c} = P_{lo}\frac{P_r}{P_{r+1}} \tag{3}$$

$$P^{(2)}_{r,c} = P_{ri}\frac{P_c}{P_{c+1}} \tag{4}$$

$$P^{(3)}_{r,c} = P_{up}\frac{P_r}{P_{r-1}} \tag{5}$$

$$P^{(4)}_{r,c} = P_{le}\frac{P_c}{P_{c-1}} \tag{6}$$

$$P^{(5)}_{r,c} = P_{lo}\frac{1-P_r}{1-P_{r+1}} - P_c\frac{P_{r+1}-P_r}{1-P_{r+1}} \tag{7}$$

$$P^{(6)}_{r,c} = P_{ri}\frac{1-P_c}{1-P_{c+1}} - P_r\frac{P_{c+1}-P_c}{1-P_{c+1}} \tag{8}$$

$$P^{(7)}_{r,c} = P_{up}\frac{1-P_r}{1-P_{r-1}} + P_c\frac{P_r-P_{r-1}}{1-P_{r-1}} \tag{9}$$

$$P^{(8)}_{r,c} = P_{le}\frac{1-P_c}{1-P_{c-1}} + P_r\frac{P_c-P_{c-1}}{1-P_{c-1}} \tag{10}$$

Joint cumulative probability $\pi_{g(i),h(i)}$ is then estimated by $\overline{P}_{r,c}$, the mean of all existing estimates in Equations 3 to 10. For the first numerical example, it may be noted that

$$
\begin{aligned}
&P_{1,5}^{(1)} = .3 \times \frac{.2}{.3} = .2\\
&P_{1,5}^{(2)} = .2 \times \frac{.6}{.7} = .1714\\
&P_{1,5}^{(3)} \;\; \text{does not exist}\\
&P_{1,5}^{(4)} = .2 \times \frac{.6}{.5} = .24\\
&P_{1,5}^{(5)} = .3 \times \frac{1-.2}{1-.3} - .6 \times \frac{.3-.2}{1-.3} = .2571\\
&P_{1,5}^{(6)} = .2 \times \frac{1-.6}{1-.7} - .2 \times \frac{.7-.6}{1-.7} = .2\\
&P_{1,5}^{(7)} \;\; \text{does not exist}\\
&P_{1,5}^{(8)} = .2 \times \frac{1-.6}{1-.5} + .2 \times \frac{.6-.5}{1-.5} = .2
\end{aligned}
$$

As a result

$$\overline{P}_{1,5} = \frac{.2 + .1714 + .24 + .2571 + .2 + .2}{6} = .2114$$

It was noted by [6] that $\overline{P}_{r,c}$ should lie in the interval

$$P_r P_c \leq \overline{P}_{r,c} \leq \min(P_r, P_c).$$

For $\overline{P}_{1,5}$, the lower bound equals $.2 \times .6 = .12$ and the upper bound equals $min(.2, .6) = .2$. Hence, the final estimate for $\pi_{1(4),2(4)} = .2$. Table 3 shows the joint cumulative proportions of the first numerical example, with all estimated unobservable joint cumulative proportions underlined. The joint cumulative proportions in Table 3 are plugged into Equation 1. Suppose that $S^2(X) = 9$. Using the values in Table 3 it may then be verified that

$$MS = \sum_{i=1}^{J} \sum_{g=1}^{m} \sum_{j=1}^{J} \sum_{h=1}^{m} \left[ \frac{P_{g(i),h(j)} - P_{g(i)} \times P_{h(j)}}{S^2(X)} \right] = \frac{7.137}{9} = .793.$$

## 3 Case II: The Computation of MS When Provisional Measures Are Needed

The following citation taken from [6] illustrates that Sect. 2 is not sufficient for computing MS in all cases.

> Furthermore, alternative approximations methods are used when $P_{g(i)}$ or $P_{h(j)}$ or both, belong to a string of identical proportions. In such cases the choice of adjacent elements becomes problematic. Since the discussion of the solutions to this problem would take much space, we prefer to give only a brief outline

**Table 3.** Marginal cumulative proportions (boldface) and joint cumulative proportions of the first numerical example. Estimated unobservable joint cumulative probabilities (accuracy in three digits) are underlined.

| | | $P_{2(4)}$ | $P_{2(3)}$ | $P_{2(2)}$ | $P_{2(1)}$ | $P_{1(4)}$ | $P_{1(3)}$ | $P_{1(2)}$ | $P_{1(1)}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | **.20** | **.30** | **.40** | **.50** | **.60** | **.70** | **.80** | **.90** |
| $P_{2(4)}$ | **.20** | $\underline{.167}$ | .20 | .20 | .20 | $\underline{.200}$ | .20 | .20 | .20 |
| $P_{2(3)}$ | **.30** | .20 | $\underline{.259}$ | .30 | .30 | .30 | $\underline{.300}$ | .30 | .30 |
| $P_{2(2)}$ | **.40** | .20 | .30 | $\underline{.359}$ | .40 | .40 | .40 | $\underline{.400}$ | .40 |
| $P_{2(1)}$ | **.50** | .20 | .30 | .40 | $\underline{.458}$ | .50 | .50 | .50 | $\underline{.500}$ |
| $P_{1(4)}$ | **.60** | $\underline{.200}$ | .30 | .40 | .50 | $\underline{.559}$ | .60 | .60 | .60 |
| $P_{1(3)}$ | **.70** | .20 | $\underline{.300}$ | .40 | .50 | .60 | $\underline{.659}$ | .70 | .70 |
| $P_{1(2)}$ | **.80** | .20 | .30 | $\underline{.400}$ | .50 | .60 | .70 | $\underline{.761}$ | .80 |
| $P_{1(1)}$ | **.90** | .20 | .30 | .40 | $\underline{.500}$ | .60 | .70 | .80 | $\underline{.875}$ |

A detailed discussion of the solutions is presented here. In Case II, $P_{g(i)}$ or $P_{h(j)}$ or both may belong to a string of identical proportions Case II is explained using a the second numerical example consisting of four items, each with three ordered categories. The second numerical example contains two strings of identical marginal cumulative proportions (Table 4).

**Table 4.** Marginal cumulative proportions of the second numerical example

| | $i = 1$ | $i = 2$ | $i = 3$ | $i = 4$ |
|---|---|---|---|---|
| $P_{0(i)}$ | 1.00 | 1.00 | 1.00 | 1.00 |
| $P_{1(i)}$ | .60 | .60 | .60 | .50 |
| $P_{2(i)}$ | .40 | .30 | .20 | .20 |

As in Case I, the marginal cumulative proportions are put in an ascending order. For the second numerical example the rank order of the cumulative marginal proportions is

$$\{P_{2(4)}, P_{2(3)}\} < P_{2(2)} < P_{2(1)} < P_{1(4)} < \{P_{1(3)}, P_{1(2)}, P_{1(1)}\}.$$

There is no unique order of the cumulative marginal proportions and, therefore, there is no unique order of the rows and columns of the matrix of joint cumulative proportions (Table 5). The order of rows 1 and 2; the order of columns 1 and 2; the order of rows 6, 7, and 8; and the order of columns 6, 7, and 8 are undetermined. As a result, the neighboring joint cumulative proportions $P_{lo}$, $P_{le}$, $P_{ri}$, and $P_{up}$ cannot be estimated unambiguously.

In general, four types of cells can be distinguished in the matrix of joint cumulative probabilities.

1: A cell whose row and column are in an arbitrary order.

**Table 5.** Marginal cumulative proportions (boldface) and joint cumulative proportions of the second numerical example

| | | $P_{2(4)}$ | $P_{2(3)}$ | $P_{2(2)}$ | $P_{2(1)}$ | $P_{1(4)}$ | $P_{1(3)}$ | $P_{1(2)}$ | $P_{1(1)}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | **.20** | **.20** | **.30** | **.40** | **.50** | **.60** | **.60** | **.60** |
| $P_{2(4)}$ | **.20** | NA | .20 | .20 | .20 | NA | .20 | .20 | .20 |
| $P_{2(3)}$ | **.20** | .20 | NA | .20 | .20 | .20 | NA | .20 | .20 |
| $P_{2(2)}$ | **.30** | .20 | .20 | NA | .30 | .30 | .30 | NA | .30 |
| $P_{2(1)}$ | **.40** | .20 | .20 | .30 | NA | .40 | .40 | .40 | NA |
| $P_{1(4)}$ | **.50** | NA | .20 | .30 | .40 | NA | .50 | .50 | .50 |
| $P_{1(3)}$ | **.60** | .20 | NA | .30 | .40 | .50 | NA | .60 | .60 |
| $P_{1(2)}$ | **.60** | .20 | .20 | NA | .40 | .50 | .60 | NA | .60 |
| $P_{1(1)}$ | **.60** | .20 | .20 | .30 | NA | .50 | .60 | .60 | NA |

2: A cell whose column is in an arbitrary order and whose row is in a unique order.
3: A cell whose row is in an arbitrary order and whose column is in a unique order.
4: A cell whose row and column are in a unique order.

Table 6 shows the type of cell for each cumulative joint proportion in Table 5. If two or more adjacent joint cumulative proportions have the same marginal cumulative proportions, then we say that the corresponding cells in the matrix of joint cumulative proportions belong to the same *set*. In Table 6, if two cells are not separated by a line, then the cells belong to the same set. For example, $P_{1,1}$, $P_{1,2}$, $P_{2,1}$, and $P_{2,2}$ belong to the same set.

**Table 6.** Types and sets of cells of Table 5. Cells pertaining to unobservable joint cumulative proportions are underlined. Marginal cumulative proportions are in boldface

| | | $P_{2(4)}$ | $P_{2(3)}$ | $P_{2(2)}$ | $P_{2(1)}$ | $P_{1(4)}$ | $P_{1(3)}$ | $P_{1(2)}$ | $P_{1(1)}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | **.20** | **.20** | **.30** | **.40** | **.50** | **.60** | **.60** | **.60** |
| $P_{2(4)}$ | **.20** | $\underline{1}$ | 1 | 3 | 3 | $\underline{3}$ | 1 | 1 | 1 |
| $P_{2(3)}$ | **.20** | 1 | $\underline{1}$ | 3 | 3 | 3 | $\underline{1}$ | 1 | 1 |
| $P_{2(2)}$ | **.30** | 2 | 2 | $\underline{4}$ | 4 | 4 | 2 | $\underline{2}$ | 2 |
| $P_{2(1)}$ | **.40** | 2 | 2 | 4 | $\underline{4}$ | 4 | 2 | 2 | $\underline{2}$ |
| $P_{1(4)}$ | **.50** | $\underline{2}$ | 2 | 4 | 4 | $\underline{4}$ | 2 | 2 | 2 |
| $P_{1(3)}$ | **.60** | 1 | $\underline{1}$ | 3 | 3 | 3 | $\underline{1}$ | 1 | 1 |
| $P_{1(2)}$ | **.60** | 1 | 1 | $\underline{3}$ | 3 | 3 | 1 | $\underline{1}$ | 1 |
| $P_{1(1)}$ | **.60** | 1 | 1 | 3 | $\underline{3}$ | 3 | 1 | 1 | $\underline{1}$ |

In the computation of neighboring joint cumulative proportions, sets rather than cells are considered. The neighboring joint cumulative proportions are defined differently for different types of cells.

If an unobserved cumulative joint probability is in a cell of Type 1, $P_{up}$, $P_{lo}$, $P_{ri}$, and $P_{le}$ are undetermined and set equal to the mean of all observed joint cumulative proportions in the set.

If an unobserved cumulative joint probability is in a cell of Type 2, $P_{ri}$ and $P_{le}$ are set equal to the mean of all observed joint cumulative proportions in the set. $P_{up}$ is set equal to the mean of all observed joint cumulative proportions in the set above the cell, and $P_{lo}$ is set equal to the mean of all observed joint cumulative proportions in the set below the cell. If a set does not exist, the corresponding neighboring joint cumulative proportion does not exist.

If an unobserved cumulative joint probability is in a cell of Type 3, $P_{up}$ and $P_{lo}$ are set equal to the mean of all observed joint cumulative proportions in the set. $P_{ri}$ is set equal to the mean of all observed joint cumulative proportions in the set right of the cell, and $P_{le}$ is set equal to the mean of all observed joint cumulative proportions in the set left of the cell. If a set does not exist, the corresponding neighboring joint cumulative proportion does not exist.

If an unobserved cumulative joint probability is in a cell of Type 4, $P_{up}$ is set equal to the mean of all observed joint cumulative proportions in the set above the cell, $P_{lo}$ is set equal to the mean of all observed joint cumulative proportions in the set below the cell, $P_{ri}$ is set equal to the mean of all observed joint cumulative proportions in the set right of the cell, $P_{le}$ is set equal to the mean of all observed joint cumulative proportions in the set left of the cell. If a set does not exist, the corresponding neighboring joint cumulative proportion does not exist.

Applying these rules to the second numerical example yields the following neighboring joint cumulative proportions.

For $P_{2(4),2(4)}$, $P_{up} = P_{lo} = P_{ri} = P_{le} = .2$;
for $P_{1(4),2(4)}$, $P_{up} = P_{lo} = .2$, $P_{ri} = .2$, and $P_{le} = .2$;
for $P_{2(3),2(3)}$, $P_{up} = P_{lo} = P_{ri} = P_{le} = .2$;
for $P_{1(3),2(3)}$, $P_{up} = P_{lo} = P_{ri} = P_{le} = .2$;
for $P_{2(2),2(2)}$, $P_{up} = .2$, $P_{lo} = .3$, $P_{ri} = .3$, and $P_{le} = .2$;
for $P_{1(2),2(2)}$, $P_{up} = .2$ and $P_{lo} = .4$, $P_{ri} = P_{le} = .3$;
for $P_{2(1),2(1)}$, $P_{up} = .3$, $P_{lo} = .4$, $P_{ri} = .4$, and $P_{le} = .3$;
for $P_{1(1),2(1)}$, $P_{up} = .3$, and $P_{lo} = .5$, $P_{ri} = P_{le} = .4$;
for $P_{1(4),1(4)}$, $P_{up} = .4$, $P_{lo} = .5$, $P_{ri} = .5$, and $P_{le} = .4$;
for $P_{1(3),1(3)}$, $P_{up} = P_{lo} = P_{ri} = P_{le} = .6$;
for $P_{1(2),1(2)}$, $P_{up} = P_{lo} = P_{ri} = P_{le} = .6$; and
for $P_{1(1),1(1)}$, $P_{up} = P_{lo} = P_{ri} = P_{le} = .6$.

Once the neighboring joint cumulative proportions have been computed, the unobservable joint cumulative proportions can be estimated using Equations 3 through 10. and the same procedure as in Case I (Sect. 2) can be used to compute MS.

## 4 Estimation of the Unobservable Joint Cumulative Probabilities in MSP5.0

The third numerical example, containing four items with two ordered answer categories, shows that the provisional measures described in this note are not applied in MSP5.0 [7]. The matrix of joint cumulative proportions is shown in Table 7.

**Table 7.** Marginal cumulative proportions (boldface) and joint cumulative proportions of the third numerical example. The proportions rounded in two decimals were taken from MSP5.0 output. Unobservable proportions are underlined.

| | | $P_{1(1)}$ | $P_{1(2)}$ | $P_{1(3)}$ | $P_{1(4)}$ | $P_{1(5)}$ |
|---|---|---|---|---|---|---|
| | | **.40** | **.60** | **.60** | **.60** | **.70** |
| $P_{1(1)}$ | **.40** | $\underline{.33}$ | .40 | .30 | .40 | .30 |
| $P_{1(2)}$ | **.60** | .40 | $\underline{.40}$ | .30 | .50 | .40 |
| $P_{1(3)}$ | **.60** | .30 | .30 | $\underline{.36}$ | .40 | .50 |
| $P_{1(4)}$ | **.60** | .40 | .50 | .40 | $\underline{.45}$ | .50 |
| $P_{1(5)}$ | **.70** | .30 | .40 | .50 | .50 | $\underline{.57}$ |

Applying the rules computing the neighboring joint cumulative proportions (Sect. 3, p. 7) to $P_{1,1}$ yields the following results. $P_{up}$ and $P_{le}$ do not exist, $P_{lo} = \frac{.4+.4+.3}{3} = .367$. $P_{ri} = \frac{.4+.4+.3}{3} = .367$. Hence,

$$P^{(1)}_{1,1} = P^{(2)}_{1,1} = .367 \times \frac{.4}{.6} = .244,$$

$$P^{(5)}_{1,1} = P^{(6)}_{1,1} = .367 \times \frac{1-.4}{1-.6} - .4 \times \frac{.6-.4}{1-.6} = .35,$$

and $P^{(3)}_{1,1}$, $P^{(4)}_{1,1}$, $P^{(7)}_{1,1}$, and $P^{(8)}_{1,1}$ do not exist. Thus, the correct estimate of $\pi_{1(1),1(1)}$ is $\overline{P}_{1,1} = \frac{.244+.35}{2} = .297$. The incorrect value, $\overline{P}_{1,1} = .33$ reported by MSP5.0 (Table 7), is obtained when one ignores that $P_{1(2)} = P_{1(3)} = P_{1(4)}$ and, in addition, when $P_{1(2)}$ is treated as the only neighboring marginal cumulative proportion (cf. Case I, Sect. 2). This results in $P_{ri} = P_{lo} = .40$. Applying Equations 3 and 4 yield $P^{(1)}_{1,1} = P^{(2)}_{1,1} = .26667$, and applying Equations 7 and 8 yield $P^{(5)}_{1,1} = P^{(6)}_{1,1} = .4$. The average of these four estimates equals the value .33 produced by MSP5.0. It may be noted this problem occurs for all unobservable joint cumulative proportions in Table 7.

## 5 Discussion

The description of the provisional measures required for the computation of MS in special cases fills a gap in the literature on this statistic. The provisional measures are required [5, 6, 9] but were not discussed in detail, and were not incorporated in the software. The explanations in this paper make it possible to compute MS correctly for future applications. For example, as of 2009, the function `check.reliability` in the R package `mokken` [10] computes MS correctly.

There are two reasons to assume that effect of the flaw in the MSP5.0 software is very small or negligible for practical situations. First, it only applies to the special case where two or more marginal cumulative proportions are identical. If sample sizes are sufficiently large, the probability that this happens is rather small. Second, only one or a few of all the joint cumulative proportions needed for the computation of MS are affected by this flaw.

## References

1. L. Cronbach. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16:297–334, 1951.
2. L. Guttman. A basis for analyzing test-retest reliability. *Psychometrika*, 10:255–282, 1945.
3. P.H. Jackson and C.C. Agunwamba. Lower bounds for the reliability of total scores on a test composed of nonhomogeneous items: I: Algebraic lower bounds. *Psychometrika*, 42:567–578, 1977.
4. F.M. Lord and M.R. Novick. *Statistical Theories of Mental Test Scores.* Addison-Wesley, Reading, MA, 1968.
5. I.W. Molenaar and K. Sijtsma. Internal consistency and reliability in Mokken's nonparametric item response model. *Tijdschrift voor onderwijsresearch*, 9:257–268, 1984.
6. I.W. Molenaar and K. Sijtsma. Mokken's approach to reliability estimation extended to multicategory items. *Kwantitatieve methoden*, 9(28):115–126, 1988.
7. I.W. Molenaar and K. Sijtsma. MSP5.0 for Windows [computer software and manual] IEC ProGAMMA, Groningen, 2000.
8. K. Sijtsma. Contributions to Mokken's nonparametric item response theory. Unpublished doctoral dissertation, Vrije Universiteit, Amsterdam, 1988.
9. K. Sijtsmaand I.W. Molenaar. Reliability of test scores in nonparametric item response theory. *Psychometrika*, 52:79–97, 1987.
10. L.A. van der Ark. Mokken scale analysis in R. *Journal of Statistical Software*, 20 (11):1–19, 2007.
11. L.A. van der Ark and D.W. van der Palm. A new reliability coefficient based on latent class analysis. Paper presented at the International Meeting of the Psychometric Society 2008, Durham, July, 2007.