


Psychometric properties of the SCREESCO questionnaire used in a colorectal cancer screening programme—A Rasch analysis

Anna Jervaeus RN PhD¹  | Kaisa Fritzell RN PhD¹ | Rolf Hultcrantz MD PhD Professor² | Yvonne Wengström RN PhD Professor^{1,3} | Anders Kottorp OT PhD Professor^{4,5}

¹Department of Neurobiology, Care Sciences and Society, Division of Nursing, Karolinska Institutet, Stockholm, Sweden

²Department of Medicine, Karolinska Institutet, Stockholm, Sweden

³Theme cancer, Radiumhemmet, Karolinska University Hospital, Stockholm, Sweden

⁴Department of Occupational Therapy, College of Applied Health Sciences, University of Illinois, Chicago, Illinois

⁵Faculty of Health and Society, Malmö University, Malmö, Sweden

Correspondence

Anna Jervaeus, Department of Neurobiology, Care Sciences and Society, Division of Nursing, Alfred Nobels allé 23, Huddinge 141 83, Sweden.

Email: anna.jervaeus@ki.se

Funding information

Funding information This study was supported by grants from the Swedish Associations of Local Authorities and Regions.

Abstract

Introduction: Colorectal cancer (CRC) is one of the most common cancer diagnoses among both sexes. Sweden has not yet implemented any CRC screening programme, but a study, Screening of Swedish Colons (SCREESCO), is ongoing. The movement within the health care sector towards a more participatory perspective has led to the increased importance of shared decision making (SDM), and this is suggestively applied when deciding upon screening participation. There is no Swedish questionnaire for assessing the level of SDM in relation to CRC screening. Therefore, the CRC screening module of the National Survey of Medical Decisions was translated and culturally adapted into a Swedish context: the SCREESCO questionnaire.

Aim: The SCREESCO questionnaire requires further evaluation, and therefore, the aim, by using the Rasch approach, was to evaluate the psychometric properties of the SCREESCO questionnaire.

Methods: A Rasch partial credit model was chosen to investigate the psychometric properties. The sample consisted of individuals invited to the SCREESCO programme, who have answered the SCREESCO questionnaire.

Results: Rating scale structures indicated stability for the response structure used. Satisfactory evidence for validity of internal structure was also shown for the whole questionnaire and two of three concepts/subscales, after deletion of a few items. Validity in response processes indicated acceptable, or close to acceptable, findings, while the results for unidimensionality and differential item functioning (DIF) were somewhat mixed. Separation index revealed less satisfactory results, both for the whole questionnaire and the concepts/subscales.

Conclusion: This Rasch analysis of the SCREESCO questionnaire revealed that the questionnaire in its current form has difficulties to assess the level of SDM in relation to CRC screening. The achieved results will guide further evaluation and development with the long-term goal of having a Swedish questionnaire, to be used in the health care sector, assessing the level of SDM in relation to CRC screening.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors. Journal of Evaluation in Clinical Practice published by John Wiley & Sons Ltd.

KEYWORDS

colorectal cancer screening, psychometric properties, Rasch analysis, SCREESCO questionnaire

1 | INTRODUCTION

Based on GLOBOCAN worldwide estimates, colorectal cancer (CRC) is one of the most common cancer diagnoses among both sexes.¹ The lifetime risk of developing CRC is relatively low while mortality rates are considerably high.² Worldwide estimates from GLOBOCAN show mortality rates around 50% for both men and women,¹ and the latest data from Sweden display a similar trend.³ Modelling studies have shown that screening efforts are cost saving, and together with other aspects, such as the high incidence, the detectable and treatable precursor and the high cost of treatment makes CRC suitable for screening efforts.² However, for screening to be effective, it needs to be widely accepted in the target population, and therefore, it is important to study reasons for nonparticipation.⁴ Previous results from CRC screening in the Stockholm and Gotland counties revealed an adherence rate of around 60%.⁵

Since 2003, the Council of the European Union (EU) have recommended CRC screening with faecal occult blood in men and women aged 50 to 74,⁶ and over the past two decades, many organized, population-based CRC screening programmes have been implemented in Europe.² Sweden has not yet implemented any nationwide CRC screening programme but a randomized controlled study, Screening of Swedish Colons (SCREESCO), with one primary endpoint to investigate what method is most suitable in Sweden, is ongoing. Other aims include exploring adherence, health economic aspects, and perceived experiences from both participants and nonparticipants, which the present study is part of.

The movement within the health care sector towards a more participatory perspective, where individual preferences and autonomy are taken into account, has led to the increased importance of shared decision making (SDM)⁷ and is suggestively applied when deciding upon cancer screening participation.⁸ In screening programs, where healthy individuals are approached, it is vital to ensure the autonomy of participants.⁹ Therefore, it is desired that individuals make an informed decision, whether it is to participate or not, on knowledge rather than ignorance, misconceptions, or fear. The Informed Medical Decisions Foundation describes SDM as a process that makes it possible for individuals and health care providers to make decisions about health care issues together. In addition, the best scientific evidence and the individual's *Values and preferences* should be considered when making such decisions.¹⁰ Although there have been several attempts to define SDM, consensus has not yet been reached. The SDM definition in the present study has its foundation in the approach where information about benefits and risks of treatment options and individual values/preference are essential.^{7,11,12} The definition is based on three concepts: *Values/preferences*, *Involvement*, and *Information/Knowledge*. *Values and preferences* include attitudes, behavior, and beliefs from both the individual and the health care professional.

Involvement encompasses engagement from the individual and the health care professional, communication and relationship, and individuals and health care providers' common ground for decision making. *Information/Knowledge* includes disease and treatment/health prevention knowledge and treatment/health prevention options.¹³

In order to be able to assess the level of SDM in a sample, a valid and sensitive outcome measure is needed.¹⁴ Such an outcome measure should demonstrate evidence of validity, as in this study indicated by including items that reflect a similar underlying unidimensional construct (validity in relation to internal structure) as well as monitoring the level of systematic response processes among the participants (validity in response processes).¹⁴ Finally, evidence of reliability, as in this study indicated by being sensitive enough to detect various levels of SDM, in a sample. To our knowledge, there is no Swedish questionnaire designed to assess the level of SDM in relation to CRC screening. Therefore, with permission, the CRC screening module of the National Survey of Medical Decisions (the DECISIONS study)⁸ was translated and culturally adapted into a Swedish context, labelled the SCREESCO questionnaire.¹⁵ The questionnaire requires further evaluation, if it has the ability to assess the level of SDM, with regard to measurement consistency and since questionnaire data are sample dependent.¹⁶ Therefore, the aim of this study, by using the Rasch approach, was to evaluate the psychometric properties of the SCREESCO questionnaire.

Specific research questions were as follows:

1. What are the psychometric properties of the different rating scales used in the SCREESCO questionnaire?
2. Is there satisfactory evidence of validity of internal structure and in response processes in the generated SCREESCO questionnaire measures?
3. Is there additional empirical evidence supporting unidimensionality within the different concepts/subscales, as indicated by principal component analysis (PCA)?
4. Is it possible to separate distinct groups among the respondents, ie, the ability of the test to separate people into different levels of SDM?
5. Do the items in the SCREESCO questionnaire function in the same way, indicated by no presence of differential item functioning (DIF), in relation to gender as well as screening participants compared with nonparticipants?

2 | METHODS

The Rasch approach is named after the Danish mathematician George Rasch. The theory was published in 1960, and it is an approach to assess human performance, attitudes, and perceptions.^{17,18} It was chosen because it provides the ability to transform cumulative raw scores

(persons across items or item across persons) into independent linear continuous measures of ability (persons) and difficulty (items).¹⁸ Further, it provides the ability to perform more detailed validity (eg, fit statistics and the investigation of violation to local independence) and reliability (standard errors) assessments in the same analysis, for both items and persons.¹⁹ Whether rating scales are judged to yield valid and reliable measurement depends on the extent to the fit between the empirical data and the Rasch model assertions.²⁰ The Rasch approach demands an acceptable level of unidimensionality (evidence that a single construct being measured) and then result in additivity (properties of measurement units). These units (logits) should be of the same size (interval), if the data fit the model.¹⁹

2.1 | Sample

In the large nationwide SCREESCO programme, individuals were randomly assigned to one cohort: (a) colonoscopy; (b) stool sample (fecal immunochemical test [FIT]) year 1 and year 3; or (c) control group. For the present study, $n = 2748$ were randomly sampled from the larger cohort population and by each of the following groups: nonparticipants ($n = 749$); participants randomized to FIT with a positive test result ($n = 750$); participants randomized to FIT with a negative test result ($n = 749$); and participants randomized to and had participated in a colonoscopy ($n = 500$). The sampling period was between October 2015 and June 2016, and individuals were born in 1954 or 1955; no additional stratification was done, and the whole of Sweden was represented.

2.2 | Measure

The SCREESCO questionnaire encompasses 24 items aiming at measuring SDM: *Values and preferences* (nine items), *Involvement* (three items), and *Information/Knowledge* (12 items).¹⁵ Some items, encompassing all three concepts/subscales, are answered on a 0 to 10 scale from either *took no consideration* to *took great consideration*; or *not at all informed* to *very much informed* and of *no importance* to *of great importance*. Other items were dichotomized before the analyses started, such as those aiming at assessing *Information/Knowledge* (eg, deciding among factors of importance for the development of CRC; deciding among symptoms of CRC or specify the expected number of individuals being diagnosed in Sweden each year) (Table 1). Upon our decision, these were dichotomized according to “having knowledge” vs “not having knowledge” based on, at the time, current facts of risk factors and incidence and prevalence of CRC according to relevant statistics.²¹ Having knowledge was considered if at least one factor/symptom was correct, regardless of other

responses to the same item. The classification was completed together with a specialist in gastroenterology (RH). One item (*Information/Knowledge*) regarded personal risk of developing CRC, and this item was also dichotomized (low/medium high risk vs high risk).

2.3 | Data collection

An information letter, including login details, was sent to eligible individuals, according to the sample description. Each participant responded to the questionnaire online. If participants wished, for instance, due to lack of computer experience, they could answer the questionnaire by telephone interview with one of the authors, and 77 participants requested this. No reminder was sent out. In total, 1498 agreed to participate, while completion of the questionnaire differed among participants at item level.

2.4 | Analyses

Descriptive statistics (demographics) were calculated using the IBM SPSS Statistics Version 23 software package. The concepts/subscales of the SCREESCO questionnaire were examined using a partial credit model (PCM), a Rasch model designed for polytomous data, meaning that item categories and/or thresholds can vary across items.²⁰ The Rasch analysis software program, WINSTEPS Rasch Measurement version 3.92.1.0 (Copyright John M. Linacre), was used for the Rasch analysis.

We conducted the analysis by applying a previously used step-by-step procedure^{22,23} and according to the specific research questions presented, to evaluate the psychometric properties of the SCREESCO questionnaire, using the Rasch approach. Since this is the first evaluation of the newly translated and culturally adapted SCREESCO questionnaire¹⁵ and investigation on its ability to assess the level of SDM or not, we analysed both the whole questionnaire together but also each concept/subscale separately. Step 1: Psychometric properties of rating scales were investigated by analysing, the category function (as labelled in Winsteps), expressed by observed average and outfit mean square values (*MnSq*). Guidelines were adjusted to as follows: (a) monotonic advancement of rating scale categories, if not, the meaning of that rating scale is deemed not stable for the particular dataset; and (b) rating scale category outfit $MnSq < 2.0$. Higher values indicate that the observations contain more misinformation than information.²⁴

Steps 2 and 3: Validity of internal structure and response processes were investigated by calculation of item and person goodness of fit statistics. Item fit statistics is used to find items that might not contribute to the measure or if local independence is violated (ie, when response to an item is depending on response to another item).¹⁹ The calculations were displayed with *MnSq* residuals and standardized z values indicating to what level responses match the expectations in the chosen Rasch model. Chosen threshold values and considered as a rule of thumb for item infit statistics were chosen to be 0.7 to 1.3 (Infit *MnSq*). Variations above 1.0 indicate more variation than predicted by the chosen Rasch model (misfitting), and values

TABLE 1 Number of items divided by concept/subscale

	Items (n)	
	Rating scale (0-10)	Items (n) Dichotomized
<i>Values and preferences</i>	8	1
<i>Involvement</i>	1	2
<i>Information/Knowledge</i>	3	9

below 1.0 indicate less variation in the data compared with what the model predicts (overfitting).^{20,25} Item fit was investigated in an iterative process by excluding item by item those items that did not display Infit *MnSq* within the range of 0.7 to 1.3. This procedure was performed until all items showed acceptable goodness of fit.²⁶ In order to ensure that the item deletion process was not biased due to sample, the item deletion process was also repeated using a random sampling of participants ($n = 750$) and then confirmed in the second random sample. Person fit statistics assess to what extent a person's set of responses correspond to what is predicted by the model.¹⁹ For person infit statistics, the following thresholds have been used in previous research, with similar analytical approaches^{23,26} and were therefore chosen. A value of ≥ 1.4 Infit *MnSq* associated with a z value equal to or larger than 2.0 in not meeting the criteria of acceptable fit to the model. A general acceptance level is that up to 5% of respondents can display nonsatisfactory goodness of fit without violating validity in individual response processes.²² A further advantage with the Rasch approach is the possibility to display items and persons on the same linear scale,¹⁹ and therefore, a person versus item map was included in the analysis to visually present where items versus persons are located along the logit scale (with equal intervals). The proportion of maximum and minimum scores in the SCREESCO questionnaire was evaluated, as this is an indication of floor and ceiling effects, something that can compromise validity and reliability. It was accepted that up to 15% of the sample could demonstrate minimum or maximum scores without being a threat to validity.

Step 4: The Rasch approach demands that a single construct is being measured at a time and that construct are to be underlying the items, ie, unidimensionality.²⁰ This was investigated by calculating variance explained by measures by a principal component analysis (PCA) for each of the concepts/subscales: *Values and preferences*, *Involvement*, and *Information/Knowledge*. The chosen thresholds were set to exceed 50% (raw variance explained by measures)²⁷ and to not exceed 5% for unexplained variance in the first contrast, a possible secondary dimension in the data.

Step 5: A reliable measure should be able to separate according to ability/difficulty etc, across the sample and the person separation index, ie, the degree of separation among individuals for the variable under measurement, can be used for this purpose. It is measured by standard error units: adjusted person standard deviation divided by the average measurement error.²⁰ In general, 1.5 is considered a threshold capable of separating two groups in the sample, while 2.0 is adequate for three groups.²⁸

Step 6: Uniform DIF was performed to explore the stability of item difficulty related to gender as well as between those who participated in the screening versus those who did not. The magnitude of DIF was investigated using the Mantel-Haenszel statistics for polytomous scales using log-odds estimators^{29,30} as reported from the WINSTEPS program, using a P value of $<.01$.

Joint analyses of the whole questionnaire and separate analysis for each concept/subscale have been performed. Firstly, under each

subheading, findings from the whole questionnaire are reported and thereafter by each concept/subscale.

2.5 | Ethical considerations

The study was approved by the regional Ethics Review Board at Karolinska Institutet, Stockholm, Sweden (No. 2012/2058-31/3). All participants gave their informed consent.

3 | RESULTS

Participant characteristics are presented in Table 2 and details of the Rasch analysis in Table 3. In the sample, a higher proportion of women belonged to those declining participation compared with men (P value .003), while women, to a larger extent, had a higher educational level compared with men (P value .001). For individuals giving a "valid" response and a "Don't know" response, the valid response was chosen. "Don't know" responses are not included in the Rasch analysis, except for those items aiming at assessing "having knowledge" vs "not having knowledge."

3.1 | Rating scale/category function

The rating scale structure for each item in the SCREESCO questionnaire met all set criteria: (a) monotonic advancement of rating scale categories and (b) rating scale category outfit *MnSq* < 2.0 ; so we concluded that the questionnaire (all 24 items) and the three concepts/subscales: *Values and preferences* (nine items), *Involvement* (three items), and *Information/Knowledge* (12 items), indicated stability for the response structure used. All concepts/subscales met the criteria also after item deletion.

3.2 | Validity of internal structure

The whole SCREESCO questionnaire displayed two items demonstrating misfit to the Rasch model: finding bowel cancer early and examination is free of charge, both included in the concept/subscale *Values and preferences*. After deleting those items, the remaining 22 displayed values within the range of 0.7 to 1.3 (Table 3).

For the concept/subscale *Values and preferences* (nine items), the item infit statistics displayed almost similar misfits (Table 3) to those of the entire questionnaire. After deleting the three items initially demonstrating misfit (finding bowel cancer early, examination free of charge, and risk for complications), the remaining six displayed values within the range of 0.7 to 1.3. The concept/subscale *Involvement* (three items) revealed two items demonstrating misfit. No further iterative analyses were conducted due to the high proportion of misfitting items. For *Information/Knowledge* (12 items), item infit statistics revealed no items demonstrating misfit.

The conclusion was that the whole questionnaire and the concepts/subscales, except one (*Involvement*), did generate quite well-matched items, after the deletion of a few items.

TABLE 2 Background characteristics of participants, by self-reported gender, total sample N = 1498

	Women	Men	P ^a
	n = 772	n = 726	
	n (%)	n (%)	
Declined screening participation	103 (13)	61 (8)	0.003 ^b
Positive FIT	225 (29)	260 (36)	
Negative FIT	287 (37)	248 (34)	
Colonoscopy	157 (21)	157 (22)	
Self-reported living situation (n = 1485) ^c			0.113 ^d
Living together with someone	585 (77)	579 (80)	
Living alone	178 (23)	143 (20)	
Self-reported highest level of education (n = 1484) ^c			0.001 ^e
Compulsory School	118 (15)	163 (23)	
High school	306 (40)	338 (47)	
Vocational high school	29 (4)	25 (3)	
University	311 (41)	194 (27)	
Self-reported occupational status (n = 1475) ^c			0.060 ^f
Working in a profession	561 (73)	553 (76)	
Studying	1		
Seeking employment	26 (3)	23 (3)	
Retired	51 (7)	51 (7)	
On disability living allowance	70 (9)	40 (6)	
On sick leave	35 (5)	10 (2)	
Other ^g	20 (3)	34 (6)	

^aTested for differences in proportions by Chi-square test. Degree of freedom: 1. P value ≤ .05 was considered statistically significant.

^bTested between those who participated vs those who did not.

^cValid responses.

^dTested between those who lived together with someone vs living alone.

^eTested between those who had completed compulsory or senior high school vs those who had completed vocational high school or university.

^fTested between those working in a profession vs all other alternatives.

^gFor those reporting "other" as the main alternative and had added text, it could encompass a combination of given alternatives such as working and/or retired, working and/or on sick leave, being a housewife or being self-employed.

3.3 | Validity in individual response processes

Person misfit displayed acceptable, or close to acceptable, results for both the whole SCREESCO questionnaire, *Values and preferences*, and *Information/Knowledge*, ie, 7.6%, 5.4%, and 3.9%, respectively (Table 3).

A person versus item map is shown in Table 4. The logit scale is shown on the far left and person measures on the left, indicating each individual's ability or, here, the anticipated level of

SDM. The item difficulty calibration is displayed on the right, ie, how challenging each item is. For *Values and preferences*, items are clustered together around the mean (M) corresponding to the person measure, ie, where the majority of respondents are located, while for *Information/Knowledge*, items are more spread out.

The item deletion process was further confirmed using a random split-half technique with two samples, where the item deletion process was initially performed in the first subsample, and the reduced item pool was tested in the second subsample. This approach confirmed that the remaining items demonstrating acceptable goodness-of-fit in the first subsample, also demonstrated acceptable goodness-of-fit in the second subsample for the whole SCREESCO questionnaire and in the subscales as well.

3.4 | Unidimensionality by PCA

The raw variance explained by measure was below 50%, and the unexplained variance in first contrast did not exceed 5% for the whole SCREESCO questionnaire (Table 3).

For *Values and preferences* (six items), raw variance explained by measure was just over 50%, unexplained variance in first contrast exceeded 5%, while *Information/Knowledge* (nine items) displayed 73.7% and 3.8%, respectively.

3.5 | Person separation index

The separation index was lower than 1.5 for the whole SCREESCO questionnaire, for *Values and preferences* and *Information/Knowledge*, ie, neither the whole questionnaire nor the separate concepts/subscales could separate the sample in a minimum of two distinct groups, indicating a low level of sensitivity.

3.6 | Differential item functioning

The DIF analysis indicated that seven of the items demonstrated DIF in relation to gender and nine of the 24 items demonstrated DIF in relation to participants participating in the screening programme or not. Those items demonstrating DIF in relation to gender belonged to *Values and preferences* (three items: Taking into consideration the risk of false alarm; Did you use someone close to you as information source about bowel cancer screening? And if so, how important was that person as an information source?), *Involvement* (one item: Before your decision, did you discuss with any care giver?), and *Information/Knowledge* (three items: Before this study, what screening examinations for bowel cancer did you know of? What symptoms of bowel cancer are you familiar with? What percentage of the Swedish population do you think will be diagnosed with bowel cancer?). The items demonstrating DIF in relation to participation or not belonged to *Values and preferences* (six items related to bowel cancer worry; importance of finding it early; risk of complications in relation to the examination; risk of discomfort

TABLE 3 Results from the Rasch analysis of psychometric properties of the SCREESCO questionnaire and by concepts/subscales

	SCREESCO questionnaire 24 items Reported n = 1480	SCREESCO questionnaire 22 items Reported n = 1480	Values and preferences 9 items Reported n = 1476	Values and preferences 6 items Reported n = 1476	Involvement 3 items Reported n = 1479	Information/ Knowledge 12 items Reported n = 1464	Values and preferences + Involvement 12 items Reported n = 1480	Values and preferences + Involvement 9 items Reported n = 1480
Item misfit, n	2 ^a	0	3 ^b	0	2 ^c	0	3 ^d	0
Person misfit, n (%)	113 (7.6)	113 (7.6)	79 (5.4)	79 (5.4)	57 (3.9)	57 (3.9)	83 (5.6)	83 (5.6)
Ceiling, n (%)	1 (0.07)	1 (0.07)	15 (1.0)	15 (1.0)	2 (0.14)	2 (0.14)	6 (0.41)	6 (0.41)
Floor, n (%)	5 (0.34)	5 (0.34)	55 (3.7)	55 (3.7)	2 (0.14)	2 (0.14)	51 (3.4)	51 (3.4)
Raw variance explained by measures, observed %	49.8	49.8	51.1	51.1	73.7	73.7	51.6	51.6
Unexplained variance in 1st contrast, %	4.9	4.9	14.0	14.0	3.8	3.8	10.5	10.5
Person Separation Index (including extreme and non-extreme)	1.27	1.27	1.03	1.03	1.27	1.27	1.10	1.10
Real RSME								

^aItem 11: Taking into account finding bowel cancer early (MnSq 1.57); Item 21: Taking into account that the test was free of charge (MnSq 1.45).

^bItem 11 (MnSq 1.58); Item 15 Taking into account risk for complications (MnSq 0.68); Item 21 (MnSq 1.47).

^cItem 8 Before decision, did you discussed with any health care personnel (MnSq 1.49); Item 26 (follow up item to item 25) How important was that person/s (health care personnel) as an information source (MnSq 0.35).

^dItem 11 (1.59), 15 (0.68), 21 (1.48).

variation than predicted can be interpreted as a threat to unidimensionality of the Rasch model, ie, in this case, not contributing to the assessment of the unidimensional concept of SDM. On the other hand, a single item with low *MnSq* value does not have to be a major threat to validity. Still the DIF performed in the present study revealed somewhat mixed results; seven of the items displayed DIF in relation to gender and nine items in relation to participating versus not participating in the SCREESCO programme. Although these differentiated item findings between people participating vs not are not surprising from a clinical perspective and also to some extent visible in our qualitative findings,¹³ especially regarding *Values and preferences*. The findings revealed different reasoning among participants and nonparticipants on how the decision to participate in screening was made and how family experience of CRC influenced this decision. However, in this qualitative study, we did not specifically study gender differences or similarities¹³; still, women in the present sample have a significantly higher educational level compared with men (Table 2), something that might influence DIF in relation to gender and the subscale *Information/Knowledge*. The results presented here may warrant further evaluation in relation to the development and use of a questionnaire aiming at assessing the level of SDM, eg, adding and/or modifying included items regarding varying decision-making styles and gender. Based upon our findings, future studies should therefore further explore the impact of DIF onto the measures generated, in order to be able to make valid and unbiased comparisons between groups. Research has also shown that having a more rational decision-making style significantly corresponded to having read more of a bowel-screening leaflet among men and women aged 60 to 70 years.³¹ Knowing someone with experience of CRC, having a higher educational level and having attended screening before, also corresponded to having read more of the leaflet³¹. Because the goal is to have a well-designed screening programme with a high uptake of individuals who have made an informed decision, it is important to know how people make their decisions and what influences the decision-making process. This will assist in directing relevant resources and in determining how information is distributed, and in this way, public knowledge will hopefully increase. Furthermore, whether, and in that case to what extent, health literacy (HL), anxiety, educational level, and other factors play a role in the decision-making process regarding the SCREESCO programme have been investigated within our research group.^{32,33}

Regarding unidimensionality, the somewhat mixed results indicate that the underlying concepts were measured to a satisfactory or almost satisfactory extent, according to the set thresholds, both for the whole questionnaire and the separate concepts/subscales. However, these similar findings make it difficult to interpret whether SDM is the concept being assessed or it is something else. *Information/Knowledge* displayed the best figures (73.7%), indicating that this concept/subscale is the most well defined in the SCREESCO questionnaire. A solid instrument should be able to capture different parts of a concept; however, those parts are to be considered as being interrelated with each other.²² Unexplained variance in first contrast (a possible secondary

dimension) revealed acceptable results for those concepts/subscales encompassing more items (12 and 22, respectively).

Regarding separation index, neither the questionnaire as a whole nor the single concepts/subscales have the ability to separate into at least two distinct groups. However, person misfit and individuals with maximum or minimum person measure (floor and ceiling effects) were limited in this sample and overall; person misfit displayed acceptable results and is therefore considered a strength. Possible explanations for the inability to separate into distinct groups might be that the range among persons and items do not fully match each other and that there is a need for a broader range of items to be able to target a greater proportion of individuals in a sample and thereby their level of SDM. One solution could therefore be to add more items and/or scale steps to the SCREESCO questionnaire. Still, the concept/subscale *Information/Knowledge* displays a separation index of 1.27, and the person-item map revealed that these items are the ones most spread out along the scale (Table 4). However, there is a drawback with items already having fixed answering options or a list of options for the respondent to choose from and also the fact that many of these items were dichotomized. The ideal situation would be that items and persons correspond to each other,²⁰ as for *Values and preferences*, while for *Information/Knowledge*, it could be considered good that those items are more spread out. Relating to our qualitative findings,¹³ it is not surprising that the three most difficult items include at what age to stop screening; estimating personal risk of developing CRC or what percentage of individuals diagnosed in Sweden will die from the disease. The qualitative findings revealed that lack of knowledge was prominent among both participants and nonparticipants in the SCREESCO programme, and this was further discussed among them in terms of lack of information in society regarding CRC.¹³ The lack of knowledge could of course also be because Sweden, as of yet, has not introduced public screening for CRC. Still, this relates to the discussion regarding how, and in what ways, information is distributed, where different approaches recently have been explored, such as using simple text messages, narrative films, and spoken text messages.³⁴⁻³⁶ Moreover, discussing personal risk of developing a disease is not very common in Swedish society or in the Swedish health care sector and therefore is most probably considered difficult to answer for many respondents. This was also evident because $n = 374$ answered "Don't know" to this item. The two least difficult items according to the person-item map were as follows: what factor/s are important to develop CRC and what symptom/s of CRC are known. However, these items were dichotomized before analysis and at least one correct factor or symptom was considered as having knowledge, regardless of the number of incorrect ones. For *Values and preferences*, items and persons are quite well matched, indicating that these items correspond to this particular sample; still, items appear not to be able to identify a variation in *Values and preferences*.

Because *Involvement* displayed two items demonstrating misfit out of three, no further analyses were performed. However, *Values and preferences* analysed together with *Involvement* displayed almost

similar results as *Values and preferences* alone, indicating that these two concepts/subscales may measure a similar concept and fail to discriminate between the two. Still, further work is needed to investigate this and perhaps develop more items to assess *Involvement* and *Values and preferences*. Worth mentioning is that it appeared difficult to discriminate between the three concepts in the qualitative analysis as well, and *Values and preferences* and *Involvement* were most closely connected.¹³

5 | STRENGTHS AND LIMITATIONS

Data collection was primarily completed in an online format. There are pros and cons with this approach; on the one hand, it is convenient for the respondent and cost effective, but on the other hand, there is a higher risk of more missing data, and the researcher has less control over the process. However, for the Rasch approach, individuals can theoretically be included in an analysis if they have at least one response; still, "Don't know" responses are not included here, except for items included in the *Information/Knowledge* concept/subscale, but not the item regarding risk. There is also a risk that those who are not comfortable with the online format choose not to participate. However, eligible participants had the opportunity to answer the questionnaire through a telephone interview, but the vast majority did it online. The present study represents the first step towards the development of a valid and reliable questionnaire aiming at assessing the level of SDM in relation to CRC screening; however, further evaluations and modifications are required.

6 | CONCLUSION

This Rasch analysis of the SCREESCO questionnaire revealed that the questionnaire in its current design have difficulties assessing the level of SDM in relation to CRC screening. According to the findings with two concepts/subscales possibly measuring the same concept, DIF in relation to gender and participation vs nonparticipation and less satisfactory results regarding separation index, further analyses are warranted. The achieved results have the potential to guide further evaluation and development, such as DIF analyses between other subgroups and adding/modifying items, with the a long-term goal of having a Swedish questionnaire, to be used in the health care sector, assessing the level of SDM in relation to CRC screening.

ACKNOWLEDGEMENTS

The authors are grateful to all the participants in this study. The authors would also like to acknowledge Kerstin Stake Nilsson (PhD) who took part in data collection.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

TRIAL REGISTRATION

Colonoscopy and FIT as Colorectal Cancer Screening Test in the Average Risk Population

<https://clinicaltrials.gov/ct2/show/NCT02078804?term=NCT02078804&rank=1>

ID NCT02078804 Protocol ID 2012/2058-31/3

First submitted: February 11, 2014; last updated October 2, 2017.

FUNDING INFORMATION

This study was supported by grants from the Swedish Associations of Local Authorities and Regions.

ORCID

Anna Jervaeus  <https://orcid.org/0000-0002-5582-8142>

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65(2):87-108.
2. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. *Gut*. 2015;64(10):1637-1649.
3. The National Board of Health and Welfare. *Official statistics of Sweden. Statistics of new cancer cases 2015.*;2017.
4. Blom J, Yin L, Liden A, et al. Toward understanding nonparticipation in sigmoidoscopy screening for colorectal cancer. *Int J Cancer*. 2008;122(7):1618-1623.
5. Blom J, Kilpelainen S, Hultcrantz R, Törnberg S. Five-year experience of organized colorectal cancer screening in a Swedish population—increased compliance with age, female gender, and subsequent screening round. *J Med Screen*. 2014;21(3):144-150.
6. The Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). *The Official Journal of the European Union*; 2003.
7. Sandman L, Granger BB, Ekman I, Munthe C. Adherence, shared decision-making and patient autonomy. *Med Health Care Philos*. 2012;15(2):115-127.
8. Zikmund-Fisher BJ, Couper MP, Singer E, et al. The DECISIONS study: a nationwide survey of United States adults regarding 9 common medical decisions. *Med Decis Making*. 2010;30(5 Suppl):20S-34S.
9. Hofmann B. Ethical issues with colorectal cancer screening—a systematic review. *J Eval Clin Pract*. 2017;23(3):631-641.
10. Informed Medical Decisions Foundation Healthwise Research and Advocacy. What is Shared Decision Making? <http://www.informedmedicaldecisions.org/what-is-shared-decision-making/>. Accessed October 27, 2015.
11. Charles C, Gafni A. The vexing problem of defining the meaning, role and measurement of values in treatment decision-making. *J Comp Eff Res*. 2014;3(2):197-209.
12. Charles C, Gafni A, Whelan T, O'Brien MA. Cultural influences on the physician-patient encounter: the case of shared treatment decision-making. *Patient Educ Couns*. 2006;63(3):262-267.
13. Fritzell K, Stake Nilsson K, Jervaeus A, Hultcrantz R, Wengström Y. The importance of people's *Values and preferences* for colorectal cancer screening participation. *Eur J Public Health*. 2017;27(6):1079-1084.
14. American Educational Research Association, APA, National Council on Measurement in Education. *Standards for Educational and Psychological*

- Testing. Washington DC: American Educational Research Association; 2014.
15. Fritzell K, Jervaeus A, Stake Nilsson K, Hultcrantz R, Wengström Y. Translation and cultural adaptation of the decisions module for colorectal cancer screening into a Swedish version—the SCREESCO questionnaire. *Scand J Gastroenterol* 2017;1-5.
 16. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales a practical guide to their development and use*. Oxford university press; 2015.
 17. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Danish Institute for Educational Research; 1960. Expanded edition with foreword and afterword by Wright BD. The University of Chicago Press; 1980 Reprinted in 1993 by MESA Press, Chicago.
 18. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. *J Rehabil Med*. 2003;35(3):105-115.
 19. Conrad KJ, Smith EV Jr. International conference on objective measurement: applications of Rasch analysis in health care. *Med Care*. 2004;42(1 Suppl):1-16.
 20. Bond TG, Fox CM. *Applying The Rasch Model Fundamental Measurement in the Human Sciences*. 2nd ed. New York, USA: Routledge; 2007.
 21. The National Board of Health and Welfare. *Official statistics of Sweden. Statistics- Health and Medical Care. Cancer Incidence in Sweden 2014*. 2015.
 22. Jervaeus A, Kottorp A, Wettergren L. Psychometric properties of KIDSCREEN-27 among childhood cancer survivors and age matched peers: a Rasch analysis. *Health Qual Life Outcomes*. 2013;11(1):96.
 23. Lerdal A, Kottorp A, Gay C, Aouizerat BE, Lee KA, Miaskowski C. A Rasch analysis of assessments of morning and evening fatigue in oncology patients using the Lee Fatigue Scale. *J Pain Symptom Manage*. 2016;51(6):1002-1012.
 24. Linacre JM. Optimizing Rating Scale Category Effectiveness. In: Smith EV, Smith RM, eds. *Introduction to Rasch Measurement Theory, Models and Applications*. Minnesota, USA: JAM Press Publisher; 2004:258-278.
 25. Smith AB, Rush R, Fallowfield LJ, Velikova G, Sharpe M. Rasch fit statistics and sample size considerations for polytomous data. *BMC Med Res Methodol*. 2008;8(1):33.
 26. Lerdal A, Opheim R, Gay CL, Moum B, Fagermoen MS, Kottorp A. Psychometric limitations of the 13-item Sense of Coherence Scale assessed by Rasch analysis. *BMC psychology*. 2017;5(1):18.
 27. *A user's guide to Winsteps Ministep Rasch-Model computer programs Program Manual 3.92.0*. [computer program]. 2016.
 28. Fisher W. Reliability, separation, strata statistics. *Rasch Meas Trans*. 1992;6(3):238.
 29. Mantel N. Chi-Square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure. *J Am Stat Assoc*. 1963;58(303):690-700.
 30. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959;22(4):719-748.
 31. Ghanouni A, Renzi C, Waller J. A cross-sectional survey assessing factors associated with reading cancer screening information: previous screening behaviour, demographics and decision-making style. *BMC Public Health*. 2017;17(1):327.
 32. Wangmar J, Jervaeus A, Fritzell K, Wangdahl J, Hultcrantz R, Wengstrom Y. Health literacy levels and views about being invited to a colorectal cancer screening program. *Acta Oncol*. 2018;1-7.
 33. Wangmar J, von Vogelsang AC, Hultcrantz R, Fritzell K, Wengstrom Y, Jervaeus A. Are anxiety levels associated with the decision to participate in a Swedish colorectal cancer screening programme? A nationwide cross-sectional study. *BMJ Open*. 2018;8(12):e025109.
 34. Meppelink CS, Smit EG, Buurman BM, van Weert JC. Should we be afraid of simple messages? The effects of text difficulty and illustrations in people with low or high health literacy. *Health Commun*. 2015;30(12):1181-1189.
 35. Meppelink CS, van Weert JC, Haven CJ, Smit EG. The effectiveness of health animations in audiences with different health literacy levels: an experimental study. *J Med Internet Res*. 2015;17(1):e11.
 36. Moran MB, Frank LB, Chatterjee JS, Murphy ST, Baezconde-Garbanati L. A pilot test of the acceptability and efficacy of narrative and non-narrative health education materials in a low health literacy population. *J Commun Healthc*. 2016;9(1):40-48.

How to cite this article: Jervaeus A, Fritzell K, Hultcrantz R, Wengström Y, Kottorp A. Psychometric properties of the SCREESCO questionnaire used in a colorectal cancer screening programme—A Rasch analysis. *J Eval Clin Pract*. 2020;26:541-550. <https://doi.org/10.1111/jep.13301>