

## **Measuring maternal health literacy in adolescents attending antenatal care in Uganda**

### **Exploring the dimensionality of the “health literacy” concept studying a composite scale**

**Authors:** Guttersrud, Øystein; Naigaga, Mpolampola Desire Alice Sandra; Pettersen, Kjell Sverre

**Source:** Journal of Nursing Measurement, Volume 23, Number 2, 2015, pp. 50E-66E(17)

**Publisher:** Springer Publishing Company

#### **Abstract**

Many scales are constructed as a composition of subscales. The purpose of this paper is to test the assumption of “local independence” in the newly developed “Maternal Health Literacy” (MaHeLi) scale measuring health literacy in pregnant adolescents attending antenatal care. The 20-item scale was administered to 384 adolescents aged 15–19 years attending antenatal care in Uganda during the period July 2013 – December 2013. Rasch analysis was conducted using RUMM2030. Differential item functioning was observed for three items. One item had disordered response categories. One subscale brought substantial multidimensionality into the MaHeLi scale. Results support the use of a 12-item version of the MaHeLi scale. The paper shows how Rasch analyses help us to identify violations of local independence in scales.

**Keywords:** Rasch analysis, violations of local independence in composite scales, developing valid and reliable indicators, maternal health literacy, midwifery, antenatal care, pregnant adolescents

## Introduction

Scales constructed in the field of caring sciences are intended to characterize more or less of some construct. Such scales or “indicators” are usually claimed to be unidimensional as responses to all the items are summed.

To increase the validity of a scale beyond the validity that could be achieved if only one aspect was measured, some scales operationalise different aspects of the construct being measured, with a subscale identified with each aspect. When a scale is constructed by a composition of subscales with the aim of increasing validity, some multidimensionality is expected in the data. Therefore, some violations of unidimensionality might be seen as a positive property of the scale rather than evidence of some fault.

The newly developed “Maternal Health Literacy” (MaHeLi) scale, explored in this study, consists of 20 polytomous items with six response categories. Seven of the items cover the health seeking behaviour (HSB) aspect, six items cover the competence and coping skills (CCS) aspect, and seven items cover the appraisal of health information (AHI) aspect. The aim of this paper is to test the following six folded hypothesis:

*Applied to adolescents aged 15–19 years attending antenatal care in the Jinja and Iganga districts of Uganda the MaHeLi scale represent 1) a well-targeted scale with 2) acceptable person separation that consists of 3) locally independent items that 4) have ordered response categories, 5) are functioning in the same way for different levels of relevant person factors, and 6) show acceptable individual fit to the unidimensional polytomous Rasch model.*

This paper contributes to the international knowledge base on the motivations and abilities of pregnant adolescents to gain access to, understand, and use information in ways that promote and maintain their health and that of their children. This issue is of great international interest and concern and relevant to a wide spectre of clinical practice.

Applying the unidimensional polytomous Rasch model, this paper aims to advance the quality of clinical health research methods by showing how violations of local independence can be identified in the data from composite questionnaire scales where items are deliberately developed to assess different aspects of a construct, with a subscale identified with each aspect.

The paper makes an important contribution to the discussion of the dimensionality of the “health literacy” concept in general; an issue of great interest to researchers in the whole health profession community and with important implications for measurement and the use of questionnaires in numerous clinical samples around the world also beyond the health related professions.

## **Background and Conceptual Framework**

### **Realms of health literacy from an antenatal perspective**

The cognitive and social skills which determine the motivation and ability of women to gain access to, understand and use information in ways to promote and maintain their health and that of their children are collectively referred to as “maternal health literacy” (MHL) (Renkert and Nutbeam, 2001). This definition gives an insight into the crucial link between health literacy and maternal health. Subsequently, three subscales, applicable in the appraisal of maternal health literacy can be derived from the definition above, these are; health seeking behaviour (HSB), appraisal of health information (AHI) and competence and coping skills (CCS).

Health seeking behaviour (HSB) is defined as any activity undertaken by individuals, who perceive themselves to have a health problem or to be ill for the purpose of finding an appropriate remedy (Ward, 1997). Health seeking behaviour has been regarded as a combination of both cognitive and non-cognitive factors with determinant models of health-

seeking behaviours including demographics such as literacy factors (Olenja, 2003). Additionally, studies have indicated that women that are more health literate are more capable of seeking appropriate and timely medical care (Hana and Abdulla, 2013).

The self-evaluative judgments that mothers have about their ability to accomplish their maternal role tasks – their maternal role competence (Pintrich and Schunk D, 1996) – emphasizes the link between societal expectations and capabilities of women as mothers. Thusly, this focuses on the cognitive and social skills that mothers find necessary in order to fulfil these tasks. Collectively, these constitute the suggested Competence and Coping Skills (CCS) sub-construct of MHL. Competence and coping abilities (CCS), particularly perceived maternal role competence refer to the self-evaluative judgements that mothers have about their ability to accomplish their maternal role tasks in pursuit of attaining adequate health status for themselves and their children (Pintrich & Schunk, 1996). In addition, well balanced and interdependent family relationships have been associated with increased wellbeing and positive health behaviour among pregnant adolescents (Stevenson et al., 1999).

Appraisal of health information (AHI) as a principal of maternal health literacy highlights the cognitive and literacy skills required understanding and interpreting the health information availed. Consequently, how health literate one is can be deduced from one's ability to read, filter and understand the health information in order to make sound judgments concerning one's health (European Commission, 2007). The appraisal of health information (AHI) encompasses cognitive skills like comprehension, analysis, synthesis and evaluation that enable one to understand the relevance and application of information. Inadequate health literacy has been found to compound the issue of lacking information appraisal skills (Rosann O'Dell, 2012).

## **The unidimensional polytomous Rasch model (PRM), thresholds, parameterisations and the likelihood ratio test**

The Rasch model (Rasch, 1960), introduced to nursing research by Hagquist et al. (2009), is increasingly used in health studies (Belvedere & de Morton, 2010). Examples of Rasch analyses can be found in Holloway, Blackman and Flynn (2014) and Gonzalez-de Paz et al. (2012).

Rasch models operationalise the formal axioms of measurements (Perline et al., 1979) as they produce linear and invariant measurements, offer strategies of detecting misfit, give estimates of precision, overcome missing data, and rely on a variable – the unidimensional underlying variable describing the quantity being measured – that is independent on the measurement instrument (Wright & Mok, 2004).

Rasch models assume that the raw scores contain all of the information on a person's attitude (sufficiency), that items monotonically relate to one dimension – that the response probability increases with higher attitude (monotonicity) and depends on a dominant dimension (unidimensionality), that the set of items yield locally independent data (local independence) – i.e., that only random, normally distributed “noise” is left after the contribution of the measures (“the Rasch factor”) to the data has been removed (see, for example, Smith, 2002 and Linacre, 1998), and that the items do not show DIF (Brodersen et al., 2007). Only Rasch models provide invariant measurements, reliability, sufficiency and support construct validity if the data fit the model.

In the mathematical representation of the (PRM),  $P\{X_{ni} = x\} = \frac{1}{\gamma} e^{(\kappa_x + x(\beta_n - \delta_i))}$ , where  $\gamma = \sum_{k=0}^m e^{(\kappa_k + k(\beta_n - \delta_i))}$  is a normalisation factor ensuring  $\int_{-\infty}^{\infty} P \cdot d\beta = 1$ , a person's attitude is described by a single, unidimensional latent variable  $\beta_n$  (Andrich, 1988 and Rasch, 1960). The PRM relates the probability ( $P$ ) of person  $n$  with attitude  $\beta_n$  ticking off response category  $x$  on a

polytomous item  $i$  with affective level  $\delta_i$  (Andrich, 1978). The different  $\kappa$  refer to category coefficients.

A threshold is defined as the location at which the probability of responding in one of two adjacent response categories is equal. A polytomous item with an  $m + 1$  number of response categories has  $m$  thresholds ( $\tau_k$ ), where the index  $k$  takes on values from 1 to  $m$  and  $x$  takes on values from 0 to  $m + 1$ . The score  $x$  indicates the number of  $m$  thresholds a respondent has passed (Andrich et al., 1997).

When respondents use the rating scale as expected, the observed succeeding thresholds should reflect successively more of the latent attitude and hence be ordered (Andrich, 1995). Disordered thresholds in the data violate the hypothesised ordering of response categories, meaning that respondents have not used the scales as expected. If so, the variables cannot be treated as interval variables (Singh, 2004). The Rasch model provides the necessary and sufficient means to transform ordinal counts, e.g. levels of attitude, into meaningful linear measurement based on the arithmetical properties of interval scales (Wright & Linacre, 1989).

When the observed distance between the response categories on a rating scale is restricted across the items, the data fit “the rating scale parameterisation” best (Andrich, 1978). If the distance is unrestricted across the items “the partial credit parameterisation” (Wright & Masters, 1982) is indicated. The Fisher’s likelihood ratio test (LRT) might be used to assess the efficiency of the different parameterisations.

### **Tests of fit to the Rasch model**

A Rasch *analysis* means to test the data from a set of items against a Rasch model and check whether the data conform to the model (Tennant & Conaghan, 2007). If the item specific data do not fit the model but rather approach a step function (i.e. a Guttman item), the item is said to over-discriminate. If the data approach a constant function the item is said to under-discriminate.

Fit residuals and item chi-square values are used to test how well the data fit the model (Smith & Plackner, 2009). Positive and negative item fit residuals indicate whether items under- or over-discriminate. A person fit residual indicates how well a person's response pattern matches the expectation under the model (Andrich, 1982 and Andrich, 1985).

Large item chi-squares indicate that persons with different attitudes do not "agree on" item affective estimates, thus compromising the required property of invariance. To adjust chi-square probabilities for the number of significant tests performed, the probabilities are Bonferroni-adjusted (Bland & Altman, 1995) using RUMM2030.

Instruments can be validated using Rasch analysis of data collected in *non-representative samples*. This can be justified by the fact that the person parameter and the item parameter of unidimensional Rasch models are independent (see, for example, Andrich, 2010), and that the Rasch models do not assume anything about the distribution of the sample of persons (see, for example, Andrich, 2004).

### **Tests of local independence**

The local independence assumption implies that there are no dependencies among items other than those that are attributable to the latent trait (persons' attitude). That is, after taking into account the persons' attitude the responses to the questionnaire items should be independent. Violations of local independence have been formalised as "response dependence" and "trait dependence" or "multidimensionality" (Marais & Andrich, 2008a).

Response dependence between items is observed when two items share something more in common than can be accounted for by the latent trait. One example is when two items ask for more or less the same information causing redundancy in the data (Andrich, Humphry & Marais, 2012 and Smith, 2005). Response dependence violates statistical independence and causes "response violations" of local independence (Andrich & Kreiner, 2010, Marais & Andrich, 2008b and Marais & Andrich, 2008a), meaning that the entire correlation between

the items is not captured by the latent trait. A high correlation between a pair of item residuals is one way of generating a “post-hoc” hypothesis of response dependence (Marais & Andrich, 2008a and Smith, 2002). When two questionnaire items ask for corresponding information causing redundancy in the data one would form a subtest i.e., merge the two items into one composite item to account for dependence.

Multidimensionality or trait dependence means that there are multiple latent variables and that some items measure one latent variable and other items measure another latent variable. When an overarching dimension like “maternal health literacy” (MaHeLi) is measured using several subsets of items assessing different aspects of the overarching dimension, with a subscale identified with each aspect, each aspect might represent a latent variable. If, e.g. the items assessing the HSB aspect and the items measuring the CCS aspect rank the respondents quite differently, the different aspects might form subscales that contribute with unique variance to the distribution of respondents’ score sums on the MaHeLi scale. Then the composite construct “maternal health literacy” is not sufficiently unidimensional, and we cannot report one score on the overarching dimension (Wright, 1999).

Unidimensionality means the presence of a dominant dimension or “underlying trait” with possible minor dimensions (see, for example, Wright & Linacre, 1989, Hambleton et al., 1991 and Keeves & Masters, 1999). If a theoretical composite construct is not sufficiently unidimensional, the theoretical composite construct does not find support in the empirical evidence, and one might want to split the assessment instrument into as many parts as there are latent variables or subscales and do separate analyses of the subscales. Principal component analysis (PCA) of residuals might help investigate the dimensionality of the data.

A PCA converts a set of observations (the data) of correlated variables (the items) into a set of linearly uncorrelated variables called principal components. The first principal component (PC 1) has the largest possible variance, i.e. accounts for as much of the

variability in the data as possible, and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to or uncorrelated with the other components.

In the MaHeLi scale the cluster of items assessing the HSB aspect and the cluster of items assessing the CCS aspect might have “maternal health literacy” in common. If we remove the common latent trait from the data we are left with the residuals or the deviations from the Rasch model. If the residuals of the HSB items correlate positively with PC 1 while the CCS items correlate negatively, the HSB (or the CCS) items might share something else in common than the general underlying variable “maternal health literacy” can explain. If so, the HSB items represent an additional latent trait that might violate the hypothesis of unidimensional data and hence violate local independence (Andrich & Kreiner, 2010, Marais & Andrich, 2008b, Marais & Andrich, 2008a and Ryan, 1983). Large variations in the percentage variance explained by each principal component (PC) is one way of generating a “post-hoc” hypothesis about multidimensionality in the data (Linacre, 1998 and Smith, 2002). The hypothesis might be tested by applying the equating tests and the t-test procedures in RUMM2030 and by estimating fractal indices based on a subtest analysis.

A set of  $k$  items can be analysed either as  $k$  distinct items or as two composite items (two subscales) where each subscale takes on the role of an item. A subtest analysis takes account of multidimensionality in the data and fractal indices ( $A$ ,  $c$  and  $r$ ) are estimated specific to the subtest structure. The value  $A$  describes the variance common to all subscales, the value  $c$  characterises the variance that is unique to the subscales and the variable  $r$  is the correlation between the two subscales (RUMM, 2009). A subtest analysis performed on an approximate unidimensional scale will return a high value for both  $A$  and  $r$  and a low value for  $c$ .

### **Reliability and targeting**

Reliability indices (see, for example, Traub & Rowley, 1981) as the Person Separation Index (PSI) and Cronbach's alpha provide estimates of the upper limit of the scale's reliability on the assumption of unidimensionality. In the presence of a multidimensional subscale structure, the variance of person estimates inflates and hence the reliability estimates inflate (Marais & Andrich, 2008b). When the distribution of the items' threshold estimates – the affective levels – match the distribution of the persons' attitude estimates an instrument is well “targeted” (Tennant & Conaghan, 2007).

### **Differential item functioning (DIF)**

DIF means that respondents with identical attitude levels, belonging to different categories or levels of a person factor (e.g. the levels first and second for the person factor pregnancy order), have unequal probabilities of giving the same response to an item (Andrich & Hagquist, 2004). This violates the property of invariance as the estimate of any item's affective level should be the same for each group of respondents (Thurstone, 1928, pp. 547, cited in Andrich & Hagquist, 2004, Rasch, 1961 pp. 332 and Engelhard, 2013 pp. 247). Common procedures for detecting DIF is the Mantel-Haenszel (MH) approach (Mantel & Haenszel, 1959 and Marascuio & Slaughter, 1981).

Uniform DIF means that the magnitude of conditional dependency is relatively invariant across the latent trait continuum, and an item mainly showing uniform DIF therefore consistently gives one person factor level an “advantage” across all levels of attitude (Walker, 2011 and Mellenbergh, 1982). Non-uniform DIF means that the conditional dependency alter direction along the latent trait (Walker et al., 2001) – that there is an interaction between the class intervals (groups of individuals based on attitude level) and the person factor levels. Non-uniform DIF is an important factor for non-invariant measures.

## **Procedures for Instrument Development**

The latent variable MHL was operationalized as a composite of the three aspects “health seeking behaviour” (HSB), “appraisal of health information” (AHI) and the closely linked “competence and coping skills” (CCS). These aspects of MHL were developed based on theoretical models relevant to the field of health literacy namely, “The health belief model” (HBM) and “The integrated model of health literacy” (IHL) (Glanz et al., 2002; Sørensen et al., 2012).

The items in the HSB and AHI subscales were founded on the domain of health promotion and the four dimensions of the IHL model, which combines the main dimensions of health literacy, the factors that impact on health literacy and pathways linking health literacy to health outcomes. This ensured that the public health aspect of health literacy as a continuum of interaction and interplay of social and environmental determinants, as described by (Sørensen et al., 2012) was considered.

To derive the items of the CCS subscale, certain appendage variables to the Health Belief Model (HBM) were used. The HBM is a value-expectancy theory which states that an individual’s behaviour can be predicted based upon certain issues that an individual may consider when making a decision about a particular behaviour concerning their health (Glanz et al., 2002). These variables included demographic, socio-psychological, motivation and perceived efficacy factors. The items in the CCS subscale aim at evaluating the girls’ cognitive abilities and their perception of social support accorded to them during pregnancy. The cognitive skills included problem-solving skills, coping skills, decision making skills, and social skills as interpersonal and communication skills. The CCS items thereby are reflective of the maternal competence and perceived social support.

In addition, demographic, socio-cultural and situational person factors reflective of respondent characteristics, found in literature to be antecedent to adolescent pregnancy, were incorporated. These included age, education level, pregnancy order (gravidity), pre-pregnancy knowledge on reproductive health and access to contraceptive services.

### **Description, Administration, and Scoring of the Instrument**

All the items in the HSB, CCS and AHI scales are reported in Table 1. The HSB scale probes for access to information (item 1, 3 and 4), assesses perception of competence to interpret and understand the information (item 2 and 5), and asks for elaboration strategies as initiative to seek for further information (item 7). The CCS scale emphasizes skills typically associated with the “cognitive domains” as “knowing” (item 8, 9 and 12), “applying” (item 11 and 13) and “reasoning” (item 10 and 14). Item 14 brings the critical aspect of health literacy into the CCS scale. The AHI scale seeks a self-evaluation of the individuals’ perceived maternal role competence (item 15, 18 and 20) and their perceived social support (item 16, 17 and 19). For practical reasons the data had to be collected using a paper-and-pencil questionnaire

Six-point rating scales with the extreme response categories anchored with the phrases “strongly disagree” (1) and “strongly agree” (6) were applied for all the items in the HSB, AHI and CCS scale.

Person factors addressing demographics and socioeconomic/cultural-factors were included in the questionnaire. These were age, level of education, pregnancy order, awareness about how pregnancy happens and frequency of condom use. In order to improve the power of test of fit person factors were recoded to dichotomous variables.

**TABLE 1 IN HERE**

## **Methods**

### **Sampling**

The target population was pregnant adolescents aged 15–19 years that attended antenatal care in the health centres in the Jinja and Iganga districts of the Busoga region of Uganda. This part of Uganda is of particular interest as it has the highest percentage (above 30%) of pregnant adolescents and adolescent mothers in Uganda (Rutaremwya, 2013).

In each district, different categories of health centres were randomly selected based on the structure of the Ugandan health care system described by (Kavuma, 2009). Nurses at the health centres provided lists of pregnant adolescents aged 15–19 years to the research assistants. The adolescents gave written consent before participation in the study.

### **Data collection**

The scale was administered to 384 adolescents aged 15–19 on the first antenatal visit at the health centre (i.e., the visit at which the female is clinically diagnosed as being pregnant). For practical reasons the data had to be collected at ten health centres during the period July 2013–December 2013. Trained research assistants, speaking fluently English and Lusoga read the items for the respondents. To ensure uniform interpretation and translation of the statements from English to Lusoga, guiding examples were issued for selected items.

## **Results**

### **Parameterisation, item discrimination, item fit and person fit**

Using the LRT, the conclusion was that the partial credit parameterisation of the PRM encloses most information about the MaHeLi scale data. The chi square probabilities in table 2 indicate that item 1 and item 20 do not fit the model, and that item 8 under-discriminates in

the 20-item construct. Person fit residuals showed that eight respondents had a positive z-fit residual outside the range ( $z = 2.5$ ) and that 69 respondents had a negative fit residual ( $z = -2.5$ ). Persons with negative residuals had response patterns that matched “too” well under the expectations of the model.

## **DIF**

Table 2 indicates that, based on the 20-item scale, item 5 and item 19 showed uniform DIF associated with the person factor “age”, and that item 20 showed non-uniform DIF associated with the person factors “age” and “pregnancy order”.

## **TABLE 2 IN HERE**

### **Reliability, targeting and overall statistics**

The PSI estimate for the composite scale was 0.91 and Cronbach’s alpha was 0.92. The scale’s total Chi square statistic was 191.5 ( $df = 100$ ). The average mean person location (attitude) value for the 20-item scale was 0.06. As the standard deviation for the distribution of attitude levels was 0.70 the variation in attitude levels was lower than expected (i.e., 1.00).

### **Ordering of thresholds**

Figure 1 shows the probability of ticking off in the six different response categories 0–5 for item 8 as functions of the person attitude level on the 20-item scale. The dotted line in Figure 1 is the upper limit asymptote (the probability equals 100%). Figure 1 indicates that item 8 had disordered thresholds as the response categories 2 and 3 on the six-point rating scale (curves marked “1” and “2” in Figure 1) were not the most likely for any attitude level.

## **FIGURE 1 IN HERE**

## **Violations of local independence**

A residual correlation above 0.3 between a pair of items was interpreted as a sign of response dependence in the data. Response dependence was observed between three pair of items: 1 and 2, 2 and 3, and 5 and 6.

Principal component analysis (PCA) of residuals were performed for each pair of subscales – the HSB subscale (items 1–7), the CCS subscale (items 8–14) and the AHI subscale (items 15–20). PCA of items 1–14 indicated that the HSB items correlated positively with the first principal component (PC 1) and that the CCS items correlated negatively with PC 1, except for item 4, 7 and 8 which all had correlation coefficients close to zero. PCA of items 1–7 and items 15–20 indicated that the HSB items correlated positively with PC1 and that the AHI items correlated negatively with PC 1, except for item 4 which had a correlation coefficient close to zero and item 7. PCA of items 8–20 indicated that the CCS items correlated positively with PC1 and that the AHI items correlated negatively with PC 1, except for item 13 which had a correlation coefficient close to zero.

Subtest structures of pairs of the three subscales were formed. There was a noticeable drop in the PSI for the subtest structure consisting of the HSB subscale and the CCS subscale (Table 2), and PSI dropped even more for the subtest structure consisting of the HSB subscale and the AHI subscale. When the subtest structure consisting of the CCS subscale and the AHI subscale was formed, the PSI only dropped from 0.90 to 0.82.

The fractal index  $c$  was quite high for the subtest structure consisting of the HSB subscale and the CCS subscale and for the subtest structure consisting of the HSB subscale and the AHI subscale (Table 2). The correlation  $r$  between the HSB subscale and the CCS subscale was rather low, and it was noticeably low between the HSB subscale and the AHI subscale. The index  $A$  was high especially for the subtest structure consisting of the CCS subscale and the AHI subscale.

The t-test procedure in RUMM showed that 44 respondents had “significantly” different scores on the HSB subscale and the CCS subscale to a 5% level, that 61 respondents had “significantly” different scores on the HSB subscale and the AHI subscale to a 5% level, and that 37 persons had “significantly” different scores on the CCS subscale and the AHI subscale to a 5% level. These results point to that the proportions of significant different scores were above 0.05 for all the three subtest pairs and consequently that all subtest pairs, as expected, had some multidimensionality.

Based on these results, a revised composite scale consisting of the CCS subscale, the AHI subscale and item 7 from the HSB subscale was formed (i.e., items 7–20). Based on this revised scale, item 8 still had disordered response categories and showed non-uniform DIF associated with the person factor “age”, and item 20 still showed non-uniform DIF associated with the person factors “age” and “pregnancy order”. Further, item 8 and item 20 did not fit the Rasch model as they under-discriminated “significantly” based on the revised latent trait. As the psychometric properties of item 20 did not improve after discarding item 8, both item 8 and item 20 were discarded from the revised scale.

### **TABLE 3 IN HERE**

## **Discussion**

Residual correlations between pair of questionnaire items are typically a sign of response dependence and hence redundancy in the data. Redundancy is caused by items collecting too similar information. As item 2 (“I understand the importance of getting correct information”) seems to collect information equal to item 1 and item 3 (“I know where to find health information” and “I have physical access to more than one source of information”), respondents might interpret the term “information” in item 1 and item 3 as “scientifically

based health information” – here expressed as “correct information”. By replacing “information” in item 1 and item 3 by “correct health information” and discarding item 2, the response violations of local independence observed in items 1–3 might be avoided. As scientifically based health information only is “correct” until evidence prove the opposite, the phrase “scientifically based health information” is preferred. As adolescents not are expected to have complete understanding of “the nature of natural science and its methods”, they might not fully grasp such an expression. Hence, the phrase “correct health information” actually might be preferred.

The residual correlation between item 5 (“There is always someone to avail to me the health information when I need it”) and item 6 (“All the information is available in a form that I am familiar with”) points to item 6 being dependent on the former item 5. When there is someone to avail the health information we can anticipate that there is someone available to explain the information. Those who do not have someone to avail the health information will not receive the information in an adapted form. Hence, most persons’ answers to item 6 will depend on their responses to item 5.

Principal component analyses of residuals implied that item 4 and item 13 had neglectable “deviations” from the expected pattern. However, item 7 (“In addition to scheduled hospital visits, I take initiative to seek for health information”) in the HSB subscale seemed to tap into the CCS subscale. This might be explained by item 7 assessing elaborating strategies associated with self-regulated learning (initiative to seek for information to clarify).

It was evident that the PSI index for the subtest structure consisting of the HSB subscale and the CCS subscale (items 1–14), and the PSI index for the subtest structure consisting of the HSB subscale and the AHI subscale (items 1–7 and items 15–20), both dropped “significantly” when the subtest structure was formed, and therefore that there is noticeable multidimensionality in the MaHeLi scale. However, a shortened version of the

MaHeLi scale consisting of the CCS subscale and the AHI subscale seems to represent a rather valid composite construct measuring maternal health literacy. Though, the PSI value 0.82 (see Table 2) is a more accurate estimate of the unidimensional reliability for this shortened version of the MaHeLi scale (consisting of all the CCS items and all the AHI items) than 0.90, which is inflated by multidimensionality. As mentioned, the results from the PCA indicated that item 7 from the HSB subscale perhaps could be included in this composite construct as part of the CCS subscale.

According to the c index the HSB subscale and the CCS subscale contribute with subscale specific variance to the common composite scale of these two subscales. This point is even more expressed when we combine the HSB subscale and the AHI subscale in a subtest structure. Again, the HSB subscale seems to “point in another direction” than the two other subscales of the composite MaHeLi scale, because the index c from the subtest structure of the CCS subscale and the AHI subscale is lower than for the other subtest structures that were formed.

Taking a cursory glance at the correlation coefficients between the subscales (Table 2) makes it evident that the HSB subscale seems to measure something “different” than the CCS subscale and the AHI subscale. Hence, the HSB subscale should be reported separately and not as part of the MaHeLi scale.

The index A was high especially for the subtest structure consisting of the CCS subscale and the AHI subscale, showing that of the systematic variance, most was left as true score variance common to both subscales. This again implies that the CCS subscale and the AHI subscale assess aspects of a construct with some but not too noticeable multidimensionality. The t-test procedure referred to also indicate some multidimensionality in all three pairs of subtests. *Some* multidimensionality may actually be seen as a positive property of a conceptually composite scale such as the MaHeLi scale.

Finally, based on our analysis we believe that a revised MaHeLi scale should be based on the CCS subscale and the AHI subscale. However, item 8 should be discarded due to disordered response categories, item 20 should be discarded due to non-uniform DIF, and item 7 from the HSB subscale should be included in the CCS subscale both from a conceptual and a psychometric point of view. These adjustments seem to remove the observed uniform DIF in item 19. The person separation index of this 12-item version of the MaHeLi scale was 0.90 (0.85 when multidimensionality was taken account of) and the mean attitude level was 0.00.

### **Conclusions**

The HSB *aspect* does not destroy unidimensionality of the concept “maternal health literacy” at a theoretical level, but the HSB *subscale* brings multidimensionality into the composite 20-item version of the MaHeLi scale from an empirical point of view. The suggested 12-item version of the MaHeLi scale, mainly leaving out the HSB aspect, forms a unidimensional scale where the revised CCS subscale and the revised AHI subscale can be added to a total “maternal health literacy” score. The 12-item version of the MaHeLi scale was well-targeted and had acceptable person separation in the Ugandan sample. The scale seems to consist of locally independent items that have ordered response categories, are functioning in the same way for different levels of relevant person factors, and show acceptable individual fit to the partial credit parameterisation of the unidimensional polytomous Rasch model.

Based on the available empirical evidence, we can conclude that the revised 12-item version of the MaHeLi scale represent 1) a well-targeted scale with 2) acceptable person separation that consists of 3) locally independent items that 4) have ordered response categories, 5) are functioning in the same way for different levels of relevant person factors, and 6) show acceptable individual fit to the unidimensional polytomous Rasch model. Hence, our earlier stated hypothesis is strengthened for the 12-item version of the MaHeLi scale.

This paper has showed how Rasch analyses might help us to identify violations of local independence in composite scales and identify the subscale that contributes the most to multidimensionality in the data. Composite scales are rather frequent applied in health related research, so our study might influence how composite scales within health related professions are structured and validated. Further analyses using Rasch analyses on similar and other clinical samples are warranted to explore the psychometric properties of the revised version of the MaHeLi scale.

This paper contributes to the international knowledge base on the motivations and abilities of pregnant adolescents to gain access to, understand, and use information in ways that promote and maintain their health and that of their children. This issue is of great international interest and concern and relevant to a wide spectre of clinical practice.

The paper aims to advance the quality of clinical health research methods by showing how violations of local independence can be identified in the data from composite questionnaire scales where items are deliberately developed to assess different aspects of a construct, with a subscale identified with each aspect. Further, the paper makes an important contribution to the discussion of the dimensionality of the “health literacy” concept in general; an issue of great interest to researchers in the whole health profession community and with important implications for measurement and the use of questionnaires in numerous clinical samples around the world.

### **Acknowledgements**

The authors thank the Ugandan health stations, the research assistants and the adolescents who participated in the study.

### **Funding**

This research received no specific grant, consulting honorarium, support for travel to meetings, fees for participation in review activities, payment for writing or reviewing the manuscript, or provision of writing assistance from any funding agency in the public, commercial or not-for-profit sectors.

### **Conflict of interest**

No conflict of interest has been declared by the authors.

### **References**

Andrich, D. (2010). Sufficiency and Conditional Estimation of Person Parameters in the Polytomous Rasch Model. *Psychometrika*, 75, 292–308.

Andrich, D. (2004). Controversy and the Rasch model: a characteristic of incompatible paradigms? *Medical Care*, 42(1) supplement, I7–I16.

Andrich, D. (1995). Models for Measurement, Precision and the Non-Dichotomization of Graded Responses. *Psychometrika*, 60, 7–26.

Andrich, D. (1988). *Rasch Models for Measurement*. Beverly Hills: Sage Publications.

Andrich, D. (1985). An elaboration of Guttman scaling with Rasch models for measurement. In: Brandon-Tuma N, editor. *Sociological Methodology*. San Francisco: Jossey-Bass, 33–80.

Andrich, D. (1982). An Index of Person Separation in Latent Trait Theory, the Traditional KR-20 Index, and the Guttman Scale Response Pattern. *Educational Research and Perspectives*, 9, 95–104.

Andrich, D. (1978). A rating scale formulation for ordered response categories.

*Psychometrika*, 43, 561–573.

Andrich, D. & Hagquist, C. (2004). Detection of Differential Item Functioning Using Analysis of Variance. Paper presented at the Second International Conference on Measurement in Health, Education, Psychology and Marketing: Developments with Rasch Models. Murdoch University, Perth, Australia.

Andrich, D., Humphry, S.M. & Marais, I. (2012). Quantifying Local, Response Dependence between Two Polytomous Items Using the Rasch Model. *Applied Psychological Measurement*, 36, 309–324.

Andrich, D., de Jong, J.H.A.L. & Sheridan, B.E. (1997). Diagnostic opportunities with the Rasch model for ordered response categories. In: Rost J, Langeheine R, editors. *Applications of Latent Trait and Latent Class Models in the Social Sciences*. New York: Waxmann, 59–70.

Andrich, D. & Kreiner, S. (2010). Quantifying response dependence between two dichotomous items using the Rasch model. *Applied Psychological Measurement*, 34, 181–192.

Belvedere, S.L. & de Morton, N.A. (2010). Application of Rasch analysis in health care is increasing and is applied for variable reasons in mobility instruments. *Journal of Clinical Epidemiology*, 63(12), 1287–1297.

Bland, J.M. & Altman, D.G. (1995). Multiple significance tests: the Bonferroni method. *British Medical Journal*, 310, 170.

Brodersen, J., Meads, D., Kreiner, S., Thorsen, H., Doward, L. & McKenna, S. (2007). Methodological aspects of differential item functioning in the Rasch model. *Journal of Medical Economics*, 10, 309–324.

Engelhard, G. (2013). *Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences*. New York: Routledge.

European Commission. (2007). Together For Health: A Strategic Approach for the EU 2008-2013. Brussels.

Glanz, K.R., Rimer, B.K. and Lewis, F.M. (2002). *Health Behavior and Health Education: Theory, Research and Practice*, San Francisco: Jossey-Bass.

González-de Paz L, Kostov B, Sisó-Almirall A, Zabalegui-Yárnoz A. (2012). A Rasch analysis of nurses' ethical sensitivity to the norms of the code of conduct. *Journal of Clinical Nursing* 21, 2747–2760.

Hagquist, C., Bruce, M. & Gustavsson P. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies*, 46, 380–393. Available online at: [http://ac.els-cdn.com/S0020748908002721/1-s2.0-S0020748908002721-main.pdf?\\_tid=d239e98e-a86b-11e3-9c9e-00000aacb35d&acdnat=1394466859\\_491540b15161da83417274af018659ac](http://ac.els-cdn.com/S0020748908002721/1-s2.0-S0020748908002721-main.pdf?_tid=d239e98e-a86b-11e3-9c9e-00000aacb35d&acdnat=1394466859_491540b15161da83417274af018659ac) (accessed 20 February 2014).

Holloway, A., Blackman, I. & Flynn, F. (2014). A study examining the appropriateness of a self-rated alcohol-related clinical confidence tool as a method of measurement among registered hospital nurses using Rasch analysis, *Journal of Advanced Nursing*, 70(4), 937–945.

Hambleton, R.K., Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of Item Response Theory*, Thousand Oaks, CA, Sage.

Hana H.W. & Abdulla S.B. (2013). Factors affecting Health Seeking Behavior for Common Childhood Illnesses in Yemen. *Dove Medical Press*, 7, 1129-1138. doi: <http://dx.doi.org/10.2147/PPA.S51124>.

Kavuma, R. M. (2009). Katine: It Starts With A Village, The Guardian. Retrieved from: <http://www.theguardian.com/katine/2009/apr/01/uganda-healthcare-system-explained> (accessed 9 March 2014).

Keeves, J. P. & Masters, G. N. (1999). Issues in educational measurement. In G. N. Masters & J.P. Keeves (Eds.), *Advances in measurement in educational research and assessment*, Amsterdam, Pergamon, 268–282.

Linacre, J.M. (1998). Detecting multidimensionality: which residual data-type works best?. *Journal of Outcome Measurement*, 2, 266–283.

Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–748.

Marasculio, L.A. & Slaughter, R.E. (1981). Statistical procedures for identifying possible sources of item bias based on 2 x 2 statistics. *Journal of Educational Measurement*, 18, 229–248.

Marais, I. & Andrich, D. (2008a). Formalising Dimension and Response Violations of Local Independence in the Unidimensional Rasch model. *Journal of Applied Measurement*, 9(3), 1–16.

Marais, I. & Andrich, D. (2008b). Effects of Varying Magnitude and Patterns of Response: Dependence in the Unidimensional Rasch Model. *Journal of Applied Measurement*, 9(2), 1–20.

Mellenbergh, G.J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, 7, 105–118.

Olenja, J. (2003). Health Seeking Behaviour in Context. *East African Medical Journal*, 80(2), 61–62.

Osborne, R., Buchbinder, R., Batterham, R., and Elsworth, G. (2012). Health Literacy Questionnaire.

Perline, R., Wright, B.D. & Wainer, H. (1979). The Rasch model as additive conjoint measurement. *Applied Psychological Measurement*, 3, 237–256.

Pintrich P. & Schunk D. (1996). *The Role of Expectancy and Self efficacy Beliefs. Motivation in Education*. New Jersey: Englewood Cliffs.

Renkert, S. & Nutbeam, D. (2001). Opportunities to improve maternal health literacy through antenatal education: an exploratory study. *Health Promotion International*, 16(4), 381-388. doi: 10.1093/heapro/16.4.381.

RUMM (2009). *Interpreting RUMM2030 Part IV Multidimensionality and subtests in RUMM*. 1. ed: RUMM Laboratory Pty Ltd.

Rutaremwu. G. (2013). Factors associated with Adolescent Pregnancy and Fertility in Uganda: Analysis of the 2011 Demographic and Health Survey Data. *American Journal of Sociological Research*, 3. 7–13 doi: 10.5923/j.sociology.20130302.03.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Achievement Tests*. Copenhagen: Danish Institute for Educational Research. Expanded edition 1983. Chicago: MESA Press.

Rasch, G. (1961). On general laws and the meaning of measurement in psychology, pp. 321–334 in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, IV. Berkeley, California: University of California Press. Available online at: [http://projecteuclid.org/download/pdf\\_1/euclid.bsmsp/1200512895](http://projecteuclid.org/download/pdf_1/euclid.bsmsp/1200512895) (accessed 20 February 2014).

Rosann O'Dell. (2012). Appraisal Skills, Health Literacy and the Patient-Provider Relationship: Considerations as the Health Care Consumer Turns to the Internet to Inform their Care. *Online Journal of Public Health Informatics*, 4(1). doi: 10.5210/ojphi.v4i13684

Ryan, J.P. (1983). Introduction to latent trait analysis and item response theory. In: Hathaway WE, editor. *Testing in the Schools New Directions for Testing and Measurement*. San Francisco: Jossey-Bass.

Singh, J. (2004). Tackling problems with item response theory: Principles, characteristics and assessment, with an illustrative example. *Journal of Business Research*, 57, 184–208.

Smith, E.V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, 3, 205–231.

Smith, E.V. (2005). Effect of Item redundancy on Rasch Item and Person Estimates. *Journal of Applied Measurement*, 6, 147–63.

Smith, R.M. & Plackner, C. (2009). The Family Approach to Assessing Fit in Rasch Measurement. *Journal of Applied Measurement* , 10, 424–437.

Stevenson, W., Maton, K. I., and Teti, D. M. (1999). Social support, relationship quality, and well-being among pregnant adolescents. *Journal of Adolescence*, 22(1), 109–121. doi: <http://dx.doi.org/10.1006/jado.1998.0204>.

Sørensen et al. (2012). Health Literacy and Public Health: A systematic review and integration of definitions and models. *BMC Public Health*, 12(80). doi: doi:10.1186/1471-2458-12-80.

Tennant, A. & Conaghan, P.G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research*, 57(8), 1358–1362. DOI: 10.1002/art.23108.

Thurstone, L.L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529–554.

Traub, R.E. & Rowley, G.L. (1981). Understanding reliability. *Educational Measurement: Issues and Practice*, 10, 37–45.

Walker, C. (2011). What's the DIF? Why differential item functioning analyses are an important part of instrument development and validation. *Journal of Psychoeducational Assessment*, 29, 364–376.

Walker, C.M., Beretvas, S.N. & Ackerman, T.A. (2001). An examination of conditioning variables used in computer adaptive testing for DIF. *Applied Measurement in Education*, 14, 3–16.

Ward, H., Mertens, T. E., and Thomas, C. (1997). Health seeking behavior and the control of sexually transmitted disease. *Health Policy Plan*, 12(1), 19–28.

Wright, B.D. (1999). Fundamental measurement for psychology. In S.E. Embretson, S.L. Hershberger (Eds.). *The New Rules of Measurement: What Every Psychologist and Educator Should Know*, 65–104. Mahwah, New Jersey and London: Lawrence Erlbaum Associates.

Wright, B.D. & Linacre, J.M. (1989). Observations are always ordinal; measurements, however, must be interval, *Archives of Physical Medicine and Rehabilitation*, 70(12), 857–860. Available online at: <http://www.rasch.org/memo44.htm> (accessed 20 February 2014).

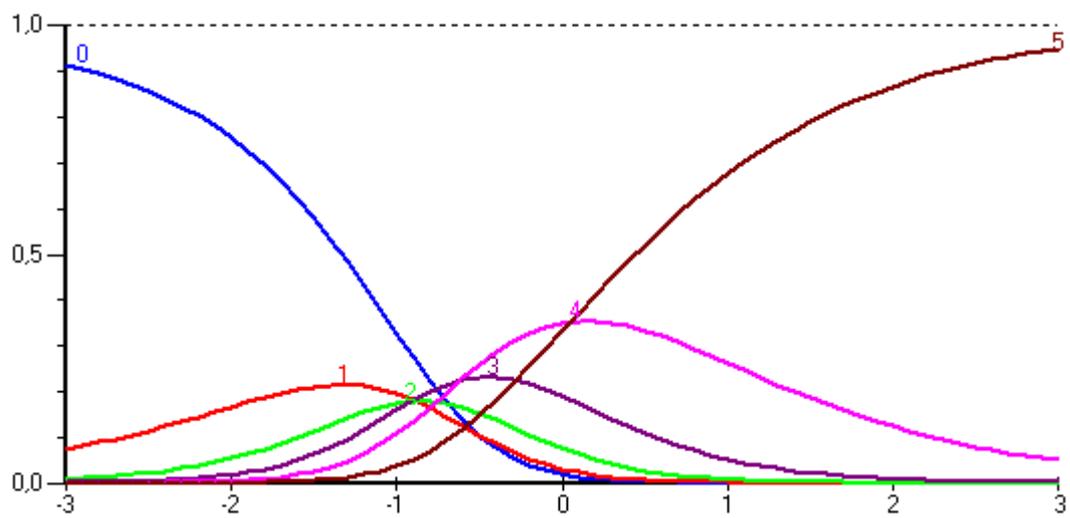
Wright, B.D. & Masters, G.N. (1982). *Rating Scale Analysis: Rasch Measurement*. Chicago: MESA press.

Wright, B.D. & Mok, M.M.C. (2004). An overview of the family of Rasch measurement models. In E. V. Smith Jr & R.M. Smith (Eds.). *Introduction to Rasch measurement*, Maple Grove, MN, JAM Press, 1-24, Available online at: <http://www.jampress.org/irmch1.pdf> (accessed 20 February 2014).

## Figures and tables

### List of figures (all to be printed in black and white)

Figure 1. Category probability curves for item 8 “I understand the language in which the health information is given”. Probability of ticking off in category 1–6 (rescored to 0–5) as a function of attitude level.



## List of tables

Table 1. The wording of the items in the HSB scale (items 1–7), the CCS scale (items 8–14) and the CCS scale (items 15–20). Items originally stated in English.

Item	Wording
1	I know where to find health information
2	I understand the importance of getting correct information
3	I have physical access to more than one source of information
4	I have the economic resources to access the information
5	There is always someone to avail to me the health information when I need it
6	All the information is available in a form that I am familiar with
7	In addition to scheduled hospital visits, I take initiative to seek for health information
8	I understand the language in which the health information is given
9	I can explain with ease health information received to others
10	Based on the health information received, I can identify and interpret symptoms of health risk
11	I am able to follow written health instructions
12	I have an idea about what to expect during pregnancy and postnatal period
13	I can discern pregnancy-related myths from accurate health information
14	I find it easy to recall health information previously received
15	I feel confident that I can independently follow health recommendations correctly
16	I feel that my input into the planning for the care of the child is valued
17	I can comfortably rely my health concerns to the people around me

- 18 I can use the health information to monitor and self-regulate my health
  - 19 I have remained active in social gatherings as I did before the pregnancy
  - 20 I have the ability to take care of my health and that of the child
-

Table 2. Initial analysis of the MaHeLi scale (items 1–20) treating the three aspects (HSB, AHI and CCS) as subscales of a possible unidimensional MaHeLi scale.

Item	Loc	SE	Res	df	ChiSq (df=5)	ChiSqProb	Disord	Uniform DIF (MS)	Non-uniform DIF (MS)
1	0.26*	0.042	<b>6.04</b>	360.85	31.7	<b>0.0000</b>			
2	0.29*	0.044	1.18	360.85	4.2	0.5165			
3	0.12	0.048	0.15*	360.85	5.0	0.4117			
4	0.45	0.046	0.52	360.85	2.2	0.8157			
5	0.14*	0.045	2.30	360.85	7.0	0.2173		Age (24.3)	
6	0.01	0.046	0.04	360.85	16.1	0.0064			
7	0.11	0.042	0.65	360.85	1.5	0.9086			
8	0.57*	0.043	<b>3.41</b>	360.85	15.8	0.0074	x		
9	0.35*	0.048	1.17*	360.85	5.3	0.3852			
10	0.06	0.045	1.99*	360.85	7.7	0.1709			
11	0.01*	0.049	2.08*	360.85	13.5	0.0190			
12	0.01	0.049	0.04*	360.85	6.7	0.2431			
13	0.23	0.046	0.33*	360.85	3.4	0.6369			
14	0.39	0.043	0.98*	360.85	6.2	0.2903			
15	0.11*	0.045	1.01*	360.85	3.3	0.6502			
16	0.12	0.046	1.15*	360.85	7.5	0.1879			
17	0.16	0.045	1.84*	360.85	10.3	0.0664			
18	0.12*	0.048	1.9*	360.85	11.2	0.0481			
19	0.23	0.044	0.77	360.85	3.3	0.6583		Age (11.5)	
20	0.03*	0.041	<b>4.46</b>	359.91	29.4	<b>0.0000</b>			Age (7.9), Preg.ord (8.6)

Negative values are marked (\*) and “significant” deviations are bold.

Table 3. Estimated values of reliability (PSI) and fractal indices ( $c$ ,  $r$  and  $A$ ) associated with the subtest structures performed of the MaHeLi subscales.

Subscales	Set of items	PSI	PSI (taking account of multidimensionality)	$c$	$r$	$A$
HSB and CCS	1–14	0.88	0.78	0.53	0.78	0.89
HSB and CCS	1–7 and 15–20	0.87	0.71	0.67	0.69	0.82
CCS and CCS	8–20	0.90	0.82	0.43	0.85	0.92