

Forbedring af skala-items, som ikke fitter Rasch-modellen

Udviklingen af test-/måleskalaer til at måle latente psykologiske, og andre, egenskaber og tilstande er en både tids- og ressourcekrævende proces med mange trin. Uanset hvilken model der trækkes på i udviklingen af en sådan skala, ligger det største tids- og ressourceforbrug i starten af udviklingsprocessen, dvs. frem til, at den første datasample er klar til at foretage den indledende psykometriske analyse.

Det at der anvendes så forholdsvis mange ressourcer på at nå frem til en første skala/test-version, står i skarp kontrast til den hyppigst anvendte strategi til at bestemme den endelige skala/test; den såkaldte salamimetode, hvor de items (dvs. opgaver eller spørgsmål, mv.), der på forskellig vis ikke lever op til de opstillede krav, simpelt hen fjernes fra skalaen og smides væk. Salamimetoden leder da indimellem også til, at den resulterende skala består af så få items, at det er nødvendigt at udvikle nye items med samme niveau af tids- og ressourceforbrug som ved de oprindelige items.

For at assistere skala-udviklere i forhold til at bevare så meget som muligt af den oprindelige pool af items har vi udviklet en systematisk strategi til at forbedre items af utilstrækkelig psykometrisk kvalitet, der kan identificere, hvilke items der kun behøver en let modificering, og hvilke items der bør udskiftes med nye items, og som også giver information om, hvordan items skal modificeres, henholdsvis udskiftes. Strategien kombinerer resultater fra grafiske loglineære Rasch-analyser med resultaterne fra teorinære indholdsanalyser til et hierarki af systematiske beslutninger om, hvilke items der skal modificeres, hhv. udskiftes for at forbedre skalaens psykometriske egenskaber (dvs. fitte Rasch-modellen uden afvigelse).

Artiklen præsenterer først strategien til item-forbedring og eksemplificerer derefter strategiens anvendelse ved at se på resultaterne af grafiske loglineære Rasch-analyser og indholdsanalyser af items i den fysiske funktionsskala i SF-36, samt på basis af dette at fremsætte forslag til forbedring af items i denne skala.

Strategien blev oprindeligt udviklet og afprøvet i forbindelse med udviklingen af en dansk læringsstiltest (The Danish Self-Assessment Learning Styles Inventory, D-SA-LSI), baseret på Sternbergs *theory of mental self-government* og de indeholdte læringsstile. Anvendelsen af strategien førte i læringsstilsprojektet til, at ud af 13 læringsstilskalaer blev 10 skalaer psykometrisk væsentligt forbedret, én skala viste ikke nogen psykometrisk forbedring, og to skalaer viste lettere forringet psykometrisk kvalitet – den ene på grund af en fejl i strategianvendelsen.

Strategien har desuden været anvendt i forbindelse med udviklingen af seks skalaer til semesterevaluering af cand.merc.-programmerne ved Copenhagen Business School. Resultaterne var i dette tilfælde også positive, men kan ikke udelukkende

tilskrives anvendelsen af strategien i item-revisionen, idet den første version af skalaerne ikke var indholdsmæssigt skarpt defineret, og vi således måtte foretage revisioner af items, som ikke var bestemt udelukkende af strategien.

Et detaljeret studie af udgangspunkterne, revisionerne og resultaterne i læringsstilsprojektet gav desuden anledning til at skelne mellem tre typer af resultater: de stærkt forudsigelige, de moderat forudsigelige og de dårligt forudsigelige. Og det blev i dette studie tentativt konkluderet, at jo bedre en basis der var for at lave items-forbedringer, jo mere forudsigeligt var resultatet, og denne konklusion blev senere støttet af de relativt mindre forudsigelige resultater i evalueringsprojektet.

Strategien bliver p.t. anvendt og således yderligere afprøvet i to projekter; et, som har til formål at udvikle et dansk spørgeskema, der måler *statistics anxiety* baseret på *the Statistical Anxiety Rating Scale* (STARS), og et, der har til formål at udvikle en dansk tilpasset version af *the Dallas Pain Questionnaire* (DPQ).

Reference: Nielsen, T. & Kreiner S. (2013). Improving items that do not fit the Rasch model: exemplified with the physical functioning scale of the SF-36. *Annales de l'I.S.U.P. Publications de l'Institut de Statistique de l'Université de Paris, Numéro Special*, 57(1-2), 91-108.

Forfatter/kontakt: Tine Nielsen, Institut for Psykologi, Københavns Universitet, tine.nielsen@psy.ku.dk