

PhD Thesis

# Measuring socio-emotional and pre-math skills in early childhood

NINA MADSEN SJÖ



# English Summary

## Background and objective

Researchers in small, non-native English-speaking countries like Denmark experience psychometric and adaptation challenges when planning to use validated assessment tools in intervention studies. Many assessment tools have been validated in English, but only some have been specifically validated and normed for a Danish pediatric population, which makes the assessment results in Danish more uncertain and less trustworthy.

The overall aim of this PhD was to adapt and evaluate two assessment tools intended for use as outcome measures in a child care intervention. One of the tools, a strength-based questionnaire, was further compared with two well-established questionnaires to gain more knowledge about how strength-based and problem-focused questionnaires differ in their evaluation of children, and whether child care providers have different perspectives on the various types of questionnaires.

The two tools that were adapted and further developed for use in an intervention study with a nationally representative sample were: 1) the Social Emotional Assessment Measure – Research Edition (*SEAM*), a questionnaire that assesses social-emotional skills and competences, and 2) a short version of the pre-math test, Tools for Early Assessment in Math (*TEAM*).

## Research questions

The three research aims of the PhD study were:

1. To adapt the assessment tools *SEAM* and *TEAM* to culturally sound tools in Danish for assessment of socio-emotional skills and pre-math skills.
2. To investigate the validity and reliability of the Danish versions of the two assessment tools.
3. To examine the perspectives of the child care providers on use of the strength-based *SEAM* in comparison with two other questionnaires for assessment of child behavior: the diagnostic

instrument Caregiver-Teacher Report Form, (*C-TRF*) and the psychopathological screener Strengths and Difficulties Questionnaire (*SDQ*).

## Methods

The first research aim concerned the adaptation process to Danish, including specific adjustments to fit the needs of the outcome study. *SEAM* was first forward-translated into Danish, then the age range was expanded upwards by one year (from 5½ to 6½), interviews with end-users were conducted, and the Danish *SEAM* was tested in a pilot study. Similarly, the adaptation process for the *TEAM* consisted of forward-translation into Danish, change of the age range (from 4–5/4–7 years to 3–6 years), pilot-testing of the Danish version, and final item selection.

The second research aim addressed the quality of the Danish versions of *SEAM* and *TEAM*, i.e. their construct validity by Rasch analysis and reliability coefficients based on data from a large-scale intervention study (*SEAM*:  $n = 9,827$  aged 0–6, *TEAM*:  $n = 5,621$  aged 3–6).

The third research aim examined the responses given by child care providers who were asked to assess the *SEAM* questionnaire alongside the *SDQ* and *C-TRF* tools with respect to clarity, relevance, and acceptability for children from a low-risk sample ( $n = 292$  aged 2–5).

## Results

The evaluation of the adaptation process showed that the adaptations of *SEAM* and *TEAM* into Danish could be considered successful, and both assessment tools were later used in the large-scale intervention study assessing the socio-emotional and pre-math skills of Danish children.

The results from examination of the construct validity of *SEAM* showed that it was possible to establish a valid tool by applying a scoring method that included all items and facilitated the combination of *SEAM*'s 10 benchmarks into two indexes. In addition, the reliability coefficients for the two new *SEAM* indexes were categorized as adequate, good, or excellent. With regard to the *TEAM*, examination of construct validity suggested a final version – the *DK-TEAM (final)* – that consisted of 19 items with no

differential item functioning relative to age and gender. The reliability coefficients for the Numeracy and Geometry scales within this Danish version were quite low.

The results of the third research aim showed that the *SEAM* was rated more positively than the two other questionnaires on three out of five questions.

## **Perspectives**

The studies included in this thesis contribute new knowledge at both national and international level. For Danish researchers and practitioners, two newly adapted tools have been tested, and information on their psychometric properties is now available. At the international level, the suggested index model for *SEAM* is worth consideration by other researchers. The child care providers' preference for a strength-based tool will hopefully inspire other researchers to include a strength-based tool in their studies, and clinicians to include a strength-based tool in their assessment package. Future studies should address some of the shortcomings of the Danish versions of *SEAM* and *TEAM*, e.g. both tools need studies of inter-rater reliability and of their ability to measure changes in children's skills as a results of an intervention.

Title

Measurement properties of the Social Emotional Assessment Measure (*SEAM*) questionnaire using Rasch analysis on a representative Danish sample of 0–6-year-olds

Short title

Measurement properties of *SEAM*

Authors

Nina Madsen Sjøe<sup>1+2</sup>, Dorthe Bleses<sup>1</sup>, Line Dybdal<sup>2</sup>, Hanne Nielsen<sup>2</sup>, Karen Krag Sehested<sup>2</sup>, Henriette Kirkeby<sup>3</sup>, Svend Kreiner<sup>4</sup>, Peter Jensen<sup>1</sup>

Affiliations

<sup>1</sup>Trygfonden's Centre for Child Research, Aarhus University, <sup>2</sup>Ramboll Management Consulting, Copenhagen, Denmark, <sup>3</sup>Department of Organization, Copenhagen Business School, Denmark, <sup>4</sup>Department of Biostatistics, University of Copenhagen, Denmark

Keywords

Social Emotional Assessment Measure (*SEAM*), Questionnaire, Early Childhood, Infant, Toddler, Preschool, Outcome measure.

## Abstract

Questionnaires used to assess socio-emotional development in children are typically developed for clinical assessment and designed to detect deviant or problematic behavior. This can be a challenge in studies of the effectiveness of child care or school interventions that target diverse populations, and where the whole range of abilities needs to be captured.

In the current study, a strength-based questionnaire developed for monitoring progress of socio-emotional skills was for the first time applied in a nationally representative sample of 0–6-year-old children as an outcome measure in a child care intervention.

This multisite study investigated the measurement properties of the newly released questionnaire Social Emotional Assessment Measure – Research Edition (*SEAM*). Specifically, criterion-related construct validity using Rasch analysis was examined, as well as accuracy and reliability. One of the primary purposes was to determine whether it was possible to merge the ten *SEAM* benchmarks (domains) into one total score or into a limited number of subscores. The data were collected by child care providers in non-parental child care facilities, and the response rate was high (95%) ( $n = 9,827$ ).

The results showed that the validity of *SEAM* was hampered when the US scoring guidelines were applied to the Danish sample. Conversion of the *SEAM*'s 10 benchmark scores and use of graphical loglinear Rasch models led to a successful combination of the 10 benchmarks to two overall indexes: Empathy and Self-regulation & Cooperation. The reliability coefficients for the two indexes ranged from adequate (0.79) to excellent (0.94), while standard errors of measurements were relatively large for the two indexes.

The study findings suggest that the adapted Danish version of *SEAM* using a revised scoring system may be appropriate as an outcome measure to assess socio-emotional development in young children, though there are tendencies to ceiling effects, and additional validation studies are warranted.

## Title

Adaptation and further development in Danish of a short version of the Tools for Early Assessment in Math (*TEAM*) for 3–6-year-olds

## Short title

Adaptation and development of Danish *TEAM*

## Authors

Nina Madsen Sjøe<sup>1+2</sup>, Dorthe Bleses<sup>1</sup>, Line Dybdal<sup>2</sup>, Eva Tideman<sup>3</sup>, Henriette Kirkeby<sup>4</sup>, Karen Krag Sehested<sup>2</sup>, Hanne Nielsen<sup>2</sup>, Svend Kreiner<sup>5</sup>, Peter Jensen<sup>1</sup>

## Affiliation

<sup>1</sup>TrygFonden's Centre for Child Research, Aarhus University, Denmark, <sup>2</sup>Ramboll Management Consulting, Copenhagen, Denmark, <sup>3</sup>Department of Psychology, Lund University, Sweden, <sup>4</sup>Department of Organization, Copenhagen Business School, Denmark, <sup>5</sup>Department of Biostatistics, University of Copenhagen, Denmark

## Keywords

Tools for Early Assessment in Math (*TEAM*), Pre-math, Test, Early Childhood, Preschool, Math, Assessment, Measure, Numeracy, Numbers, Geometry, Short form, *REMA*

## Abstract

Children's pre-math skills are strong predictors of how they later succeed in math in elementary, middle, and even high school. Children with little experience with basic pre-math skills are limited in their ability to learn more complex math, and a knowledge gap continues to exist after school entry. As a result pre-math interventions have caught the attention of researchers, teachers, and politicians with the aim of familiarizing all children with pre-math concepts before they enter the formal school system. Sound assessment tools are needed to evaluate different pre-math interventions, but assessment of pre-math skills is complex. This study explores the adaptation and further development in Danish of a short version of the Tools for Early Assessment in Math (*TEAM*) and presents results of a test based on a nationally representative sample of 3–6-year-olds.

The Danish adaptation and development process was carried out in four steps: 1) Choosing and translating relevant items, 2) Conducting a pilot study ( $n = 170$ ) and quantitative and qualitative evaluation, 3) Testing the most suitable items in a representative sample of Danish children aged 3–6 years ( $n = 5,621$ ), and 4) Analysis of the data based on Rasch models.

The final test resulted in a 19-item version – the *DK-TEAM (final)* – which has no differential item functioning relative to age and gender. The 19 items were capable of measuring different levels of skills for most of the children in the representative sample, even though a floor effect was seen for 3-year-olds. User perspectives indicated that the majority of children viewed the test as fun. The study confirms the applicability of *DK-TEAM (final)* in a Danish child care setting for research use. However, further investigation of validity and reliability measures is recommended.



Title

Further validation of a Danish version of the Social Emotional Assessment Measure (*SEAM*): Comparison with the Strengths and Difficulties Questionnaire (*SDQ-T*) and the Caregiver-Teacher Report Form from ASEBA (*C-TRF*)

Short title

Study of three questionnaires for 2–5-year-olds

Authors

Nina Madsen Sjøe<sup>1+2</sup>, Astrid Kiil<sup>1+3</sup>, Dorthe Bleses<sup>1</sup>, Line Dybdal<sup>2</sup>, Svend Kreiner<sup>4</sup>, Peter Jensen<sup>1</sup>

<sup>1</sup>TrygFonden's Centre for Child Research, Aarhus University, Denmark, <sup>2</sup>Ramboll Management Consulting, Copenhagen, Denmark, <sup>3</sup>KORA, the Danish Institute for Local and Regional Government Research, Copenhagen, Denmark, <sup>4</sup>Department of Biostatistics, University of Copenhagen, Denmark

Keywords

Validity, ASEBA, CBCL, C-TRF, SDQ, SEAM, Questionnaire, Early Childhood, Toddler, Preschool, Teacher, Strength-based Assessment

## Abstract

Questionnaires assessing socio-emotional development and behavior have different purposes, lengths, and wording. There is limited knowledge of how strength-based and problem-focused questionnaires differ in their evaluation of children, and whether child care providers have different perspectives on the various types of questionnaires. This validation study compared assessments of children made with a diagnostic instrument (*C-TRF*), a psychopathological screener (*SDQ*), and a strength-based assessment measure (*SEAM*). We also examined the perspectives of the child care providers on the questionnaires' clarity, relevance, and acceptability.

A low-risk sample of 2–5-year-old children ( $n = 292$ ) were assessed by 59 staff in Danish child care settings including preschools. A graphical model was used to identify associations between *SEAM*, *SDQ-T* and *C-TRF*. Scatterplots were used to explore the variation of scores. The child care providers' perspectives on clarity and face validity were shown graphically and analyzed using Wilcoxon signed-rank tests.

The two *SEAM* indexes were strongly and positively related to each other. The *SEAM* Empathy Index was directly related to the *SDQ-T* Prosocial subscale, while the *SEAM* Self-regulation & Cooperation Index was directly related to three externalizing subscales. The scatterplots revealed that while the subscales of *C-TRF* and *SDQ-T* did not identify any variation, the *SEAM* did. The child care providers rated the *SEAM* more positively than the two other questionnaires.

We conclude that in this low-risk sample, *SEAM* was better able to differentiate between children with different socio-emotional skills than the subscales from *C-TRF* and *SDQ-T*. *SEAM* measures some of the same phenomena as *SDQ-T* and *C-TRF*, but describes additional aspects of socio-emotional behavior.