Assessing health-related stigma and social participation: research methods are coming of age

This thesis addresses the measurement of stigma and participation (restriction) in people affected by disabling health conditions in a low and middle-income country setting. People affected by health conditions are often stigmatized and experience restrictions in social participation due to the particular health condition.

Chapter 1 introduces the two key topics of my thesis, stigma and participation. After a conceptualization of both concepts in the International Classification of Functioning, Health and Disability, I explore these in detail. Stigma may negatively impact the life of the person affected, but also negative effects have been reported on public health programmes and interventions. Also constrains in the performance of daily activities and in social life, such as mobility, community life, work and domestic are often reported. Much progress has been made in understanding stigma and participation and, as a result, in our ability to measure these phenomena. A valid measurement of both stigma and participation is essential to gain insight in underlying determinants and dynamics, to provide an overview of the extent and the severity of the problems in a given target group and to assess the effectiveness of interventions. For the accurate assessment of both concepts, it is important that the measures used are valid. Two developments in the field of health measurement may facilitate this validation process, namely the development of quality criteria that provide benchmarks for what constitute ‘good’ psychometric properties, and of a framework for the assessment of cultural equivalence. The latter refers to the process of assessing the suitability of an instrument for use in a different cultural setting than it was initially developed for, whereas the former describes the assessment of psychometric properties such as content and construct validity, internal consistency and reliability. This thesis aims to apply these new developments in the field of participation and stigma measures in low and middle income settings.

The research questions underlying this thesis are:

1) What are the differences and commonalities in the level of stigma and participation restrictions among persons living with HIV/AIDS and persons affected by leprosy in Southern India?
2) What are the psychometric properties of the Participation Scale in a study population consisting of people with various disabilities living in Eastern Nepal?
3) How is internalized stigma conceptualized and to what extent is it being measured adequately?
4) Can the Participation Scale be shortened, without negatively affecting the psychometric properties of the scale?

5) What factor structure best fits the Participation Scale data from various study populations?

6) To what extent is cultural equivalence testing addressed in cross-cultural validation studies of participation measures?

Chapter 2 presents the results from a stigma comparison study among persons affected by leprosy and persons living with HIV/AIDS in Southern India. Levels of internalized and perceived stigma were assessed, as well as restrictions in participation using 190 questionnaire-based interviews. Persons living with HIV/AIDS reported a significantly higher level of stigma and participation restrictions. However, also the persons affected by leprosy faced a substantial burden of stigma and restrictions in participation. In both respondents groups, especially restrictions in work-related participation were identified, whereas social domains were less affected. Based on the commonalities found in this study, we suggested the possibility to develop joint interventions to combat stigma and participation restrictions in leprosy and HIV/AIDS.

Chapter 3 reports on the further validation of the Participation Scale among 153 people with disabling conditions in south-east Nepal, using the new quality criteria proposed by Terwee et al. The study focused on the assessment of various psychometric properties such as structural validity, internal consistency, inter-tester reliability, construct validity and floor and ceiling effects. A two-factor structure of the Participation Scale showed the best model fit. These two factors were conceptualized as ‘work-related participation’ and ‘general participation’. Internal consistency was found to be good, namely $\alpha = 0.93$ for the whole scale and $\alpha = 0.78$ and $\alpha = 0.93$ for the ‘work-related’ and ‘general participation’ subscales, respectively. The inter-tester reliability coefficient was excellent, namely 0.90. Our a priori hypotheses were confirmed; we found a positive association with a perceived stigma score and a negative association with a self-reported health score. These findings confirmed the construct validity of the Participation Scale. No floor and ceiling effects were identified for the whole scale. In conclusion, the psychometric properties of the Participation Scale were found to be good in this particular cultural setting.

Chapter 4 presents results from a systematic review assessing the psychometric properties of internalized stigma measures. The review included 33 studies, covering 21 instruments to measure internalized stigma. The psychometric properties of these instruments were rated as ‘positive’, ‘negative’, ‘indeterminate’, or as ‘no information available’. The rating was done using a list of quality criteria that included the following properties: content, criterion and construct validity, internal
consistency, reliability, agreement, responsiveness, interpretability and floor and ceiling effects. The results showed that the properties content and construct validity, and internal consistency were addressed often, whereas only few studies investigated agreement and responsiveness. We concluded that the majority of these instruments need further testing of their psychometric properties.

Chapter 5 describes the development of the Participation Scale Short (PSS) and the structural validation of both versions of the Participation Scale, using a multi-country and multi-culture database of 5,125 respondents. Five items of the Participation Scale were deleted, based on a combination of low endorsement, low factor-loadings, substantial cross-loadings, low item-total correlations or negative effects on the scale internal reliability, resulting in the 13-item PSS. Findings from explanatory and confirmatory factor analysis suggested best model fit for a two factor model, named ‘work-related participation’ (3 items) and ‘general participation’ (10 items) (CFI = 0.983, TLI = 0.979, RMSEA = 0.061). Correlation between the two factors was high ($r = 0.75$) as well as between the full Participation Scale and the PSS ($r = 0.99$). The Cronbach’s alphas for the shortened scale were above the threshold of $\alpha \geq 0.70$ (0.78 and 0.93, respectively). Additional field testing of the PSS is necessary to confirm its reliability, validity and usefulness.

Chapter 6 describes the findings from a review of the process of cross-cultural equivalence testing of participation measures. This process was assessed using a cultural equivalence framework that defined five categories of cultural equivalence: conceptual, item, semantic, operational and measurement equivalence. I adapted this framework, thereby integrating the quality criteria of Terwee and the COSMIN consortium. A rating system was used, comprising ‘extensive’, ‘partial’ and ‘none/minimal’ ratings, to address the different categories of equivalence of five different participation instruments in eight cross-validation studies. Only two out of eight studies scored two ‘extensive’ ratings and five studies scored no ‘extensive’ rating at all. Three studies were assigned at least three ‘partial’ ratings. Conceptual, item and semantic equivalence were most often rated ‘extensive’, whereas the majority of the ‘none/minimal’ ratings were given for item and measurement equivalence. In conclusion, our review suggested that the cultural equivalence of participation instruments has been tested inadequately. We proposed a framework that can be used to facilitate adequate reporting and testing of cultural equivalence.

Chapter 7 is a general discussion of the main findings. A reflection on the methodology is provided, after which this chapter closes with the take-home messages and future research priorities.