BMJ Open

# Translation, adaption and psychometric testing of the German version of the Organizational Readiness for Implementing Change measure (ORIC)



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2019-034380 |
| Article Type: | Original research |
| Date Submitted by the Author: | 17-Sep-2019 |
| Complete List of Authors: | Lindig, Anja; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Hahlweg, Pola; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Christalle, Eva; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Scholl, Isabelle; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology |
| Keywords: | Organizational readiness for change, Psychometrics, Translation, Implementation, Shared decision-making, Measurement |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Translation, adaption and psychometric testing of the German version of the Organizational Readiness for Implementing Change measure (ORIC)

Anja Lindig[1], Pola Hahlweg[1], Eva Christalle[1], Isabelle Scholl[1]

[1] *Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany*

Email addresses:
Anja Lindig: a.lindig@uke.de
Pola Hahlweg: p.hahlweg@uke.de
Eva Christalle: e.christalle@uke.de
Isabelle Scholl: i.scholl@uke.de

Corresponding author:
Anja Lindig
University Medical Center Hamburg-Eppendorf
Department of Medical Psychology
Martinistr. 52
20246 Hamburg, Germany
Email: a.lindig@uke.de
Phone: +49-(0)40-7410-56206

Word count: 3844

## ABSTRACT

**Objectives** to translate the measure Organizational Readiness for Implementing Change (ORIC) into German, adapt it for the context of SDM, and assess its psychometric properties.

**Design** psychometric study based on secondary analysis of baseline data from a SDM implementation study.

**Setting** Three departments within one academic cancer center in Hamburg, Germany

**Participants** For comprehensibility assessment of the translated ORIC version we conducted cognitive interviews with n=11 healthcare professionals (HCPs). Afterwards, n=230 HCPs filled out the measure.

**Primary and Secondary Outcome Measures** The original English version of the ORIC was translated into German using a team translation protocol. Based on comprehensibility assessment via cognitive interviews with HCPs, the translated version was revised. For psychometric evaluation we conducted a secondary analysis of baseline data from a SDM implementation study. We analyzed acceptance (response rate), structural validity (exploratory factor analysis (EFA), confirmatory factor analysis (CFA)), item characteristics (item difficulties, corrected item-total correlations, inter-item correlations), and reliability (Cronbach's α).

**Results** Translation and cognitive testing of the ORIC was successful except for item 10, which showed low comprehensibility. Response rate was > 97%. Structural validity analysis provided a one factorial structure. Item difficulties ranged between 55.98 and 65.32, corrected item-total-correlation ranged between .66 and .74, inter-item correlations ranged between .43 and .72, and Cronbach's α was .93.

**Conclusions** The translated and adapted German ORIC is a highly accepted and reliable measure with satisfying psychometric properties and a one factorial structure. To increase comprehensibility and therefore content validity of the measure we suggest to remove item 10. The German ORIC can be used to analyze organizational readiness for change as a precursor for implementation success of various interventions.



**Keywords:**

Organizational readiness for change, Psychometrics, Translation, Implementation, Shared decision-making, Measurement

## ARTICLE SUMMARY

### Strengths and limitations of this study

- Appropriate qualitative methods were used to provide a German version of the ORIC which is comprehensible to healthcare professionals (HCPs).
- The sample size was large enough to robustly perform the psychometric analyses on the German version of the ORIC measure.
- Due to the design of the study as a secondary analysis, it was not possible to calculate some psychometric parameters.
- Data was collected at a single academic cancer center. Thus, psychometric properties of the German ORIC need replication in other settings.

## INTRODUCTION

Implementing interventions in healthcare systems is an important and widely discussed topic [1–3] and often mediated by public policies, market forces, or new technologies [4]. The intention to implement new interventions might be to reduce costs, improve quality, increase efficiency, or patient satisfaction [5]. Nevertheless, implementing change in healthcare organizations can be challenging [6–10].

In the German healthcare system, the implementation of shared decision-making (SDM) has received much attention [11]. SDM can be described as an interactional process on the basis of information exchange. Patients and healthcare professionals (HCPs) are equally and actively involved and jointly responsible for the decision [12–14]. SDM has been supported by health policy [15–17] and research [18–21] during the last decades. However, SDM is currently not widely implemented in the German healthcare system [11,22–24].

When implementing SDM or other interventions in organizations, several barriers on different levels of the organization (individual members of the organization, teams, organizational setting, wider environment) need consideration [1,5–7,10,25–28]. Barriers for implementing SDM in the clinical setting often address both the organizational setting (e.g. lack of resources and lack of management support) and the individual level (e.g. resistance to change or negative attitudes towards SDM) [7,10,11,29–31].

When implementing SDM or other interventions in healthcare systems, the clinical members' perspective on organizational readiness for change is a critical precursor to successfully implement empirical

knowledge [5,28,32–34]. Armenakis et al. [35] describes organizational readiness for change as the degree to which organizational members are prepared to participate in change processes. This is characterized by the belief that the change is needed and that the organization is capable of changing. Weiner et al. [36] differentiate between change commitment, where organizational members have positive attitudes towards implementing a change, and change efficiency, where they belief in capability to change. If readiness for change is high, organizational members invest more in the change effort and exhibit greater persistence to overcome obstacles and setbacks [36,37].

To analyze effects on change implementation success and to find suitable strategies to best implement changes, specific measures for assessing organizational readiness for change are needed [5]. However, only a few validated measures exist [5,38–40] and none were available in German. One of those measures is the Organizational Readiness for Implementing Change (ORIC) [39]. The ORIC is the only available measure that is brief, easy to administer and theoretically grounded [36]. It was previously translated into Danish and French [41,42]. The ORIC has been psychometrically tested, revealing good acceptability (response rate > 72%), good reliability (Cronbach's α > .80) and good validity with a two factor structure [39,41,42]. Due to the described properties, the ORIC seemed well-suited to measure organizational readiness for implementing SDM in Germany [5,39,43].

Therefore, the aim of the study was to translate the ORIC into German, adapt it for the context of SDM, and assess its psychometric properties.

## METHODS

### Measure

The ORIC measures organizational readiness for implementing change. It uses a 5-point Likert scale ranging from 0 „disagree“ to 4 „agree“ [39]. In the original English version, two subscales were described: „change commitment“ (items 1 to 5) and “change efficiency” (items 6 to 10). Sum scores were calculated for both subscales separately with higher scores indicating higher organizational readiness for change. By using the phrases “to implement this change” or “implementing this change”, the original scale does not specify, which change is addressed. The items can be specified to adapt to a specific research question and an introductory description can be added [36,41] . English items are displayed in Table 5.

### Translation

Translation followed the team translation protocol TRAPD (Translation, Review, Adjudication, Pretesting and Documentation) [44]. First, two team members (AL, SZ, cp. list of abbreviations) proficient in German and English independently translated the original ORIC into German. Second, a third bilingual team member (IS) suggested a third version based on the first two translations. Finally, we discussed all versions until reaching consensus on a final version. To find consensus on item 10 we additionally consulted an official translator (MM, cp. list of abbreviations) and an additional team member (PH).

**Assessment of comprehensibility as part of content validity and subsequent adaption of the scale**

Content validity is the degree to which the content of the measure and its items is adequately reflecting the measured construct [45]. According to the COSMIN criteria (Consensus-based Standards for the selection of health Measurement Instruments) [46], content validity includes the relevance of the items and scales, their comprehensiveness, and comprehensibility. As this study aimed at testing the translation of an existing measure, we focused on assessment of comprehensibility, which includes that items are appropriately worded and understood by participants as intended.

To do so, we conducted two rounds of cognitive interviews with a convenience sample of HCPs (nurses, physicians, and psychooncologists), working in a comprehensive cancer center in Germany. Two female psychologists experienced in interviewing (AL, PH) conducted the interviews. A compensation fee of 25 Euro was offered to the participants. We developed an interview guide based on recommendations of Willis et al. [47] and conducted interviews until reaching theoretical saturation. Interviews were audio-recorded and transcribed verbatim. After each round of cognitive interviews, we extracted and discussed comments and suggestions from the transcripts (AL, PH, IS) and adapted the German ORIC scale accordingly. As a further step to enhance comprehensibility, we discussed the items in the study team (AL, PH, IS, SZ, cp. list of abbreviations), with a second bilingual researcher in the field (DF, cp. list of abbreviations), the original author (CS, cp. list of abbreviations) as well as French (MR, cp. list of abbreviations) and Norwegian researchers (AH, cp. list of abbreviations) who have been working on translations of the ORIC into their languages.

We calculated descriptive statistics of demographic characteristics of participants using SPSS (IBM SPSS Statistics, Version 23).

**Psychometric evaluation**

**Data collection.** The psychometric evaluation of the ORIC measure is a secondary analysis of data gathered in a SDM implementation study [48]. Data from baseline assessment of the SDM implementation study were included. The ORIC was part of a three-page questionnaire measuring HCPs' attitudes regarding SDM and its implementation. Besides the ORIC, it contained the Control Preference Scale [49] and the IcanSDM [50], demographic questions (e.g. gender, age, profession, work experience) as well as several self-generated questions [51]. Results of these additional measures will be published as part of the primary evaluation of the SDM implementation study.

Participants were physicians and nurses, who worked at one of three departments within the University Cancer Center Hamburg at the University Medical Center Hamburg-Eppendorf. Eligible HPCs were identified through employee lists provided by department managers. The measure was handed out either (1) by a member of the study team (e.g. during a regular physician meeting), (2) by the supervising nurses, or (3) via employees' mailboxes. Participants returned the questionnaire personally to a study team member or by mail.

Data were entered into SPSS (IBM SPSS Statistics, Version 23) including blinded double entry of 20% of the data for quality control.

**Data Analysis**. Descriptive statistics were calculated for demographic characteristics. Cases were excluded if more than 30% of the ORIC items were missing. For all other cases, missing data were replaced with item means. We evaluated the response rate and therefore the acceptance of the measure by calculating frequencies of missing data per item. For analysis of acceptance, we also included the cases with more than 30% of ORIC items missing.

We a priori hypothesized to replicate the two-dimensional structure of the original English ORIC version. Two factors "change commitment" (item 1 to 5) and "change efficiency" (item 6 to 10) were postulated. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett`s test of sphericity were performed to test prerequisites for factor analysis [52,53]. A confirmatory factor analysis (CFA) with Maximum Likelihood Estimates and two factors was applied for the whole data set as a first step. Because the two-factor model could not be confirmed, the data set was randomized and split into two subsets (each n=115). An exploratory factor analysis (EFA) with orthogonal varimax rotation and extraction of components with eigenvalues > 1 was applied for the first subset. A CFA was again calculated for the second subset. A range of global goodness of fit indices were used to assess the degree to which observed data were accounted for by the proposed models: discrepancy chi-squared statistic ($Chi^2$), degree of freedom (df),

normed chi-squared statistic ($Chi^2/df$), comparative fit index (CFI), Tucker-Lewis Index (TLI), root mean square error of approximation (RMSEA) as well as Akaike Information Criterion (AIC) and Parsimonious Normed Fit Index (PNFI) for analysing model complexity. Established rules to estimate the model fit were used [54–57].

Item analysis was performed for the one-factor model. It included calculation of item means and standard deviation as well as observation of floor and ceiling effects [58], calculation of corrected item-total correlations [52,59], inter-item correlations [52,59], and item difficulties [53]. Internal consistency of the scale was assessed by Cronbach's alpha coefficient (α) [53,59,60]. For a detailed overview on performed data analyses see Table 1.

Because of low content validity of item 10, the use of item 10 for the German ORIC needs to be evaluated. Therefore, we also conducted psychometric analysis (EFA with varimax rotation and extraction of components with eigenvalues > 1, corrected item-total correlations, Cronbach's α, and goodness of fit indexes) for the 9-item version of the ORIC (see Supplementary File 1).

Analysis of demographic data, analysis of response rate, item analysis and EFA were performed using SPSS (IBM SPSS Statistics, Version 23). CFA and calculation of model fit indices were performed using Amos (IBM SPSS Amos 22.0.0).

Table 1: Psychometric analyses conducted.

| **Psychometric measure** | **Criteria** |
|---|---|
| Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett`s test of sphericity | These tests ensure that correlations between variables can be accounted for by a smaller set of factors [53]. KMO value should be higher than .05 and Bartlett`s test value should be less than .05 to fulfil the criteria for calculating a factor analysis [53,54]. |
| Analysis of frequencies for item response distributions | Floor and ceiling effects were assumed present if more than 15% of participants choose the lowest or highest possible score [59]. |
| Corrected item-total correlations | If items correlate with the total score of above .30, they measure the same underlying concept. Items with lower correlations should be removed because they do not add exploratory power to the measure [53,60]. |
| Item difficulties | Item difficulties are calculated by dividing item means by the maximal value of the answer range (0-4) and multiplying it with 100. Item difficulty should be near to 50%, and items should not differ much in their difficulty level [54]. |

| | |
|---|---|
| Inter-item correlations | Inter-item correlations ensure association between items. High inter-item correlations of above .80 indicate that items ask the same questions and might be redundant [53,60]. |
| Cronbach's α | Cronbach's α is a measure for reliability and internal consistency. A value of at least .70 is acceptable and higher coefficients indicate a more stable measure [53,54,61]. |
| Normed chi-squared statistic ($Chi^2$/df) | $Chi^2$/df is an indicator for model fit, dependent on sample size and should be as small as possible. A ratio between 2 and 3 indicate a good data fit [55]. |
| Comparative fit indexes (CFI) | CFIs is an indicator for model fit. It ranges from 0 to 1 and higher values indicate better fit. Values above .95 indicate a good model fit [58,62]. |
| Tucker-Lewis Index (TLI) | TLI is an indicator for model fit. It corrects for complexity of the model and is sensitive to small sample sizes. Values above .95 indicate good fit [56]. |
| Root mean square error of approximation (RMSEA) | RMSEA is an absolut index which describes closeness to fit. Values below .05 indicate a good fit, values between .05 and .08 indicate an adequate fit, values between .08 and 1 indicate a moderate fit and values above 1 are unacceptable [63]. |
| Akaike Information Criterion (AIC) | AIC is a parsimony model fit index. It can be used to compare fit of competing models with smaller values indicating better fit [55,58]. |
| Parsimonious Normed Fit Index (PNFI) | PNFI is a parsimony model fit index. It ranges between 0 and 1 and higher values indicate a more parsimonious fit [55]. No threshold levels are recommended and it has to be analysed in combination with other goodness of fit indices [58]. |

## Patient and public involvement

The measure ORIC preliminary addresses HCPs, therefore physicians, nurses and psychooncologists were involved in the adaption of the measure by taking part in cognitive interviews. Our cooperation partner as well as stakeholders of the three participating clinics (physicians and nurses) supported the recruitment process by handing over surveys to their colleagues or supporting us to reach eligible participants. Patients were not involved in this study.

# RESULTS

## Translation

We adapted the ORIC to fit our research question of SDM implementation. Therefore "to implement this change" / "implementing this change" was rephrased into "to implement shared decision-making" / "implementing shared decision-making". Additionally we added an introductory description. Within the first

round of discussion, we reached consensus for items 1 to 9. We struggled to translate the phrase “manage the politics” in item 10 into German. Therefore, we discussed item 10 with additional colleagues (cp. methods section) until consensus was found.

**Assessment of comprehensibility as part of content validity and subsequent adaption of the scale**

To test the German ORIC for comprehensibility, cognitive interviews with N=11 participants (nurses, physicians, and psychooncologists) were conducted. Cognitive interviews lasted about one hour. For demographic data of participants see Table A in Supplementary File 2.

After two rounds of cognitive interviews and slight modifications of single words or phrases, items 1 to 9 were well understood by all participants. Item 10 could not be translated successfully. After the first round of cognitive interviews with n=7, we adapted item 10 after discussions within the study team (AL, PH, IS) as well as with other international researchers working with the ORIC (CS, AH, MR, cp. list of abbreviations). A second round of cognitive interviews with n=4 showed that comprehension of the German translation of the phrase “manage the politics” did not picture the correct English meaning. We therefore involved another bilingual SDM implementation researcher (DF, cp. list of abbreviations). We found consensus for an adapted version of item 10 but it was still not satisfying from the study team and experts view. Item 10 was found to have low comprehensibility as part of content validity according to COSMIN criteria [46]. The final German ORIC measure, used in this study, is presented in Supplementary File 3.

During cognitive interviews some nurses reported that they had not heard about the term “shared decision-making” (German: “Partizipative Entscheidungsfindung”) prior to participation. Thus, we provided a definition of SDM in the introduction part of the questionnaire within the SDM implementation study.

**Psychometric evaluation**

**Sample characteristics.** Data of 235 HCPs were available for this secondary analysis. Five cases were excluded (except for assessment of response rate), because all items of the ORIC were missing. For all other cases <30% of the items were missing. Thus, data of 230 HCPs could be included into analysis.

Table 2 provides an overview of demographic characteristics of participants. Most of the 230 HCPs were between 31 and 40 years old (36.96%), female (70.43%), worked as a nurse (56.96%), and had a work experience of < 5 years (43.91%).

Table 2: Demographic characteristics of participants (n=230).

| | | N | % |
|---|---|---|---|
| Age | < 30 years | 72 | 31.3 |
| | 31-40 years | 85 | 36.96 |
| | 41-50 years | 42 | 18.26 |
| | > 50 years | 26 | 11.30 |
| | Missings | 5 | 2.17 |
| Gender | Female | 162 | 70.43 |
| | Male | 59 | 25.65 |
| | Different gender or preferred not to answer questions | 5 | 2.1 |
| | Missings | 4 | 1.74 |
| Profession | Nurse | 131 | 56.96 |
| | Junior physician | 69 | 30.00 |
| | Senior physician | 27 | 11.74 |
| | Missings | 3 | 1.30 |
| Work experience in healthcare | < 5 years | 101 | 43.91 |
| | 5-10 years | 48 | 20.87 |
| | 11-20 years | 46 | 20.00 |
| | > 20 years | 28 | 12.17 |
| | Missings | 7 | 3.04 |

**Structural validity.** KMO measure was .933 and Barlett´s test of sphericity yielded $X^2 = 1485.11$, $p < .001$, indicating that a factor analysis of the data was appropriate. CFA for the hypothesised two-factor model showed a high correlation of .87 between the two components (see Supplementary File 4). Therefore, we postulated a one factorial structure and conducted a post-hoc EFA. Since the main component explains 67% of the variance, a one-factor model was assumed (see Table 3). The factor loading for the first component is shown in Table B of Supplementary File 2.

Table 3: Results of EFA.

| | Eigenvalue | | |
|---|---|---|---|
| | Total | % of variance | Cumulative % |
| Component 1 | 6.72 | 67.23 | 67.23 |
| Component 2 | 0.83 | 8.30 | 75.53 |
| Component 3 | 0.47 | 4.75 | 80.28 |
| Component 4 | 0.41 | 4.08 | 84.36 |
| Component 5 | 0.39 | 3.91 | 88.28 |
| Component 6 | 0.32 | 3.24 | 91.52 |
| Component 7 | 0.27 | 2.74 | 94.26 |
| Component 8 | 0.23 | 2.34 | 96.60 |
| Component 9 | 0.17 | 1.74 | 98.35 |
| Component 10 | 0.16 | 1.65 | 100.00 |

Notes: Eigenvalues of the ten components of the German ORIC, percentage of explained variance, and cumulative percentage of explained variance of each component. For EFA, half of the data set (n=115) was used.

A CFA was performed again with the one-factor model to analyse its fit indices. Indices of the two-factor model and the one-factor model are compared in Table 4.

Table 4: Fit indices of two calculated models for factor analysis of the German ORIC.

| | Chi² [1] | df [2] | Chi²/df [3] | CFI [4] | TLI [5] | RMSEA [6] | AIC [7] | PNFI [8] |
|---|---|---|---|---|---|---|---|---|
| Two-factor model | 81.71* | 34 | 2.40 | .968 | .947 | .078 | 143.71 | .585 |
| One-factor model | 77.19* | 35 | 2.20 | .928 | .907 | .103 | 117.19 | .682 |

Notes: Two-factor model was calculated for the whole data set (n=230): factor 1 includes item 1 to 5, factor 2 includes item 6 to 10; One-factor model was calculated for half of the data set (n=115): includes items 1 to 10.
[1] discrepancy chi-squared statistic, [2] degrees of freedom, [3] normed chi-squared statistic, [4] comparative fit indexes, [5] Tucker-Lewis Index, [6] root mean square error of approximation, [7] Akaike Information Criterion, [8] Parsimonious Normed Fit Index (PNFI). * $p$ = .000

Results of factor analysis for the 9-item version of the ORIC (without item 10) were similar. For the 9-item version also, a one-factor model was assumed by exploratory factor analysis. Goodness of fit indices of the one-factor model of the 9-item ORIC version showed similar values compared to the 10-item ORIC version (see Supplementary File 1).

**Analysis of the ORIC items and reliability.** Table 5 shows response distribution, means, standard deviations, acceptance, corrected item-total correlation and item difficulty of the ten items. Means range between 2.24 (item 9) and 2.61 (item 5). Most people responded in the middle of the scale with a slight shift to more agreement. Missing values for items 1 to 9 ranged from four to six. For item 10, nine missing values were found. Taking all items into account, more than 97% of the measure were answered. Corrected item-total correlations ranged from .66 (item 9) to .74 (item 3), and item difficulties ranged from 55.98 (item 9) to 65.32 (item 5). Inter-item correlations range from .434 (item2/item9) to .723 (item3/item 5) (see Table C of Supplementary File 2). Internal consistency yielded a Cronbach's α of .93.

Results of reliability and item analysis of the 9-item ORIC version were similar to results for the 10-item version (see Supplementary File 2).

278 Table 5: Response distribution, means, standard deviation, acceptance, discrimination and item difficulty of the German ORIC.

| | Items | Disagree N (%) | Somewhat Disagree N (%) | Neither Agree nor Disagree N (%) | Somewhat Agree N (%) | Agree N (%) | Mean (SD) | Acceptance (Response in %) * | Item discrimination (corrected item-total correlation) | Item difficulty |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | People who work here are committed to implementing shared decision-making. | 1 (0.4) | 22 (9.6) | 109 (47.0) | 73 (30.9) | 25 (10.9) | 2.42 (.826) | 97.43 | .744 | 60.58 |
| 2 | People who work here will do whatever it takes to implement shared decision-making. | 4 (1.7) | 37 (16.1) | 103 (44.8) | 68 (29.6) | 18 (7.4) | 2.25 (.878) | 98.29 | .689 | 56.36 |
| 3 | People who work here want to implement shared decision-making. | 0 (0.0) | 15 (6.5) | 107 (46.5) | 84 (35.7) | 24 (10.4) | 2.50 (.768) | 97.43 | .774 | 62.61 |
| 4 | People who work here are determined to implement shared decision-making. | 2 (9.0) | 38 (16.5) | 107 (46.5) | 67 (29.1) | 16 (7.0) | 2.25 (.843) | 98.29 | .758 | 56.19 |
| 5 | People who work here are motivated to implement shared decision-making. | 1 (0.4) | 16 (7.0) | 85 (37.0) | 97 (42.2) | 31 (13.5) | 2.61 (.821) | 98.29 | .764 | 65.32 |
| 6 | People who work here feel confident that they can handle the challenges that might arise in implementing shared decision-making. | 2 (0.9) | 20 (8.7) | 93 (40.4) | 93 (40.4) | 22 (9.6) | 2.49 (.819) | 98.29 | .760 | 62.28 |
| 7 | People who work here feel confident that they can keep track of progress in implementing shared decision-making. | 1 (0.4) | 26 (11.3) | 93 (40.4) | 92 (40.0) | 18 (7.8) | 2.43 (.811) | 98.29 | .725 | 60.87 |
| 8 | People who work here feel confident that they can coordinate tasks so that implementation goes smoothly. | 5 (2.2) | 24 (10.4) | 107 (46.5) | 78 (33.5) | 16 (6.5) | 2.32 (.833) | 99.56 | .697 | 58.13 |
| 9 | People who work here feel confident that the organization can support people as they adjust to shared decision-making. | 6 (2.6) | 43 (18.7) | 89 (38.7) | 74 (32.2) | 18 (7.8) | 2.24 (.934) | 97.86 | .665 | 55.98 |
| 10 | People who work here feel confident that they can manage the politics of implementing shared decision-making. | 3 (1.3) | 24 (10.04) | 122 (50.9) | 65 (28.3) | 16 (7.0) | 2.29 (.796) | 96.15 | .714 | 57.44 |

Notes: Items could be answered on a 5-step Likert scale rating from 0 „disagree" to 4 „agree". SD = standard deviation. * For calculation of response rate, four additional cases were included because these participants only skipped the ORIC but filled out the rest of the questionnaire.

279

## DISCUSSION

The original English ORIC measure is a brief measure with good psychometric properties [39], which were confirmed in Danish [41] and French [42] validation studies. We chose the ORIC to assess organizational readiness for implementing SDM in a implementation study in a hospital setting in Germany [48]. As a first step, the study at hand aimed to translate the ORIC into German, adapt it for the context of SDM, and assess its psychometric properties.

### Assessment of comprehensibility as part of content validity

Items 1 to 9 could be translated and adapted successfully. Feedback by participants, members of the study team, and external experts as well as response rates suggest that item 10 could not be translated satisfyingly and comprehensibility of this item is low [46]. A reason might be the translation of the phrase "manage the politics" into "Machenschaften". We assume that the term "manage the politics" has a strong cultural connotation because no equivalent phrase in German language exists. The German phrase "Machenschaften" might be confusing for some participants and they decided to rather skip the item. Ruest et al. [42], who translated the English ORIC into French, also identified several differences in cultural concepts during their adaptation process but could translate all items successfully. They concluded that limitations in linguistic validation could decrease comparability of psychometric results of the translated measure. In our sample, item 10 comes along with similar and inconspicuous item characteristics compared to other items. When repeating factor and item analyses for the 9-item ORIC version including only item 1 to 9, very similar results were observed compared to the 10-item version. Thus, we would recommend to use the 9-item ORIC version to increase comprehensibility and therefore content validity of the scale.

### Structural validity

We could not confirm the two-factor structure of the English and the translated Danish and French versions of ORIC [39,41,42]. Results of EFA indicated a one-factor structure while the two-factor model showed better fit indices. Both models have acceptable values for $Chi^2/df$ [54] and CFI [57,61] but only the two-factor model has acceptable values for TLI [55] and RMSEA [62]. When involving parsimony of the models by calculating AIC and PNFI, the one-factor model fits better to the data [54,57]. Therefore, the more parsimonious one-factor model should be preferred.

**Analysis of ORIC items and reliability**

Missing value rates for single items and the overall measure were quite low. Therefore, the German version of the ORIC was found to be a well-accepted measure. Observations of floor and ceiling effects, corrected item-total correlations, item difficulties, and inter-item correlations met the criteria for good quality of measures [52,53,58–60]. Cronbach's α (α=.93) suggest excellent reliability [52,53,60].

According to Streiner and Norman [53] a Cronbachs' α above .90 might also indicate item redundancy. On the other hand, inter-item correlations of the German ORIC are in an acceptable range and all items seem to add additional information to the underlying concept [52,59]. In implementation research there is a need for preferably brief measures, which can be applied in diverse settings with high work-load. Thus, future research could investigate the possibility to further reduce the number of items.

**Strength and Limitations**

This study has some limitations. First, several psychometric parameters are not analysable because of the secondary analysis. Due to the design of the study, it was not possible to calculate e.g. convergent or divergent validity yet. Second, we applied the ORIC only in three departments of one University Medical Center in Germany. Further validation in different organizational settings is needed to ensure generalizability.

A major strength of this study is that we provided the first measure to assess organizational readiness for change in German language for use in implementation studies. Furthermore, we assessed the ORIC in a large sample including physicians and nurses.

## CONCLUSION

Organizational readiness is a crucial indicator to successfully implement change and a possible barrier if missing. For implementation studies, it is essential to measure organizational readiness with valid and reliable measures. We provide the first German measure for organizational readiness for implementing change. The German ORIC is a brief measure and highly accepted. We found satisfying psychometric properties in a German hospital setting. To increase content validity of the measure, we suggest to leave out item 10. As the ORIC targets the attitude of organizational members, it can detect reduced or missing readiness for implementing a change on the individuals' level. Therefore, the German ORIC can be used

to analyse organizational readiness as a possible barrier for implementing various interventions in the healthcare setting.

## LIST OF ABBREVIATIONS

AH: Anne Haugstvedt

AIC: Akaike Information Criterion

AL: Anja Lindig

AMOS: Analysis of Moment Structure, Statistical Package for the Social Sciences, International Business Machines Corporation

$Chi^2$ : Discrepancy chi-squared statistic

Chi2/df: Normed chi-squared statistic

CFI: Confirmatory fit index

CS: Christopher M. Shea

Df: Degree of freedom

DF: Dominick Frosch

EC: Eva Christalle

HCP: Healthcare professional

IS: Isabelle Scholl

KMO: Kaiser-Meyer-Olkin criterion

MM: Marcel Machalski

MR: Mélanie Ruest

ORIC: Organizational Readiness for Implementing Change

PH: Pola Hahlweg

PNFI: Parsimonious Normed Fit Index

RMSEA: Root mean square error of approximation

SDM: Shared decision-making

SPSS: Statistical Package for the Social Sciences, International Business Machines Corporation

SZ: Stefan Zeh

TLI: Tucker-Lewis Index

TRAPD: Translation, Review, Adjudication, Pretesting, Documentation

## DECLARATIONS

### Ethics approval and consent to participate

The study was approved by the Ethics Committee of the Medical Association Hamburg (Germany, study ID PV5368). The study was carried out in accordance to the latest version of the Helsinki Declaration of the World Medical Association. Principles of good clinical practice were respected. Data protection requirements were met. Study participation was voluntary. A waiver of consent for HCPs was obtained from the Ethics Committee, as proposed by current statements on ethical designs of implementation research [63]. HCPs were able to decline participation in the study.

### Funding

This study is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, grant number 232160533).

### Competing interests

PH gave one scientific presentation on shared decision-making during a lunch symposium, for which she received compensation and travel compensation from GlaxoSmithKline GmbH in 2018. AL, EC and IS declared to not have any competing interests.

### Author contributions

AL, PH and IS made substantial contributions to the design and preparation of the study. AL and PH collected the data. AL conducted the analysis in collaboration with EC. All authors contributed to the interpretation of results. AL drafted the manuscript and PH, IS and EC were involved in critically revising the manuscript for important intellectual content. All authors gave final approval of the version to be published.

### Checklist for reporting statement

To report the results of this validation study, we used the Authors Guidelines for Reporting Scale Development and Validation Results by Cabrera-Nguyen [64]. The Checklist can be found in Supplementary File 5.

**Patient consent form**

Not applicable

**Data sharing statement**

The dataset collected and analyzed during this study is available from the corresponding author on reasonable request.

**Acknowledgements**

We thank our student assistants Anastasia Izotova, Sophia Schulte and Nicolai Pergande for their help preparing the study and the analysis. We thank Stefan Zeh for his part the team translation process. We thank Dominick Frosch, Anne Haugstvedt, Marcel Machalski, Mélanie Ruest and Christopher M. Shea for helpful comments and suggestions in the translation and adaption process of item 10. We also would like to thank Christopher M. Shea for giving the opportunity to translate the ORIC into German and to use it in our implementation study.

**Supplementary Files**

Supplementary File 1: Results of psychometric analysis of the 9-item German ORIC version: Factor analysis, corrected item-total correlations, Cronbach's α

Supplementary File 2: Additional tables for psychometric evaluation of the 10-item ORIC version

Supplementary File 3: Organizational Readiness for Implementing Change (ORIC) – German version

Supplementary File 4: Confirmatory factor analysis model of the 10-item German ORIC

Supplementary File 5: Reporting Checklist: Authors Guidelines for Reporting Scale Development and Validation Results by Cabrera-Nguyen [64]

**REFERENCES**

1 Øvretveit J, Andreen-Sachs M, Carlsson J, et al. Implementing organisation and management innovations in Swedish healthcare. *J Health Organ Manag* 2012;26:237–57.

2 Montgomery A, Doulougeri K, Panagopoulou E. Implementing action research in hospital settings: a systematic review. *J Health Organ Manag* 2015;29:729–49.

3 Rye CB, Kimberly JR. The Adoption of Innovations by Provider Organizations in Health Care. *Med Care Res Rev* 2007;64:235–78.

4 Bazzoli GJ, Dynan L, Burns LR, et al. Two Decades of Organizational Change in Health Care: What Have we Learned? *Med Care Res Rev* 2004;61:247–331.

5 Weiner BJ, Amick H, Lee S-YD. Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields. *Med Care Res Rev* 2008;65:379–436.

6 Batras D, Duff C, Smith BJ. Organizational change theory: Implications for health promotion practice. *Health Promot* 2016:231–41.

7 Yazdani A, Wells R. Barriers for implementation of successful change to prevent musculoskeletal disorders and how to systematically address them. *Appl Ergon*. 2018;73:122–40.

8 Schelvis RMC, Wiezer NM, Blatter BM, et al. Evaluating the implementation process of a participatory organizational level occupational health intervention in schools. *BMC Public Health* 2016;16:1212.

9 Scott DAH, Grant SM. A meta-ethnography of the facilitators and barriers to successful implementation of patient complaints processes in health-care settings. *Heal Expect* 2018;21:508–17.

10 Williams B, Perillo S, Brown T. What are the factors of organisational culture in health care settings that act as barriers to the implementation of evidence-based practice? A scoping review. N*urse Educ Today* 2015;35:e34–41.

11 Härter M, Dirmaier J, Scholl I, et al. The long way of implementing patient-centered care and shared decision making in Germany. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:46–51. A

12 Bieber C, Loh A, Ringel N, et al. Patientenbeteiligung bei medizinischen Entscheidungen. Manual zur partizipativen Entscheidungsfindung (Shared Decision-Making). Freiburg: Universitätsklinikum Freiburg 2007.

13 Härter M, Buchholz A, Nicolai J, et al. Shared decision making and the use of decision aids. *Dtsch Arztebl Int* 2015;112:672–9.

14 Salzburg Global Seminar. Salzburg statment on shared decision making. *BMJ* 2011;342:1745.

15 Härter M, Müller H, Dirmaier J, et al. Patient participation and shared decision making in Germany - History, agents and current transfer to practice. *Z Evid Fortbild Qual Gesundhwes* 2011;105:263–70.

16 Härter M, Moumjid N, Cornuz J, et al. Shared decision making in 2017: International accomplishments in policy, research and implementation. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:1–5.

17 Härter M, van der Weijden T, Elwyn G. Policy and practice developments in the implementation of shared decision making: An international perspective. *Z Evid Fortbild Qual Gesundhwes* 2011;105:229–33.

18 Fujimori M, Uchitomi Y. Preferences of Cancer Patients Regarding Communication of Bad News: A Systematic Literature Review. *Jpn J Clin Oncol* 2009;39:201–16.

19 Ernst J, Brahler E, Weissflog G. [Patient involvement in medical decision making--an overview on patient preferences and impacting factors]. *Gesundheitswesen* 2014;76:187–92.

20 Chewning B, Bylund CL, Shah B, et al. Patient preferences for shared decisions: A systematic review. *Patient Educ Couns* 2012;86:9–18.

21 Brom L, Hopmans W, Pasman HR, et al. Congruence between patients' preferred and perceived participation in medical decision-making: A review of the literature. *BMC Med Inform Decis Mak* 2014;14:25.

22 Légaré F, Stacey D, Turcotte S, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2014;9:CD006732.

23 Légaré F, Witteman HO. Shared decision making: Examining key elements and barriers to adoption into routine clinical practice. *Health Aff* 2013;32:276–84.

24 Frosch D, Tietbohl C, Scholl I. Overcoming implementation challenges to advance shared-decision making in routine practice. In: Elwyn G, Thompson R, Edwards A, editors. Shar Decis Mak Heal Care Achiev Evidence-Based Patient Choice, 3rd Ed. Oxford: Oxford University Press 2016.

25 Grol R, Wensing M, Eccles M, et al. Improving patient care: The implementation of change in health care. John Wiley & Sons 2013.

26 Grol R, Grimshaw J. From best evidence to best practice: Effective implementation of change in patients' care. *Lancet* 2003;362:1225–30.

27 Cummings GG, Estabrooks CA, Midodzi WK, et al. Influence of Organizational Characteristics and Context on Research Utilization. *Nurs Res* 2007;56:S24–39.

28 Li S-A, Jeffs L, Barwick M, et al. Organizational contextual features that influence the implementation of evidence-based practices across healthcare settings: a systematic integrative review. *Syst Rev* 2018;7:72.

29 Légaré F, Ratté S, Gravel K, et al. Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions. *Patient Educ Couns*.

2008;73:526–35.

30 Frerichs W, Hahlweg P, Müller E, et al. Shared decision-making in oncology – a qualitative analysis of healthcare providers' views on current practice. *PLoS One* 2016;11:e0149789.

31 Scholl I, LaRussa A, Hahlweg P, et al. Organizational- and system-level characteristics that influence implementation of shared decision-making and strategies to address them - a scoping review. *Implement Sci* 2018;13:40.

32 André B, Sjøvold E. What characterizes the work culture at a hospital unit that successfully implements change - a correlation study. *BMC Health Serv Res* 2017;17:486.

33 Armenakis AA, Harris SG, Mossholder KW. Creating Readiness for Organizational Change. *Hum Relations* 1993;46:681–703.

34 Kash BA, Spaulding A, Gamm LD, et al. Leadership, culture, and organizational technologies as absorptive capacity for innovation and transformation in the healthcare sector: A framework for research. *Chang Manag* 2014;13:1–13.

35 Armenakis AA, Harris SG, Mossholder KW. Creating Readiness for Organizational Change. *Hum Relations* 1993;46:681–703.

36 Weiner BJ. A theory of organizational readiness for change. *Implement Sci* 2009;4:67.

37 Armenakis AA, Harris SG. Crafting a change message to create transformational readiness. *J Organ Chang Manag* 2002;15:169–83.

38 Holt D, Armenakis A, Harris S, et al. Toward a Comprehensive Definition of Readiness for Change: A Review of Research and Instrumentation. *Res Organ Chang Dev*. 2007;16:289–336.

39 Shea CM, Jacobs SR, Esserman DA, et al. Organizational readiness for implementing change: A psychometric assessment of a new measure. I*mplement Sci* 2014;9:7.

40 Gagnon M-P, Attieh R, Ghandour EK, et al. A Systematic Review of Instruments to Assess Organizational Readiness for Knowledge Translation in Health Care. *PLoS One* 2014;9:e114338.

41 Storkholm MH, Mazzocato P, Tessma MK, et al. Assessing the reliability and validity of the Danish version of Organizational Readiness for Implementing Change (ORIC). *Implement Sci* 2018;13:78.

42 Ruest M, Leonard G, Thomas A, et al. French cross-cultural adaptation of the Organizational Readiness for Implementing Change (ORIC). *BMC Health Serv Res* 2019;19:535.

43 Alexander JA, Hearld LR. Methods and metrics challenges of delivery-system research. *Implement Sci* 2012;7:15.

44 Harkness J, Pennell BE, Schoua-Glusberg A. Methods for Testing and Evaluating Survey Questionnaires. In: Presser S, Rothgeb J, Couper M, Lessler J, Martin E, Martin J, et al., editors. Surv Quest Transl Assess. Hoboken (NJ): JohnWiley & Sons; 2004:453–73.

45 Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737–45.

46 Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27:1159–70.

47 Willis GB, Research Triangle Institute. Cognitive interviewing - a “how to” guide. 1999.

48 Scholl I, Hahlweg P, Lindig A, et al. Evaluation of a program for routine implementation of shared decision-making in cancer care : study protocol of a stepped wedge cluster randomized trial. *Implement Sci* 2018:1–10.

49 Degner LF, Sloan JA, Venkatesh P. The Control Preferences Scale. *Can J Nurs Res* 1997;29:21–43.

50 Giguere A, Bogza LM, Coudert L, et al. Validation of the IcanSDM instrument to assess clinicians’ perceptions of their ability to adopt shared decision making [in preparation].

51 Hahlweg P, Härter M, Nestoriuc Y, et al. How are decisions made in cancer care? A qualitative study using participant observation of current practice. *BMJ Open* 2017;7:e016360.

52 Rattray J, Jones MC. Essential elements of questionnaire design and development. *J Clin Nurs* 2007;16:234–43.

53 Streiner DL, Norman GR, Cairney J. Health measurement scales: A practical guide to their development and use. 2015.

54 Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Meth Psychol Res Online* 2003;8:23–74.

55 Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Eq Model* 1999;6:1–55.

56 Hair J, Black W, Babin B, et al. Multivariate data analysis. Upper Saddle River, NJ: Pearson Prentice Hall 2009.

57 Hooper D, Coughlan J, Mullen M. Structural Equation Modelling: Guidelines for Determining Model Fit. *Articles* 2008.

58 McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995;4:293–307.

59 Ferketich S. Focus on psychometrics. Aspects of item analysis. *Res Nurs Health* 1991;14:165–8.

60 Bannigan K, Watson R. Reliability and validity in a nutshell. *J Clin Nurs* 2009;18:3237–43.

61 Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238–46.

62 Browne MW, Cudeck R. Single Sample Cross-Validation Indices for Covariance Structures. *Multivariate Behav Res* 1989;24:445–55.

63 Hutton J, Eccles M, Grimshaw J. Ethical issues in implementation research: A discussion of the problems in achieving informed consent. *Implement Sci* 2008;3:52.

64 Cabrera-Nguyen P. Author Guidelines for Reporting Scale Development and Validation Results. *J Soc Social Work Res* 2010;1:99–103.

**Supplementary File 1: Results of psychometric analysis of the 9-item German ORIC: Factor analysis, corrected item-total correlations, Cronbach's α**

**Structural validity**

To get information about the factorial structure of the 9-item German ORIC, an exploratory factor analysis was conducted. KMO measure was .926 and Barlett's test of sphericity was significant ($X^2$ = 1302.78, $p < .001$), indicating that a factor analysis of the data was appropriate to observe data. Table A shows results of the exploratory factor analysis. Since the main component explains 66% of the variance, a one-factor model could be assumed. The factor loading for the first component can be observed in table B.

Table A: Results of exploratory factor analysis of the 9-item German ORIC.

| | Eigenvalue | | |
|---|---|---|---|
| | Total | % of variance | Cumulative % |
| Component 1 | 6.02 | 66.85 | 66.85 |
| Component 2 | 0.802 | 8.91 | 75.76 |
| Component 3 | 0.472 | 5.25 | 81.01 |
| Component 4 | 0.392 | 4.35 | 85.37 |
| Component 5 | 0.366 | 4.06 | 89.43 |
| Component 6 | 0.304 | 3.38 | 92.81 |
| Component 7 | 0.250 | 2.78 | 95.59 |
| Component 8 | 0.232 | 2.57 | 98.16 |
| Component 9 | 0.165 | 1.84 | 100.00 |

Notes: Eigenvalues of the nine components of the German ORIC, percentage of explained variance and cumulative percentage of explained variance of each component. For EFA, the sample was split and n=115.

Table B: Factor loadings on the first component.

| | Component 1 |
|---|---|
| Item 1 | 0.860 |
| Item 2 | 0.856 |
| Item 3 | 0.837 |
| Item 4 | 0.836 |
| Item 5 | 0.835 |
| Item 6 | 0.799 |
| Item 7 | 0.797 |
| Item 8 | 0.791 |
| Item 9 | 0.739 |

Notes: For EFA, the sample was split and n=115.

A confirmatory factor analysis was performed with the one-factor model to analyse its fit indices. Indices of the one-factor model are presented in table C.

Table C: Fit indices of the one-factor model of the 9-item German ORIC.

| | Chi² [1] | df [2] | Chi²/df [3] | CFI [4] | TLI [5] | RMSEA [6] | AIC [7] | PNFI [8] |
|---|---|---|---|---|---|---|---|---|
| One-factor model | 56.04 | 27 | 2.08 | .945 | .927 | .097 | 92.04 | .676 |

Notes: One-factor model (including a split data set of n=115): includes items 1 to 9.
[1] discrepancy chi-squared statistic, [2] degrees of freedom, [3] normed chi-squared statistic, [4] comparative fit indexes, [5] Tucker-Lewis Index, [6] root mean square error of approximation, [7] Akaike Information Criterion, [8] Parsimonious Normed Fit Index (PNFI). * $p$ = .000

**Descriptive analysis of items of the 9-item ORIC version and reliability analysis**

Table D shows the corrected item-total correlation. Corrected item-total correlation ranged from .665 (item 9) to .744 (item 3).

Table D: Discrimination of the nine items of the German ORIC.

| | Item discrimination (corrected item-total correlation) |
|---|---|
| Item 1 | .747 |
| Item 2 | .696 |
| Item 3 | .777 |
| Item 4 | .760 |
| Item 5 | .762 |
| Item 6 | .759 |
| Item 7 | .720 |
| Item 8 | .691 |
| Item 9 | .638 |

Notes: Items could be answered on a 5-step Likert scale rating from 0 „ disagree“ to 4 „agree“.

Cronbach’s α with α = .924 showed the measure to reach excellent reliability.

**Supplementary File 2: Additional tables for psychometric evaluation of 10-item ORIC version**

Table A: Demographic data of participants of cognitive interviews (n=11).

| | | Frequencies for n=11 |
|---|---|---|
| Age | < 30 years | 1 |
| | 31-40 years | 3 |
| | 41-50 years | 3 |
| | > 50 years | 4 |
| Gender | Female | 9 |
| | Male | 2 |
| Profession* | Nurse | 8 |
| | Physician | 2 |
| | Psychooncologist | 3 |
| Work experience in health care | > 5 years | 3 |
| | 5-10 years | 4 |
| | 11-20 years | 3 |
| | > 20 years | 1 |

* multiple answers possible

Table B: Factor loadings on the first component.

| | Component 1 |
|---|---|
| Item 1 | 0.790 |
| Item 2 | 0.788 |
| Item 3 | 0.833 |
| Item 4 | 0.846 |
| Item 5 | 0.831 |
| Item 6 | 0.851 |
| Item 7 | 0.839 |
| Item 8 | 0.800 |
| Item 9 | 0.754 |
| Item 10 | 0.860 |

For EFA, the sample was split and n=115.

Table C: Inter-item correlation matrix for the German ORIC.

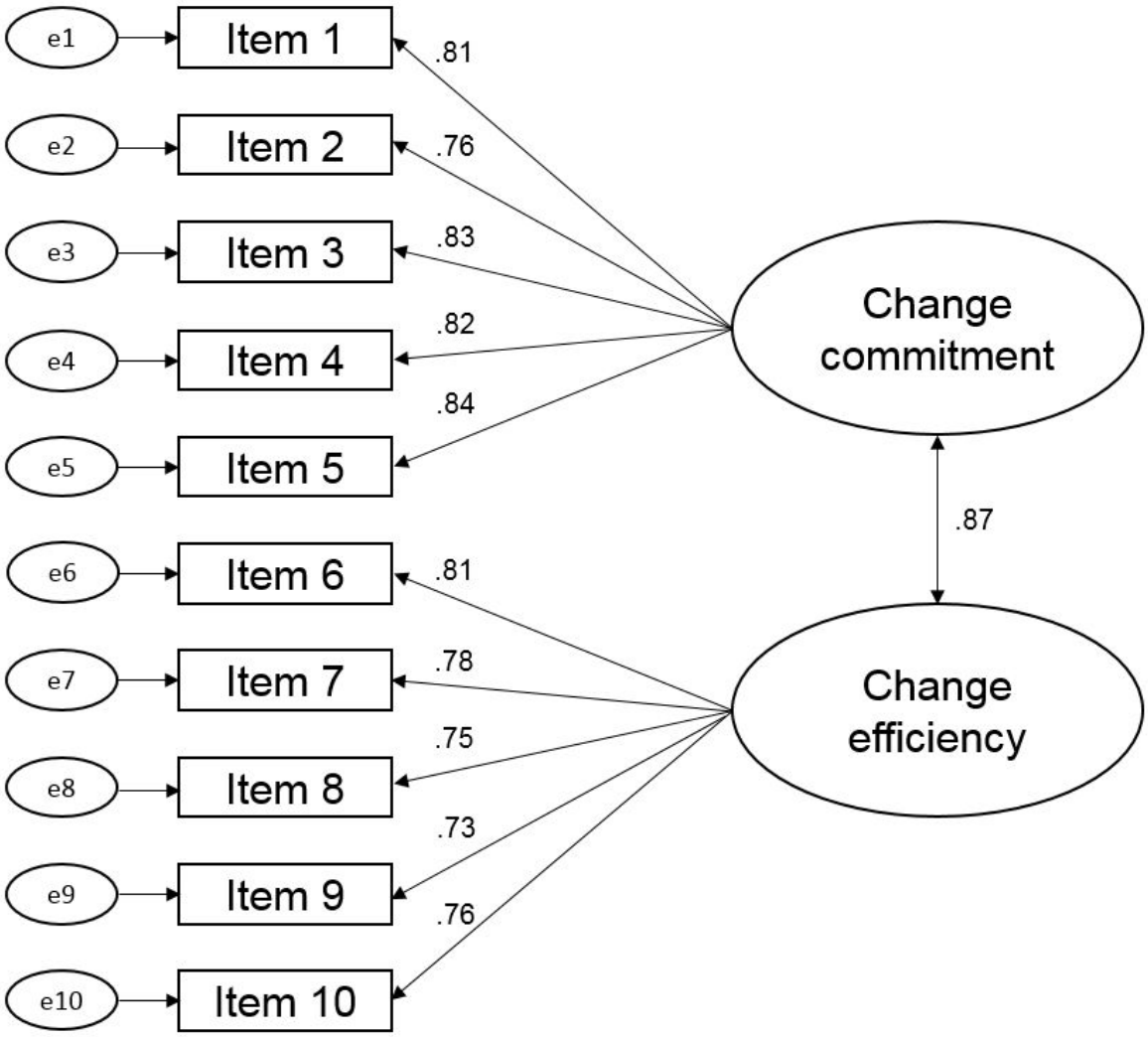
| | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 | Item 6 | Item 7 | Item 8 | Item 9 | Item 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Item 1 | 1.000 | .634 | .650 | .611 | .652 | .587 | .558 | .538 | .507 | .542 |
| Item 2 | .634 | 1.000 | .586 | .681 | .607 | .518 | .504 | .485 | .434 | .478 |
| Item 3 | .650 | .586 | 1.000 | .673 | .723 | .636 | .600 | .558 | .492 | .560 |
| Item 4 | .611 | .681 | .673 | 1.000 | .675 | .588 | .531 | .541 | .512 | .551 |
| Item 5 | .652 | .607 | .723 | .675 | 1.000 | .622 | .621 | .466 | .474 | .578 |
| Item 6 | .587 | .518 | .636 | .588 | .622 | 1.000 | .677 | .590 | .588 | .573 |
| Item 7 | .558 | .504 | .600 | .531 | .621 | .677 | 1.000 | .612 | .491 | .567 |
| Item 8 | .538 | .485 | .558 | .541 | .466 | .590 | .612 | 1.000 | .611 | .560 |
| Item 9 | .507 | .434 | .492 | .512 | .474 | .588 | .491 | .611 | 1.000 | .659 |
| Item 10 | .542 | .478 | .560 | .551 | .578 | .573 | .567 | .560 | .659 | 1.000 |

**Supplementary File 3: Organizational Readiness for Implementing Change (ORIC) – German version**

Bitte geben Sie an, wie sehr Sie den folgenden Aussagen zur Umsetzung von partizipativer Entscheidungsfindung an Ihrem aktuellen Arbeitsplatz zustimmen. Falls partizipative Entscheidungsfindung aktuell nicht umgesetzt wird, wie wäre es im Falle der Umsetzung?

| | | stimme nicht zu | stimme eher nicht zu | teils teils | stimme eher zu | stimme zu |
|---|---|---|---|---|---|---|
| 1 | Personen, die hier arbeiten, zeigen hohes Engagement bei der Umsetzung von partizipativer Entscheidungsfindung. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2 | Personen, die hier arbeiten, werden tun, was auch immer nötig ist, um partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3 | Personen, die hier arbeiten, wollen partizipative Entscheidungsfindung umsetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4 | Personen, die hier arbeiten, sind fest entschlossen, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 5 | Personen, die hier arbeiten, sind motiviert, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 6 | Eventuell entstehen bei der Umsetzung von partizipativer Entscheidungsfindung Herausforderungen. Personen, die hier arbeiten, sind zuversichtlich, diese zu meistern. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 7 | Personen, die hier arbeiten, sind zuversichtlich, dass sie den Verlauf der Umsetzung von partizipativer Entscheidungsfindung überblicken können. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 8 | Personen, die hier arbeiten, sind zuversichtlich, dass sie Aufgaben so koordinieren können, dass die Umsetzung reibungslos abläuft. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 9 | Personen, die hier arbeiten, sind zuversichtlich, dass die Klinik sie dabei unterstützten kann, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 10 | Personen, die hier arbeiten, sind zuversichtlich, Machenschaften bei der Umsetzung von partizipativer Entscheidungsfindung bewältigen zu können. | ☐ | ☐ | ☐ | ☐ | ☐ |

**Supplementary File 4: Confirmatory factor analysis model of the 10-item German ORIC**



**Supplementary File 5: Checklist for reporting standards. Authors guideline for Scale Development and validation results by Cabrera-Nguyen [1].**

| | Guidelines by Cabrera-Nguyen [1] | Transfer to our study |
|---|---|---|
| 1 | Precisely define the target construct. | See page 3 and 4 (Introduction section) |
| 2 | Justify the need for your new measure. For example, if measures of the construct exist in the literature, explain the value added by your new scale. How might the new measure enhance the substantive knowledge base or social work practice? | See page 3 and 4 (Introduction section) |
| 3 | Indicate that you have submitted your initial pool of items to expert review (Worthington & Whittaker, 2006). Report (a) the number of items in the preliminary pool; (b) the number of expert reviewers and their qualifications; and (c) any major changes to your initial item pool following the review (e.g., a substantial decrease in the number of items, changes to the original item response format, overhaul of item pool due to experts' assessment regarding content validity). | See page 4 and 5 (Methods section: Measure, Translation, Assessment of comprehensibility and adaption of the scale) and page 8 and 9 (Results section: Translation, Assessment of comprehensibility and adaption of the scale) |
| 4 | Report the name and version of the statistical software package used for all analyses. | See page 5 to 7 (Methods section: Assessment of comprehensibility and adaption of the scale, Psychometric evaluation) |
| 5 | Identify and justify the sampling strategy (e.g., convenience, snowball) and sampling frame. Report standard sample demographic characteristics as well as other salient sample characteristics. | See page 5 and 6 (Methods section: Assessment of comprehensibility and adaption of the scale, Psychometric evaluation) and page 8 to 10 (Results section: Assessment of comprehensibility and adaption of the scale, Psychometric evaluation) and Additional file 3 |
| 6 | Discuss relevant data preparation and screening procedures. For instance, do the data meet the appropriate assumptions for factor analysis? If not, what actions were taken? Report tests of factorability if appropriate (e.g., report Bartlett's test of sphericity). | See page 9 and 10 (Results section: Psychometric evaluation) |
| 7 | Provide all dates of data collection. | See page 8 to 12 (Results section) |
| 8 | Avoid use of principal components analysis (PCA) as a precursor to CFA (Costello & Osborne, 2005; Worthington & Whittaker, 2006). Instead, start with EFA to assess the underlying factor structure and refine the item pool. EFA should be followed by CFA using a different sample (or samples) to evaluate the EFA-informed a priori theory about the measure's factor-structure and psychometric properties. (Costello & Osborne, 2005; Henson & Roberts, 2006; Worthington & Whittaker, 2006). For CFA, authors | An a priory hypothesized model for CFA was specified based on the original measures structure. The model could not be confirmed and an EFA was calculated afterwards. The same data set was used for EFA and CFA. For more |

| | | |
|---|---|---|
| | should specify an a priori hypothesized model and a priori competing models (Jackson, Gillaspy, & Purc-Stephenson, 2009). | information see page 6 and 7 (Results section). |
| 9 | Guidelines for reporting EFA results. | |
| | How large is a sample? One common rule of thumb is to ensure a person-to-item ratio of 10:1. Another rule of thumb is that N= 300 is usually acceptable (Worthington & Whittaker, 2006). However, some researchers have criticized these sample size rules of thumb, noting the appropriate sample size is dependent on the features of the gathered data. These researchers recommend obtaining the largest possible sample because the adequacy of the sample size cannot be determined until after the data have been analyzed (Henson & Roberts, 2006). | Sample size is n=115, person-to-item ratio 11.5:1 (see page 6: Methods section, Psychometric analysis) |
| | Run EFA . . . or not. Run a preliminary EFA to determine if further data collection is required based on the following criteria: (a) If communalities are greater than .50 or there are 10:1 items per factor with factor loadings of roughly \|.4\|, then a sample size of 150 to 200 is likely to be adequate; (b) If communalities are all at least .60 or there are a minimum of 4:1 items per factor with factor loadings above \|.6\|, then even smaller sample sizes may suffice; (Worthington & Whittaker, 2006). Report if additional data collection was necessary due to inadequate sample size. If so, report the new participants' sociodemographic characteristics and test for differences between groups using standard statistical procedures (e.g., ttests). | EFA communalities are all above .70, therefore the sample size of n=115 can be determined as adequate and no additional data collection was necessary. For more information, see Additional file 3). |
| | Give EFA details. Report the specific rotation strategy used (e.g. varimax, geomin). Justify the decision to use an orthogonal or oblique solution. One recommendation is to always begin with an oblique rotation, empirically assess factor intercorrelations, and report them before deciding upon a final rotation solution (Henson & Roberts, 2006; Worthington &Whittaker, 2006). Some researchers argue oblique rotation is always the best approach because (a) factor intercorrelations are the norm in social sciences and (b) both approaches yield the same result if the factors happen to be uncorrelated (Costello & Osborne, 2005). Conversely, other researchers contend that orthogonal rotation is preferable because fewer parameters are estimated—orthogonal rotation is more parsimonious and amenable to replication (Henson &Roberts, 2006). Similarly, some researchers warn against relying on a statistical software package's default settings to determine the appropriate type of oblique rotation (Henson & Roberts, 2006; Worthington &Whittaker, 2006). Others state that doing so is —finell (Costello & Osborne, 2005, p.3). Given the lack of consensus, it is probably best to describe what you do and defend your approach on substantive grounds, if possible. | See page 6 and 7 (Methods section: Psychometric evaluation) |
| | Report the whole factor pattern/structure. Always report the whole factor pattern/structure matrix, including all of the items in the analysis. It is recommended that authors report | See page 10 (Results section: Psychometric evaluation) and Additional file 3 |

| | | |
|---|---|---|
| | this information in a chart following the example provided by Henson and Roberts (2006) on page 411. | |
| | Criteria for deleting (crossloaded) items. Report any deleted items and the criteria used for deletion. Crossloading items with values ≥ .32 on at least two factors should generally be candidates for deletion, especially if there are other items with factor loadings of .50 or greater (Costello & Osborne,2005). Rerun the EFA each time an item is deleted. | No crossloading items could be observed, so item deletion was not necessary. |
| | Criteria for number of factors. Report the number of factors retained and justify this decision using multiple criteria (eigenvalue > 1, scree test, parallel analysis, rejection of a factor with fewer than 3 items, etc). Reporting the eigenvalue > 1 rule alone is inadequate because it has been shown to among the least accurate criteria for assessing factor retention (Costello & Osborne, 2005; Henson & Roberts, 2006) | See page 6 and 7 (Methods section: Psychometric evaluation) |
| | Explained variance. Report the variance explained by the factors. | See page 10 (Results section: Psychometric evaluation) and Additional file 3 |
| | In general, describe your decisions. | See page 10 (Results section: Psychometric evaluation) and Additional file 3 |
| 10 | Guidelines for reporting CFA results. | |
| | Describe and justify the theoretical model. Report hypothesized factor structure. Provide theoretical and empirical justification (e.g., results of preliminary EFAs) for your hypothesis. In addition, report a priori competing models. | See page 6 and 7 (Methods section: Psychometric evaluation) |
| | Describe the parameterization. Provide a comprehensive description of the a priori parameter specification. Identify fixed parameters, free parameters, and constrained parameters. For example, indicate if you freed the errors of any items to correlate. | One factor loading was constrained to equal 1, the corresponding intercept was constrained equal to zero. The other factor loadings and intercepts were estimated. Errors of items were not freed to correlate. |
| | Include a figure. Include a figure of each CFA model being tested using Kline's (2005) graphical conventions if feasible. | See Additional file 5 |
| | Identification. Demonstrate model identification (e.g., df > 0; scaling of factors; assess and report the —t-rule; the two-indicator rule). Necessary and sufficient conditions for model identification may vary for certain types of CFA models. When in doubt, authors should consult Brown's (2006) CFA textor Kline's (2005) SEM text for guidance. | See page 10 and 11 (Results section: Psychometric evaluation) |
| | Select an estimator based on distributional patterns and assumptions. Report the estimator used (e.g., ML, WLSMV) and justify your choice based on distributional | See page 6 (Methods section: Psychometric evaluation) |

| | | |
|---|---|---|
| | assumptions. It is not appropriate to report that you relied on your statistical software's default setting. | |
| | Use multiple fit indices. After estimating a model, always report multiple fit indices (e.g., model X2, df, p, CFI/TLI, RMSEA, SRMR). Report all appropriate fit indices, not just those favorable to your hypotheses (Jackson et al., 2009). For example, do not report acceptable CFI and TLI scores while omitting a relevant fit index with a subobtimal value. | See page 6 and 7 (Methods section: Psychometric evaluation) and page 10 and 11 (Results section: Psychometric evaluation) |
| | What is acceptable fit? For model fit indices, authors should generally use the cut-off values recommended by Hu and Bentler (1999) and endorsed by Brown (2006), assuming ML estimation: a.CFI/TLI ≥ .95 b. RMSEA ≤ .06 c. SRMR ≤ .08 | See page 8 (Methods section: Psychometric evaluation) |
| | Localized strain? When reporting model fit, include an assessment for localized areas of strain by examining standardized residuals. Standardized residuals greater than 1.96 (for p< .05) indicate areas of strain (Harrington, 2009). Report the absence of localized strain, if appropriate; otherwise, note localized areas of strain by reporting the relevant standardized residuals. | Standardized residuals do not indicate localized strains. |
| | Parameter estimates and SEs. When reporting factor loadings and other parameter estimates, always report the unstandardized estimates, their p values, and the standard errors. In addition, include the standardized estimates when appropriate. Be sure to report all parameter estimates, even those that are non significant (Brown, 2006; Jackson et al., 2009). | See Additional file 5 |
| | Assessing the validity of the factor solution. Comment on the new measure's convergent and discriminant validity based on parameter estimates. For instance, factor correlations ≥ .80 may indicate poor discriminant validity (Brown, 2006). In addition, strong factor loadings that do not crossload may indicate good convergent validity. One rule of thumb is that factor loadings < .40 are weak and factor loadings ≥ .60 are strong (Garson, 2010). | Exploratory factor analysis revealed a one-factor model, so there are no factor correlations. |
| | Other measures. Report squared multiple correlations and comment on the measure's reliability (e.g., report Raykov's Rho if appropriate) | See page 11 (Results section: Psychometric analysis) |
| | Respecification: Caution! Report any post-hoc respecifications to improve model fit based on modification indices. Justify the respecifications on theoretical or conceptual grounds (Jackson et al., 2009). Respecification to allow for correlated errors is not supportable without strong pragmatic justification(e.g., items contain similar words or phrases). Note that respecification precludes comparing the model with your a priori specified competing models. Report improvements in appropriate model fit indices for respecified models (e.g., chi-square difference test) | No respecifications to improve model fit were applied. |
| 10 | Describe the matrix (or matrices) you analysed (e.g., covariance, correlation). Include matrices in the manuscript | See Additional file 3 |

| | | |
|---|---|---|
| | if feasible; otherwise, indicate these data are available upon request. | |
| 11 | Report the amount of missing data and describe how missing data were handled. For a review of practices for handling missing data, see Sterne and colleagues (2009), Rose and Fraser (2008), and Horton and Kleinman (2007).Provide a rationale for your approach to handling missing data. Authors are encouraged to consider using multiple imputation or model estimation with full-information maximum likelihood (FIML; Rose & Fraser, 2008). | See page 6 (Methods section: Psychometric evaluation) and page 11 (Results section: Psychometric evaluation) |
| 12 | Compare your CFA model with the alternative or competing models. Do competing models fit the data better or worse than your model (e.g., does your four-factor model of acculturation fit the data better than a two-factor model or a one-factor model)? Identify the preferable model based on appropriate fit statistics (e.g. chi-square difference test for nested models, Akaike information criterion for non-nested models), parsimony, and relevant theory | See page 13 (Discussions sections) |
| 13 | Include your scale (items and response options) in an appendix. | See Additional file 2 |
| 14 | Report how methodological limitations may have impacted findings regarding your measure's psychometric properties (e.g., note potential repercussions of suboptimal sampling techniques, discuss implications of using listwise deletion to handle missing data instead of multiple imputation or FIML). | See page 14 (Discussions sections) |
| 15 | Discuss directions for future research (e.g., if appropriate, testing your scale for measurement invariance by conducting CFA on different populations). | See page 13 and 14 (Discussions sections) |

1 Cabrera-Nguyen P. Author Guidelines for Reporting Scale Development and Validation Results. *J Soc Social Work Res*. University of Chicago PressChicago, IL 2010;1:99–103.

BMJ Open

# Translation and psychometric evaluation of the German version of the Organizational Readiness for Implementing Change measure (ORIC) – a cross-sectional study

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2019-034380.R1 |
| Article Type: | Original research |
| Date Submitted by the Author: | 26-Feb-2020 |
| Complete List of Authors: | Lindig, Anja; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Hahlweg, Pola; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Christalle, Eva; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Scholl, Isabelle; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology |
| Primary Subject Heading: | Research methods |
| Secondary Subject Heading: | Communication, Evidence based practice, Patient-centred medicine |
| Keywords: | Organizational readiness for change, Psychometrics, Translation, Implementation, Shared decision-making, Measurement |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Translation and psychometric evaluation of the German version of the Organizational Readiness for Implementing Change measure (ORIC) – a cross-sectional study

Anja Lindig[1], Pola Hahlweg[1], Eva Christalle[1], Isabelle Scholl[1]

[1] *Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany*

Email addresses:
Anja Lindig: a.lindig@uke.de
Pola Hahlweg: p.hahlweg@uke.de
Eva Christalle: e.christalle@uke.de
Isabelle Scholl: i.scholl@uke.de

Corresponding author:
Anja Lindig
University Medical Center Hamburg-Eppendorf
Department of Medical Psychology
Martinistr. 52
20246 Hamburg, Germany
Email: a.lindig@uke.de
Phone: +49-(0)40-7410-56206

Word count: 4186

## ABSTRACT

**Objectives:** To translate the measure Organizational Readiness for Implementing Change (ORIC) into German and assess its psychometric properties.

**Design:** Cross-sectional psychometric study based on secondary analysis of baseline data from a shared decision-making (SDM) implementation study.

**Setting:** Three departments within one academic cancer center in Hamburg, Germany.

**Participants:** For comprehensibility assessment of the translated ORIC version, we conducted cognitive interviews with n=11 healthcare professionals (HCPs). Afterwards, n=230 HCPs filled out the measure.

**Primary and Secondary Outcome Measures:** The original English version of the ORIC was translated into German using a team translation protocol. Based on comprehensibility assessment via cognitive interviews with HCPs, the translated version was revised. We analyzed acceptance (completion rate), factorial structure (exploratory factor analysis (EFA), confirmatory factor analysis (CFA), model fit), item characteristics (item difficulties, corrected item-total correlations, inter-item correlations), and internal consistency (Cronbach's α).

**Results:** Translation and cognitive testing of the ORIC was successful except for item 10, which showed low comprehensibility as part of content validity in cognitive interviews. Completion rate was > 97%. EFA and CFA provided a one-factorial structure. Item difficulties ranged between 55.98 and 65.32, corrected item-total-correlation ranged between .665 and .774, inter-item correlations ranged between .434 and .723, and Cronbach's α was .93.

**Conclusions:** The German ORIC is a reliable measure with high completion rates and satisfying psychometric properties. A one-factorial structure of the German ORIC was confirmed. Item 10 showed limited comprehensibility and therefore reduces content validity of the measure. The German ORIC can be used to analyze organizational readiness for change as a precursor for implementation success of various interventions.



**Keywords:**

Organizational readiness for change, Psychometrics, Translation, Implementation, Shared decision-making, Measurement

## ARTICLE SUMMARY

### Strengths and limitations of this study

- Appropriate qualitative methods were used to provide a German version of the ORIC to be assessed for healthcare professionals (HCPs).
- The sample size was large enough to robustly perform the psychometric analyses on the German version of the ORIC measure.
- Due to the design of the study as a secondary analysis, it was not possible to calculate psychometric parameters like convergent and divergent validity.
- Data were collected at a single academic cancer center in the context of a shared decision-making implementation study. Thus, psychometric properties of the German ORIC need replication in other settings.

## INTRODUCTION

Implementing interventions in healthcare systems is an important and widely discussed topic [1–3] and often mediated by public policies, market forces, or new technologies [4]. The intention to implement new interventions might be to reduce costs, improve quality, increase efficacy or patient satisfaction [5]. Nevertheless, implementing change in healthcare organizations can be challenging [6–10]. In the German healthcare system, the implementation of shared decision-making (SDM) has received much attention [11]. SDM can be described as an interactional process on the basis of information exchange. Patients and healthcare professionals (HCPs) are equally and actively involved and jointly responsible for the decision [12–14]. This is especially important in situations with complex treatment options and high impact on patients' quality of life [15]. Patients want to be actively involved in decision-making [16] and benefit from SDM by developing better knowledge about their disease and treatment options, better risk perception, and less insecurity and decisional conflict [17,18]. SDM has been supported by health policy [19–21] and research [22–25] over the last decades. However, SDM is currently not routinely implemented in the German healthcare system [11,26–28].

When implementing SDM or other interventions in organizations, several barriers on different levels of the organization (i.e., individual members of the organization, teams, organizational setting, system level) need consideration [1,5–7,10,29–32]. Barriers for implementing SDM in the clinical setting often address both the organizational setting (e.g. lack of resources and lack of management support) and the individual level

(e.g. resistance to change or negative attitudes towards SDM) [7,10,11,33–35]. When implementing SDM or other interventions in healthcare systems, the clinical employees' perspective on organizational readiness for change is a critical precursor for successful implementation [5,32,36–38]. Armenakis et al. [39] describe organizational readiness for change as the degree to which organizational members are prepared to participate in change processes. This is characterized by the belief that the change is needed and that the organization is capable of changing. In their theory of organizational readiness for change, Weiner et al. [40] differentiate between change commitment (i.e., organizational members' attitudes towards implementing a change) and change efficacy (i.e., organizational members' belief in their capability to implement a change). If readiness for change is high, organizational members invest more in the change effort and exhibit greater persistence to overcome barriers and setbacks [40,41].

To analyze effects of organizational readiness on implementation success, specific measures for assessing organizational readiness for change are needed [5]. However, only a few validated measures exist [5,42–44] and none were available in German. One of those measures is the Organizational Readiness for Implementing Change (ORIC) [43]. The ORIC is brief, easy to administer, and theoretically and psychometrically well-grounded [40]. It was previously translated into Danish and French [45,46]. The ORIC has been psychometrically tested, revealing a completion rate of > 72%, a Cronbach's α of > .80 and two correlating factors [43,45,46]. Due to the described properties, the ORIC seemed well-suited to measure organizational readiness for implementing SDM in Germany [5,43,47].

Therefore, the aim of the study was to translate the measure ORIC into German and assess its psychometric properties.

## METHODS

### Measure

The ORIC measures organizational readiness for implementing change. It uses a 5-point Likert scale ranging from „disagree" to „agree" [43]. In the original English version, two subscales were described based on Weiner et al. [40]: „change commitment" (items 1 to 5) and "change efficacy" (items 6 to 10). Sum scores were calculated for both subscales separately with higher scores indicating higher organizational readiness for change. By using the phrases "to implement this change" or "implementing this change", the original scale does not specify which change is addressed. The items can be specified to adapt to a specific

research question and a survey instruction can be added [40,45]. The English items are displayed in the results section.

**Translation**

Translation followed the team translation protocol TRAPD (Translation, Review, Adjudication, Pretesting and Documentation), a method with growing recognition within translation research [48]. Thereby an optimal translation is facilitated by discussions between members of the translation team with different expertise in translation. First, two team members (AL, SZ, cp. list of abbreviations) proficient in German and English, but little experienced in survey translation, independently translated the original ORIC into German. Second, a third bilingual team member (IS) with experience in survey translation, reviewed both versions and suggested a third version based on the first two translations. Finally, IS, AL, and SZ discussed all versions until reaching consensus on a final version. To find consensus on item 10 we additionally consulted an official translator (MM, cp. list of abbreviations) and an additional team member (PH), who is proficient in German and English and experienced in translation. During the translation process, we changed the phrases “to implement this change” and “implementing this change” into “to implement shared decision-making” and “implementing shared decision-making” to address our specific research question. Additionally, we added a survey instruction in German which motivated participants to think about the clinic, they are working in, when answering the item.

**Assessment of comprehensibility as part of content validity and subsequent adaptation of the scale**

Content validity is the degree to which the content of the measure and its items adequately reflect the measured construct [49]. According to the COSMIN criteria (Consensus-based standards for the selection of health measurement instruments) [50], content validity includes the relevance of the items and scales, their comprehensiveness, and comprehensibility. As this study aimed to evaluate the translation of an existing measure, we focused on the assessment of comprehensibility (i.e., items being appropriately worded and understood by participants as intended).

To do so, we conducted two rounds of cognitive interviews with a convenience sample of HCPs (nurses, physicians, and psychooncologists), working in a comprehensive cancer center in Germany. Two female researchers and psychologists experienced in interviewing (AL, PH) conducted the interviews. We developed an interview guide based on recommendations by Willis et al. [51]. We used verbal probing

techniques like comprehension probes (e.g. “What does the term ‘organization’ mean to you?”) and paraphrasing (e.g. “Can you repeat this sentence in your own words?”). We conducted interviews until reaching theoretical saturation. Interviews were audio-recorded and transcribed verbatim. After the first round of cognitive interviews, we extracted and discussed comments and suggestions from the transcripts (AL, PH, IS). As a further step to enhance comprehensibility, we discussed the items with the original author (CS, cp. list of abbreviations) as well as French (MR, cp. list of abbreviations) and Norwegian researchers (AH, cp. list of abbreviations), who translated the ORIC into their languages. We adapted items of the German ORIC, which were not well understood by participants of the first round of cognitive interviews, according to these discussions. We tested these items in a second round of cognitive interviews. After the second round, we discussed further adaptions of the items and involved another bilingual researcher in the field (DF, cp. list of abbreviations).

We calculated descriptive statistics of participants’ demographic characteristics using SPSS (IBM SPSS Statistics, Version 23).

**Psychometric evaluation**

**Data collection.** For psychometric evaluation of the ORIC measure we conducted a secondary analysis of cross-sectional data gathered in a SDM implementation study [52]. Data from baseline assessment of the SDM implementation study were included. The ORIC was the last questionnaire of a three-page survey measuring HCPs’ attitudes regarding SDM and its implementation. Besides the ORIC, it contained the Control Preference Scale [53] and the IcanSDM [54], demographic questions (e.g. gender, age, profession, work experience) as well as several questions that have been used in previous studies in cancer care [55]. Results of these additional measures will be published as part of the primary evaluation of the SDM implementation study [52].

Participants were part of a convenience sample of physicians and nurses. Since this is a secondary analysis, inclusion criteria were identical to inclusion criteria of the SDM implementation study [52]. We included physicians and nurses who worked at one of three departments within the University Cancer Center Hamburg at the University Medical Center Hamburg-Eppendorf during baseline evaluation of the SDM implementation study [52]. Eligible HPCs were identified through employee lists provided by department managers. The measure was handed out to eligible HCPs either (1) by a member of our study

team (e.g. during a regular physician meeting), (2) by the supervising nurses, or (3) via employees' mailboxes. Participants returned the questionnaire personally to a study team member or by mail.

Data were entered into SPSS (IBM SPSS Statistics, Version 23) including blinded double entry of 20% of the data for quality control.

**Patient and public involvement.** The measure ORIC preliminary addresses HCPs, therefore physicians, nurses and psychooncologists were involved in the adaptation of the measure by taking part in cognitive interviews. Patients were not involved in this study.

**Data Analyses**. Descriptive statistics were calculated for demographic characteristics. Cases were excluded if more than 30% of the ORIC items were missing [56]. For all other cases, missing data were replaced with item means. We evaluated the completion rate and therefore the acceptance of the measure by calculating frequencies of missing data per item as well as for the overall measure. For this analysis, we also included cases with more than 30% of ORIC items missing because these values are part of completion rate and relevant for interpretation of acceptance.

We a priori hypothesized to replicate the theory-based two-dimensional structure of the original English ORIC version. Two correlating factors "change commitment" (item 1 to 5) and "change efficacy" (item 6 to 10) were postulated. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett`s test of sphericity were performed to test prerequisites for factor analysis [57,58]. A confirmatory factor analysis (CFA) with Maximum Likelihood Estimates and two factors was applied for the whole data set as a first step. Because the two-factor model could not be confirmed, we decided to calculate an exploratory factor analysis (EFA) and afterwards an additional CFA to check for model fit. It is recommended to not calculate EFA and CFA with the same data set so the data set was randomized by AL and split into two subsets [59,60]. The first 115 randomized cases including all data of participants were added to EFA, the second 115 cases were added to CFA. An EFA with oblique rotation was calculated for the first subset. The non-orthogonal rotation was chosen according to Weiner et al. [40]. In their theory, organizational readiness for change consists of two interrelated dimensions, therefore the two factors are expected to be correlated. Analogue to analyses done by authors of the English ORIC [43,61], we extracted components based on parallel analysis. The criterion of parallel analysis was shown to be superior to other statistic criteria like the Kaiser criterion [62,63]. It compares the eigenvalues of the data to the eigenvalues based on random data with equivalent sample size and number of variables and chooses only factors with eigenvalues higher than for random data [61,62]. A CFA was calculated for the second subset. A range of global goodness of

fit indices were used to assess the degree to which observed data were accounted for by the proposed models: discrepancy chi-squared statistic ($Chi^2$), degree of freedom (df), normed chi-squared statistic ($Chi^2/df$), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA) as well as Akaike Information Criterion (AIC) and Parsimonious Normed Fit Index (PNFI) for analysing model complexity. Established rules to estimate the model fit were used [64–67].

Item analyses was performed for the one-factor model. It included calculation of item means and standard deviation as well as observation of floor and ceiling effects [68], calculation of corrected item-total correlations [57,69], inter-item correlations [57,69], and item difficulties [58]. Internal consistency of the scale was assessed by Cronbach's alpha coefficient (α) [58,69,70]. For a detailed overview on performed data analyses, see Table 1.

During the translation process and cognitive interviews we found low content validity for item 10 (see results section), therefore the use of item 10 for the German ORIC needs to be evaluated. Accordingly, we also conducted psychometric analyses (EFA with oblique rotation and extraction of components based on parallel analysis, corrected item-total correlations, Cronbach's α, and goodness of fit indexes) for the 9-item version of the ORIC.

Analysis of demographic data, analysis of completion rate, item analysis and EFA were performed using SPSS (IBM SPSS Statistics, Version 23). CFA and calculation of model fit indices were performed using Amos (IBM SPSS Amos 22.0.0).

Table 1: Psychometric analyses conducted.

| **Psychometric measure** | **Criteria** |
|---|---|
| Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett`s test of sphericity | These tests ensure that correlations between variables can be accounted for by a smaller set of factors [57]. KMO value should be higher than .05 and Bartlett`s test value should be less than .05 to fulfil the criteria for calculating a factor analysis [57,58]. |
| Normed chi-squared statistic ($Chi^2/df$) | $Chi^2/df$ is an indicator for model fit, dependent on sample size and should be as small as possible. A ratio between 2 and 3 indicate a good data fit [64]. |
| Comparative fit indexes (CFI) | CFIs is an indicator for model fit. It ranges from 0 to 1 and higher values indicate better fit. Values above .95 indicate a good model fit [67,71]. |
| Tucker-Lewis Index (TLI) | TLI is an indicator for model fit. It corrects for complexity of the model and is sensitive to small sample sizes. Values above .95 indicate good fit [65]. |

| | |
|---|---|
| Root mean square error of approximation (RMSEA) | RMSEA is an absolute index which describes closeness to fit. Values below .05 indicate a good fit, values between .05 and .08 indicate an adequate fit, values between .08 and 1 indicate a moderate fit and values above 1 are unacceptable [72]. |
| Akaike Information Criterion (AIC) | AIC is a parsimony model fit index. It can be used to compare fit of competing models with smaller values indicating better fit [64,67]. |
| Parsimonious Normed Fit Index (PNFI) | PNFI is a parsimony model fit index. It ranges between 0 and 1 and higher values indicate a more parsimonious fit [64]. No threshold levels are recommended and it has to be analysed in combination with other goodness of fit indices [67]. |
| Analysis of frequencies for item response distributions | Floor and ceiling effects were assumed present if more than 15% of participants choose the lowest or highest possible score [68]. |
| Corrected item-total correlations | If items correlate with the total score of above .30, they measure the same underlying concept. Items with lower correlations should be removed because they do not add exploratory power to the measure [57,69]. |
| Item difficulties | Item difficulties are calculated by dividing item means by the maximal value of the answer range (0-4) and multiplying it with 100. Item difficulty should be near to 50%, and items should not differ much in their difficulty level [58]. |
| Inter-item correlations | Inter-item correlations ensure association between items. High inter-item correlations of above .80 indicate that items ask the same questions and might be redundant [57,69]. |
| Cronbach's α | Cronbach's α is a measure for reliability and internal consistency. A value of at least .70 is acceptable and higher coefficients indicate a more stable measure [57,58,70]. |

## RESULTS

### Translation

Both translators (AL and SZ) and the reviewer (IS) did not differ much in their translations of items 2 to 5 and items 8. Greater translation differences were found for items 1, 6, 7, 9, and 10, mainly related to differences in sentence structure or single words. Within the first round of team discussion, we reached consensus for items 2 to 9, the translation of the response scale and the survey introduction. For item 1 we suggested two versions to be further tested in subsequent cognitive interviews. We struggled to translate the phrase "manage the politics" in item 10 into German. Therefore, we discussed item 10 with additional colleagues (cp. methods section) until consensus was found.

### Assessment of comprehensibility as part of content validity and subsequent adaptation of the scale

To test the German ORIC for comprehensibility, cognitive interviews with n=11 participants (nurses, physicians, and psychooncologists) were conducted. Cognitive interviews lasted about one hour. For demographic data of participants see Table A in Supplementary File 1.

After the first round of cognitive interviews (n=7), no changes have to be made to the response scale as well as for item 2 to 5 and 8 because these items were already well understood by participants. Participants made some minor suggestions for modifications for the introductory description and items 1 and 6. Additionally, some participants did not understand the correct meaning of items 7, 9 and 10 in general or of single words or phrases of these items. After discussions and modifications of these items, we tested alternative versions of the survey introduction, for items 1, 6, 9 and 10 as well as two alternative versions of item 7. After the second round of cognitive interviews (n=4), items 1, 6 and 9 were now understood well by all participants. We had to slightly modify the survey introduction again and decided to use the version of item 7 which was understood best. After all, item 10 could not be translated successfully. Both rounds of cognitive interviews showed that comprehension of the German translation of the phrase “manage the politics” did not picture the correct English meaning. So n a next step we consulted with DF (cp. list of abbreviations) and reached consensus on a final version. Nevertheless, the final version of item 10 was still not satisfying from the study team and experts view. Item 10 was found to have low comprehensibility as part of content validity according to COSMIN criteria [50]. The final German ORIC measure, used in this study, is presented in Supplementary File 2.

During cognitive interviews some nurses reported that they had not heard about the term “shared decision-making” (German: “Partizipative Entscheidungsfindung”) prior to participation. Thus, we provided a definition of SDM in the introduction part of the questionnaire within the SDM implementation study [52].

**Psychometric evaluation**

**Sample characteristics.** Data of 235 HCPs were available for this secondary analysis. In line with recommendations of Bannon [56], five cases (0.02% of all cases) were excluded (except for assessment of completion rate), because all items of the ORIC were missing. Missing values were replaced by means and data of 230 HCPs could be included into analyses.

Table 2 provides an overview of participants’ demographic characteristics. Most of the 230 HCPs were between 31 and 40 years old (37.0%), female (70.4%), worked as a nurse (57.0%), and had a work experience of < 5 years (43.9%).

Table 2: Demographic characteristics of participants (n=230).

| | | N | % |
|---|---|---|---|
| Age | < 30 years | 72 | 31.3 |
| | 31-40 years | 85 | 37.0 |
| | 41-50 years | 42 | 18.3 |
| | > 50 years | 26 | 11.3 |
| | Missings | 5 | 2.2 |
| Gender | Female | 162 | 70.4 |
| | Male | 59 | 25.7 |
| | Different gender or preferred not to answer this question | 5 | 2.1 |
| | Missings | 4 | 1.7 |
| Profession | Nurse | 131 | 57.0 |
| | Junior physician | 69 | 30.0 |
| | Senior physician | 27 | 11.7 |
| | Missings | 3 | 1.3 |
| Work experience in healthcare | < 5 years | 101 | 43.9 |
| | 5-10 years | 48 | 20.9 |
| | 11-20 years | 46 | 20.0 |
| | > 20 years | 28 | 12.2 |
| | Missings | 7 | 3.0 |

**Factor analysis.** Sample size was large enough (>100) to allow factor analysis [50], even for a split data set with n=115. Furthermore, there were no outliers and values were approximately normally distributed. KMO measure was .933 and Barlett´s test of sphericity yielded $X^2 = 1485.11$, $p < .001$. This indicates that a factor analysis of the data was appropriate [57,58]. CFA for the hypothesised two-factor model showed a high correlation of .87 between the two components (see Supplementary File 3). Therefore, we postulated a one-factorial structure and conducted a post-hoc EFA. As shown in Table 3, only the first component had an eigenvalue higher than 95% percentile of the eigenvalues of corresponding random data and the main component explains 67.23% of the variance. Thus, according to parallel analysis, a one-factor model was assumed. The factor loadings for the first component were above 0.754 for all items (see Table B of Supplementary File 1).

Table 3: Results of EFA with oblique rotation and parallel analysis: eigenvalues of the ten components of the German ORIC and eigenvalues for corresponding random data.

| | Eigenvalues | | | Eigenvalues for random data | |
|---|---|---|---|---|---|
| | Total | % of variance | Cumulative % | Means | 95% percentile |
| Component 1 | 6.72 | 67.23 | 67.23 | 1.49 | 1.65 |
| Component 2 | 0.83 | 8.30 | 75.53 | 1.33 | 1.44 |
| Component 3 | 0.47 | 4.75 | 80.28 | 1.21 | 1.30 |
| Component 4 | 0.41 | 4.08 | 84.36 | 1.11 | 1.19 |

| | | | | | |
|---|---|---|---|---|---|
| Component 5 | 0.39 | 3.91 | 88.28 | 1.02 | 1.08 |
| Component 6 | 0.32 | 3.24 | 91.52 | 0.93 | 1.00 |
| Component 7 | 0.27 | 2.74 | 94.26 | 0.85 | 0.92 |
| Component 8 | 0.23 | 2.34 | 96.60 | 0.77 | 0.84 |
| Component 9 | 0.17 | 1.74 | 98.35 | 0.68 | 0.75 |
| Component 10 | 0.16 | 1.65 | 100.00 | 0.58 | 0.66 |

Notes: For EFA, half of the data set (n=115) was used.

A second CFA was performed with the one-factor model to analyse its fit indices. Indices of the two-factor model and the one-factor model are compared in Table 4.

Table 4: Fit indices of two calculated models for factor analysis of the German ORIC.

| | Chi² [1] | df [2] | Chi²/df [3] | CFI [4] | TLI [5] | RMSEA [6] | AIC [7] | PNFI [8] |
|---|---|---|---|---|---|---|---|---|
| Two-factor model | 81.71* | 34 | 2.40 | .968 | .947 | .078 | 143.71 | .585 |
| One-factor model | 77.19* | 35 | 2.20 | .928 | .907 | .103 | 117.19 | .682 |

Notes: Two-factor model was calculated for the whole data set (n=230): factor 1 includes item 1 to 5, factor 2 includes item 6 to 10; One-factor model was calculated for half of the data set (n=115): includes items 1 to 10.
[1] discrepancy chi-squared statistic, [2] degrees of freedom, [3] normed chi-squared statistic, [4] comparative fit indexes, [5] Tucker-Lewis Index, [6] root mean square error of approximation, [7] Akaike Information Criterion, [8] Parsimonious Normed Fit Index (PNFI). * $p$ = .000

Results of factor analysis for the 9-item version of the ORIC (without item 10) were similar. Also for the 9-item version, a one-factor model was assumed by exploratory factor analysis. Only the first component had an eigenvalue higher than 95% percentile of the eigenvalues of corresponding random data and the main component explains 66.85% of the variance. Factor loadings of the first component are above 0.739 for all items (see Supplementary File 4, Table A and B). Goodness of fit indices of the one-factor model of the 9-item ORIC version showed similar values compared to the 10-item ORIC version (see Supplementary File 4, Table C).

**Analysis of the ORIC items and internal consistency.** Table 5 shows response distribution, means, standard deviations, acceptance, corrected item-total correlation, and item difficulty of the ten items. Means ranged between 2.24 (item 9) and 2.61 (item 5). Most participants responded in the middle of the scale with a slight shift to more agreement. For items 1 to 9 between four and six missing values could be detected. For item 10, nine missing values were found. Taking all items into account, more than 97% of the measure were answered. Corrected item-total correlations ranged from .665 (item 9) to .774 (item 3), item difficulties

from 55.98 (item 9) to 65.32 (item 5), and inter-item correlations from .434 (item2/item9) to .723 (item3/item 5) (see Table C of Supplementary File 1). Internal consistency yielded a Cronbach's α of .93. Additionally, corrected item-total correlations and internal consistency were calculated for the 9-item ORIC version (see Supplementary File 4, Table D). They were similar to the results for the 10-item version with corrected item-total correlations between .638 (item 9) and .777 (item 3) and a Cronbach's α of .92.

Table 5: Response distribution, means, standard deviation, acceptance, discrimination and item difficulty of the German ORIC.

| | Items | Disagree N (%) | Somewhat Disagree N (%) | Neither Agree nor Disagree N (%) | Somewhat Agree N (%) | Agree N (%) | Mean (SD) | Acceptance (Comple-tion rate in %) * | Item discrimination (corrected item-total correlation) | Item difficulty |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | People who work here are committed to implementing shared decision-making. | 1 (0.4) | 22 (9.6) | 109 (47.0) | 73 (30.9) | 25 (10.9) | 2.42 (.826) | 97.43 | .744 | 60.58 |
| 2 | People who work here will do whatever it takes to implement shared decision-making. | 4 (1.7) | 37 (16.1) | 103 (44.8) | 68 (29.6) | 18 (7.4) | 2.25 (.878) | 98.29 | .689 | 56.36 |
| 3 | People who work here want to implement shared decision-making. | 0 (0.0) | 15 (6.5) | 107 (46.5) | 84 (35.7) | 24 (10.4) | 2.50 (.768) | 97.43 | .774 | 62.61 |
| 4 | People who work here are determined to implement shared decision-making. | 2 (9.0) | 38 (16.5) | 107 (46.5) | 67 (29.1) | 16 (7.0) | 2.25 (.843) | 98.29 | .758 | 56.19 |
| 5 | People who work here are motivated to implement shared decision-making. | 1 (0.4) | 16 (7.0) | 85 (37.0) | 97 (42.2) | 31 (13.5) | 2.61 (.821) | 98.29 | .764 | 65.32 |
| 6 | People who work here feel confident that they can handle the challenges that might arise in implementing shared decision-making. | 2 (0.9) | 20 (8.7) | 93 (40.4) | 93 (40.4) | 22 (9.6) | 2.49 (.819) | 98.29 | .760 | 62.28 |
| 7 | People who work here feel confident that they can keep track of progress in implementing shared decision-making. | 1 (0.4) | 26 (11.3) | 93 (40.4) | 92 (40.0) | 18 (7.8) | 2.43 (.811) | 98.29 | .725 | 60.87 |
| 8 | People who work here feel confident that they can coordinate tasks so that implementation goes smoothly. | 5 (2.2) | 24 (10.4) | 107 (46.5) | 78 (33.5) | 16 (6.5) | 2.32 (.833) | 99.56 | .697 | 58.13 |
| 9 | People who work here feel confident that the organization can support people as they adjust to shared decision-making. | 6 (2.6) | 43 (18.7) | 89 (38.7) | 74 (32.2) | 18 (7.8) | 2.24 (.934) | 97.86 | .665 | 55.98 |
| 10 | People who work here feel confident that they can manage the politics of implementing shared decision-making. | 3 (1.3) | 24 (10.04) | 122 (50.9) | 65 (28.3) | 16 (7.0) | 2.29 (.796) | 96.15 | .714 | 57.44 |

Notes: Items could be answered on a 5-step Likert scale rating from 0 „disagree" to 4 „agree". SD = standard deviation. * For calculation of completion rate, five additional cases were included because these participants only skipped the ORIC but filled out the rest of the survey.

## DISCUSSION

The original English ORIC measure is a brief measure with good psychometric properties [43], which were confirmed in Danish [45] and French [46] validation studies. The study at hand aimed to translate the ORIC into German and assess its psychometric properties.

### Translation and assessment of comprehensibility as part of content validity

Items 1 to 9 could be translated and adapted successfully after two rounds of cognitive interviews and several rounds of discussions within the study team and with external experts. The translation team quickly reached consensus for items 2 to 5 and 8. These items were also well understood by all participants within the first round of cognitive interviews. For items 1, 6, 7 and 9, the translation process was more complex and several adaptations and discussions were necessary. Feedback by participants, members of the study team, and external experts as well as completion rates suggest that comprehensibility of item 10 seems to be low [50]. This might be due to the translation of the phrase “manage the politics” into “Machenschaften”. The term “manage the politics” seems to have a strong cultural connotation and no equivalent phrase in German language exists. The German term “Machenschaften” might have a different connotation as the English phrase and might lead to skipping the item. Ruest et al. [46], who translated the English ORIC into French, also identified several differences in cultural concepts during their adaptation process, but could translate all items successfully. They concluded that limitations in linguistic validation could decrease comparability of psychometric results of the translated measure. However, item 10 showed similar and inconspicuous item characteristics compared to other items in our sample. When repeating factor and item analyses for the 9-item ORIC version including only item 1 to 9, very similar results were observed compared to the 10-item version. To increase comprehensibility and thereby content validity of the scale, the use of the 9-item German ORIC might be a solution and should be evaluated in future studies.

### Factor analysis

We a priory hypothesized a two-factorial structure of the German ORIC, because Shea et al. [43] described correlations between the two theory-based factors “change commitment” and “change efficacy” of .56 to .60. However, we found much higher factor correlations of 0.87. Results of the subsequent EFA clearly indicated a one-factorial structure. Thus, we could not confirm the two-factor structure of the English and the translated Danish and French versions of ORIC [43,45,46]. When comparing the two models, both

models have acceptable values for $Chi^2/df$ [64] and CFI [67,71], but only the two-factor model has acceptable values for TLI [65] and RMSEA [72]. When involving parsimony of the models by calculating AIC and PNFI, the one-factor model fits better to the data [64,67]. Therefore, we prefer the more parsimonious one-factor model. These differences to previous validation studies might be a consequence of diverse cultural connotations of the ORIC items in different languages, caused by the adaption to the context of SDM, or due to specific characteristics of the participating clinics.

**Analysis of ORIC items and internal consistency**

Since the ORIC was presented as the last measure in a three-page survey, missing values might indicate respondent fatigue. However, missing value rates for single items and the overall measure were quite low and the German ORIC was found to be a well-accepted measure. There were no floor or ceiling effects. Corrected item-total correlations of above .66 indicate that all items measure the same underlying concept [57,69]. Criteria for good item difficulties are met since item difficulties are near to 50% and do not differ much from each other [58]. Inter-item correlations are below .80, indicating that items add additional information and are not redundant [57,69]. Cronbach's α (α=.93) suggest excellent internal consistency [57,58,70]. In summary, item analysis and internal consistency of the German ORIC suggest good quality of the measure.

Nevertheless, according to Streiner and Norman [58] a Cronbachs' α above .90 might also indicate item redundancy. On the other hand, inter-item correlations and corrected item-total correlations are in an acceptable range [57,69]. In implementation research there is a need for preferably brief measures, which can be applied in diverse settings with high work-load. Thus, future research could investigate the possibility to further reduce the number of items.

**Strength and Limitations**

This study has some limitations. First, several psychometric parameters are not analysable because this study was a secondary analysis of cross-sectional data. It was not possible to calculate e.g. convergent or divergent validity yet. Second, we applied the ORIC only in three departments of one University Medical Center in Germany. Further validation in different organizational settings is needed to ensure generalizability. Third, for this psychometric evaluation we used a German ORIC, which we adapted and specified for the context of SDM implementation. Our results might not be generalizable for other

interventions in other organizations. Fourth, although SDM was not implemented to the participating clinics before, there might be participants who were more familiar with the concept of SDM than others. Fifth, item 10 was again slightly changed after finishing cognitive interviews. This item was not finally tested for comprehensibility.

A major strength of this study is that we provided the first measure to assess organizational readiness for change in German language for use in implementation studies. We conducted an elaborated translation procedure, which was recommended for survey translations. We furthermore used a qualitative approach to explore comprehensibility including discussions with international colleagues and experts outside of the study team. Furthermore, we assessed the ORIC in a sample including physicians and nurses which was large enough to robustly perform the psychometric analysis on the German version of the ORIC measure.

## CONCLUSION

Organizational readiness is a crucial indicator to successfully implement change and a possible barrier if missing. For implementation studies, it is essential to measure organizational readiness with valid and reliable measures. We provide the first German measure for organizational readiness for implementing change and validated it for the context of SDM implementation. The German ORIC is a brief measure with a high completion rate. We found satisfying psychometric properties in a German hospital setting. To increase content validity of the measure, the use of a 9-item German ORIC (without item 10) should be evaluated in future studies. As the ORIC targets the attitude of organizational members, it can detect reduced or missing readiness for implementing a change on the individuals' level. Therefore, the German ORIC can be used to analyse organizational readiness as a possible barrier for implementing various interventions in organizations.

## LIST OF ABBREVIATIONS

AH: Anne Haugstvedt

AIC: Akaike Information Criterion

AL: Anja Lindig

AMOS: Analysis of Moment Structure, Statistical Package for the Social Sciences, International Business Machines Corporation

$Chi^2$ : Discrepancy chi-squared statistic

Chi2/df: Normed chi-squared statistic

CFI: Confirmatory fit index

CS: Christopher M. Shea

Df: Degree of freedom

DF: Dominick Frosch

EC: Eva Christalle

HCP: Healthcare professional

IS: Isabelle Scholl

KMO: Kaiser-Meyer-Olkin criterion

MM: Marcel Machalski

MR: Mélanie Ruest

ORIC: Organizational Readiness for Implementing Change

PH: Pola Hahlweg

PNFI: Parsimonious Normed Fit Index

RMSEA: Root mean square error of approximation

SDM: Shared decision-making

SPSS: Statistical Package for the Social Sciences, International Business Machines Corporation

SZ: Stefan Zeh

TLI: Tucker-Lewis Index

TRAPD: Translation, Review, Adjudication, Pretesting, Documentation

## DECLARATIONS

### Ethics approval and consent to participate

The study was approved by the Ethics Committee of the Medical Association Hamburg (Germany, study ID PV5368). The study was carried out in accordance to the latest version of the Helsinki Declaration of the World Medical Association. Principles of good clinical practice were respected. Data protection requirements were met. Study participation was voluntary. A waiver of consent for HCPs was obtained from the Ethics Committee, as proposed by current statements on ethical designs of implementation research [73]. HCPs were able to decline participation in the study.

**Funding**

This study is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, grant number 232160533).

**Competing interests**

PH gave one scientific presentation on shared decision-making during a lunch symposium, for which she received compensation and travel compensation from GlaxoSmithKline GmbH in 2018. AL, EC and IS declared to not have any competing interests.

**Author contributions**

AL, PH and IS made substantial contributions to the design and preparation of the study. AL and PH collected the data. AL conducted the analysis in collaboration with EC. All authors contributed to the interpretation of results. AL drafted the manuscript and PH, IS and EC were involved in critically revising the manuscript for important intellectual content. All authors gave final approval of the version to be published.

**Checklist for reporting statement**

To report the results of this validation study, we used the Authors' Guidelines for Reporting Scale Development and Validation Results by Cabrera-Nguyen.

**Patient consent form**

Not applicable

**Data sharing statement**

The dataset collected and analyzed during this study is available from the corresponding author on reasonable request.

**Acknowledgements**

We thank our student assistants Anastasia Izotova, Sophia Schulte and Nicolai Pergande for their help preparing the study and the analysis. We thank Stefan Zeh for his part the team translation process. We

thank Dominick Frosch, Anne Haugstvedt, Marcel Machalski, Mélanie Ruest and Christopher M. Shea for helpful comments and suggestions in the translation and adaptation process of item 10. We also would like to thank Christopher M. Shea for giving the opportunity to translate the ORIC into German and to use it in our implementation study.

**Supplementary Files**

Supplementary File 1: Additional tables for psychometric evaluation of the 10-item ORIC version

Supplementary File 2: Organizational Readiness for Implementing Change (ORIC) – German version

Supplementary File 3: Confirmatory factor analysis model of the 10-item German ORIC

Supplementary File 4: Results of psychometric evaluation of the 9-item German ORIC version

Supplementary File 5: Reporting Checklist: Authors Guidelines for Reporting Scale Development and Validation Results

**REFERENCES**

1 Øvretveit J, Andreen-Sachs M, Carlsson J, et al. Implementing organisation and management innovations in Swedish healthcare. *J Health Organ Manag* 2012;26(2):237-57.

2 Montgomery A, Doulougeri K, Panagopoulou E. Implementing action research in hospital settings: a systematic review. *J Health Organ Manag* 2015;29:729–49.

3 Rye CB, Kimberly JR. The Adoption of Innovations by Provider Organizations in Health Care. *Med Care Res Rev* 2007;64:235–78.

4 Bazzoli GJ, Dynan L, Burns LR, et al. Two Decades of Organizational Change in Health Care: What Have we Learned? *Med Care Res Rev* 2004;61:247–331.

5 Weiner BJ, Amick H, Lee S-YD. Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields. *Med Care Res Rev* 2008;65:379–436.

6 Batras D, Duff C, Smith BJ. Organizational change theory: Implications for health promotion practice *Health Promot Int* 2016;31(1):231–41.

7 Yazdani A, Wells R. Barriers for implementation of successful change to prevent musculoskeletal disorders and how to systematically address them. *Appl Ergon* 2018;73:122–40.

8 Schelvis RMC, Wiezer NM, Blatter BM, et al. Evaluating the implementation process of a participatory

organizational level occupational health intervention in schools. *BMC Public Health* 2016;16(1):1212.

9 Scott DAH, Grant SM. A meta-ethnography of the facilitators and barriers to successful implementation of patient complaints processes in health-care settings. *Heal Expect* 2018;21:508–17.

10 Williams B, Perillo S, Brown T. What are the factors of organisational culture in health care settings that act as barriers to the implementation of evidence-based practice? A scoping review. *Nurse Educ Today* 2015;35(2):e34–41.

11 Härter M, Dirmaier J, Scholl I, et al. The long way of implementing patient-centered care and shared decision making in Germany. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:46–51.

12 Bieber C, Loh A, Ringel N, et al. Patientenbeteiligung bei medizinischen Entscheidungen. Manual zur partizipativen Entscheidungsfindung (Shared Decision-Making). Freiburg: Universitätsklinikum Freiburg; 2007.

13 Härter M, Buchholz A, Nicolai J, et al. Shared decision making and the use of decision aids. *Dtsch Arztebl Int* 2015;112:672–9.

14 Salzburg Global Seminar. Salzburg statment on shared decision making. *BMJ* 2011;342:1745.

15 Reuter K, Loh A, Härter M. Patient as partner in oncology - Opportunities for shared decision-making. In: Koch U, Weis J, eds. Psychoonkologie - Eine Disziplin in der Entwicklung. Göttingen: Hogrefe; 2009:131–45.

16 Hahlweg P, Kriston L, Scholl I, Bet al. Cancer patients' preferred and perceived level of involvement in treatment decision-making. [submitted]

17 Hauser K, Koerfer A, Kuhr K, et al. Outcome-relevant effects of shared decision making. *Dtsch Arztebl Int* 2015;112:665–71.

18 Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. *Med Decis Mak* 2015;35:114–31.

19 Härter M, Müller H, Dirmaier J, et al.. Patient participation and shared decision making in Germany - History, agents and current transfer to practice. *Z Evid Fortbild Qual Gesundhwes* 2011;105:263–70.

20 Härter M, Moumjid N, Cornuz J, et al.Shared decision making in 2017: International accomplishments in policy, research and implementation. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:1–5.

21 Härter M, van der Weijden T, Elwyn G. Policy and practice developments in the implementation of shared decision making: An international perspective. *Z Evid Fortbild Qual Gesundhwes* 2011;105:229–33.

22 Fujimori M, Uchitomi Y. Preferences of Cancer Patients Regarding Communication of Bad News: A Systematic Literature Review. *Jpn J Clin Oncol* 2009;39:201–16.

23 Ernst J, Brahler E, Weissflog G. Patient involvement in medical decision making--an overview on patient preferences and impacting factors. *Gesundheitswesen* 2014;76:187–92.

24 Chewning B, Bylund CL, Shah B, et al Patient preferences for shared decisions: A systematic review. *Patient Educ Couns* 2012;86:9–18.

25 Brom L, Hopmans W, Pasman HR, et al. Congruence between patients' preferred and perceived participation in medical decision-making: A review of the literature. *BMC Med Inform Decis Mak* 2014;14:25.

26 Légaré F, Stacey D, Turcotte S, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2014;9:CD006732.

27 Légaré F, Witteman HO. Shared decision making: Examining key elements and barriers to adoption into routine clinical practice. *Health Aff* 2013;32:276–84.

28 Frosch D, Tietbohl C, Scholl I. Overcoming implementation challenges to advance shared-decision making in routine practice. In: Elwyn G, Thompson R, Edwards A, eds. Shared Decision Making Health Care Achieve Evidence-Based Patient Choice, 3rd Ed. Oxford: Oxford University Press; 2016.

29 Grol R, Wensing M, Eccles M, et al. Improving patient care: The implementation of change in health care. Hoboken (NJ): John Wiley & Sons; 2013.

30 Grol R, Grimshaw J. From best evidence to best practice: Effective implementation of change in patients' care. *Lancet* 2003;362:1225–30.

31 Cummings GG, Estabrooks CA, Midodzi WK, et al. Influence of Organizational Characteristics and Context on Research Utilization. *Nurs Res* 2007;56:S24–39.

32 Li S-A, Jeffs L, Barwick M, Stevens B. Organizational contextual features that influence the implementation of evidence-based practices across healthcare settings: a systematic integrative review. *Syst Rev* 2018;7:72.

33 Légaré F, Ratté S, Gravel K, et al. Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions. *Patient Educ Couns*. 2008;73:526–35.

34 Frerichs W, Hahlweg P, Müller E, et al. Shared decision-making in oncology – a qualitative analysis of healthcare providers' views on current practice. *PLoS One* 2016;11:e0149789.

35 Scholl I, LaRussa A, Hahlweg P, et al. Organizational- and system-level characteristics that influence implementation of shared decision-making and strategies to address them - a scoping review. *Implement Sci* 2018;13:40.

36 André B, Sjøvold E. What characterizes the work culture at a hospital unit that successfully implements change - a correlation study. *BMC Health Serv Res* 2017;17:486.

37 Armenakis AA, Harris SG, Mossholder KW. Creating Readiness for Organizational Change. *Hum Relations* 1993;46:681–703.

38 Kash BA, Spaulding A, Gamm LD, et al. Leadership, culture, and organizational technologies as absorptive capacity for innovation and transformation in the healthcare sector: A framework for research. Chang Manag. *Common Ground Publishing* 2014;13:1–13.

39 Armenakis AA, Harris SG, Mossholder KW. Creating Readiness for Organizational Change. *Hum Relations* 1993;46:681–703.

40 Weiner BJ. A theory of organizational readiness for change. *Implement Sci* 2009;4:67.

41 Armenakis AA, Harris SG. Crafting a change message to create transformational readiness. *J Organ Chang Manag* 2002;15:169–83.

42 Holt D, Armenakis A, Harris S, et al. Toward a Comprehensive Definition of Readiness for Change: A Review of Research and Instrumentation. *Res Organ Chang Dev* 2007;16:289–336.

43 Shea CM, Jacobs SR, Esserman DA, et al. Organizational readiness for implementing change: A psychometric assessment of a new measure. *Implement Sci* 2014;9:7.

44 Gagnon M-P, Attieh R, Ghandour EK, et al. A Systematic Review of Instruments to Assess Organizational Readiness for Knowledge Translation in Health Care. *PLoS One* 2014;9:e114338.

45 Storkholm MH, Mazzocato P, Tessma MK, et al. Assessing the reliability and validity of the Danish version of Organizational Readiness for Implementing Change (ORIC). *Implement Sci* 2018;13:78.

46 Ruest M, Leonard G, Thomas A, et al. French cross-cultural adaptation of the Organizational Readiness for Implementing Change (ORIC). *BMC Health Serv Res* 2019;19:535.

47 Alexander JA, Hearld LR. Methods and metrics challenges of delivery-system research. *Implement Sci* 2012;7:15.

48 Harkness J, Pennell BE, Schoua-Glusberg A. Survey Questionnaire Translation and Assessment. In: Presser S, Rothgeb J, Couper M, et al., eds. Methods for Testing and Evaluating Survey Questionnaires. Hoboken (NJ): JohnWiley & Sons; 2004:453–73.

49 Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737–45.

50 Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27:1159–70.

51 Willis GB, Research Triangle Institute. Cognitive interviewing - a “how to” guide. 1999.

52 Scholl I, Hahlweg P, Lindig A, et al. Evaluation of a program for routine implementation of shared decision-making in cancer care : study protocol of a stepped wedge cluster randomized trial. *Implement Sci* 2018;1–10.

53 Degner LF, Sloan JA, Venkatesh P. The Control Preferences Scale. *Can J Nurs Res* 1997;29:21–43.

54 Coudert L, Carmichael P-H, Renaud J-S, et al. Validation of the IcanSDM instrument to assess clinicians’ perceptions of their ability to adopt shared decision making [in preparation].

55 Hahlweg P, Härter M, Nestoriuc Y, et al. How are decisions made in cancer care? A qualitative study using participant observation of current practice. *BMJ Open* 2017;7:e016360.

56 Bannon WJ. Missing data within a quantitative research study: How to assess it, treat it, and why you should care. *J Am Assoc Nurse Pr* 2015;27:230–2.

57 Rattray J, Jones MC. Essential elements of questionnaire design and development. *J Clin Nurs* 2007;16:234–43.

58 Streiner DL, Norman GR, Health measurement scales: A practical guide to their development and use. Oxford, UK: Oxford University Press 2008.

59 Fabrigar L, Wegener DT, MacCallum RC, et al. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods* 1999;4:272.

60 Worthington R. Scale development research. *Couns Psychol* 2006;34:806–38.

61 O’Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer’s MAP test. *Behav Res Methods, Instruments, Comput* 2000;23:396–402.

62 Hayton JC, David GA, Scarpello V. Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. *Organ Res Methods* 2004;7:191–205.

63 Zwick WR, Velicer WF. Comparison of five rules for determining the number of components to retain. *Psychol Bull* 1986;99:432–42.

64 Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the Fit of Structural Equation Models:

Tests of Significance and Descriptive Goodness-of-Fit Measures. *Meth Psychol Res Online* 2003;8:23–74.

65 Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Eq Model* 1999;6:1–55.

66 Hair J, Black W, Babin B, RE Anderson. Multivariate data analysis. Upper Saddle River, NJ: Pearson Prentice Hall; 2009.

67 Hooper D, Coughlan J, Mullen M. Structural Equation Modelling: Guidelines for Determining Model Fit. *Articles* 2008.

68 McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1996;4:293–307.

69 Ferketich S. Focus on psychometrics. Aspects of item analysis. *Res Nurs Health* 1991;14:165–8.

70 Bannigan K, Watson R. Reliability and validity in a nutshell. *J Clin Nurs* 2009;18:3237–43.

71 Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238–46.

72 Browne MW, Cudeck R. Single Sample Cross-Validation Indices for Covariance Structures. *Multivariate Behav Res* 1989;24:445–55.

73 Hutton J, Eccles M, Grimshaw J. Ethical issues in implementation research: A discussion of the problems in achieving informed consent. *Implement Sci* 2008;3:52.

**Supplementary File 1: Additional tables for psychometric evaluation of 10-item ORIC version**

Table A: Demographic data of participants of cognitive interviews (n=11).

| | | Frequencies for n=11 |
|---|---|---|
| Age | < 30 years | 1 |
| | 31-40 years | 3 |
| | 41-50 years | 3 |
| | > 50 years | 4 |
| Gender | Female | 9 |
| | Male | 2 |
| Profession* | Nurse | 8 |
| | Physician | 2 |
| | Psychooncologist | 3 |
| Work experience in health care | > 5 years | 3 |
| | 5-10 years | 4 |
| | 11-20 years | 3 |
| | > 20 years | 1 |

* multiple answers possible

Table B: Factor loadings on the first component.

| | Component 1 |
|---|---|
| Item 1 | 0.790 |
| Item 2 | 0.788 |
| Item 3 | 0.833 |
| Item 4 | 0.846 |
| Item 5 | 0.831 |
| Item 6 | 0.851 |
| Item 7 | 0.839 |
| Item 8 | 0.800 |
| Item 9 | 0.754 |
| Item 10 | 0.860 |

For EFA, the sample was split and n=115.

Table C: Inter-item correlation matrix for the German ORIC.

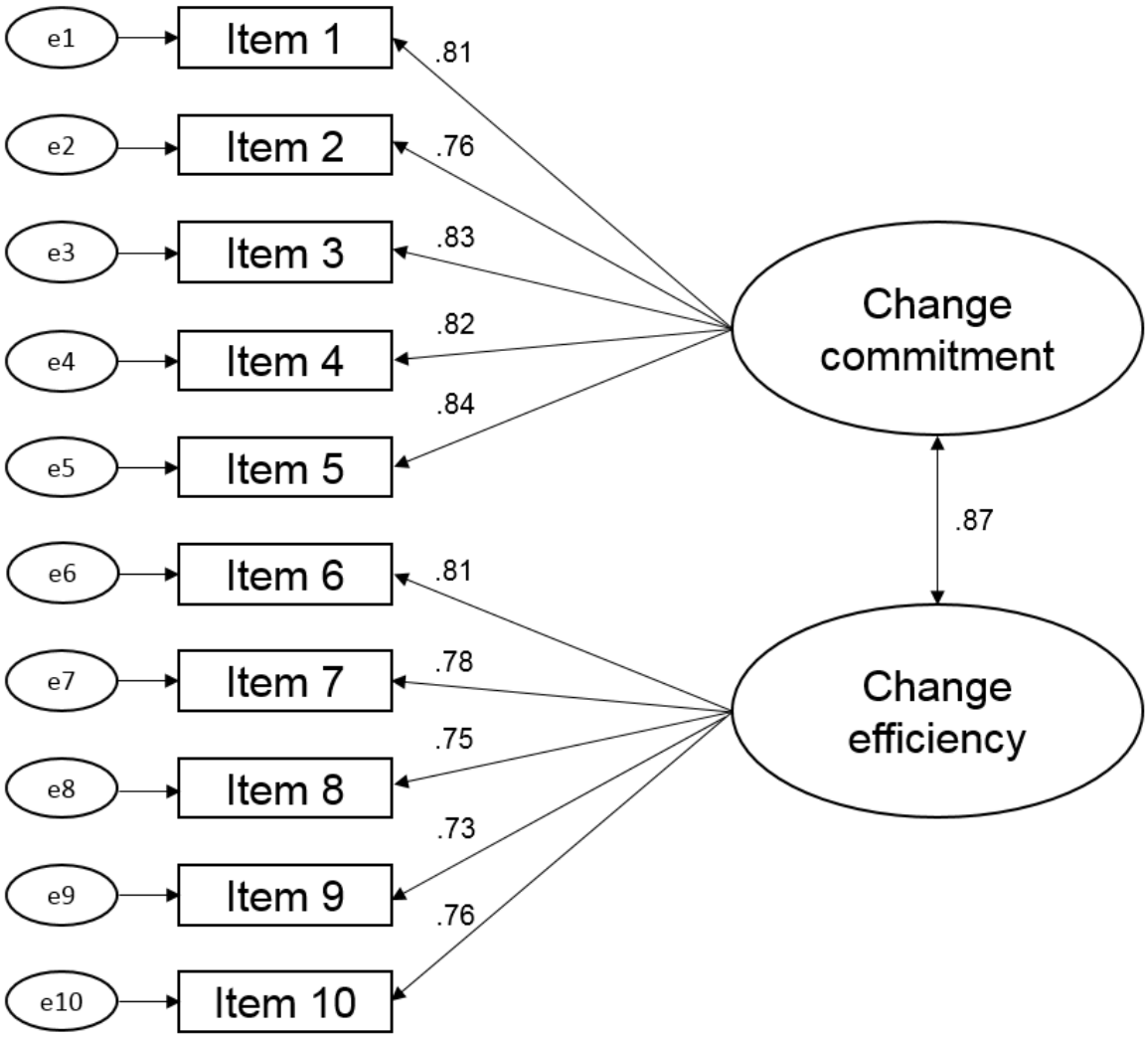
| | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 | Item 6 | Item 7 | Item 8 | Item 9 | Item 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Item 1 | 1.000 | .634 | .650 | .611 | .652 | .587 | .558 | .538 | .507 | .542 |
| Item 2 | .634 | 1.000 | .586 | .681 | .607 | .518 | .504 | .485 | .434 | .478 |
| Item 3 | .650 | .586 | 1.000 | .673 | .723 | .636 | .600 | .558 | .492 | .560 |
| Item 4 | .611 | .681 | .673 | 1.000 | .675 | .588 | .531 | .541 | .512 | .551 |
| Item 5 | .652 | .607 | .723 | .675 | 1.000 | .622 | .621 | .466 | .474 | .578 |
| Item 6 | .587 | .518 | .636 | .588 | .622 | 1.000 | .677 | .590 | .588 | .573 |
| Item 7 | .558 | .504 | .600 | .531 | .621 | .677 | 1.000 | .612 | .491 | .567 |
| Item 8 | .538 | .485 | .558 | .541 | .466 | .590 | .612 | 1.000 | .611 | .560 |
| Item 9 | .507 | .434 | .492 | .512 | .474 | .588 | .491 | .611 | 1.000 | .659 |
| Item 10 | .542 | .478 | .560 | .551 | .578 | .573 | .567 | .560 | .659 | 1.000 |

**Supplementary File 2: Organizational Readiness for Implementing Change (ORIC) – German version**

Bitte geben Sie an, wie sehr Sie den folgenden Aussagen zur Umsetzung von partizipativer Entscheidungsfindung an Ihrem aktuellen Arbeitsplatz zustimmen. Falls partizipative Entscheidungsfindung aktuell nicht umgesetzt wird, wie wäre es im Falle der Umsetzung?

| | | stimme nicht zu | stimme eher nicht zu | teils teils | stimme eher zu | stimme zu |
|---|---|---|---|---|---|---|
| 1 | Personen, die hier arbeiten, zeigen hohes Engagement bei der Umsetzung von partizipativer Entscheidungsfindung. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2 | Personen, die hier arbeiten, werden tun, was auch immer nötig ist, um partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3 | Personen, die hier arbeiten, wollen partizipative Entscheidungsfindung umsetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4 | Personen, die hier arbeiten, sind fest entschlossen, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 5 | Personen, die hier arbeiten, sind motiviert, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 6 | Eventuell entstehen bei der Umsetzung von partizipativer Entscheidungsfindung Herausforderungen. Personen, die hier arbeiten, sind zuversichtlich, diese zu meistern. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 7 | Personen, die hier arbeiten, sind zuversichtlich, dass sie den Verlauf der Umsetzung von partizipativer Entscheidungsfindung überblicken können. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 8 | Personen, die hier arbeiten, sind zuversichtlich, dass sie Aufgaben so koordinieren können, dass die Umsetzung reibungslos abläuft. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 9 | Personen, die hier arbeiten, sind zuversichtlich, dass die Klinik sie dabei unterstützten kann, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 10 | Personen, die hier arbeiten, sind zuversichtlich, Machenschaften bei der Umsetzung von partizipativer Entscheidungsfindung bewältigen zu können. | ☐ | ☐ | ☐ | ☐ | ☐ |

**Supplementary File 3: Confirmatory factor analysis model of the 10-item German ORIC**



**Supplementary File 4: Results of psychometric evaluation of the 9-item German ORIC version: Factor analysis, corrected item-total correlations, Cronbach's α**

## Factor analysis

To get information about the factorial structure of the 9-item German ORIC, an exploratory factor analysis was conducted. KMO measure was .926 and Barlett's test of sphericity was significant ($X^2$ = 1302.78, $p < .001$), indicating that a factor analysis of the data was appropriate to observe data. Table A shows results of the exploratory factor analysis. Since the main component explains 66% of the variance, a one-factor model could be assumed. The factor loading for the first component can be observed in table B.

Table A: Results of exploratory factor analysis of the 9-item German ORIC with oblique rotation and parallel analysis: eigenvalues of the ten components of the German ORIC and eigenvalues for corresponding random data.

| | Eigenvalue | | | Eigenvalues for random data | |
|---|---|---|---|---|---|
| | Total | % of variance | Cumulative % | Means | 95% percentile |
| Component 1 | 6.02 | 66.85 | 66.85 | 1.45 | 1.59 |
| Component 2 | 0.802 | 8.91 | 75.76 | 1.29 | 1.39 |
| Component 3 | 0.472 | 5.25 | 81.01 | 1.17 | 1.25 |
| Component 4 | 0.392 | 4.35 | 85.37 | 1.07 | 1.14 |
| Component 5 | 0.366 | 4.06 | 89.43 | 0.98 | 1.05 |
| Component 6 | 0.304 | 3.38 | 92.81 | 0.89 | 0.96 |
| Component 7 | 0.250 | 2.78 | 95.59 | 0.81 | 0.87 |
| Component 8 | 0.232 | 2.57 | 98.16 | 0.71 | 0.79 |
| Component 9 | 0.165 | 1.84 | 100.00 | 0.61 | 0.69 |

Notes: For EFA, half of the data set (n=115) was used.

Table B: Factor loadings on the first component.

| | Component 1 |
|---|---|
| Item 1 | 0.860 |
| Item 2 | 0.856 |
| Item 3 | 0.837 |
| Item 4 | 0.836 |
| Item 5 | 0.835 |
| Item 6 | 0.799 |
| Item 7 | 0.797 |
| Item 8 | 0.791 |
| Item 9 | 0.739 |

Notes: For EFA, the sample was split and n=115.

A confirmatory factor analysis was performed with the one-factor model to analyse its fit indices. Indices of the one-factor model are presented in table C.

Table C: Fit indices of the one-factor model of the 9-item German ORIC.

| | Chi² [1] | df [2] | Chi²/df [3] | CFI [4] | TLI [5] | RMSEA [6] | AIC [7] | PNFI [8] |
|---|---|---|---|---|---|---|---|---|
| One-factor model | 56.04 | 27 | 2.08 | .945 | .927 | .097 | 92.04 | .676 |

Notes: One-factor model (including a split data set of n=115): includes items 1 to 9.
[1] discrepancy chi-squared statistic, [2] degrees of freedom, [3] normed chi-squared statistic, [4] comparative fit indexes, [5] Tucker-Lewis Index, [6] root mean square error of approximation, [7] Akaike Information Criterion, [8] Parsimonious Normed Fit Index (PNFI). * $p$ = .000

**Corrected item-total correlations and Cronbach's α of the 9-item ORIC version**

Table D shows the corrected item-total correlation. Corrected item-total correlation ranged from .665 (item 9) to .744 (item 3).

Table D: Discrimination of the nine items of the German ORIC.

| | Item discrimination (corrected item-total correlation) |
|---|---|
| Item 1 | .747 |
| Item 2 | .696 |
| Item 3 | .777 |
| Item 4 | .760 |
| Item 5 | .762 |
| Item 6 | .759 |
| Item 7 | .720 |
| Item 8 | .691 |
| Item 9 | .638 |

Notes: Items could be answered on a 5-step Likert scale rating from 0 „ disagree" to 4 „agree".

Cronbach's α with α = .924 showed the measure to reach excellent reliability.

**Supplementary File 5: Checklist for reporting standards. Authors' guideline for Scale Development and validation results by Cabrera-Nguyen.**

| | **Guidelines by Cabrera-Nguyen** [1] | **Transfer to our study** |
|---|---|---|
| 1 | Precisely define the target construct. | See page 3 and 4 (Introduction section) |
| 2 | Justify the need for your new measure. For example, if measures of the construct exist in the literature, explain the value added by your new scale. How might the new measure enhance the substantive knowledge base or social work practice? | See page 4 (Introduction section) |
| 3 | Indicate that you have submitted your initial pool of items to expert review (Worthington & Whittaker, 2006). Report (a) the number of items in the preliminary pool; (b) the number of expert reviewers and their qualifications; and (c) any major changes to your initial item pool following the review (e.g., a substantial decrease in the number of items, changes to the original item response format, overhaul of item pool due to experts' assessment regarding content validity). | See page 5 and 6 (Methods section: Measure, Translation, Assessment of comprehensibility and adaption of the scale) and page 9 and 10 (Results section: Translation, Assessment of comprehensibility and adaption of the scale) |
| 4 | Report the name and version of the statistical software package used for all analyses. | See page 6 to 8 (Methods section: Assessment of comprehensibility and adaption of the scale, Psychometric evaluation) |
| 5 | Identify and justify the sampling strategy (e.g., convenience, snowball) and sampling frame. Report standard sample demographic characteristics as well as other salient sample characteristics. | See page 6 and 7 (Methods section: Assessment of comprehensibility and adaption of the scale, Psychometric evaluation) and page 8 to 10 (Results section: Assessment of comprehensibility and adaption of the scale, Psychometric evaluation) and Additional file 3 |
| 6 | Discuss relevant data preparation and screening procedures. For instance, do the data meet the appropriate assumptions for factor analysis? If not, what actions were taken? Report tests of factorability if appropriate (e.g., report Bartlett's test of sphericity). | See page 10 to 12 (Results section: Psychometric evaluation) |
| 7 | Provide all dates of data collection. | See page 9 to 13 (Results section) |
| 8 | Avoid use of principal components analysis (PCA) as a precursor to CFA (Costello & Osborne, 2005; Worthington & Whittaker, 2006). Instead, start with EFA to assess the underlying factor structure and refine the item pool. EFA should be followed by CFA using a different sample (or samples) to evaluate the EFA-informed a priori theory about the measure's factor-structure and psychometric properties. (Costello & Osborne, 2005; Henson & Roberts, 2006; Worthington & Whittaker, 2006). For CFA, authors | An a priory hypothesized model for CFA was specified based on the original measures structure. The model could not be confirmed and an EFA was calculated afterwards. The same data set was used for EFA and CFA. For more |

| | | |
|---|---|---|
| | should specify an a priori hypothesized model and a priori competing models (Jackson, Gillaspy, & Purc-Stephenson, 2009). | information see page 10 to 12 (Results section). |
| 9 | Guidelines for reporting EFA results. | |
| | How large is a sample? One common rule of thumb is to ensure a person-to-item ratio of 10:1. Another rule of thumb is that N= 300 is usually acceptable (Worthington & Whittaker, 2006). However, some researchers have criticized these sample size rules of thumb, noting the appropriate sample size is dependent on the features of the gathered data. These researchers recommend obtaining the largest possible sample because the adequacy of the sample size cannot be determined until after the data have been analyzed (Henson & Roberts, 2006). | Sample size is n=115, person-to-item ratio 11.5:1 (see page 7: Methods section, Psychometric analysis) |
| | Run EFA . . . or not. Run a preliminary EFA to determine if further data collection is required based on the following criteria: (a) If communalities are greater than .50 or there are 10:1 items per factor with factor loadings of roughly .4, then a sample size of 150 to 200 is likely to be adequate; (b) If communalities are all at least .60 or there are a minimum of 4:1 items per factor with factor loadings above .6, then even smaller sample sizes may suffice; (Worthington & Whittaker, 2006). Report if additional data collection was necessary due to inadequate sample size. If so, report the new participants' sociodemographic characteristics and test for differences between groups using standard statistical procedures (e.g., t-tests). | EFA communalities are all above .70, therefore the sample size of n=115 can be determined as adequate and no additional data collection was necessary. For more information, see Supplementary File1). |
| | Give EFA details. Report the specific rotation strategy used (e.g. varimax, geomin). Justify the decision to use an orthogonal or oblique solution. One recommendation is to always begin with an oblique rotation, empirically assess factor intercorrelations, and report them before deciding upon a final rotation solution (Henson & Roberts, 2006; Worthington &Whittaker, 2006). Some researchers argue oblique rotation is always the best approach because (a) factor intercorrelations are the norm in social sciences and (b) both approaches yield the same result if the factors happen to be uncorrelated (Costello & Osborne, 2005). Conversely, other researchers contend that orthogonal rotation is preferable because fewer parameters are estimated—orthogonal rotation is more parsimonious and amenable to replication (Henson &Roberts, 2006). Similarly, some researchers warn against relying on a statistical software package's default settings to determine the appropriate type of oblique rotation (Henson & Roberts, 2006; Worthington &Whittaker, 2006). Others state that doing so is fine (Costello & Osborne, 2005, p.3). Given the lack of consensus, it is probably best to describe what you do and defend your approach on substantive grounds, if possible. | See page 6 to 8 (Methods section: Psychometric evaluation) |
| | Report the whole factor pattern/structure. Always report the whole factor pattern/structure matrix, including all of the items in the analysis. It is recommended that authors report | See page 10 to 12 (Results section: Psychometric evaluation) and Supplementary File 1 |

| | | |
|---|---|---|
| | this information in a chart following the example provided by Henson and Roberts (2006) on page 411. | |
| | Criteria for deleting (crossloaded) items. Report any deleted items and the criteria used for deletion. Crossloading items with values ≥ .32 on at least two factors should generally be candidates for deletion, especially if there are other items with factor loadings of .50 or greater (Costello & Osborne,2005). Rerun the EFA each time an item is deleted. | No crossloading items could be observed, so item deletion was not necessary. |
| | Criteria for number of factors. Report the number of factors retained and justify this decision using multiple criteria (eigenvalue > 1, scree test, parallel analysis, rejection of a factor with fewer than 3 items, etc). Reporting the eigenvalue > 1 rule alone is inadequate because it has been shown to among the least accurate criteria for assessing factor retention (Costello & Osborne, 2005; Henson & Roberts, 2006) | See page 6 to 8 (Methods section: Psychometric evaluation) |
| | Explained variance. Report the variance explained by the factors. | See page 10 to 12 (Results section: Psychometric evaluation) and Supplementary File 1 |
| | In general, describe your decisions. | See page 10 to 12 (Results section: Psychometric evaluation) and Supplementary File 1 |
| 10 | Guidelines for reporting CFA results. | |
| | Describe and justify the theoretical model. Report hypothesized factor structure. Provide theoretical and empirical justification (e.g., results of preliminary EFAs) for your hypothesis. In addition, report a priori competing models. | See page 6 to 8 (Methods section: Psychometric evaluation) |
| | Describe the parameterization. Provide a comprehensive description of the a priori parameter specification. Identify fixed parameters, free parameters, and constrained parameters. For example, indicate if you freed the errors of any items to correlate. | One factor loading was constrained to equal 1, the corresponding intercept was constrained equal to zero. The other factor loadings and intercepts were estimated. Errors of items were not freed to correlate. |
| | Include a figure. Include a figure of each CFA model being tested using Kline's (2005) graphical conventions if feasible. | See Supplementary File 3 |
| | Identification. Demonstrate model identification (e.g., df > 0; scaling of factors; assess and report the —t-rule; the two-indicator rule). Necessary and sufficient conditions for model identification may vary for certain types of CFA models. When in doubt, authors should consult Brown's (2006) CFA textor Kline's (2005) SEM text for guidance. | See page 10 to 12 (Results section: Psychometric evaluation) |
| | Select an estimator based on distributional patterns and assumptions. Report the estimator used (e.g., ML, | See page 6 (Methods section: Psychometric evaluation) |

| | | |
|---|---|---|
| | WLSMV) and justify your choice based on distributional assumptions. It is not appropriate to report that you relied on your statistical software's default setting. | |
| | Use multiple fit indices. After estimating a model, always report multiple fit indices (e.g., model X2, df, p, CFI/TLI, RMSEA, SRMR). Report all appropriate fit indices, not just those favorable to your hypotheses (Jackson et al., 2009). For example, do not report acceptable CFI and TLI scores while omitting a relevant fit index with a subobtimal value. | See page 8 (Methods section: Psychometric evaluation) and page 10 and 11 (Results section: Psychometric evaluation) |
| | What is acceptable fit? For model fit indices, authors should generally use the cut-off values recommended by Hu and Bentler (1999) and endorsed by Brown (2006), assuming ML estimation: a.CFI/TLI ≥ .95 b. RMSEA ≤ .06 c. SRMR ≤ .08 | See page 9, Table 1 (Methods section: Psychometric evaluation) |
| | Localized strain? When reporting model fit, include an assessment for localized areas of strain by examining standardized residuals. Standardized residuals greater than 1.96 (for p< .05) indicate areas of strain (Harrington, 2009). Report the absence of localized strain, if appropriate; otherwise, note localized areas of strain by reporting the relevant standardized residuals. | Standardized residuals do not indicate localized strains. |
| | Parameter estimates and SEs. When reporting factor loadings and other parameter estimates, always report the unstandardized estimates, their p values, and the standard errors. In addition, include the standardized estimates when appropriate. Be sure to report all parameter estimates, even those that are non significant (Brown, 2006; Jackson et al., 2009). | See Supplementary File 3 |
| | Assessing the validity of the factor solution. Comment on the new measure's convergent and discriminant validity based on parameter estimates. For instance, factor correlations ≥ .80 may indicate poor discriminant validity (Brown, 2006). In addition, strong factor loadings that do not crossload may indicate good convergent validity. One rule of thumb is that factor loadings < .40 are weak and factor loadings ≥ .60 are strong (Garson, 2010). | Exploratory factor analysis revealed a one-factor model, so there are no factor correlations. |
| | Other measures. Report squared multiple correlations and comment on the measure's reliability (e.g., report Raykov's Rho if appropriate) | See page 10 to 12 (Results section: Psychometric analysis) |
| | Respecification: Caution! Report any post-hoc respecifications to improve model fit based on modification indices. Justify the respecifications on theoretical or conceptual grounds (Jackson et al., 2009). Respecification to allow for correlated errors is not supportable without strong pragmatic justification(e.g., items contain similar words or phrases). Note that respecification precludes comparing the model with your a priori specified competing models. Report improvements in appropriate model fit indices for respecified models (e.g., chi-square difference test) | No respecifications to improve model fit were applied. |
| 10 | Describe the matrix (or matrices) you analysed (e.g., covariance, correlation). Include matrices in the manuscript | See Supplementary File 1 |

| | | |
|---|---|---|
| | if feasible; otherwise, indicate these data are available upon request. | |
| 11 | Report the amount of missing data and describe how missing data were handled. For a review of practices for handling missing data, see Sterne and colleagues (2009), Rose and Fraser (2008), and Horton and Kleinman (2007).Provide a rationale for your approach to handling missing data. Authors are encouraged to consider using multiple imputation or model estimation with full-information maximum likelihood (FIML; Rose & Fraser, 2008). | See page 7 (Methods section: Psychometric evaluation) and page 10 (Results section: Psychometric evaluation) |
| 12 | Compare your CFA model with the alternative or competing models. Do competing models fit the data better or worse than your model (e.g., does your four-factor model of acculturation fit the data better than a two-factor model or a one-factor model)? Identify the preferable model based on appropriate fit statistics (e.g. chi-square difference test for nested models, Akaike information criterion for non-nested models), parsimony, and relevant theory | See page 15 and 16 (Discussions sections) |
| 13 | Include your scale (items and response options) in an appendix. | See Supplementary File 2 |
| 14 | Report how methodological limitations may have impacted findings regarding your measure's psychometric properties (e.g., note potential repercussions of suboptimal sampling techniques, discuss implications of using listwise deletion to handle missing data instead of multiple imputation or FIML). | See page 16 and 17 (Discussions sections) |
| 15 | Discuss directions for future research (e.g., if appropriate, testing your scale for measurement invariance by conducting CFA on different populations). | See page 16 and 17 (Discussions sections) |

1 Cabrera-Nguyen P. Author Guidelines for Reporting Scale Development and Validation Results in the *J Soc Social Work Res* 2010;1:99–103.

BMJ Open

# Translation and psychometric evaluation of the German version of the Organizational Readiness for Implementing Change measure (ORIC) – a cross-sectional study



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2019-034380.R2 |
| Article Type: | Original research |
| Date Submitted by the Author: | 01-Apr-2020 |
| Complete List of Authors: | Lindig, Anja; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Hahlweg, Pola; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Christalle, Eva; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology<br>Scholl, Isabelle; University Medical Center Hamburg-Eppendorf, Department of Medical Psychology |
| Primary Subject Heading: | Research methods |
| Secondary Subject Heading: | Communication, Evidence based practice, Patient-centred medicine |
| Keywords: | Organizational readiness for change, Psychometrics, Translation, Implementation, Shared decision-making, Measurement |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Translation and psychometric evaluation of the German version of the Organizational Readiness for Implementing Change measure (ORIC) – a cross-sectional study

Anja Lindig[1], Pola Hahlweg[1], Eva Christalle[1], Isabelle Scholl[1]

[1] *Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany*

Email addresses:
Anja Lindig: a.lindig@uke.de
Pola Hahlweg: p.hahlweg@uke.de
Eva Christalle: e.christalle@uke.de
Isabelle Scholl: i.scholl@uke.de

Corresponding author:
Anja Lindig
University Medical Center Hamburg-Eppendorf
Department of Medical Psychology
Martinistr. 52
20246 Hamburg, Germany
Email: a.lindig@uke.de
Phone: +49-(0)40-7410-56206

Word count: 4186

## ABSTRACT

**Objectives:** To translate the Organizational Readiness for Implementing Change (ORIC) measure into German and assess its psychometric properties.

**Design:** Cross-sectional psychometric study based on secondary analysis of baseline data from a shared decision-making implementation study.

**Setting:** Three departments within one academic cancer center in Hamburg, Germany.

**Participants:** For comprehensibility assessment of the translated ORIC version, we conducted cognitive interviews with healthcare professionals (HCPs, n=11). Afterwards, HCPs (n=230) filled out the measure.

**Primary and Secondary Outcome Measures:** The original English version of the ORIC was translated into German using a team translation protocol. Based on comprehensibility assessment via cognitive interviews with HCPs, the translated version was revised. We analyzed acceptance (completion rate), factorial structure (exploratory factor analysis (EFA), confirmatory factor analysis (CFA), model fit), item characteristics (item difficulties, corrected item-total correlations, inter-item correlations), and internal consistency (Cronbach's α).

**Results:** Translation and cognitive testing of the German ORIC was successful except for item 10, which showed low comprehensibility as part of content validity in cognitive interviews. Completion rate was > 97%. EFA and CFA provided a one-factorial structure. Item difficulties ranged between 55.98 and 65.32, corrected item-total-correlation ranged between .665 and .774, inter-item correlations ranged between .434 and .723, and Cronbach's α was .93.

**Conclusions:** The German ORIC is a reliable measure with high completion rates and satisfying psychometric properties. A one-factorial structure of the German ORIC was confirmed. Item 10 showed limited comprehensibility and therefore reduces content validity of the measure. The German ORIC can be used to analyze organizational readiness for change as a precursor for implementation success of various interventions.



**Keywords:**

Organizational readiness for change, Psychometrics, Translation, Implementation, Shared decision-making, Measurement

## ARTICLE SUMMARY

### Strengths and limitations of this study

- Appropriate qualitative methods were used to provide a German version of the ORIC to be assessed for healthcare professionals (HCPs).
- The sample size was large enough to robustly perform the psychometric analyses on the German version of the ORIC measure.
- Due to the design of the study as a secondary analysis, it was not possible to calculate psychometric parameters like convergent and divergent validity.
- Data were collected at a single academic cancer center in the context of a shared decision-making implementation study. Thus, psychometric properties of the German ORIC need replication in other settings.

## INTRODUCTION

Implementing interventions in healthcare systems is an important and widely discussed topic [1–3] and often mediated by public policies, market forces, or new technologies [4]. The intention to implement new interventions might be to reduce costs, improve quality, increase efficacy or patient satisfaction [5]. Nevertheless, implementing change in healthcare organizations can be challenging [6–10]. In the German healthcare system, the implementation of shared decision-making (SDM) has received much attention [11]. SDM can be described as an interactional process on the basis of information exchange. Patients and healthcare professionals (HCPs) are equally and actively involved and jointly responsible for the decision [12–14]. This is especially important in situations with complex treatment options and high impact on patients' quality of life [15]. Patients want to be actively involved in decision-making [16] and benefit from SDM by developing better knowledge about their disease and treatment options, better risk perception, and less insecurity and decisional conflict [17,18]. SDM has been supported by health policy [19–21] and research [22–25] over the last decades. However, SDM is currently not routinely implemented in the German healthcare system [11,26–28].

When implementing SDM or other interventions in organizations, several barriers on different levels of the organization (i.e., individual level, group level, organizational and system level) need to be considered [1,5–7,10,29–32]. Barriers for implementing SDM in the clinical setting often address both the organizational setting (e.g. lack of resources and lack of management support) and the individual level (e.g.

resistance to change or negative attitudes towards SDM) [7,10,11,33–35]. When implementing SDM or other interventions in healthcare systems, the clinical employees' perspective on organizational readiness for change is a critical precursor for successful implementation [5,32,36–38]. Armenakis et al. [39] describe organizational readiness for change as the degree to which organizational members are prepared to participate in change processes. This is characterized by the belief that the change is needed and that the organization is capable of changing. In their theory of organizational readiness for change, Weiner et al. [40] differentiate between change commitment (i.e., organizational members' attitudes towards implementing a change) and change efficacy (i.e., organizational members' belief in their capability to implement a change). If readiness for change is high, organizational members invest more in the change effort and exhibit greater persistence to overcome barriers and setbacks [40,41].

To analyze effects of organizational readiness on implementation success, specific measures for assessing organizational readiness for change are needed [5]. However, only a few validated measures exist [5,42–44] and none were available in German. One of those measures is the Organizational Readiness for Implementing Change (ORIC) [43]. The ORIC is brief, easy to administer, and theoretically and psychometrically well-grounded [40]. It was previously translated into Danish and French [45,46]. The ORIC has been psychometrically tested, revealing a completion rate of more than 72%, a Cronbach's α of above .80 and two correlating factors [43,45,46]. Due to the described properties, the ORIC seemed well-suited to measure organizational readiness for implementing SDM in Germany [5,43,47].

Therefore, the aim of the study was to translate the ORIC measure into German and assess its psychometric properties.

## METHODS

### Measure

The ORIC measures organizational readiness for implementing change. It uses a 5-point Likert scale ranging from „disagree" to „agree" [43]. In the original English version, two subscales were described based on Weiner et al. [40]: „change commitment" (items 1 to 5) and "change efficacy" (items 6 to 10). Sum scores were calculated for both subscales separately with higher scores indicating higher organizational readiness for change. By using the phrases "to implement this change" or "implementing this change", the original scale does not specify which change is addressed. The items can be specified to adapt to a specific

research question and a survey instruction can be added [40,45]. The English items are displayed in the results section.

**Translation**

Translation followed the team translation protocol TRAPD (Translation, Review, Adjudication, Pretesting and Documentation), a method with growing recognition within translation research [48–52]. Thereby an optimal translation is facilitated by discussions between members of the translation team with different expertise in translation. First, two team members (AL, SZ, cp. list of abbreviations) proficient in German and English, but little experienced in survey translation, independently translated the original ORIC into German. Second, a third bilingual team member (IS) with experience in survey translation, reviewed both versions and suggested a third version based on the first two translations. Finally, IS, AL, and SZ discussed all versions until reaching consensus on a final version. To find consensus on item 10 we additionally consulted an official translator (MM, cp. list of abbreviations) and an additional team member (PH), who is proficient in German and English and experienced in translation. During the translation process, we changed the phrases “to implement this change” and “implementing this change” into “to implement shared decision-making” and “implementing shared decision-making” to address our specific research question. Additionally, we added a survey instruction in German which motivated participants to think about the clinic, they are working in, when answering the item. As a next step we pretested the translated measure by conducting cognitive interviews and thereby assessed comprehensibility as part of content validity.

**Assessment of comprehensibility as part of content validity and subsequent adaptation of the scale**

Content validity is the degree to which the content of the measure and its items adequately reflect the measured construct [53]. According to the COSMIN criteria (Consensus-based standards for the selection of health measurement instruments) [54], content validity includes the relevance of the items and scales, their comprehensiveness, and comprehensibility. As this study aimed to evaluate the translation of an existing measure, we focused on the assessment of comprehensibility (i.e., items being appropriately worded and understood by participants as intended).

To do so, we conducted two rounds of cognitive interviews with a convenience sample of HCPs (nurses, physicians, and psychooncologists), working in a comprehensive cancer center in Germany. Two female researchers and psychologists experienced in interviewing (AL, PH) conducted the interviews. We

developed an interview guide based on recommendations by Willis et al. [55]. We used verbal probing techniques like comprehension probes (e.g. “What does the term ‘organization’ mean to you?”) and paraphrasing (e.g. “Can you repeat this sentence in your own words?”). We conducted interviews until reaching theoretical saturation. Interviews were audio-recorded and transcribed verbatim. After the first round of cognitive interviews, we extracted and discussed comments and suggestions from the transcripts (AL, PH, IS). As a further step to enhance comprehensibility, we discussed the items with the original author (CS, cp. list of abbreviations) as well as French (MR, cp. list of abbreviations) and Norwegian researchers (AH, cp. list of abbreviations), who translated the ORIC into their languages. We adapted items of the German ORIC, which were not well understood by participants of the first round of cognitive interviews, according to these discussions. We tested these items in a second round of cognitive interviews. After the second round, we discussed further adaptations of the items and involved another bilingual researcher in the field (DF, cp. list of abbreviations).

We calculated descriptive statistics of participants’ demographic characteristics using SPSS (IBM SPSS Statistics, Version 23).

**Psychometric evaluation**

**Data collection.** For psychometric evaluation of the ORIC measure we conducted a secondary analysis of cross-sectional data gathered in a SDM implementation study [56]. Data from baseline assessment of the SDM implementation study were included. The ORIC was the last questionnaire of a three-page survey measuring HCPs’ attitudes regarding SDM and its implementation. Besides the ORIC, it contained the Control Preference Scale [57] and the IcanSDM [58], demographic questions (e.g. gender, age, profession, work experience) as well as several questions that have been used in previous studies in cancer care [59]. Results of these additional measures will be published as part of the primary evaluation of the SDM implementation study [56].

Participants were part of a convenience sample of physicians and nurses. Since this is a secondary analysis, inclusion criteria were identical to inclusion criteria of the SDM implementation study [56]. We included physicians and nurses who worked at one of three departments within the University Cancer Center Hamburg at the University Medical Center Hamburg-Eppendorf during baseline evaluation of the SDM implementation study [56]. Eligible HPCs were identified through employee lists provided by department managers. The measure was handed out to eligible HCPs either (1) by a member of our study

team (e.g. during a regular physician meeting), (2) by the supervising nurses, or (3) via employees' mailboxes. Participants returned the questionnaire personally to a study team member or by mail.

Data were entered into SPSS (IBM SPSS Statistics, Version 23) including blinded double entry of 20% of the data for quality control.

**Patient and public involvement.** The ORIC measure preliminary addresses HCPs. Physicians, nurses and psychooncologists were involved in the adaptation of the measure by taking part in cognitive interviews. Patients were not involved in this study.

**Data Analyses**. Descriptive statistics were calculated for demographic characteristics. Cases were excluded if more than 30% of the ORIC items were missing [60]. For all other cases, missing data were replaced with item means. We evaluated the completion rate and therefore the acceptance of the measure by calculating frequencies of missing data per item as well as for the overall measure. For this analysis, we also included cases with more than 30% of ORIC items missing because these values are part of completion rate and relevant for interpretation of acceptance.

We a priori hypothesized to replicate the theory-based two-dimensional structure of the original English ORIC version. Two correlating factors "change commitment" (item 1 to 5) and "change efficacy" (item 6 to 10) were postulated. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett`s test of sphericity were performed to test prerequisites for factor analysis [61,62]. A confirmatory factor analysis (CFA) with Maximum Likelihood Estimates and two factors was applied for the whole data set as a first step. Because the two-factor model could not be confirmed, we decided to calculate an exploratory factor analysis (EFA) and afterwards an additional CFA to check for model fit. It is recommended to not calculate EFA and CFA with the same data set so the data set was randomized by AL and split into two subsets [63,64]. The first 115 randomized cases including all data of participants were added to EFA, the second 115 cases were added to CFA. An EFA with oblique rotation was calculated for the first subset. The non-orthogonal rotation was chosen according to Weiner et al. [40]. In their theory, organizational readiness for change consists of two interrelated dimensions, therefore the two factors are expected to be correlated. Analogue to analyses done by authors of the English ORIC [43,65], we extracted components based on parallel analysis. The criterion of parallel analysis was shown to be superior to other statistic criteria like the Kaiser criterion [66,67]. It compares the eigenvalues of the data to the eigenvalues based on random data with equivalent sample size and number of variables and chooses only factors with eigenvalues higher than for random data [65,66]. A CFA was calculated for the second subset. A range of global goodness of

fit indices were used to assess the degree to which observed data were accounted for by the proposed models: discrepancy chi-squared statistic ($Chi^2$), degree of freedom (df), normed chi-squared statistic ($Chi^2/df$), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA) as well as Akaike Information Criterion (AIC) and Parsimonious Normed Fit Index (PNFI) for analysing model complexity. Established rules to estimate the model fit were used [68–71].

Item analyses was performed for the one-factor model. It included calculation of item means and standard deviation as well as observation of floor and ceiling effects [72], calculation of corrected item-total correlations [61,73], inter-item correlations [61,73], and item difficulties [62]. Internal consistency of the scale was assessed by Cronbach's alpha coefficient (α) [62,73,74]. For a detailed overview on performed data analyses, see Table 1.

During the translation process and cognitive interviews we found low content validity for item 10 (see results section). Thus, the use of item 10 for the German ORIC needs to be evaluated. Accordingly, we also conducted psychometric analyses (EFA with oblique rotation and extraction of components based on parallel analysis, corrected item-total correlations, Cronbach's α, and goodness of fit indexes) for the 9-item version of the ORIC.

Analysis of demographic data, analysis of completion rate, item analysis and EFA were performed using SPSS (IBM SPSS Statistics, Version 23). CFA and calculation of model fit indices were performed using Amos (IBM SPSS Amos 22.0.0).

Table 1: Psychometric analyses conducted.

| **Psychometric measure** | **Criteria** |
|---|---|
| Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett`s test of sphericity | These tests ensure that correlations between variables can be accounted for by a smaller set of factors [61]. KMO value should be higher than .05 and Bartlett`s test value should be less than .05 to fulfil the criteria for calculating a factor analysis [61,62]. |
| Normed chi-squared statistic ($Chi^2/df$) | $Chi^2/df$ is an indicator for model fit, dependent on sample size and should be as small as possible. A ratio between 2 and 3 indicate a good data fit [68]. |
| Comparative fit indexes (CFI) | CFIs is an indicator for model fit. It ranges from 0 to 1 and higher values indicate better fit. Values above .95 indicate a good model fit [71,75]. |
| Tucker-Lewis Index (TLI) | TLI is an indicator for model fit. It corrects for complexity of the model and is sensitive to small sample sizes. Values above .95 indicate good fit [69]. |

| | |
|---|---|
| Root mean square error of approximation (RMSEA) | RMSEA is an absolute index which describes closeness to fit. Values below .05 indicate a good fit, values between .05 and .08 indicate an adequate fit, values between .08 and 1 indicate a moderate fit and values above 1 are unacceptable [76]. |
| Akaike Information Criterion (AIC) | AIC is a parsimony model fit index. It can be used to compare fit of competing models with smaller values indicating better fit [68,71]. |
| Parsimonious Normed Fit Index (PNFI) | PNFI is a parsimony model fit index. It ranges between 0 and 1 and higher values indicate a more parsimonious fit [68]. No threshold levels are recommended and it has to be analyzed in combination with other goodness of fit indices [71]. |
| Analysis of frequencies for item response distributions | Floor and ceiling effects were assumed present if more than 15% of participants choose the lowest or highest possible score [72]. |
| Corrected item-total correlations | If items correlate with the total score of above .30, they measure the same underlying concept. Items with lower correlations should be removed because they do not add exploratory power to the measure [61,73]. |
| Item difficulties | Item difficulties are calculated by dividing item means by the maximal value of the answer range (0-4) and multiplying it with 100. Item difficulty should be near to 50%, and items should not differ much in their difficulty level [62]. |
| Inter-item correlations | Inter-item correlations ensure association between items. High inter-item correlations of above .80 indicate that items ask the same questions and might be redundant [61,73]. |
| Cronbach's α | Cronbach's α is a measure for reliability and internal consistency. A value of at least .70 is acceptable and higher coefficients indicate a more stable measure [61,62,74]. |

## RESULTS

To report the results of this validation study, we used the Authors' Guidelines for Reporting Scale Development and Validation Results by Cabrera-Nguyen (see Supplementary File 1).

### Translation

Both translators (AL and SZ) and the reviewer (IS) did not differ much in their translations of items 2 to 5 and 8 as well as the response scale. For these items and the response scale, only the choice of single words differed without differences in meaning. Greater translation differences were found for items 1, 6, 7, 9, and 10. For item 1, the word "committed" was differently translated. For items 6 and 7, differences were found in the translation of the phrase "feel confident" and the sentence structure. For item 9, differences mainly addressed translation of the phrases "feel confident" and "adjust to this change". For item 10,

differences occurred in the translation of the term “manage the politics” and the sentence structure. Within the first round of team discussion, we reached consensus for items 2 to 9, the translation of the response scale and the survey introduction. For item 1 we suggested two versions to be further tested in subsequent cognitive interviews. We struggled to translate the phrase “manage the politics” in item 10 into German. Therefore, we discussed item 10 with additional colleagues (cp. methods section) until consensus was found.

**Assessment of comprehensibility as part of content validity and subsequent adaptation of the scale**

To test the German ORIC for comprehensibility, cognitive interviews with n=11 participants (nurses, physicians, and psychooncologists) were conducted. Cognitive interviews lasted about one hour. For demographic data of participants see Table A in Supplementary File 2.

After the first round of cognitive interviews (n=7), no changes have to be made to the response scale as well as for items 2 to 5 and 8 because these items were already well understood by participants. Participants made some minor suggestions for modifications for the introductory description and items 1 and 6. Additionally, some participants did not understand the correct meaning of items 7, 9 and 10 in general or of single words or phrases of these items. After discussions and modifications of these items, we tested alternative versions of the survey introduction, for items 1, 6, 9 and 10 as well as two alternative versions of item 7. After the second round of cognitive interviews (n=4), items 1, 6 and 9 were now understood well by all participants. We had to slightly modify the survey introduction again and decided to use the version of item 7 which was understood best. After all, item 10 could not be translated successfully. Both rounds of cognitive interviews showed that comprehension of the German translation of the phrase “manage the politics” did not picture the correct English meaning. Thus, in a next step we consulted with DF (cp. list of abbreviations) and reached consensus on a final version. Nevertheless, the final version of item 10 was still not satisfying from the study team and experts view. Item 10 was found to have low comprehensibility as part of content validity according to COSMIN criteria [54]. The final German ORIC measure, used in this study, is presented in Supplementary File 3.

During cognitive interviews some nurses reported that they had not heard about the term “shared decision-making” (German: “Partizipative Entscheidungsfindung”) prior to participation. Thus, we provided a definition of SDM in the introduction part of the questionnaire within the SDM implementation study [56].

## Psychometric evaluation

**Sample characteristics.** Data of 235 HCPs were available for this secondary analysis. In line with recommendations of Bannon [60], five cases (0.02% of all cases) were excluded (except for assessment of completion rate), because all items of the ORIC were missing. Missing values were replaced by means and data of 230 HCPs could be included into analyses.

Table 2 provides an overview of participants' demographic characteristics. Most of the 230 HCPs were between 31 and 40 years old (37.0%), female (70.4%), worked as a nurse (57.0%), and had a work experience of < 5 years (43.9%).

Table 2: Demographic characteristics of participants (n=230).

| | | N | % |
|---|---|---|---|
| Age | < 30 years | 72 | 31.3 |
| | 31-40 years | 85 | 37.0 |
| | 41-50 years | 42 | 18.3 |
| | > 50 years | 26 | 11.3 |
| | Missings | 5 | 2.2 |
| Gender | Female | 162 | 70.4 |
| | Male | 59 | 25.7 |
| | Different gender or preferred not to answer this question | 5 | 2.1 |
| | Missings | 4 | 1.7 |
| Profession | Nurse | 131 | 57.0 |
| | Junior physician | 69 | 30.0 |
| | Senior physician | 27 | 11.7 |
| | Missings | 3 | 1.3 |
| Work experience in healthcare | < 5 years | 101 | 43.9 |
| | 5-10 years | 48 | 20.9 |
| | 11-20 years | 46 | 20.0 |
| | > 20 years | 28 | 12.2 |
| | Missings | 7 | 3.0 |

**Factor analysis.** Requirements for factor analysis were met [54]. Sample size was large enough (>100), even for a split data set with n=115. Furthermore, no outliers were found and data values were approximately normally distributed. KMO measure was .933 and Barlett´s test of sphericity yielded $X^2$ = 1485.11, p < .001. This indicates that a factor analysis of the data was appropriate [61,62]. CFA for the hypothesised two-factor model showed a high correlation of .87 between the two components (see Supplementary File 4). Therefore, we postulated a one-factorial structure and conducted a post-hoc EFA. As shown in Table 3, only the first component had an eigenvalue higher than 95% percentile of the eigenvalues of corresponding random data and the main component explains 67.23% of the variance.

Thus, according to parallel analysis, a one-factor model was assumed. The factor loadings for the first component were above 0.754 for all items (see Table B of Supplementary File 2).

Table 3: Results of EFA with oblique rotation and parallel analysis: eigenvalues of the ten components of the German ORIC and eigenvalues for corresponding random data.

| | Eigenvalues | | | Eigenvalues for random data | |
|---|---|---|---|---|---|
| | Total | % of variance | Cumulative % | Means | 95% percentile |
| Component 1 | 6.72 | 67.23 | 67.23 | 1.49 | 1.65 |
| Component 2 | 0.83 | 8.30 | 75.53 | 1.33 | 1.44 |
| Component 3 | 0.47 | 4.75 | 80.28 | 1.21 | 1.30 |
| Component 4 | 0.41 | 4.08 | 84.36 | 1.11 | 1.19 |
| Component 5 | 0.39 | 3.91 | 88.28 | 1.02 | 1.08 |
| Component 6 | 0.32 | 3.24 | 91.52 | 0.93 | 1.00 |
| Component 7 | 0.27 | 2.74 | 94.26 | 0.85 | 0.92 |
| Component 8 | 0.23 | 2.34 | 96.60 | 0.77 | 0.84 |
| Component 9 | 0.17 | 1.74 | 98.35 | 0.68 | 0.75 |
| Component 10 | 0.16 | 1.65 | 100.00 | 0.58 | 0.66 |

Notes: For EFA, half of the data set (n=115) was used.

A second CFA was performed with the one-factor model to analyze its fit indices. Indices of the two-factor model and the one-factor model are compared in Table 4.

Table 4: Fit indices of two calculated models for factor analysis of the German ORIC.

| | Chi² [1] | df [2] | Chi²/df [3] | CFI [4] | TLI [5] | RMSEA [6] | AIC [7] | PNFI [8] |
|---|---|---|---|---|---|---|---|---|
| Two-factor model | 81.71* | 34 | 2.40 | .968 | .947 | .078 | 143.71 | .585 |
| One-factor model | 77.19* | 35 | 2.20 | .928 | .907 | .103 | 117.19 | .682 |

Notes: Two-factor model was calculated for the whole data set (n=230): factor 1 includes item 1 to 5, factor 2 includes item 6 to 10; One-factor model was calculated for half of the data set (n=115): includes items 1 to 10.
[1] discrepancy chi-squared statistic, [2] degrees of freedom, [3] normed chi-squared statistic, [4] comparative fit indexes, [5] Tucker-Lewis Index, [6] root mean square error of approximation, [7] Akaike Information Criterion, [8] Parsimonious Normed Fit Index (PNFI). * $p$ = .000

Results of factor analysis for the 9-item version of the ORIC (without item 10) were similar. Also for the 9-item version, a one-factor model was assumed by exploratory factor analysis. Only the first component had an eigenvalue higher than 95% percentile of the eigenvalues of corresponding random data and the main component explains 66.85% of the variance. Factor loadings of the first component are above 0.739 for all items (see Supplementary File 5, Tables A and B). Goodness of fit indices of the one-factor model of the

9-item ORIC version showed similar values compared to the 10-item ORIC version (see Supplementary File 5, Table C).

**Analysis of the ORIC items and internal consistency.** Table 5 shows response distribution, means, standard deviations, acceptance, corrected item-total correlation, and item difficulty of the ten items. Means ranged between 2.24 (item 9) and 2.61 (item 5). Most participants responded in the middle of the scale with a slight shift to more agreement. For items 1 to 9, between four and six missing values could be detected. For item 10, nine missing values were found. Taking all items into account, more than 97% of the measure were answered. Corrected item-total correlations ranged from .665 (item 9) to .774 (item 3), item difficulties from 55.98 (item 9) to 65.32 (item 5), and inter-item correlations from .434 (item 2 and item 9) to .723 (item 3 and item 5) (see Table C of Supplementary File 2). Internal consistency yielded a Cronbach's α of .93. Additionally, corrected item-total correlations and internal consistency were calculated for the 9-item ORIC version (see Supplementary File 5, Table D). They were similar to the results for the 10-item version with corrected item-total correlations between .638 (item 9) and .777 (item 3) and a Cronbach's α of .92.

Table 5: Response distribution, means, standard deviation, acceptance, discrimination and item difficulty of the German ORIC.

| | Items | Disagree N (%) | Somewhat Disagree N (%) | Neither Agree nor Disagree N (%) | Somewhat Agree N (%) | Agree N (%) | Mean (SD) | Acceptance (Comple-tion rate in %) * | Item discrimination (corrected item-total correlation) | Item difficulty |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | People who work here are committed to implementing shared decision-making. | 1 (0.4) | 22 (9.6) | 109 (47.0) | 73 (30.9) | 25 (10.9) | 2.42 (.826) | 97.43 | .744 | 60.58 |
| 2 | People who work here will do whatever it takes to implement shared decision-making. | 4 (1.7) | 37 (16.1) | 103 (44.8) | 68 (29.6) | 18 (7.4) | 2.25 (.878) | 98.29 | .689 | 56.36 |
| 3 | People who work here want to implement shared decision-making. | 0 (0.0) | 15 (6.5) | 107 (46.5) | 84 (35.7) | 24 (10.4) | 2.50 (.768) | 97.43 | .774 | 62.61 |
| 4 | People who work here are determined to implement shared decision-making. | 2 (9.0) | 38 (16.5) | 107 (46.5) | 67 (29.1) | 16 (7.0) | 2.25 (.843) | 98.29 | .758 | 56.19 |
| 5 | People who work here are motivated to implement shared decision-making. | 1 (0.4) | 16 (7.0) | 85 (37.0) | 97 (42.2) | 31 (13.5) | 2.61 (.821) | 98.29 | .764 | 65.32 |
| 6 | People who work here feel confident that they can handle the challenges that might arise in implementing shared decision-making. | 2 (0.9) | 20 (8.7) | 93 (40.4) | 93 (40.4) | 22 (9.6) | 2.49 (.819) | 98.29 | .760 | 62.28 |
| 7 | People who work here feel confident that they can keep track of progress in implementing shared decision-making. | 1 (0.4) | 26 (11.3) | 93 (40.4) | 92 (40.0) | 18 (7.8) | 2.43 (.811) | 98.29 | .725 | 60.87 |
| 8 | People who work here feel confident that they can coordinate tasks so that implementation goes smoothly. | 5 (2.2) | 24 (10.4) | 107 (46.5) | 78 (33.5) | 16 (6.5) | 2.32 (.833) | 99.56 | .697 | 58.13 |
| 9 | People who work here feel confident that the organization can support people as they adjust to shared decision-making. | 6 (2.6) | 43 (18.7) | 89 (38.7) | 74 (32.2) | 18 (7.8) | 2.24 (.934) | 97.86 | .665 | 55.98 |
| 10 | People who work here feel confident that they can manage the politics of implementing shared decision-making. | 3 (1.3) | 24 (10.04) | 122 (50.9) | 65 (28.3) | 16 (7.0) | 2.29 (.796) | 96.15 | .714 | 57.44 |

Notes: Items could be answered on a 5-step Likert scale rating from 0 „disagree" to 4 „agree". SD = standard deviation. * For calculation of completion rate, five additional cases were included because these participants only skipped the ORIC but filled out the rest of the survey.

## DISCUSSION

The original English ORIC measure is a brief measure with good psychometric properties [43], which were confirmed in Danish [45] and French [46] validation studies. The study at hand aimed to translate the ORIC into German and assess its psychometric properties.

### Translation and assessment of comprehensibility as part of content validity

Items 1 to 9 were translated and adapted successfully after two rounds of cognitive interviews and several rounds of discussions within the study team and with external experts. The translation team quickly reached consensus for items 2 to 5 and 8. These items were also well understood by all participants within the first round of cognitive interviews. For items 1, 6, 7 and 9, the translation process was more complex and several adaptations and discussions were necessary. Feedback by participants, members of the study team, and external experts as well as completion rates suggest that comprehensibility of item 10 seems to be low [54]. This might be due to the translation of the phrase “manage the politics” into “Machenschaften”. The term “manage the politics” seems to have a strong cultural connotation and no equivalent phrase in German language exists. The German term “Machenschaften” might have a different connotation as the English phrase and might lead to skipping the item. Ruest et al. [46], who translated the English ORIC into French, also identified several differences in cultural concepts during their adaptation process, but could translate all items successfully. They concluded that limitations in linguistic validation could decrease comparability of psychometric results of the translated measure. However, item 10 showed similar and inconspicuous item characteristics compared to other items in our sample. When repeating factor and item analyses for the 9-item ORIC version including only item 1 to 9, very similar results were observed compared to the 10-item version. To increase comprehensibility and thereby content validity of the scale, the use of the 9-item German ORIC might be a solution and should be evaluated in future studies.

### Factor analysis

We a priory hypothesized a two-factorial structure of the German ORIC, because Shea et al. [43] described correlations between the two theory-based factors “change commitment” and “change efficacy” of .56 to .60. However, we found much higher factor correlations of 0.87. Results of the subsequent EFA clearly indicated a one-factorial structure. Thus, we could not confirm the two-factor structure of the English and the translated Danish and French versions of ORIC [43,45,46]. When comparing the two models, both

models have acceptable values for $Chi^2/df$ [68] and CFI [71,75], but only the two-factor model has acceptable values for TLI [69] and RMSEA [76]. When involving parsimony of the models by calculating AIC and PNFI, the one-factor model fits better to the data [68,71]. Therefore, we prefer the more parsimonious one-factor model. These differences to previous validation studies might be a consequence of diverse cultural connotations of the ORIC items in different languages, caused by the adaptation to the context of SDM, or due to specific characteristics of the participating clinics.

**Analysis of ORIC items and internal consistency**

Since the ORIC was presented as the last measure in a three-page survey, missing values might indicate respondent fatigue. However, missing value rates for single items and the overall measure were quite low and the German ORIC was found to be a well-accepted measure. There were no floor or ceiling effects. Corrected item-total correlations of above .66 indicate that all items measure the same underlying concept [61,73]. Criteria for good item difficulties are met since item difficulties are near to 50% and do not differ much from each other [62]. Inter-item correlations are below .80, indicating that items add additional information and are not redundant [61,73]. Cronbach's α (α=.93) suggest excellent internal consistency [61,62,74]. In summary, item analysis and internal consistency of the German ORIC suggest good quality of the measure.

Nevertheless, according to Streiner and Norman [62] a Cronbachs' α above .90 might also indicate item redundancy. On the other hand, inter-item correlations and corrected item-total correlations are in an acceptable range [61,73]. In implementation research there is a need for preferably brief measures, which can be applied in diverse settings with high work-load. Thus, future research could investigate the possibility to further reduce the number of items.

**Strengths and Limitations**

This study has some limitations. First, several psychometric parameters are not analysable because this study was a secondary analysis of cross-sectional data. It was not possible to calculate e.g. convergent or divergent validity yet. Second, we applied the ORIC only in three departments of one University Medical Center in Germany. Further validation in different organizational settings is needed to ensure generalizability. Third, for this psychometric evaluation we used a German ORIC, which we adapted and specified for the context of SDM implementation. Our results might not be generalizable for other

interventions in other organizations. Fourth, although SDM was not implemented to the participating clinics before, there might be participants who were more familiar with the concept of SDM than others. Fifth, item 10 was again slightly changed after finishing cognitive interviews. This item was not finally tested for comprehensibility.

A major strength of this study is that we provided the first measure to assess organizational readiness for change in German language for use in implementation studies. We conducted an elaborated translation procedure, which was recommended for survey translations. We furthermore used a qualitative approach to explore comprehensibility including discussions with international colleagues and experts outside of the study team. Furthermore, we assessed the ORIC in a sample including physicians and nurses which was large enough to robustly perform the psychometric analysis on the German version of the ORIC measure.

## CONCLUSION

Organizational readiness is a crucial indicator to successfully implement change and a possible barrier if missing. For implementation studies, it is essential to measure organizational readiness with valid and reliable measures. We provide the first German measure for organizational readiness for implementing change and validated it for the context of SDM implementation. The German ORIC is a brief measure with a high completion rate. We found satisfying psychometric properties in a German hospital setting. To increase content validity of the measure, the use of a 9-item German ORIC (without item 10) should be evaluated in future studies. As the ORIC targets the attitude of organizational members, it can detect reduced or missing readiness for implementing a change on the individuals' level. Therefore, the German ORIC can be used to analyze organizational readiness as a possible barrier for implementing various interventions in organizations.

## LIST OF ABBREVIATIONS

AH: Anne Haugstvedt

AIC: Akaike Information Criterion

AL: Anja Lindig

AMOS: Analysis of Moment Structure, Statistical Package for the Social Sciences, International Business Machines Corporation

$Chi^2$ : Discrepancy chi-squared statistic

Chi2/df: Normed chi-squared statistic

CFI: Confirmatory fit index

CS: Christopher M. Shea

Df: Degree of freedom

DF: Dominick Frosch

EC: Eva Christalle

HCP: Healthcare professional

IS: Isabelle Scholl

KMO: Kaiser-Meyer-Olkin criterion

MM: Marcel Machalski

MR: Mélanie Ruest

ORIC: Organizational Readiness for Implementing Change

PH: Pola Hahlweg

PNFI: Parsimonious Normed Fit Index

RMSEA: Root mean square error of approximation

SDM: Shared decision-making

SPSS: Statistical Package for the Social Sciences, International Business Machines Corporation

SZ: Stefan Zeh

TLI: Tucker-Lewis Index

TRAPD: Translation, Review, Adjudication, Pretesting, Documentation

## DECLARATIONS

### Ethics approval and consent to participate

The study was approved by the Ethics Committee of the Medical Association Hamburg (Germany, study ID PV5368). The study was carried out in accordance to the latest version of the Helsinki Declaration of the World Medical Association. Principles of good clinical practice were respected. Data protection requirements were met. Study participation was voluntary. A waiver of consent for HCPs was obtained from the Ethics Committee, as proposed by current statements on ethical designs of implementation research [77]. HCPs were able to decline participation in the study.

**Funding**

This study is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, grant number 232160533).

**Competing interests**

PH gave one scientific presentation on shared decision-making during a lunch symposium, for which she received compensation and travel compensation from GlaxoSmithKline GmbH in 2018. AL, EC and IS declared to not have any competing interests.

**Author contributions**

AL, PH and IS made substantial contributions to the design and preparation of the study. AL and PH collected the data. AL conducted the analysis in collaboration with EC. All authors contributed to the interpretation of results. AL drafted the manuscript and PH, IS and EC were involved in critically revising the manuscript for important intellectual content. All authors gave final approval of the version to be published.

**Checklist for reporting statement**

To report the results of this validation study, we used the Authors' Guidelines for Reporting Scale Development and Validation Results by Cabrera-Nguyen.

**Patient consent form**

Not applicable

**Data sharing statement**

The dataset collected and analyzed during this study is available from the corresponding author on reasonable request.

**Acknowledgements**

We thank our student assistants Anastasia Izotova, Sophia Schulte and Nicolai Pergande for their help preparing the study and the analysis. We thank Stefan Zeh for his part the team translation process. We

thank Dominick Frosch, Anne Haugstvedt, Marcel Machalski, Mélanie Ruest and Christopher M. Shea for helpful comments and suggestions in the translation and adaptation process of item 10. We also would like to thank Christopher M. Shea for giving the opportunity to translate the ORIC into German and to use it in our implementation study.

**Supplementary Files**

Supplementary File 1: Reporting Checklist: Authors Guidelines for Reporting Scale Development and Validation Results

Supplementary File 2: Additional tables for psychometric evaluation of the 10-item ORIC version

Supplementary File 3: Organizational Readiness for Implementing Change (ORIC) – German version

Supplementary File 4: Confirmatory factor analysis model of the 10-item German ORIC

Supplementary File 5: Results of psychometric evaluation of the 9-item German ORIC version

**REFERENCES**

1. Øvretveit J, Andreen-Sachs M, Carlsson J, Gustafsson H, Hansson J, Keller C, et al. Implementing organisation and management innovations in Swedish healthcare. *J Health Organ Manag* 2012;26:237–57.

2. Montgomery A, Doulougeri K, Panagopoulou E. Implementing action research in hospital settings: a systematic review. *J Health Organ Manag* 2015;29:729–49.

3. Rye CB, Kimberly JR. The Adoption of Innovations by Provider Organizations in Health Care. *Med Care Res Rev* 2007;64:235–78.

4. Bazzoli GJ, Dynan L, Burns LR, Yap C. Two Decades of Organizational Change in Health Care: What Have we Learned? *Med Care Res Rev* 2004;61:247–331.

5. Weiner BJ, Amick H, Lee S-YD. Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields. *Med Care Res Rev* 2008;65:379–436.

6. Batras D, Duff C, Smith BJ. Organizational change theory: Implications for health promotion practice. *Health Promot Int* 2016;31(1):231–41.

7. Yazdani A, Wells R. Barriers for implementation of successful change to prevent musculoskeletal disorders and how to systematically address them. *Appl Ergon* 2018;73:122–40.

8. Schelvis RMC, Wiezer NM, Blatter BM, van Genabeek JAGM, Oude Hengel KM, Bohlmeijer ET, et al. Evaluating the implementation process of a participatory organizational level occupational health intervention in schools. *BMC Public Health* 2016;16:1212.

9. Scott DAH, Grant SM. A meta-ethnography of the facilitators and barriers to successful implementation of patient complaints processes in health-care settings. *Heal Expect* 2018;21:508–17.

10. Williams B, Perillo S, Brown T. What are the factors of organisational culture in health care settings that act as barriers to the implementation of evidence-based practice? A scoping review. *Nurse Educ Today* 2015;35:e34–41.

11. Härter M, Dirmaier J, Scholl I, Donner-Banzhoff N, Dierks M-L, Eich W, et al. The long way of implementing patient-centered care and shared decision making in Germany. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:46–51.

12. Bieber C, Loh A, Ringel N, Eich W, Härter M. Patientenbeteiligung bei medizinischen Entscheidungen. Manual zur partizipativen Entscheidungsfindung (Shared Decision-Making). Freiburg: Universitätsklinikum Freiburg; 2007.

13. Härter M, Buchholz A, Nicolai J, Reuter K, Komarahadi F, Kriston L, et al. Shared decision making and the use of decision aids. *Dtsch Arztebl Int* 2015;112:672–9.

14. Salzburg Global Seminar. Salzburg statment on shared decision making. *BMJ* 2011;342:1745.

15. Reuter K, Loh A, Härter M. Patient as partner in oncology - Opportunities for shared decision-making. In: Koch U, Weis J, eds. Psychoonkologie - Eine Disziplin in der Entwicklung. Göttingen: Hogrefe; 2009:131–45.

16. Hahlweg P, Kriston L, Scholl I, Brähler E, Faller H, Schulz H, et al. Cancer patients' preferred and perceived level of involvement in treatment decision-making. [submitted]

17. Hauser K, Koerfer A, Kuhr K, Albus C, Herzig S, Matthes J. Outcome-relevant effects of shared decision making. *Dtsch Arztebl Int* 2015;112:665–71.

18. Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. *Med Decis Mak* 2015;35:114–31.

19. Härter M, Müller H, Dirmaier J, Donner-Banzhoff N, Bieber C, Eich W. Patient participation and shared decision making in Germany - History, agents and current transfer to practice. *Z Evid Fortbild Qual Gesundhwes* 2011;105:263–70.

20. Härter M, Moumjid N, Cornuz J, Elwyn G, van der Weijden T. Shared decision making in 2017:

International accomplishments in policy, research and implementation. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:1–5. A

21. Härter M, van der Weijden T, Elwyn G. Policy and practice developments in the implementation of shared decision making: An international perspective. *Z Evid Fortbild Qual Gesundhwes* 2011;105:229–33.

22. Fujimori M, Uchitomi Y. Preferences of Cancer Patients Regarding Communication of Bad News: A Systematic Literature Review. *Jpn J Clin Oncol* 2009;39:201–16.

23. Ernst J, Brahler E, Weissflog G. Patient involvement in medical decision making--an overview on patient preferences and impacting factors. *Gesundheitswesen* 2014;76:187–92.

24. Chewning B, Bylund CL, Shah B, Arora NK, Gueguen JA, Makoul G. Patient preferences for shared decisions: A systematic review. *Patient Educ Couns* 2012;86:9–18.

25. Brom L, Hopmans W, Pasman HR, Timmermans D, Widdershoven G, Onwuteaka-Philipsen B. Congruence between patients' preferred and perceived participation in medical decision-making: A review of the literature. *BMC Med Inform Decis Mak* 2014;14:25.

26. Légaré F, Stacey D, Turcotte S, Cossi MJ, Kryworuchko J, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2014;9:CD006732.

27. Légaré F, Witteman HO. Shared decision making: Examining key elements and barriers to adoption into routine clinical practice. *Health Aff* 2013;32:276–84.

28. Frosch D, Tietbohl C, Scholl I. Overcoming implementation challenges to advance shared-decision making in routine practice. In: Elwyn G, Thompson R, Edwards A, eds. Shared Decision Making Health Care Achieve Evidence-Based Patient Choice, 3rd Ed. Oxford: Oxford University Press; 2016.

29. Grol R, Wensing M, Eccles M, Davis D. Improving patient care: The implementation of change in health care. Hoboken (NJ): John Wiley & Sons; 2013.

30. Grol R, Grimshaw J. From best evidence to best practice: Effective implementation of change in patients' care. *Lancet* 2003;362:1225–30.

31. Cummings GG, Estabrooks CA, Midodzi WK, Wallin L, Hayduk L. Influence of Organizational Characteristics and Context on Research Utilization. *Nurs Res* 2007;56:S24–39.

32. Li S-A, Jeffs L, Barwick M, Stevens B. Organizational contextual features that influence the implementation of evidence-based practices across healthcare settings: a systematic integrative review.

*Syst Rev* 2018;7:72.

33. Légaré F, Ratté S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions. *Patient Educ Couns* 2008;73:526–35.

34. Frerichs W, Hahlweg P, Müller E, Adis C, Scholl I. Shared decision-making in oncology – a qualitative analysis of healthcare providers' views on current practice. PLoS One 2016;11:e0149789.

35. Scholl I, LaRussa A, Hahlweg P, Kobrin S, Elwyn G. Organizational- and system-level characteristics that influence implementation of shared decision-making and strategies to address them - a scoping review. *Implement Sci* 2018;13:40.

36. André B, Sjøvold E. What characterizes the work culture at a hospital unit that successfully implements change - a correlation study. *BMC Health Serv Res* 2017;17:486.

37. Armenakis AA, Harris SG, Mossholder KW. Creating Readiness for Organizational Change. *Hum Relations* 1993;46:681–703.

38. Kash BA, Spaulding A, Gamm LD, Johnson C. Leadership, culture, and organizational technologies as absorptive capacity for innovation and transformation in the healthcare sector: A framework for research. *Chang Manag* 2014;13:1–13.

39. Armenakis AA, Harris SG, Mossholder KW. Creating Readiness for Organizational Change. *Hum Relations* 1993;46:681–703.

40. Weiner BJ. A theory of organizational readiness for change. *Implement Sci* 2009;4:67.

41. Armenakis AA, Harris SG. Crafting a change message to create transformational readiness. *J Organ Chang Manag* 2002;15:169–83.

42. Holt D, Armenakis A, Harris S, Field H. Toward a Comprehensive Definition of Readiness for Change: A Review of Research and Instrumentation. *Res Organ Chang Dev* 2007;16:289–336.

43. Shea CM, Jacobs SR, Esserman DA, Bruce K, Weiner BJ. Organizational readiness for implementing change: A psychometric assessment of a new measure. *Implement Sci* 2014;9:7.

44. Gagnon M-P, Attieh R, Ghandour EK, Légaré F, Ouimet M, Estabrooks CA, et al. A Systematic Review of Instruments to Assess Organizational Readiness for Knowledge Translation in Health Care. Jeyaseelan K, editor. *PLoS One* 2014;9:e114338.

45. Storkholm MH, Mazzocato P, Tessma MK, Savage C. Assessing the reliability and validity of the Danish version of Organizational Readiness for Implementing Change (ORIC). *Implement Sci*

2018;13:78.

46. Ruest M, Leonard G, Thomas A, Desrosiers J, Guay M. French cross-cultural adaptation of the Organizational Readiness for Implementing Change (ORIC). *BMC Health Serv Res* 2019;19:535.

47. Alexander JA, Hearld LR. Methods and metrics challenges of delivery-system research. *Implement Sci* 2012;7:15.

48. Harkness J, Pennell BE, Schoua-Glusberg A. Survey Questionnaire Translation and Assessment. In: Presser S, Rothgeb J, Couper M, Lessler J, Martin E, Martin J, et al., eds. Methods for Testing and Evaluating Survey Questionnaires. Hoboken (NJ): JohnWiley & Sons; 2004:453–73.

49. Forcino RC, Bustamante N, Thompson R, Percac-Lima S, Elwyn G, Pe´rez-Arechaederra D, et al. Developing and Pilot Testing a Spanish Translation of CollaboRATE for Use in the United States. *PLoS One* 2016;11:e0168538.

50. Epstein J, Osborne RH, Elsworth GR, Beaton DE, Guillemin F. Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *J Clin Epidemiol* 2015;68:360–9.

51. Perneger TV, Leplège A, Etter JF. Cross-cultural adaptation of a psychometric instrument: two methods compared. *J Clin Epidemiol* 1999;52:1037–46.

52. Ponce NA, Lavarreda SA, Yen W, Brown ER, DiSogra C, Satter DE. The California Health Interview Survey 2001: translation of a major survey for California's multiethnic population. *Public Heal Rep* 2004;119:388–95.

53. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737–45.

54. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27:1159–70.

55. Willis GB, Research Triangle Institute. Cognitive interviewing - a "how to" guide. 1999.

56. Scholl I, Hahlweg P, Lindig A, Bokemeyer C, Coym A, Hanken H, et al. Evaluation of a program for routine implementation of shared decision-making in cancer care : study protocol of a stepped wedge cluster randomized trial. *Implement Sci* 2018;1–10.

57. Degner LF, Sloan JA, Venkatesh P. The Control Preferences Scale. *Can J Nurs Res* 1997;29:21–43.

628 58. Giguere A, Carmichael P-C, Coudert L, Renaud J-S, Légaré F, Witteman, H, et al. Development of
629 the IcanSDM scale to assess primary care clinicians' ability to adopt shared decision making. [submitted]
630 59. Hahlweg P, Härter M, Nestoriuc Y, Scholl I. How are decisions made in cancer care? A qualitative
631 study using participant observation of current practice. *BMJ Open* 2017;7:e016360.
632 60. Bannon WJ. Missing data within a quantitative research study: How to assess it, treat it, and why you
633 should care. *J Am Assoc Nurse Pr* 2015;27:230–2.
634 61. Rattray J, Jones MC. Essential elements of questionnaire design and development. *J Clin Nurs*
635 2007;16:234–43.
636 62. Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and
637 use. Oxford, UK: Oxford University Press 2008.
638 63. Fabrigar L, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor
639 analysis in psychological research. *Psychol Methods* 1999;4:272.
640 64. Worthington R. Scale development research. *Couns Psychol* 2006;34:806–38.
641 65. O'Connor BP. SPSS and SAS programs for determining the number of components using parallel
642 analysis and Velicer's MAP test. *Behav Res Methods, Instruments, Comput* 2000;23:396–402.
643 66. Hayton JC, David GA, Scarpello V. Factor Retention Decisions in Exploratory Factor Analysis: a
644 Tutorial on Parallel Analysis. *Organ Res Methods*. 2004;7:191–205.
645 67. Zwick WR, Velicer WF. Comparison of five rules for determining the number of components to retain.
646 *Psychol Bull* 1986;99:432–42.
647 68. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the Fit of Structural Equation Models:
648 Tests of Significance and Descriptive Goodness-of-Fit Measures. *Meth Psychol Res Online* 2003;8:23–
649 74.
650 69. Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria
651 versus new alternatives. *Struct Eq Model* 1999;6:1–55.
652 70. Hair J, Black W, Babin B, RE Anderson. Multivariate data analysis. Upper Saddle River, NJ: Pearson
653 Prentice Hall; 2009.
654 71. Hooper D, Coughlan J, Mullen M. Structural Equation Modelling: Guidelines for Determining Model
655 Fit. *Articles* 2008.
656 72. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status
657 surveys adequate? *Qual Life Res* 1995;4:293–307.

73. Ferketich S. Focus on psychometrics. Aspects of item analysis. *Res Nurs Health* 1991;14:165–8.

74. Bannigan K, Watson R. Reliability and validity in a nutshell. *J Clin Nurs* 2009 ;18:3237–43.

75. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238–46.

76. Browne MW, Cudeck R. Single Sample Cross-Validation Indices for Covariance Structures. *Multivariate Behav Res* 1989;24:445–55.

77. Hutton J, Eccles M, Grimshaw J. Ethical issues in implementation research: A discussion of the problems in achieving informed consent. *Implement Sci* 2008;3:52.

**Supplementary File 1: Checklist for reporting standards. Authors' guideline for Scale Development and validation results by Cabrera-Nguyen.**

| | **Guidelines by Cabrera-Nguyen** [1] | **Transfer to our study** |
|---|---|---|
| 1 | Precisely define the target construct. | See page 3 and 4 (Introduction section) |
| 2 | Justify the need for your new measure. For example, if measures of the construct exist in the literature, explain the value added by your new scale. How might the new measure enhance the substantive knowledge base or social work practice? | See page 4 (Introduction section) |
| 3 | Indicate that you have submitted your initial pool of items to expert review (Worthington & Whittaker, 2006). Report (a) the number of items in the preliminary pool; (b) the number of expert reviewers and their qualifications; and (c) any major changes to your initial item pool following the review (e.g., a substantial decrease in the number of items, changes to the original item response format, overhaul of item pool due to experts' assessment regarding content validity). | See page 5 and 6 (Methods section: Measure, Translation, Assessment of comprehensibility and adaptation of the scale) and page 9 and 10 (Results section: Translation, Assessment of comprehensibility and adaptation of the scale) |
| 4 | Report the name and version of the statistical software package used for all analyses. | See page 6 to 8 (Methods section: Assessment of comprehensibility and adaptation of the scale, Psychometric evaluation) |
| 5 | Identify and justify the sampling strategy (e.g., convenience, snowball) and sampling frame. Report standard sample demographic characteristics as well as other salient sample characteristics. | See page 6 and 7 (Methods section: Assessment of comprehensibility and adaptation of the scale, Psychometric evaluation) and page 8 to 10 (Results section: Assessment of comprehensibility and adaptation of the scale, Psychometric evaluation) and Additional file 3 |
| 6 | Discuss relevant data preparation and screening procedures. For instance, do the data meet the appropriate assumptions for factor analysis? If not, what actions were taken? Report tests of factorability if appropriate (e.g., report Bartlett's test of sphericity). | See page 10 to 12 (Results section: Psychometric evaluation) |
| 7 | Provide all dates of data collection. | See page 9 to 13 (Results section) |
| 8 | Avoid use of principal components analysis (PCA) as a precursor to CFA (Costello & Osborne, 2005; Worthington & Whittaker, 2006). Instead, start with EFA to assess the underlying factor structure and refine the item pool. EFA should be followed by CFA using a different sample (or samples) to evaluate the EFA-informed a priori theory about the measure's factor-structure and psychometric | An a priory hypothesized model for CFA was specified based on the original measures structure. The model could not be confirmed and an EFA was calculated afterwards. The same data set was used for |

| | | |
|---|---|---|
| | properties. (Costello & Osborne, 2005; Henson & Roberts, 2006; Worthington & Whittaker, 2006). For CFA, authors should specify an a priori hypothesized model and a priori competing models (Jackson, Gillaspy, & Purc-Stephenson, 2009). | EFA and CFA. For more information see page 10 to 12 (Results section). |
| 9 | Guidelines for reporting EFA results. | |
| | How large is a sample? One common rule of thumb is to ensure a person-to-item ratio of 10:1. Another rule of thumb is that N= 300 is usually acceptable (Worthington & Whittaker, 2006). However, some researchers have criticized these sample size rules of thumb, noting the appropriate sample size is dependent on the features of the gathered data. These researchers recommend obtaining the largest possible sample because the adequacy of the sample size cannot be determined until after the data have been analyzed (Henson & Roberts, 2006). | Sample size is n=115, person-to-item ratio 11.5:1 (see page 7: Methods section, Psychometric analysis) |
| | Run EFA . . . or not. Run a preliminary EFA to determine if further data collection is required based on the following criteria: (a) If communalities are greater than .50 or there are 10:1 items per factor with factor loadings of roughly .4, then a sample size of 150 to 200 is likely to be adequate; (b) If communalities are all at least .60 or there are a minimum of 4:1 items per factor with factor loadings above .6, then even smaller sample sizes may suffice; (Worthington & Whittaker, 2006). Report if additional data collection was necessary due to inadequate sample size. If so, report the new participants' sociodemographic characteristics and test for differences between groups using standard statistical procedures (e.g., t-tests). | EFA communalities are all above .70, therefore the sample size of n=115 can be determined as adequate and no additional data collection was necessary. For more information, see Supplementary File1). |
| | Give EFA details. Report the specific rotation strategy used (e.g. varimax, geomin). Justify the decision to use an orthogonal or oblique solution. One recommendation is to always begin with an oblique rotation, empirically assess factor intercorrelations, and report them before deciding upon a final rotation solution (Henson & Roberts, 2006; Worthington &Whittaker, 2006). Some researchers argue oblique rotation is always the best approach because (a) factor intercorrelations are the norm in social sciences and (b) both approaches yield the same result if the factors happen to be uncorrelated (Costello & Osborne, 2005). Conversely, other researchers contend that orthogonal rotation is preferable because fewer parameters are estimated—orthogonal rotation is more parsimonious and amenable to replication (Henson &Roberts, 2006). Similarly, some researchers warn against relying on a statistical software package's default settings to determine the appropriate type of oblique rotation (Henson & Roberts, 2006; Worthington &Whittaker, 2006). Others state that doing so is fine (Costello & Osborne, 2005, p.3). Given the lack of consensus, it is probably best to describe what you do and defend your approach on substantive grounds, if possible. | See page 6 to 8 (Methods section: Psychometric evaluation) |
| | Report the whole factor pattern/structure. Always report the whole factor pattern/structure matrix, including all of the items in the analysis. It is recommended that authors report | See page 10 to 12 (Results section: Psychometric |

| | | |
|---|---|---|
| | this information in a chart following the example provided by Henson and Roberts (2006) on page 411. | evaluation) and Supplementary File 1 |
| | Criteria for deleting (crossloaded) items. Report any deleted items and the criteria used for deletion. Crossloading items with values ≥ .32 on at least two factors should generally be candidates for deletion, especially if there are other items with factor loadings of .50 or greater (Costello & Osborne,2005). Rerun the EFA each time an item is deleted. | No crossloading items could be observed, so item deletion was not necessary. |
| | Criteria for number of factors. Report the number of factors retained and justify this decision using multiple criteria (eigenvalue > 1, scree test, parallel analysis, rejection of a factor with fewer than 3 items, etc). Reporting the eigenvalue > 1 rule alone is inadequate because it has been shown to among the least accurate criteria for assessing factor retention (Costello & Osborne, 2005; Henson & Roberts, 2006) | See page 6 to 8 (Methods section: Psychometric evaluation) |
| | Explained variance. Report the variance explained by the factors. | See page 10 to 12 (Results section: Psychometric evaluation) and Supplementary File 1 |
| | In general, describe your decisions. | See page 10 to 12 (Results section: Psychometric evaluation) and Supplementary File 1 |
| 10 | Guidelines for reporting CFA results. | |
| | Describe and justify the theoretical model. Report hypothesized factor structure. Provide theoretical and empirical justification (e.g., results of preliminary EFAs) for your hypothesis. In addition, report a priori competing models. | See page 6 to 8 (Methods section: Psychometric evaluation) |
| | Describe the parameterization. Provide a comprehensive description of the a priori parameter specification. Identify fixed parameters, free parameters, and constrained parameters. For example, indicate if you freed the errors of any items to correlate. | One factor loading was constrained to equal 1, the corresponding intercept was constrained equal to zero. The other factor loadings and intercepts were estimated. Errors of items were not freed to correlate. |
| | Include a figure. Include a figure of each CFA model being tested using Kline's (2005) graphical conventions if feasible. | See Supplementary File 3 |
| | Identification. Demonstrate model identification (e.g., df > 0; scaling of factors; assess and report the —t-rule; the two-indicator rule). Necessary and sufficient conditions for model identification may vary for certain types of CFA models. When in doubt, authors should consult Brown's (2006) CFA textor Kline's (2005) SEM text for guidance. | See page 10 to 12 (Results section: Psychometric evaluation) |
| | Select an estimator based on distributional patterns and assumptions. Report the estimator used (e.g., ML, | See page 6 (Methods section: Psychometric evaluation) |

| | | |
|---|---|---|
| | WLSMV) and justify your choice based on distributional assumptions. It is not appropriate to report that you relied on your statistical software's default setting. | |
| | Use multiple fit indices. After estimating a model, always report multiple fit indices (e.g., model X2, df, p, CFI/TLI, RMSEA, SRMR). Report all appropriate fit indices, not just those favorable to your hypotheses (Jackson et al., 2009). For example, do not report acceptable CFI and TLI scores while omitting a relevant fit index with a subobtimal value. | See page 8 (Methods section: Psychometric evaluation) and page 10 and 11 (Results section: Psychometric evaluation) |
| | What is acceptable fit? For model fit indices, authors should generally use the cut-off values recommended by Hu and Bentler (1999) and endorsed by Brown (2006), assuming ML estimation: a.CFI/TLI ≥ .95 b. RMSEA ≤ .06 c. SRMR ≤ .08 | See page 9, Table 1 (Methods section: Psychometric evaluation) |
| | Localized strain? When reporting model fit, include an assessment for localized areas of strain by examining standardized residuals. Standardized residuals greater than 1.96 (for p< .05) indicate areas of strain (Harrington, 2009). Report the absence of localized strain, if appropriate; otherwise, note localized areas of strain by reporting the relevant standardized residuals. | Standardized residuals do not indicate localized strains. |
| | Parameter estimates and SEs. When reporting factor loadings and other parameter estimates, always report the unstandardized estimates, their p values, and the standard errors. In addition, include the standardized estimates when appropriate. Be sure to report all parameter estimates, even those that are non significant (Brown, 2006; Jackson et al., 2009). | See Supplementary File 3 |
| | Assessing the validity of the factor solution. Comment on the new measure's convergent and discriminant validity based on parameter estimates. For instance, factor correlations ≥ .80 may indicate poor discriminant validity (Brown, 2006). In addition, strong factor loadings that do not crossload may indicate good convergent validity. One rule of thumb is that factor loadings < .40 are weak and factor loadings ≥ .60 are strong (Garson, 2010). | Exploratory factor analysis revealed a one-factor model, so there are no factor correlations. |
| | Other measures. Report squared multiple correlations and comment on the measure's reliability (e.g., report Raykov's Rho if appropriate) | See page 10 to 12 (Results section: Psychometric analysis) |
| | Respecification: Caution! Report any post-hoc respecifications to improve model fit based on modification indices. Justify the respecifications on theoretical or conceptual grounds (Jackson et al., 2009). Respecification to allow for correlated errors is not supportable without strong pragmatic justification(e.g., items contain similar words or phrases). Note that respecification precludes comparing the model with your a priori specified competing models. Report improvements in appropriate model fit indices for respecified models (e.g., chi-square difference test) | No respecifications to improve model fit were applied. |
| 10 | Describe the matrix (or matrices) you analyzed (e.g., covariance, correlation). Include matrices in the manuscript | See Supplementary File 1 |

| | | |
|---|---|---|
| | if feasible; otherwise, indicate these data are available upon request. | |
| 11 | Report the amount of missing data and describe how missing data were handled. For a review of practices for handling missing data, see Sterne and colleagues (2009), Rose and Fraser (2008), and Horton and Kleinman (2007).Provide a rationale for your approach to handling missing data. Authors are encouraged to consider using multiple imputation or model estimation with full-information maximum likelihood (FIML; Rose & Fraser, 2008). | See page 7 (Methods section: Psychometric evaluation) and page 10 (Results section: Psychometric evaluation) |
| 12 | Compare your CFA model with the alternative or competing models. Do competing models fit the data better or worse than your model (e.g., does your four-factor model of acculturation fit the data better than a two-factor model or a one-factor model)? Identify the preferable model based on appropriate fit statistics (e.g. chi-square difference test for nested models, Akaike information criterion for non-nested models), parsimony, and relevant theory | See page 15 and 16 (Discussions sections) |
| 13 | Include your scale (items and response options) in an appendix. | See Supplementary File 2 |
| 14 | Report how methodological limitations may have impacted findings regarding your measure's psychometric properties (e.g., note potential repercussions of suboptimal sampling techniques, discuss implications of using listwise deletion to handle missing data instead of multiple imputation or FIML). | See page 16 and 17 (Discussions sections) |
| 15 | Discuss directions for future research (e.g., if appropriate, testing your scale for measurement invariance by conducting CFA on different populations). | See page 16 and 17 (Discussions sections) |

1 Cabrera-Nguyen P. Author Guidelines for Reporting Scale Development and Validation Results in the *J Soc Social Work Res* 2010;1:99–103.

**Supplementary File 2: Additional tables for psychometric evaluation of 10-item ORIC version**

Table A: Demographic data of participants of cognitive interviews (n=11).

| | | Frequencies for n=11 |
|---|---|---|
| Age | < 30 years | 1 |
| | 31-40 years | 3 |
| | 41-50 years | 3 |
| | > 50 years | 4 |
| Gender | Female | 9 |
| | Male | 2 |
| Profession* | Nurse | 8 |
| | Physician | 2 |
| | Psychooncologist | 3 |
| Work experience in health care | > 5 years | 3 |
| | 5-10 years | 4 |
| | 11-20 years | 3 |
| | > 20 years | 1 |

* multiple answers possible

Table B: Factor loadings on the first component.

| | Component 1 |
|---|---|
| Item 1 | 0.790 |
| Item 2 | 0.788 |
| Item 3 | 0.833 |
| Item 4 | 0.846 |
| Item 5 | 0.831 |
| Item 6 | 0.851 |
| Item 7 | 0.839 |
| Item 8 | 0.800 |
| Item 9 | 0.754 |
| Item 10 | 0.860 |

For EFA, the sample was split and n=115.

Table C: Inter-item correlation matrix for the German ORIC.

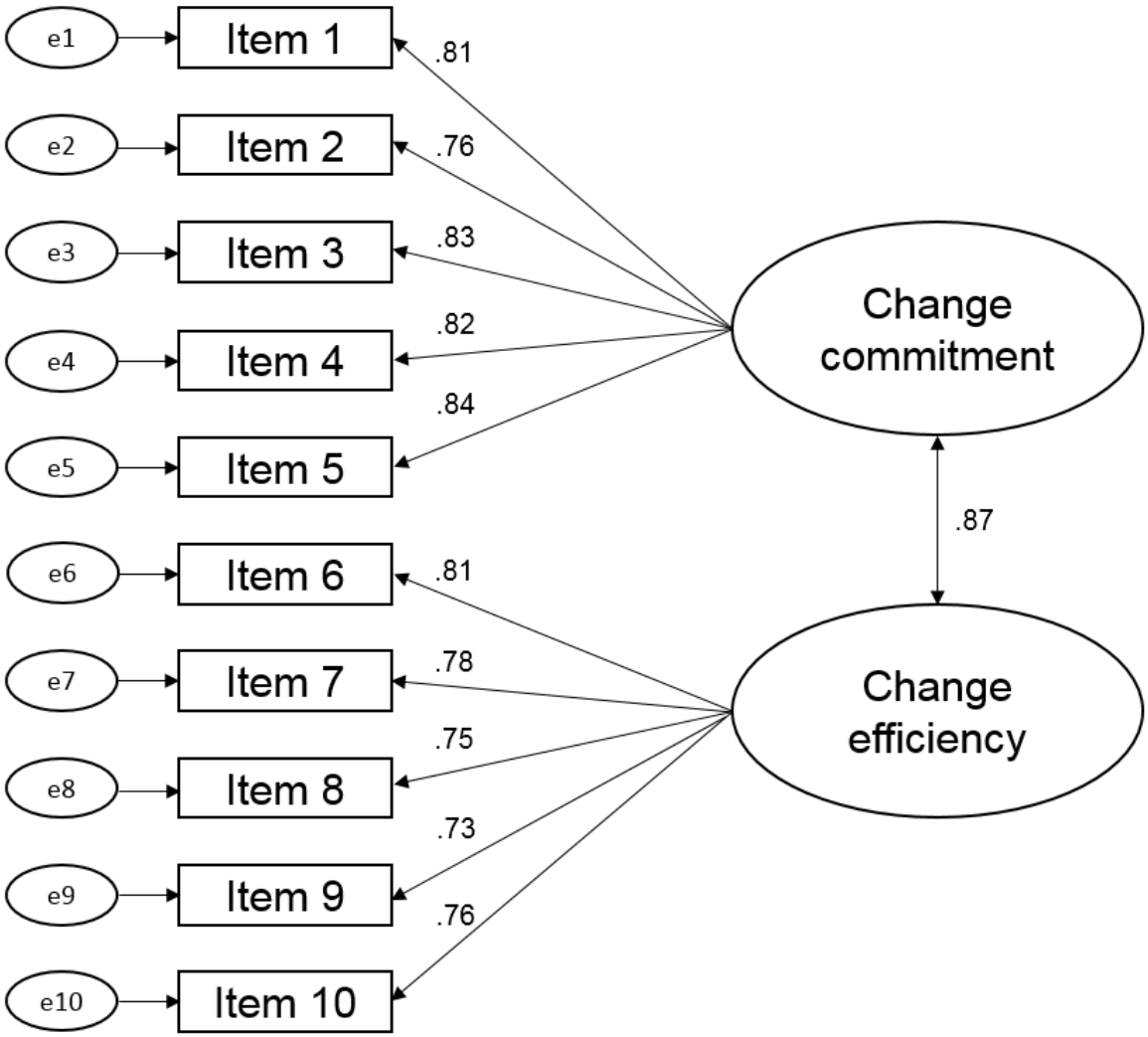
| | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 | Item 6 | Item 7 | Item 8 | Item 9 | Item 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Item 1 | 1.000 | .634 | .650 | .611 | .652 | .587 | .558 | .538 | .507 | .542 |
| Item 2 | .634 | 1.000 | .586 | .681 | .607 | .518 | .504 | .485 | .434 | .478 |
| Item 3 | .650 | .586 | 1.000 | .673 | .723 | .636 | .600 | .558 | .492 | .560 |
| Item 4 | .611 | .681 | .673 | 1.000 | .675 | .588 | .531 | .541 | .512 | .551 |
| Item 5 | .652 | .607 | .723 | .675 | 1.000 | .622 | .621 | .466 | .474 | .578 |
| Item 6 | .587 | .518 | .636 | .588 | .622 | 1.000 | .677 | .590 | .588 | .573 |
| Item 7 | .558 | .504 | .600 | .531 | .621 | .677 | 1.000 | .612 | .491 | .567 |
| Item 8 | .538 | .485 | .558 | .541 | .466 | .590 | .612 | 1.000 | .611 | .560 |
| Item 9 | .507 | .434 | .492 | .512 | .474 | .588 | .491 | .611 | 1.000 | .659 |
| Item 10 | .542 | .478 | .560 | .551 | .578 | .573 | .567 | .560 | .659 | 1.000 |

**Supplementary File 3: Organizational Readiness for Implementing Change (ORIC) – German version**

Bitte geben Sie an, wie sehr Sie den folgenden Aussagen zur Umsetzung von partizipativer Entscheidungsfindung an Ihrem aktuellen Arbeitsplatz zustimmen. Falls partizipative Entscheidungsfindung aktuell nicht umgesetzt wird, wie wäre es im Falle der Umsetzung?

| | | stimme nicht zu | stimme eher nicht zu | teils teils | stimme eher zu | stimme zu |
|---|---|---|---|---|---|---|
| 1 | Personen, die hier arbeiten, zeigen hohes Engagement bei der Umsetzung von partizipativer Entscheidungsfindung. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2 | Personen, die hier arbeiten, werden tun, was auch immer nötig ist, um partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3 | Personen, die hier arbeiten, wollen partizipative Entscheidungsfindung umsetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4 | Personen, die hier arbeiten, sind fest entschlossen, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 5 | Personen, die hier arbeiten, sind motiviert, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 6 | Eventuell entstehen bei der Umsetzung von partizipativer Entscheidungsfindung Herausforderungen. Personen, die hier arbeiten, sind zuversichtlich, diese zu meistern. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 7 | Personen, die hier arbeiten, sind zuversichtlich, dass sie den Verlauf der Umsetzung von partizipativer Entscheidungsfindung überblicken können. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 8 | Personen, die hier arbeiten, sind zuversichtlich, dass sie Aufgaben so koordinieren können, dass die Umsetzung reibungslos abläuft. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 9 | Personen, die hier arbeiten, sind zuversichtlich, dass die Klinik sie dabei unterstützten kann, partizipative Entscheidungsfindung umzusetzen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 10 | Personen, die hier arbeiten, sind zuversichtlich, Machenschaften bei der Umsetzung von partizipativer Entscheidungsfindung bewältigen zu können. | ☐ | ☐ | ☐ | ☐ | ☐ |

**Supplementary File 4: Confirmatory factor analysis model of the 10-item German ORIC**



**Supplementary File 5: Results of psychometric evaluation of the 9-item German ORIC version: Factor analysis, corrected item-total correlations, Cronbach's α**

**Factor analysis**

To get information about the factorial structure of the 9-item German ORIC, an exploratory factor analysis was conducted. KMO measure was .926 and Barlett's test of sphericity was significant ($X^2$ = 1302.78, $p < .001$), indicating that a factor analysis of the data was appropriate to observe data. Table A shows results of the exploratory factor analysis. Since the main component explains 66% of the variance, a one-factor model could be assumed. The factor loading for the first component can be observed in table B.

Table A: Results of exploratory factor analysis of the 9-item German ORIC with oblique rotation and parallel analysis: eigenvalues of the ten components of the German ORIC and eigenvalues for corresponding random data.

| | Eigenvalue | | | Eigenvalues for random data | |
|---|---|---|---|---|---|
| | Total | % of variance | Cumulative % | Means | 95% percentile |
| Component 1 | 6.02 | 66.85 | 66.85 | 1.45 | 1.59 |
| Component 2 | 0.802 | 8.91 | 75.76 | 1.29 | 1.39 |
| Component 3 | 0.472 | 5.25 | 81.01 | 1.17 | 1.25 |
| Component 4 | 0.392 | 4.35 | 85.37 | 1.07 | 1.14 |
| Component 5 | 0.366 | 4.06 | 89.43 | 0.98 | 1.05 |
| Component 6 | 0.304 | 3.38 | 92.81 | 0.89 | 0.96 |
| Component 7 | 0.250 | 2.78 | 95.59 | 0.81 | 0.87 |
| Component 8 | 0.232 | 2.57 | 98.16 | 0.71 | 0.79 |
| Component 9 | 0.165 | 1.84 | 100.00 | 0.61 | 0.69 |

Notes: For EFA, half of the data set (n=115) was used.

Table B: Factor loadings on the first component.

| | Component 1 |
|---|---|
| Item 1 | 0.860 |
| Item 2 | 0.856 |
| Item 3 | 0.837 |
| Item 4 | 0.836 |
| Item 5 | 0.835 |
| Item 6 | 0.799 |
| Item 7 | 0.797 |
| Item 8 | 0.791 |
| Item 9 | 0.739 |

Notes: For EFA, the sample was split and n=115.

A confirmatory factor analysis was performed with the one-factor model to analyze its fit indices. Indices of the one-factor model are presented in table C.

Table C: Fit indices of the one-factor model of the 9-item German ORIC.

| | Chi² [1] | df [2] | Chi²/df [3] | CFI [4] | TLI [5] | RMSEA [6] | AIC [7] | PNFI [8] |
|---|---|---|---|---|---|---|---|---|
| One-factor model | 56.04 | 27 | 2.08 | .945 | .927 | .097 | 92.04 | .676 |

Notes: One-factor model (including a split data set of n=115): includes items 1 to 9.
[1] discrepancy chi-squared statistic, [2] degrees of freedom, [3] normed chi-squared statistic, [4] comparative fit indexes, [5] Tucker-Lewis Index, [6] root mean square error of approximation, [7] Akaike Information Criterion, [8] Parsimonious Normed Fit Index (PNFI). * $p$ = .000

**Corrected item-total correlations and Cronbach's α of the 9-item ORIC version**

Table D shows the corrected item-total correlation. Corrected item-total correlation ranged from .665 (item 9) to .744 (item 3).

Table D: Discrimination of the nine items of the German ORIC.

| | Item discrimination (corrected item-total correlation) |
|---|---|
| Item 1 | .747 |
| Item 2 | .696 |
| Item 3 | .777 |
| Item 4 | .760 |
| Item 5 | .762 |
| Item 6 | .759 |
| Item 7 | .720 |
| Item 8 | .691 |
| Item 9 | .638 |

Notes: Items could be answered on a 5-step Likert scale rating from 0 „disagree" to 4 „agree".

Cronbach's α with α = .924 showed the measure to reach excellent reliability.