

## Bilaga 3: Psykometrisk kvalitetsbedömning av enkäter

*Denna bilaga hör till RISE rapport 2022:70 (ISBN 978-91-89711-10-5): "Struktur och indikatorer för uppföljning av föräldraskapsstöd", som producerats på uppdrag av Myndigheten för Familjerätt och Föräldraskapsstöd.*

Detta är ett tekniskt dokument som inte är språkligt anpassat för allmänheten, utan avsett för att personer med psykometrisk kompetens ska kunna granska RISE bedömningar och ges tillgång till en lista på de vetenskapliga källor som dessa bedömningar baseras på. Läsaren förutsätts ha förkunskaper inom psykometrisk terminologi, och ha läst avsnittet nedan om grundläggande kriterier vid psykometrisk analys.

Det förekommer granskade artiklar i denna bilaga där vår skriftliga bedömning är att ingen psykometrisk analys har gjorts, där det kan vara fallet att enbart Cronbach's alpha har redovisats i artikeln. Eftersom Cronbach's alpha inte ger någon information om dimensionalitet eller andra centrala aspekter av psykometri (se nästa avsnitt) är det på egen hand ett intetsägande värde att rapportera, sett ur psykometrisk synvinkel. En annan återkommande brist i flera artiklar är att PCA/EFA används utan att sedan följa upp med konfirmatorisk metod. PCA och EFA är explorativa metoder, och resultatet från dem kan inte användas för att dra konfirmatoriska slutsatser om faktorstruktur.

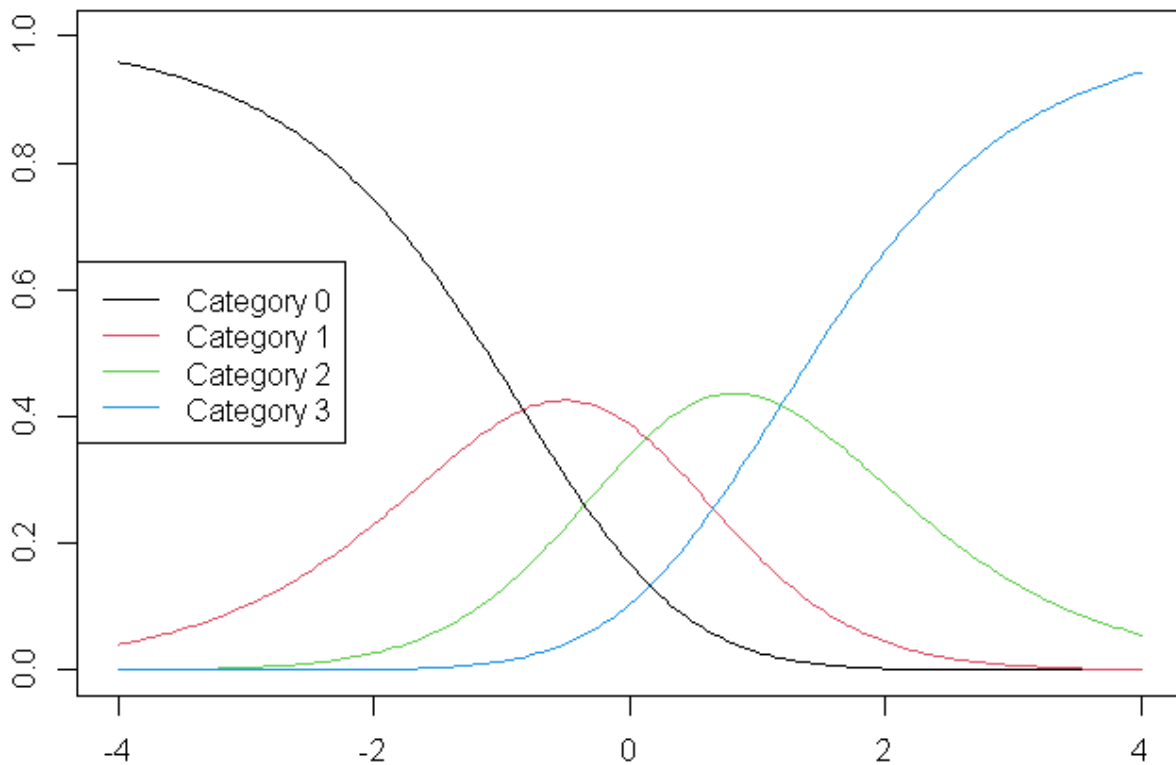
Ordlista för återkommande förkortningar som används i bilagan: CTT = klassisk testteori; DIF = differential item functioning (invariansanalys på itemnivå); PCA = principalkomponentanalys; CFA = konfirmatorisk faktoranalys; EFA = explorativ faktoranalys; WLS = weighted least squares; local dependencies = när två items residualer korrelerar kraftigt nog att skapa bias i mätningar; ML = maximum likelihood; MLR = robust maximum likelihood.

## Grundläggande kriterier vid psykometrisk analys

När vi samlar in enkätdata med svarskategorier och sätter siffror på kategorierna, exempelvis mellan 1 och 5, är detta ordinala data. Det betyder att det finns en ordning sinsemellan svarskategorierna (förutsatt att svarskategorierna fungerar korrekt, se nästa stycke), men att vi varken vet hur stora avstånd det är mellan svarskategorier eller om avstånden är lika stora, vare sig inom en och samma fråga eller mellan frågor. Om vi summerar, beräknar genomsnittsvärden, eller gör andra typer av matematiska beräkningar på ordinala svarsdata kommer det att medföra felaktigheter i varierande storleksgrad. Om samtliga nedan beskrivna psykometriska kriterier uppfylls är det möjligt att använda ordinala summerade rådata och omvandla dem till **intervalldata**. Det medför att korrekta statistiska och matematiska beräkningar kan göras. Målsättningen är alltså att summeringen från en enkät/skala ska kunna användas, och för att göra det behöver nedanstående kriterier uppfyllas.

**Svarskategorierna** är en ofta undervärderad aspekt av enkäter. Modern testteori, så som Item Response Theory och Rasch-analys, låter oss analysera om svarskategorierna fungerar som

avsett. Analyser visar ofta på problem när det finns för många svars kategorier och/eller när det saknas etiketter på flera av svars kategorierna. När svars kategorierna inte fungerar som tänkt medför det mätbrus. Ett omfattande mätbrus kan medföra att det som i den rådata ser ut som skillnader mellan grupper eller mättillfällena består av brus snarare än faktiska skillnader.

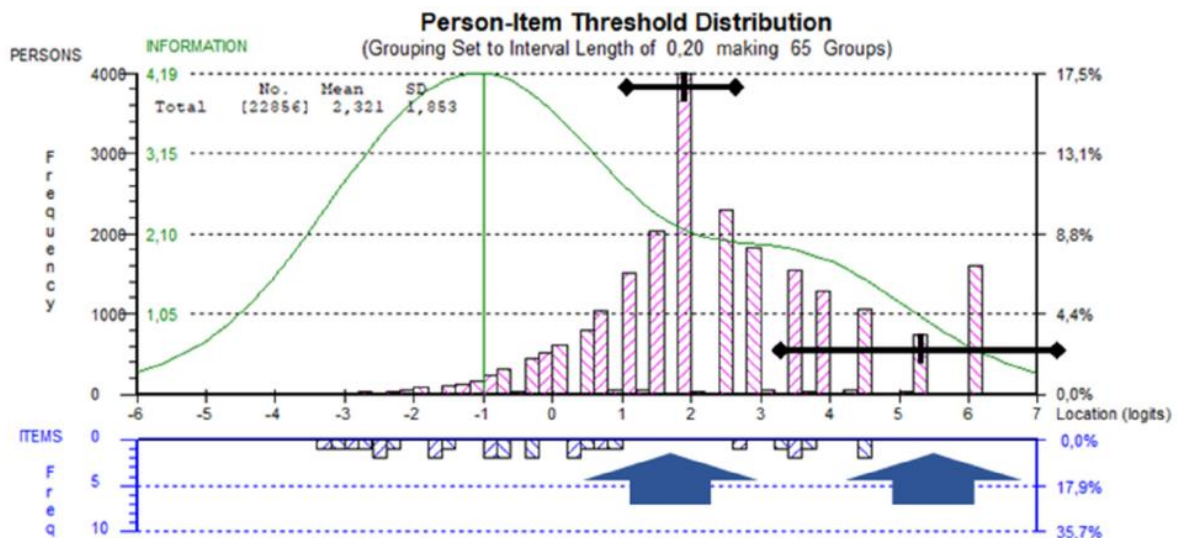


Figur 1. Visar sannolikheten för svars kategori 0-3 för en fråga på y-axeln och respondentens övergripande förmåga på x-axeln.

Mätningar görs oftast med syftet att kunna göra olika jämförelser, vanligtvis mellan grupper och/eller över tid. Det innebär att vi behöver säkerställa att enkätens frågor uppfattas likadant för de grupper som ska besvara den. Det innebär alltså inte att de svarar samma sak på frågorna, utan att frågornas egenskaper är tillräckligt konstanta mellan grupper. Detta brukar betecknas som "invarians" och är ett av de grundläggande kriterierna för jämförande mätningar. Invarians är inte relaterat till "test-retest correlation". Inom klassisk testteori innebär det att invarians av samtliga tre typer "configural, metric, scalar" uppfylls vid jämförelser av grupper. Inom modern testteori testas detta huvudsakligen med analys av **DIF**, "differential item functioning", vilket ger utförlig information om eventuella skillnader mellan grupper som beror på egenskaper hos deltagarna och hur dessa inverkar på enskilda frågor. Det kan exempelvis visa sig att män och kvinnor upplever en eller flera frågor olika och då kan inte deras svar jämföras direkt. När det gäller föräldraskattade enkäter som berör barn är detta extra viktigt, då både föräldrarnas och barnets kön och interaktionen mellan dessa två variabler kan påverka hur skattningar görs, och invarians är centralt i att säkerställa jämförbarhet. Med hjälp av Rasch-analys är det ofta möjligt att justera för DIF i framtagande av indexvärden.

Ofta ställs flera frågor i samma enkät som avser mäta aspekter av samma process eller utfall, t.ex. föräldraförmåga, vilket potentiellt medför avsevärt mycket bättre mätprecision än enstaka frågor. För att utifrån flera frågor ta fram ett indexvärde som beskriver den övergripande förmågan krävs att frågorna uppvisar **unidimensionalitet**. Det innebär att de i tillräckligt stor grad hör till samma underliggande fenomen, utan att ha för starka samband sinsemellan (**residualkorrelationer**). Ibland visar det sig att frågorna i en enkät bildar flera dimensioner eller index, och då behöver var och en av dessa uppvisa unidimensionalitet och i sin tur användas som separata indexvärden och inte slås samman till ett övergripande indexvärde. En viktig del i analys av unidimensionalitet är att undersöka huruvida något eller några par av items har för starkt korrelerade residualer. I regel sätts gränsvärdet till  $r = 0.2-0.3$  över den genomsnittsliga korrelationen mellan alla par av items, beroende på hur många items som ingår. Om gränsvärdet korsas innebär det att signifikant bias kan uppstå i indexvärdet, och åtgärder behöver vidtas för att säkerställa unidimensionalitet.

Frågorna behöver också vara **lämpliga** för målgruppens nivå på den förmåga eller upplevelse man avser mäta. Det gäller både frågeställningen och svarskategorierna. Ett grundläggande, men otillräckligt, sätt att undersöka detta är att titta efter tak- och golfeffekter genom att kartlägga hur många respondenter som får högsta eller lägsta möjliga poäng när svaren summeras utifrån ordinaldata. Med Item Response Theory eller Rasch Measurement Theory, går det att få en mera nyanserad bild av hur frågorna och svarskategorierna passar mot deltagarnas nivå/förmåga, vilket också medför att man får mått på **mätosäkerheter** över hela skalans spann. När mätosäkerheter är kända blir det också möjligt att bedöma hur stora förändringar som kan anses vara statistiskt säkerställda. Mätosäkerheter beräknas på individnivå, vilket gör att individuella förändringar kan jämföras på ett mer tillförlitligt sätt.



Figur 2. Figuren visar deltagarnas övergripande förmåga med rosa staplar, och frågornas tröskelvärden i blå staplar. Den gröna kurvan visar reliabiliteten. Blå pilar visar luckor där det saknas frågor som bidrar med information om deltagarna. Svarta horisontella streck visar exempel på mätosäkerheter vid två punkter. Bilden från Melin et al. (2022).

För mera utförlig beskrivning av ovanstående kriterier och vetenskapliga bakgrundsreferenser, se:

Andrich, D., & Marais, I. (2019). *A Course in Rasch Measurement Theory: Measuring in the Educational, Social and Health Sciences*. Springer. <https://doi.org/10.1007/978-981-13-7496-8>

Bond, T., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (3rd ed.). Routledge.

<http://ebookcentral.proquest.com/lib/hioa/detail.action?docID=3569314>

Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How? *CBE Life Sciences Education*, 15(4). <https://doi.org/10.1187/cbe.16-04-0148>

Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education*, 4(1), 1416898.

<https://doi.org/10.1080/2331186X.2017.1416898>

Hobart, J. C., & Cano, S. J. (2009). Improving the evaluation of therapeutic interventions in multiple sclerosis: The role of new psychometric methods. *Health Technology Assessment*, 13(12). <https://doi.org/10.3310/hta13120>

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research*, 57(8), 1358–1362. <https://doi.org/10.1002/art.23108>

## Granskning av mätverktyg/enkäter

### Strengths and Difficulties Questionnaire (SDQ)

#### Svensk forskning:

Självskattningsversionen för 12-18-åringar kan ej rekommenderas. Hagquist (2007) har gjort analyser med Rasch-metodik och visar stora problem. Inga delskalor uppfyller reliabilitetskriterier. Nämner problem med reverserade/negativt formulerade items och att flera items är tvetydiga (två meningar långa) och svåra att svara på. Har ej gjort DIF-analys för kön.

Lundh et al. (2008) gör PCA på självskattnings-versionen, hittar låga faktorladdningar på flera items, och undermålig reliabilitet för samtliga delskalor.

Smedje et al. (1999) har analyserat föräldraversion (barn 6-10 år) med CFA. Inga av de traditionella indexen för model fit redovisas, så underlaget är bristfälligt för kvalitetsbedömning. Däremot redovisas separata analyser för faktorladdningar (indelad på barnets kön), och där finns stora skillnader på flera items, vilket talar för att invarians inte råder och att det skulle behövas separata scoringmetoder för pojkar och flickor om resultaten ska vara jämförbara.

Dahlberg et al. (2019) har använt CFA för att analysera föräldraskattning av mammor och pappor (barn 3-5 år), samt skattning av förskolelärare. Man använder DWLS som estimator, vilket är adekvat. Model fit ger acceptabla resultat för alla tre grupperna av raters/skattare. Metrisk invarians mellan barnets kön finns, dock oklart varför man inte gjort fler invarians-analyser med ett så stort dataunderlag tillgängligt. Metrisk invarians är tyvärr otillräckligt (se Putnick & Bornstein, 2016), även scalar invariance bör testas.

Gustafsson et al. (2016) analyserar data från förskollärare till barn 1-5 år med PCA för grupperingarna 1-3 och 4-5 år. PCA visar på 2 faktorer för barn 1-3 år, och 4 faktorer för 4-5-åringar. Test-retest visar ojämna resultat för delskalorna, och av oklar anledning använder man 5 delskalor i den analysen, i stället för att utgå från PCA-resultatet. PCA är en explorativ metod, och kan inte användas för konfirmatorisk analys.

Björnsdotter et al. (2013) har analyserat föräldraversionen (barn 10-13 år) i ett representativt sampel från normalpopulation. Bra att man noterat att Cronbach's alpha är olämpligt för SDQ-data (pga skew/kurtosis i responser). Men det anges inte vilken estimator som använts i CFA. Model fit index ser bra ut efter att ha tagit bort item 22 (Problem med item 22 för pappor). 4 av 5 items i "conduct disorder" har mycket låga faktorladdningar, från 0.40 till 0.46 (OBS att Hagquist (2007) fann denna subskala sämst när självskattningsversionen av SDQ analyserades med Rasch). "Peer problems" likaså, 4 av 5 items är under 0.50. Två items har mycket låg faktorladdning (<0.4), och flera items strax över 0.4. Ingen invarianstestning, men fit index rapporteras från flera grupperingar och de skiljer sig åt tillräckligt för att det ska vara tveksamt med invarians.

#### Internationell forskning:

Keller & Langmeyer (2019) beskriver klara brister för föräldraskattningen bland subskalornas mätprecision (IRT-baserat). Korrelationer mellan faktorer är höga, både i 3 och 5-faktor-varianter. Möjligen fungerar total problem score (4 av 5 subskalor summerade, ev. som bifactor men tveksamt)

bättre, men fortfarande problem med local deps. McCarron et al. (2021) har Rasch-analyserat föräldraskattningen, med 143 föräldrar till barn med förvärvad hjärnskada, och funnit att alla delskalor har problem i varierande grad, och att ingen av dem har adekvat reliabilitet.

Essau et al. (2012) med CTT/CFA testar både 3 och 5 faktorer, med dålig model fit. Liksom Hagquist (2007) påtalas problem med reverserade/negativt formulerade items och man rekommenderar att ändra formuleringar och testa nya items.

Mellor & Stokes (2007) använder både CFA och Rasch, och hittar stora problem med samtliga versioner av SDQ (barn, förälder, lärare), och DIF mellan dem (bara 5 av 25 items är utan DIF mellan grupper). Ingen delskala uppfyller unidimensionalitet.

Kersten et al. (2018) har stora datamängder från Nya Zeeland och har Rasch-analyserat data för föräldrar till barn 4-5 år gamla, samt förskollärare när det varit aktuellt. Mycket låg samstämmighet i bedömning mellan förälder och lärare. Alla delskalor undermåliga, men man har tagit fram en modifierad Total Problems scale med transformeringstabell som verkar fungera acceptabelt. Författarna noterar liksom Mellor & Stokes (2007) att SDQ främst används för screening och att göra en bedömning om ett barn behöver stöd eller inte, vilket de tycker är rimligt, iaf med hjälp av den modifierade Total Problemsskalan utifrån transformeringstabellen som de tillhandahåller.

Bosik et al. (2021) analyserar lärarversionen i USA med CTT (estimator WLSMV, bra!) och multidimensionell IRT (2PL). Två items togs bort tidigt i analysen, och 5-faktormodellen uppvisade sedan knappt acceptabel model fit som var i princip identisk med 2- och 3-faktormodellerna (dock märkliga frihetsgrader i resultattabellen). Låg spridning av item difficulties. DIF undersöktes med CTT och problem med invarians föreligger mellan kön för hela skalan, och i synnerhet för 2 items. Man fann med samma metodik även DIF mellan olika årskurser. Targeting redovisas inte explicit, d.v.s. hur item/test information matchar mot samplets egenskaper.

Information om villkor för användning av SDQ finns på följande webbsida: <https://www.sdqinfo.org/>

Björnsdotter, A., Enebrink, P., & Ghaderi, A. (2013). Psychometric properties of online administered parental strengths and difficulties questionnaire (SDQ), and normative data based on combined online and paper-and-pencil administration. *Child and Adolescent Psychiatry and Mental Health*, 7(1), 40. <https://doi.org/10.1186/1753-2000-7-40>

Bosik, S. W., Harris, B., Alibrahim, N., & Donovan, C. (2021). Construct Validity of the Teacher Version of the Strengths and Difficulties Questionnaire in a Sample of Elementary-Aged Children in the United States. *Contemporary School Psychology*. <https://doi.org/10.1007/s40688-021-00371-8>

Dahlberg, A., Ghaderi, A., Sarkadi, A., & Salari, R. (2019). SDQ in the Hands of Fathers and Preschool Teachers—Psychometric Properties in a Non-clinical Sample of 3–5-Year-Olds. *Child Psychiatry & Human Development*, 50(1), 132–141. <https://doi.org/10.1007/s10578-018-0826-4>

Dahlberg, A., Fält, E., Sarkadi, A., Fabian, H., & Salari, R. (2020). Strengths and Difficulties Questionnaire (SDQ) as an assessment tool for measuring emotional and behavioural problems in young children. *European Journal of Public Health*, 30(Supplement\_5). <https://doi.org/10.1093/eurpub/ckaa165.418>

- Essau, C. A., Olaya, B., Anastassiou-Hadjicharalambous, X., Pauli, G., Gilvarry, C., Bray, D., O’callaghan, J., & Ollendick, T. H. (2012). Psychometric properties of the Strength and Difficulties Questionnaire from five European countries. *International Journal of Methods in Psychiatric Research*, 21(3), 232–245. <https://doi.org/10.1002/mpr.1364>
- Gustafsson, B. M., Gustafsson, P. A., & Proczkowska-Björklund, M. (2016). The Strengths and Difficulties Questionnaire (SDQ) for preschool children—A Swedish validation. *Nordic Journal of Psychiatry*, 70(8), 567–574. <https://doi.org/10.1080/08039488.2016.1184309>
- Hagquist, C. (2007). The psychometric properties of the self-reported SDQ – An analysis of Swedish data based on the Rasch model. *Personality and Individual Differences*, 43(5), 1289–1301. <https://doi.org/10.1016/j.paid.2007.03.022>
- Keller, F., & Langmeyer, A. N. (2019). An Item Response Theory Analysis of the Strength and Difficulties Questionnaire (SDQ). *European Journal of Psychological Assessment*, 35(2), 266–279. <https://doi.org/10.1027/1015-5759/a000390>
- Kersten, P., Vandal, A. C., Elder, H., & McPherson, K. M. (2018). Strengths and Difficulties Questionnaire: Internal validity and reliability for New Zealand preschoolers. *BMJ Open*, 8(4), e021551. <https://doi.org/10.1136/bmjopen-2018-021551>
- Lundh, L.-G., Wångby-Lundh, M., & Bjärehed, J. (2008). Self-reported emotional and behavioral problems in Swedish 14 to 15-year-old adolescents: A study with the self-report version of the Strengths and Difficulties Questionnaire. *Scandinavian Journal of Psychology*, 49(6), 523–532. <https://doi.org/10.1111/j.1467-9450.2008.00668.x>
- Malmberg, M., Rydell, A., & Smedje, H. (2003). Validity of the Swedish version of the Strengths and Difficulties Questionnaire (SDQ-Swe). *Nordic Journal of Psychiatry*, 57(5), 357–363. <https://doi.org/10.1080/08039480310002697>
- McCarron, R. H., Gracey, F., & Bateman, A. (2021). Detecting mental health problems after paediatric acquired brain injury: A pilot Rasch analysis of the strengths and difficulties questionnaire. *Neuropsychological Rehabilitation*, 31(7), 1048–1068. <https://doi.org/10.1080/09602011.2020.1760111>
- Mellor, D., & Stokes, M. (2007). The Factor Structure of the Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment*, 23(2), 105–112. <https://doi.org/10.1027/1015-5759.23.2.105>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review : DR*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Smedje, H., Broman, J.-E., Hetta, J., & von Knorring, A.-L. (1999). Psychometric properties of a Swedish version of the “Strengths and Difficulties Questionnaire.” *European Child & Adolescent Psychiatry*, 8(2), 63–70. <https://doi.org/10.1007/s007870050086>

## Parental Stress Scale

Två danska Rasch-analyser (Nielsen et al., 2020; Pontoppidan et al., 2018) ger item- och faktor-reduktion till 16 items och 2 subskalor. Dock har ena subskalan (Lack of parental satisfaction) mycket dålig targeting, medan subskala Parental Stress fungerar bra. Tyvärr oklart hur Rasch-analyserna gått till utifrån den information som finns. I artiklarna refereras inte till någon etablerad mjukvara, och man redogör inte för hur problemen med svarskategorierna ser ut och varför man kom till slutsatsen att dikotomisera istf att slå ihop kategorier. En kinesisk Rasch-analys med Winsteps har inga problem med svarskategorierna (Leung & Tsang, 2010)

Det finns stora problem med local dependencies och DIF (barnets ålder, föräldrarnas kön och utbildningsnivå) i både danska och kinesiska analyser. Transformeringsstabeller finns för ordinal-intervall-omvandling (danska), som påstås ta hänsyn till problemen, men känns ändå svårt att rekommendera denna.

Leung, C., & Tsang, S. K. M. (2010). The Chinese Parental Stress Scale: Psychometric Evidence Using Rasch Modeling on Clinical and Nonclinical Samples. *Journal of Personality Assessment*, 92(1), 26–34. <https://doi.org/10.1080/00223890903379209>

Nielsen, T., Pontoppidan, M., & Rayce, S. B. (2020). The Parental Stress Scale revisited: Rasch-based construct validity for Danish parents of children 2–18 years old with and without behavioral problems. *Health and Quality of Life Outcomes*, 18(1), 281. <https://doi.org/10.1186/s12955-020-01495-w>

Pontoppidan, M., Nielsen, T., & Kristensen, I. H. (2018). Psychometric properties of the Danish Parental Stress Scale: Rasch analysis in a sample of mothers with infants. *PLOS ONE*, 13(11), e0205662. <https://doi.org/10.1371/journal.pone.0205662>

## KIDSCREEN

KIDSCREEN-10: Gong et al. (2021) hittar problem i den kinesiska versionen med svarskategori 2 och 3 (seldom & sometimes), samt local deps (item 4 och 5) och DIF (item 3 och 4). Efter sammanslagning av svars kategorier och eliminering av 3 items fungerar en 7-itemversion ok, men man varnar för att den inte täcker upp det teoretiska konstruktet på ett adekvat sätt.

Barkmann et al. (2021) menar att det finns problem med faktormodellen för tyska versionen, och Rasch-analys visar på stora residualer för en del items och bristande targeting. En annan tysk artikel (Müller & Haenni Hoti, 2020) menar att data inte passar Rasch-modellen, men funkar i en "graphical loglinear Rasch model", samtidigt som det finns local dependencies för 4 item-par, och DIF för både kön och medborgarskap. Targeting/reliabilitet är också suboptimalt.

Erhart, Ottova, et al. (2009) tittade enbart på infit MNSQ (0.7-1.3) men skriver själva i limitationsdelen att de inte testade unidimensionalitet eller DIF. Oklart vad denna artikel tillför eftersom residuals i sig själv inte ger tillräcklig information för någon typ av slutsats.

Ravens-Sieberer et al. (2005; 2010) hänvisar till en separat publikation där Rasch-analys ska presenteras, men den går inte att hitta på internet(!).



Erhart et al. (2010) fokuserar på delen av KIDSCREEN om fysiskt välbefinnande, självskattad av barn 8-18 år utifrån 19 items. Analyser görs med både Rasch och CTT, med itemreduktion som primärt syfte. Åtta items tas bort utifrån dålig item fit. Två kvarvarande item-par uppvisade problematisk residuallkorrelation och två item som uppvisade DIF mellan länder. DIF analyseras enbart mellan grupperna med eller utan CP-diagnos. Generellt redovisas ovanliga indexvärden från Rasch-analysen, och några aspekter saknas helt i analysen (targeting, reliabilitet, mätosäkerheter, svars-kategorier).

RISE har inte hittat någon publicerad svensk psykometrisk studie på KIDSCREEN-10. Interna projektdata har analyserats och påtagliga problem finns med dimensionalitet och residuallkorrelationer. Efter att ha tagit bort 4 items (1,3,4,9) är unidimensionalitet och item fit adekvata, och ingen DIF mellan kön föreligger, men reliabiliteten är för låg (PSI = 0.63).

Berman et al. (2016) analyserar svenska data för KIDSCREEN-27, men helt utan psykometrisk analys.

Vélez et al. (2016) gör en Rasch-analys på KIDSCREEN-27 (Colombiansk data) och kommer fram till att de två lägsta svars-kategorierna bör slås ihop, samt att fyra items inte passar mätmodellen .

Barkmann, C., Otto, C., Meyrose, A.-K., Reiss, F., Wüstner, A., Voß, C., Erhart, M., & Ravens-Sieberer, U. (2021). Psychometrie und Normierung des Lebensqualitätsinventars KIDSCREEN in Deutschland. *Diagnostica*, 67(1), 2–12. <https://doi.org/10.1026/0012-1924/a000257>

Berman, A. H., Liu, B., Ullman, S., Jadbäck, I., & Engström, K. (2016). Children's Quality of Life Based on the KIDSCREEN-27: Child Self-Report, Parent Ratings and Child-Parent Agreement in a Swedish Random Population Sample. *PLoS ONE*, 11(3). <https://doi.org/10.1371/journal.pone.0150545>

Erhart, Ravens-Sieberer, et al. (2009) analyserar KIDSCREEN-52 för barn med och utan CP-diagnos. Självskattningsversionen uppges fungera väl, medan föräldraskattningen har problem med ett item i delskalan som har med skolan att göra.

Erhart, M., Ottova, V., Gaspar, T., Jericek, H., Schnohr, C., Alikasifoglu, M., Morgan, A., Ravens-Sieberer, U., & the HBSC Positive Health Focus Group. (2009). Measuring mental health and well-being of school-children in 15 European countries using the KIDSCREEN-10 Index. *International Journal of Public Health*, 54(2), 160–166. <https://doi.org/10.1007/s00038-009-5407-7>

Erhart, M., Hagquist, C., Auquier, P., Rajmil, L., Power, M., Ravens-Sieberer, U., & Group, the E. K. (2010). A comparison of Rasch item-fit and Cronbach's alpha item reduction analysis for the development of a Quality of Life scale for children and adolescents. *Child: Care, Health and Development*, 36(4), 473–484. <https://doi.org/10.1111/j.1365-2214.2009.00998.x>

Erhart, M., Ravens-Sieberer, U., Dickinson, H. O., & Colver, A. (2009). Rasch Measurement Properties of the KIDSCREEN Quality of Life Instrument in Children with Cerebral Palsy and Differential Item Functioning between Children with and without Cerebral Palsy. *Value in Health*, 12(5), 782–792. <https://doi.org/10.1111/j.1524-4733.2009.00508.x>

Gong, Z., Xue, J., Han, Z., & Li, Y. (2021). Validation of the Chinese Version of KIDSCREEN-10 Quality of Life Questionnaire: A Rasch Model Estimation. *Frontiers in Psychology*, 12, 647692. <https://doi.org/10.3389/fpsyg.2021.647692>

Müller, M., & Haenni Hoti, A. (2020). Item analysis of the KIDSCREEN-10 using Rasch modelling. *Health and Quality of Life Outcomes*, 18(1), 342. <https://doi.org/10.1186/s12955-020-01596-6>

Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., Auquier, P., Power, M., Abel, T., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J., & Group, E. K. (2005). KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Review of Pharmacoeconomics & Outcomes Research*, 5(3), 353–364. <https://doi.org/10.1586/14737167.5.3.353>

Ravens-Sieberer, U., Erhart, M., Rajmil, L., Herdman, M., Auquier, P., Bruil, J., Power, M., Duer, W., Abel, T., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J., & the European KIDSCREEN Group. (2010). Reliability, construct and criterion validity of the KIDSCREEN-10 score: A short measure for children and adolescents' well-being and health-related quality of life. *Quality of Life Research*, 19(10), 1487–1500. <https://doi.org/10.1007/s11136-010-9706-5>

Vélez, C.-M., Lugo-Agudelo, L.-H., Hernández-Herrera, G.-N., & García-García, H.-I. (2016). Colombian Rasch validation of KIDSCREEN-27 quality of life questionnaire. *Health and Quality of Life Outcomes*, 14(1), 67. <https://doi.org/10.1186/s12955-016-0472-0>

## Child-Parent Relationship Scale (CPRS)

Mäter föräldra-barnrelationen med 6 frågor som är identiska med delskalan ur KIDSCREEN-52. Hagquist beskriver i artikeln att CPRS fungerar väl med avseende på svars kategorier, unidimensionalitet och reliabilitet. Artikeln saknar dock ingående redovisning av psykometriska analyser, vilket gör att detta underlag skulle behöva förstärkas med en utförligare rapportering av psykometriska parametrar.

KIDSCREEN är fritt att använda för icke-kommersiella syften, men kräver registrering: <https://www.kidscreen.org/english/conditions-of-use/>

Hagquist, C. (2016). Family residency and psychosomatic problems among adolescents in Sweden: The impact of child-parent relations. *Scandinavian Journal of Public Health*, 44(1), 36–46. <https://doi.org/10.1177/1403494815610664>

## PedsQL

Svensk CTT-artikel (Petersen et al., 2009) använder ML som estimator, vilket inte är bra med den typ av skeva data som PedsQL brukar innehålla (WLSMV är att rekommendera). Invarians mellan kön påvisas med CTT.

En annan svensk artikel (Sand et al., 2012) jämför PedsQL version 3 och 4, men är psykometriskt undermålig. Cronbach's alpha och t-test/ANOVA på rådata för items ger otillräcklig information.

RISE har kunnat analysera ett mindre sampel med PedsQL-data från svenska skolbarn. Delskalorna har för låg reliabilitet var för sig. Svarskategorierna fungerar dåligt. Det finns flera item-par med för hög residualkorrelation. Det kan vara möjligt att använda PedsQL-data i analyser efter omfattande åtgärder (4 item-par som görs till superitems + svarskategorier åtgärdade), och enbart som helskala.

Landfeldt et al. (2021) hittar problem med svarskategorierna i 20 av 23 items (tyska versionen). Även problem med multidimensionalitet, första 8 items utgör egen dimension i deras analys.

Vélez et al., (2016) hittar också stora problem med svarskategorierna i den spanska versionen (Colombia), samt att targeting och reliabilitet är dåliga – ingen av delskalorna når PSI 0.7. Positivt är att ingen DIF kan påvisas mellan kön.

Även Kook & Varni (2008) identifierar låg reliabilitet på delskalorna (den bästa når PSI 0.60) för den koreanska versionen.

Amin et al. (2012) redovisar reliabilitet (PSI) 0.78 för kanadensiska föräldraversionen, och rekommenderar att ändra från 5 till 3 svarskategorier. Targeting ser bättre ut, men man använder den märkliga "subdomäner som subtest"-varianten för att få unidimensionalitet. OBS att detta inte är självrporteringsversionen.

Jafari et al. (2012) analyserar den persiska versionen för både barn och föräldrar. Även här svarskategorier med problem och låga PSI för delskalorna (0.63 som bäst).

Amin, L., Rosenbaum, P., Barr, R., Sung, L., Klaassen, R. J., Dix, D. B., & Klassen, A. (2012). Rasch analysis of the PedsQL: An increased understanding of the properties of a rating scale. *Journal of Clinical Epidemiology*, 65(10), 1117–1123. <https://doi.org/10.1016/j.jclinepi.2012.04.014>

Jafari, P., Bagheri, Z., Ayatollahi, S. M. T., & Soltani, Z. (2012). Using Rasch rating scale model to reassess the psychometric properties of the Persian version of the PedsQL™ 4.0 Generic Core Scales in school children. *Health and Quality of Life Outcomes*, 10(1), 27. <https://doi.org/10.1186/1477-7525-10-27>

Kook, S. H., & Varni, J. W. (2008). Validation of the Korean version of the pediatric quality of life inventory™ 4.0 (PedsQL™) generic core scales in school children and adolescents using the rasch model. *Health and Quality of Life Outcomes*, 6(1), 41. <https://doi.org/10.1186/1477-7525-6-41>

Landfeldt, E., Iff, J., Henricson, E., Vishwanathan, V., Chidambaranathan, S., Biggar, W. D., McAdam, L. C., Mah, J. K., Tulinius, M., Cnaan, A., Morgenroth, L. P., Leshner, R., TesiRocha, C., Thangarajh, M., Duong, T., Kornberg, A., Ryan, M., Nevo, Y., Dubrovsky, A., ... McDonald, C. M. (2021). Rasch Analysis of the Pediatric Quality of Life Inventory 4.0 Generic Core Scales Administered to Patients With Duchenne Muscular Dystrophy. *Value in Health*, 24(10), 1490–1498. <https://doi.org/10.1016/j.jval.2021.05.016>

Petersen, S., Hägglöf, B., Stenlund, H., & Bergström, E. (2009). Psychometric properties of the Swedish PedsQL, Pediatric Quality of Life Inventory 4.0 generic core scales. *Acta Paediatrica*, 98(9), 1504–1512. <https://doi.org/10.1111/j.1651-2227.2009.01360.x>

Sand, P., Kljajić, M., Schaller, J., & Forsander, G. (2012). The reliability of the Health Related Quality Of Life questionnaire PedsQL 3.0 Diabetes Module™ for Swedish children with Type 1 diabetes: Reliability in PedsQL Diabetes Module, Swedish version. *Acta Paediatrica*, 101(8), e344–e349. <https://doi.org/10.1111/j.1651-2227.2012.02706.x>

Vélez, C. M., Villada Ramírez, A. C., Arias, A. C. A., & Eslava-Schmalbach, J. H. (2016). Validación por modelo de Rasch del Cuestionario de Calidad de Vida (PedsQL 4.0®) en niños y adolescentes colombianos. *Revista Colombiana de Psiquiatría*, 45(3), 186–193. <https://doi.org/10.1016/j.rcp.2015.12.002>

## Eyberg Child Behavior Inventory (ECBI)

ECBI har 36 items som kan användas för skattning av intensitet (skala med 7 steg) eller skattning av problem (svarsalternativ ja/nej). Skalan summeras vanligtvis, vilket indikerar att man menar att ECBI är unidimensionell.

Axberg et al. (2008) har gjort CFA på svenska data från en interventionsstudie på De Otroliga Åren och använt 22-items i en 3-faktormodell som bygger på tidigare analyser (Burns & Patterson, 2000). Dock uppger man stora problem med skeva data och dikotomiserar svarskategorierna innan analys. Det hade varit lämpligare att använda WLSMV som estimator för denna typ av data utan dikotomisering, eller med justeringar av svarskategorierna. Svarskategorierna hade kunnat analyseras med IRT/Rasch för att hitta en bättre lösning än dikotomisering. Två item-par med korrelerade residualer (items 25–27 och 21–22) läggs in i CFA-modellen, vilket ger bättre model fit, men löser inte problemen med bias i indexvärden som skapas av residualkorrelationer. Delskalan Conduct Problem Behavior (CPB) har väldigt låga faktorladdningar på 5 items (0.25-0.42), varav 4 är de två itemparen med korrelerade residualer, men detta problematiseras inte i artikeln.

Abrahamse et al. (2015) analyserar den holländska versionen och har problem med svars kategorier, så man dikotomiserar data. Data passar inte Rasch-modellen, och man byter analys till OPLM i stället och viktar items. Mycket märklig analysprocedur överlag, ser inte att denna artikel tillför något utom information att svarskategorierna har problem och att data inte fungerar i Rasch-modellen. Man påstår också felaktigt att Rasch-modellen kräver dikotoma data. Faktorladdning 0.09 lyfts fram som "satisfactory", och senare i samma stycke beskrivs faktorladdning under .30 som "poor". Oklart hur denna artikel passerat peer review.

Machado (2020) studerar 221 mammor med barn (6-12 år) och hittar problem med svarskategorierna, vilket justeras från 7 till 5 kategorier för analysen. Item 7 och 10 i delskalan Oppositional Defiant Behavior towards Adults har för hög respektive för låg item fit. I delskalan CPB fann man problem med unidimensionalitet och residualkorrelationer. Alla delskalor uppvisar reliabilitet (PSI) över 0.7.

Att svarsskalan med 7 steg fungerar dåligt i flera studier är inte förvånande när två av svarskategorierna delar etiketter med en annan svarskategori: "1 (never), 2 and 3 (seldom), 4 (sometimes), 5 and 6 (often), and 7 (always)."

RISE har fått tillgång till svenska data för ECBI, både intensitets- och problemskattningar, och genomfört en psykometrisk analys. Dock var sampelstorleken inte tillräcklig för att analyserna ska kunna beskrivas som välbelagda (159 individer vid två mättillfällen), vilket innebär att resultaten ska

tolkas med försiktighet. I likhet med tidigare analyser fann vi stora problem med multidimensionalitet och svars-kategorier för intensitetsskattningarna. Efter sammanslagning av svars-kategorier från 7 till 4 (och för vissa items från 7 till 3 svars-kategorier) prövades indelningen i tre dimensioner som Axberg et al. (2008) använt för intensitetsskattningarna. Separata analyser gjordes för varje delskala. Delskalan Oppositional Defiant Behavior towards Adults (ODBTA) med 10 items fungerade acceptabelt efter eliminering av två items, men uppvisade problem med DIF kopplat till barnets kön för 6 av 8 items, vilket innebär att skattningarna för pojkar och flickor inte är direkt jämförbara. ODBTA nådde inte reliabilitet 0.7 vid någon del av skalan. Delskalan Inattentive behaviors (InattB) har god unidimensionalitet, men uppvisade problem med låg item fit för 3 av 4 items och mycket låg reliabilitet och kan därför ej rekommenderas. DIF har ej undersökts för InattB. Conduct Problem Behavior (CPB, 8 items) har problem med låg item fit för tre items, och residualkorrelationer mellan hela 4 par av items, varav två är samma som Axberg et al. (2008) också fann. Reliabiliteten för CPB når inte heller i närheten av 0.7. DIF ej analyserat för CPB.

Gällande delen med problemskattningar (ja/nej-frågor) av ECBI gjordes en explorativ analys som resulterade i två delskalor, men ingen av dem nådde reliabilitet 0.7. Delskalan med fler items uppvisade DIF för barnets kön för tre items.

Axberg, U., Johansson Hanse, J., & Broberg, A. G. (2008). Parents' description of conduct problems in their children—A test of the Eyberg Child Behavior Inventory (ECBI) in a Swedish sample aged 3-10. *Scandinavian Journal of Psychology*, 49(6), 497–505. <https://doi.org/10.1111/j.1467-9450.2008.00670.x>

Abrahamse, M. E., Junger, M., Leijten, P. H. O., Lindeboom, R., Boer, F., & Lindauer, R. J. L. (2015). Psychometric Properties of the Dutch Eyberg Child Behavior Inventory (ECBI) in a Community Sample and a Multi-Ethnic Clinical Sample. *Journal of Psychopathology and Behavioral Assessment*, 37(4), 679–691. <https://doi.org/10.1007/s10862-015-9482-1>

Burns, G. L., & Patterson, D. R. (2000). Factor Structure of the Eyberg Child Behavior Inventory: A Parent Rating Scale of Oppositional Defiant Behavior Toward Adults, Inattentive Behavior, and Conduct Problem Behavior. *Journal of Clinical Child Psychology*, 29(4), 569–577. [https://doi.org/10.1207/S15374424JCCP2904\\_9](https://doi.org/10.1207/S15374424JCCP2904_9)

Eyberg, S. M., & Robinson, E. A. (1983). Conduct Problem Behavior: Standardization of a Behavior Rating Scale with Adolescents. *Journal of Clinical Child Psychology*, 12(3), 347–354. [https://doi.org/10.1207/s15374424jccp1203\\_19](https://doi.org/10.1207/s15374424jccp1203_19)

Machado, E. (2020). *Psychometric Evaluation of the Eyberg Child Behavior Inventory in an Ethnically Diverse Sample*. Nova Southeastern University.

## Dyadic adjustment scale

Ahlborg et al. (2005) analyserar svenska data, men tar ej hänsyn till klustring i data, d.v.s. att föräldraren utgör kluster (ej oberoende data). Svaga dimensionalitetsanalyser, enbart PCA och

item-total korrelationer, inga residuallkorrelationer. Ingen invariansanalys, ingen analys av svars-kategorier.

Sabourin et al. (2005) gör en märklig invariansanalys. Intressant metodik för itemreduktion där diskrimineringskapacitet för items runt cut-offvärden används primärt. Men oklart värde i övrigt.

Ahlborg, T., Persson, L.-O., & Hallberg, L. R.-M. (2005). Assessing the Quality of the Dyadic Relationship in First-Time Parents: Development of a New Instrument. *Journal of Family Nursing*, 11(1), 19–37. <https://doi.org/10.1177/1074840704273462>

Sabourin, S., Valois, P., & Lussier, Y. (2005). Development and Validation of a Brief Version of the Dyadic Adjustment Scale With a Nonparametric Item Analysis Model. *Psychological Assessment*, 17(1), 15–27. <https://doi.org/10.1037/1040-3590.17.1.15>

## Emotion Regulation Questionnaire (ERQ)

Enebrink et al. (2013) redogör inte för vilken estimator som använts i analyserna. Det beskrivs flera problem med CFA som inte hanteras på adekvat sätt, exempelvis crossloadings och korrelerade residualer. Ingen invariansanalys, och eftersom model fit hade påtagliga skillnader mellan män och kvinnor kan man anta att invarians inte föreligger.

Brandão et al. (2017) gör CFA och IRT, dock enbart kvinnor (med cancer). Verkar fungera ok. Invarians för ålder.

Brandão, T., Schulz, M. S., Gross, J. J., & Matos, P. M. (2017). The emotion regulation questionnaire in women with cancer: A psychometric evaluation and an item response theory analysis. *Psycho-Oncology*, 26(10), 1647–1653. <https://doi.org/10.1002/pon.4356>

Enebrink, P., Björnsdotter, A., & Ghaderi, A. (2013). The Emotion Regulation Questionnaire: Psychometric Properties and Norms for Swedish Parents of Children Aged 10-13 Years. *Europe's Journal of Psychology*, 9(2), 289–303. <https://doi.org/10.5964/ejop.v9i2.535>

## Child Health Utility 9D (CHU9D)

Lindvall et al. (2021). Årskurs 1-3 borde inte använda CHU9D p.g.a. undermålig model fit. Korrelation mellan items om "worried" och "sad" kan ge bias i indexvärden. Item "Abilities to join in activities" laddar .32 för barn/unga, och borde således tas bort eller omarbetas, men det rekommenderas inte. Ingen invariansanalys finns.

Lindvall, K., Vaezghasemi, M., Feldman, I., Ivarsson, A., Stevens, K. J., & Petersen, S. (2021). Feasibility, reliability and validity of the health-related quality of life instrument Child Health Utility 9D (CHU9D) among school-aged children and adolescents in Sweden. *Health and Quality of Life Outcomes*, 19(1), 193. <https://doi.org/10.1186/s12955-021-01830-9>



## Health Behaviour in School-aged Children Symptom Checklist (HBSC-SCL)

Hagquist (2019) identifierar problem med svarskategorierna och rekommenderar att justera från 5 till 3 kategorier. Ett item ("Stomach ache") visar på DIF mellan kön, särskilt för högre årskurser.

Hagquist et al. (2019) hittar problem med svarskategorierna och gör sin analys efter reduktion från 5 till 3 kategorier för samtliga frågor. Targeting visar på takeffekter. Unidimensionalitet fungerar bra. Inga problematiska residualkorrelationer. Item fit residuals rapporteras inte.

Ravens-Sieberer et al (2008) gör en Rasch-analys av HBSC och finner att den fungerar väl på alla parametrar utom en fråga om sömnsvårigheter som uppvisar DIF mellan de länder som finns med i analysen (Sverige är en av dem).

Hagquist & Andrich (2004) gör en Rasch-analys på svenska data från 80- och 90-tal. Svarskategorierna fungerar inte som väntat, och åtgärden i denna artikel är att ta bort 3 av 8 items ur analysen. Man identifierar DIF mellan kön för två items ("huvudvärk" och "sömnsvårigheter"), vilket kan lösas med "item split", men innebär att användning av summerade ordinala rådata blir missvisande. Targeting visar på takeffekter. Item fit residuals rapporteras inte. I diskussionen tar man upp alternativet att slå ihop svars kategorier, vilket Hagquist et al. (2019) gjorde i sin senare artikel.

Hagquist, C. (2019). Explaining differential item functioning focusing on the crucial role of external information – an example from the measurement of adolescent mental health. *BMC Medical Research Methodology*, 19(1), 185. <https://doi.org/10.1186/s12874-019-0828-3>

Hagquist, C., Due, P., Torsheim, T., & Välimaa, R. (2019). Cross-country comparisons of trends in adolescent psychosomatic symptoms – a Rasch analysis of HBSC data from four Nordic countries. *Health and Quality of Life Outcomes*, 17(1), 27. <https://doi.org/10.1186/s12955-019-1097-x>

Hagquist, C., & Andrich, D. (2004). Measuring subjective health among adolescents in Sweden. *Social Indicators Research*, 68(2), 201–220. <https://doi.org/10.1023/B:SOCI.0000025593.97559.7f>

Ravens-Sieberer, U., Erhart, M., Torsheim, T., Hetland, J., Freeman, J., Danielson, M., Thomas, C., & The HBSC Positive Health Group. (2008). An international scoring system for self-reported health complaints in adolescents. *European Journal of Public Health*, 18(3), 294–299. <https://doi.org/10.1093/eurpub/ckn001>

## General Population-Clinical Outcomes in Routine Evaluation (GP-CORE)

Mäter "general psychological distress" med 14 items. Ursprungsartikeln av Evans et al. (2005) redovisar en explorativ analys med PCA där man hittar 4 faktorer, som är svagt korrelerade och

därmed är GP-CORE inte unidimensionell och bör inte summeras till ett indexvärde. Inga övriga analyser utöver dimensionalitet görs av psykometriska egenskaper.

En IRT-baserad studie (Klocek et al., 2021) kommer fram till att en bifaktormodell med två metodfaktorer (positiva/negativa items) passar data bäst, men även 2-faktormodellen har acceptabel model fit. Dock föreligger DIF mellan kön för åtminstone fyra items, vilket innebär att invarians inte råder.

RISE har inte hittat någon publicerad svensk psykometrisk studie för GP-CORE. Däremot finns en svensk studie på CORE-OM (Elfström et al., 2012), som GP-CORE hämtat sina items från. Det går dock inte att bedöma kvaliteten på GP-CORE utifrån den studien, som i sin tur saknar i princip alla centrala psykometriska analyser/parametrar.

Evans, C., Connell, J., Audin, K., Sinclair, A., & Barkham, M. (2005). Rationale and development of a general population well-being measure: Psychometric status of the GP-CORE in a student sample. *British Journal of Guidance & Counselling*, 33(2), 153–173.

<https://doi.org/10.1080/03069880500132581>

Elfström, M. L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S. G. (2012). Validation of the Swedish Version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM): Validation of the Swedish Version of CORE-OM. *Clinical Psychology & Psychotherapy*, n/a-n/a. <https://doi.org/10.1002/cpp.1788>

Klocek, A., Řiháček, T., & Cígler, H. (2021). Psychometric evaluation of the Clinical Outcome in Routine Evaluation – General Population: Czech version. *Československá Psychologie*, 65(2), 163–177. <https://doi.org/10.51561/cspsych.65.2.163>

## Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS)

Åtskilliga Rasch-analyser finns publicerade, se Stewart-Brown, 2021, för en överblick. RISE har tillsammans med Centrum för Epidemiologi och Samhällsmedicin (CES) publicerat en psykometrisk artikel på svenska data (Melin et al., 2022). Överlag fungerar SWEMWBS väl. Det finns dock svagheter i mätprecision när det gäller högre nivåer av välbefinnande, och en tydlig takeffekt.

SWEMWBS är fritt att använda för icke-kommersiellt bruk, men kräver att man ansöker om en licens: <https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/non-commercial-licence-registration/>

Melin, J., Lundin, A., & Johansson, M. (2022). An off-target scale limits the utility of Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) as a measure of well-being in public health surveys. *Public Health*, 202, 43–48. <https://doi.org/10.1016/j.puhe.2021.10.009>

Stewart-Brown, S. (2021). *15 years on: Insights and reflections on the Warwick-Edinburgh Mental Wellbeing Scales (WEMWBS)*. What Works Centre for Wellbeing.



[https://whatworkswellbeing.org/wp-content/uploads/2021/11/Measuring-Wellbeing-Series-Sarah-Stewart-Brown\\_final.pdf](https://whatworkswellbeing.org/wp-content/uploads/2021/11/Measuring-Wellbeing-Series-Sarah-Stewart-Brown_final.pdf)

## Brunnsviken Brief Quality of Life Scale (BBQ)

BBQ har sex items som mäter "satisfaction" och motsvarande sex items som mäter "importance", alltså hur viktig frågan är för individen. Bortsett från den mycket tveksamma metoden att multiplicera svarskategorier för "satisfaction" med "importance" verkar BBQ fungera acceptabelt. Artikeln av Lindner et al. (2016) använder CTT för analys av dimensionalitet, men saknar invariansanalys och analys av övriga grundläggande mätegenskaper.

RISE har kunnat analysera data från ett sampel från normalpopulation. Dimensionaliteten är bra, targeting är ok (viss takeffekt), svarskategorierna fungerar fint, det föreligger ingen DIF mellan kön. Person reliability är över 0.70 mellan ca -1.8 till +1.8 logits. RISE rekommenderar att enbart använda de sex "satisfaction"-frågorna, och inte kombinera dem med "importance"-skattningar. Omvandlingstabell till intervallskala och mätosäkerheter finns i slutet av denna bilaga. RISE vill rikta ett tack till Alexander Rozental och David Forsström för tillgängliggörande av data för analys av BBQ.

BBQ är fritt att använda för icke-kommersiella syften och finns tillgängligt på webbsidan <http://www.bbqscale.com>

Lindner, P., Frykheden, O., Forsström, D., Andersson, E., Ljótsson, B., Hedman, E., Andersson, G., & Carlbring, P. (2016). The Brunnsviken Brief Quality of Life Scale (BBQ): Development and Psychometric Evaluation. *Cognitive Behaviour Therapy*, 45(3), 182–195. <https://doi.org/10.1080/16506073.2016.1143526>

## Parenting self-efficacy (TOPSE)

Kendall & Bloomfield (2005) beskriver bara framtagandet av items och anger Cronbach's alpha. Ingen psykometrisk analys.

Uppföljningsartikeln av Bloomfield & Kendall (2007) är lika undermålig som den tidigare artikeln.

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Kendall, S., & Bloomfield, L. (2005). Developing and validating a tool to measure parenting self-efficacy. *Journal of Advanced Nursing*, 51(2), 174–181. <https://doi.org/10.1111/j.1365-2648.2005.03479.x>

Bloomfield, L., & Kendall, S. (2007). Testing a parenting programme evaluation tool as a pre- and post-course measure of parenting self-efficacy. *Journal of Advanced Nursing*, 60(5), 487–493. <https://doi.org/10.1111/j.1365-2648.2007.04420.x>

## Parents' Sense of Competence (PSOC)

Gilmore och Cuskelly (2009) gör en PCA och presenterar faktorladdningar separat för män och kvinnor. Det är tydligt att det finns skillnader mellan dem, samma gäller Cronbach's alpha, vilket innebär att invarians är mycket osannolikt. Delskalan "Control" uppvisar så låg alpha att den tas bort ur analysen, och man menar att tre delskalor finns. Dock framgår att tidigare faktoranalyser visat olika resultat (exempelvis Ohan et al. (2000) med två delskalor), så det är tydligt att PSOC i sin ursprungliga form inte har en stabil faktorstruktur.

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Gilmore, L., & Cuskelly, M. (2009). Factor structure of the Parenting Sense of Competence scale using a normative sample. *Child: Care, Health and Development*, 35(1), 48–55.

<https://doi.org/10.1111/j.1365-2214.2008.00867.x>

Ohan, J. L., Leung, D. W., & Johnston, C. (2000). The Parenting Sense of Competence scale: Evidence of a stable factor structure and validity. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, 32(4), 251–261. <https://doi.org/10.1037/h0087122>

## Parenting Scale

Salari et al. (2012) testar olika faktormodeller (med svenska data) utan att hitta någon med adekvat model fit. Man anger olika indexvärden för män och kvinnor och drar slutsatsen att enkäten förmodligen är känslig för egenskaper hos samplet. Detta är likvärdigt med att säga att DIF föreligger, vilket gör det olämpligt att använda skalan om inte DIF-egenskaperna analyseras och tas hänsyn till.

Salari, R., Terreros, C., & Sarkadi, A. (2012). Parenting Scale: Which Version Should We Use? *Journal of Psychopathology and Behavioral Assessment*, 34(2), 268–281. <https://doi.org/10.1007/s10862-012-9281-x>

## Parental Involvement

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Antipkina, I., & Ludlow, L. H. (2020). Parental Involvement as a Holistic Concept Using Rasch/Guttman Scenario Scales. *Journal of Psychoeducational Assessment*, 38(7), 846–865.

<https://doi.org/10.1177/0734282920903164>

## Alabama Parenting Questionnaire (APQ)

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Clerkin et al., (2007) skriver att Alabama Parenting Questionnaire (APQ) utvecklades av Shelton, Frick, och Wootton (1996). APQ har 42-items fördelade på 5 delskalor och använder 5 svarskategorier genomgående. Clerkins artikel studerar föräldrar till yngre barn (3-6 år), och tog bort 10 items som bedömdes olämpliga för målgruppen i denna studie.

Ward et al. (2015) har gjort en Rasch-analys, men enbart av svarskategorierna och finner att den femgradiga svarsskalan fungerar dåligt, och reducerar till tre svarskategorier. De konstaterar låg reliabilitet hos tre av fem delskalor (poor monitoring and supervision, inconsistent discipline, corporal punishment) bedömt med Cronbach's alpha.

Clerkin, S. M., Halperin, J. M., Marks, D. J., & Policaro, K. L. (2007). Psychometric Properties of the Alabama Parenting Questionnaire—Preschool Revision. *Journal of Clinical Child & Adolescent Psychology*, 36(1), 19–28. <https://doi.org/10.1080/15374410709336565>

Elgar, F. J., Waschbusch, D. A., Dadds, M. R., & Sigvaldason, N. (2007). Development and Validation of a Short Form of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies*, 16(2), 243–259. <https://doi.org/10.1007/s10826-006-9082-5>

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Psychometric Properties of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies*, 15(5), 595–614. <https://doi.org/10.1007/s10826-006-9036-y>

Ward, C. L., Gould, C., Kelly, J., & Mauff, K. (2015). Spare the rod and save the child: Assessing the impact of parenting on child behaviour and mental health. *South African Crime Quarterly*, 51, 9–22. <https://doi.org/10.4314/sacq.v51i0.2>

Zlomke, K., Bauman, S., & Lamport, D. (2015). Adolescents' Perceptions of Parenting Behavior: Validation of the Alabama Parenting Questionnaire Adolescent Self Report. *Journal of Child and Family Studies*, 24(11), 3159–3169. <https://doi.org/10.1007/s10826-015-0119-5>

Zlomke, K. R., Lamport, D., Bauman, S., Garland, B., & Talbot, B. (2014). Parenting Adolescents: Examining the Factor Structure of the Alabama Parenting Questionnaire for Adolescents. *Journal of Child and Family Studies*, 23(8), 1484–1490. <https://doi.org/10.1007/s10826-013-9803-5>

## Parents As Social Context Questionnaire (PASCQ)

Keijser et al. (2020). Översatt från engelska till svenska. Samplet bestod av 16-18 åringar och deras föräldrar (1345 dyader). Sex delskalor beskrivs som utfall av CTT-analys (Warmth, Rejection, Structure, Chaos, Autonomy support and Coercion). Ungdomsversionen har 24 items (6x4), föräldraversionen 30 items. 4 svarskategorier används. RMSEA och SRMR ser ok ut, CFI/TLI redovisas inte. Vilken estimator som används framgår inte heller. Enstaka faktorladdningar är under 0.4, utan att det uppmärksammas. Residualkorrelationer tas ej upp. Ingen invarianstestning, varken mellan kön på barn eller föräldrar. Det nämns könsskillnader mellan föräldrar, och skevhet i samplet (80% biologiska mammor). Skattningarna gjorda av förälder och barn skilde sig kraftigt åt. Samtliga

analyser av skillnader mellan grupper är gjorda med rådata, i stället för inom faktoranalyseramverket.

Keijsers, R., Nilsson, K. W., & Åslund, C. (2020). Measuring parental dimensions: A psychometric evaluation of the parents as social context questionnaire, Swedish version. *Cogent Psychology*, 7(1), 1757856. <https://doi.org/10.1080/23311908.2020.1757856>

## Adult-Child Relationship Scale (ACRS)

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Trentacosta, C. J., Criss, M. M., Shaw, D. S., Lacourse, E., Hyde, L. W., & Dishion, T. J. (2011). Antecedents and Outcomes of Joint Trajectories of Mother–Son Conflict and Warmth During Middle Childhood and Adolescence. *Child Development*, 82(5), 1676–1690. <https://doi.org/10.1111/j.1467-8624.2011.01626.x>

## Parenting Young Children (PARYC)

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Ursprungsartikel av McEachern et al. (2012). 21 frågor (7 svarskategorier) fördelade på tre faktorer. Studien gör först en PCA/EFA och sedan tre separata CFA-modeller på de tre subskalorna. Det finns signifikanta residualkorrelationer i två av subskalorna (supporting positive behavior (item 4 och 7) och proactive parenting (item 6 och 7)). Dessa korrelationer läggs in i modellen utan att problematisera vad residualkorrelationerna innebär. Oklart varför man inte gör en 3-faktormodell i CFA, för att exempelvis undersöka hur mycket faktorerna korrelerar med varandra. Ingen invariansanalys, ingen analys av svarskategorier, ingen targeting, inga mätosäkerheter. 2 items fungerar inte i 3-faktormodellen, men man verkar inte rekommendera att ta bort dem.

McEachern, A. D., Dishion, T. J., Weaver, C. M., Shaw, D. S., Wilson, M. N., & Gardner, F. (2012). Parenting Young Children (PARYC): Validation of a Self-Report Parenting Measure. *Journal of Child and Family Studies*, 21(3), 498–511. <https://doi.org/10.1007/s10826-011-9503-y>

## Ages and Stages Questionnaires: Social-Emotional (ASQ:SE)

31 items, 3 svarskategorier (scoras 0,5,10 och råpoängen summeras). IRT-analysen av Vaezghasemi et al. (2020) med svenska data dikotomiserar alla items (alternativen 5 och 10 blir samma), av oklara skäl. Ingen förklaring ges i artikeln, och det finns inga begränsningar i IRT-modeller som gör att bara dikotoma data skulle fungera. Man beskriver en 1PL-modell i metoddelen, vilket i praktiken innebär en Rasch-modell, om det inte är OPLM, men det nämns

inte. CFA används för dimensionalitetsanalys, men med "quasi maximum likelihood", vilket inte är en optimal estimator för dikotoma data (WLS/ULS att föredra). Enbart RMSEA och CD rapporteras för model fit. Residualkorrelationer analyseras inte, och tillsammans med svagheter i dimensionalitetsanalys ger det ett tveksamt intryck. TIF-figur visas, vilket är ett bra sätt att åskådliggöra reliabilitet över skalans omfång. Dock saknas matchning mot samplets egenskaper på samma skala, vilket innebär att targeting inte framgår (hur frågorna passar målgruppen). DIF för barnets kön uppvisas för ungefär hälften av frågorna, vilket innebär att jämförelser med råpoäng inte bör göras i en studie där barn av olika kön ingår. Bra att detta problem med invarians framgår, även om tillämpligheten för ASQ:SE som den brukar användas (med råpoäng) är oklar p.g.a. dikotomiseringen av data. DIF-analys för kön på förälder/vårdnadshavare som gör skattningen hade varit väldigt värdefullt, men saknas.

En uppföljande artikel av Vaezghasemi et al. (2022) undersöker cut-off-värden för svenska 3-åringar. Ingen psykometrisk analys ingår. Man verkar inte ta hänsyn till problemen med DIF mellan kön som föregående artikel tar upp.

Velikonja et al. (2017) gör en systematisk litteraturöversikt som enligt titeln värderar psykometriska egenskaper hos ASQ, men saknar helt analys av dimensionalitet och andra relevanta mätegenskaper.

Briggs et al. (2012) gör ingen psykometrisk analys.

Squires et al. (2001) menar att artikeln är en psykometri-artikel, men den saknar alla relevanta analyser av mätegenskaper.

Briggs, R. D., Stettler, E. M., Silver, E. J., Schrag, R. D. A., Nayak, M., Chinitz, S., & Racine, A. D. (2012). Social-Emotional Screening for Infants and Toddlers in Primary Care. *Pediatrics*, 129(2), e377–e384. <https://doi.org/10.1542/peds.2010-2211>

Squires, J., Bricker, D., Heo, K., & Twombly, E. (2001). Identification of social-emotional problems in young children using a parent-completed screening measure. *Early Childhood Research Quarterly*, 16(4), 405–419. [https://doi.org/10.1016/S0885-2006\(01\)00115-6](https://doi.org/10.1016/S0885-2006(01)00115-6)

Vaezghasemi, M., Eurenus, E., Ivarsson, A., Richter Sundberg, L., Silfverdal, S.-A., & Lindkvist, M. (2020). Social-emotional problems among Swedish three-year-olds: An Item Response Theory analysis of the Ages and Stages Questionnaires: Social-Emotional. *BMC Pediatrics*, 20(1), 149. <https://doi.org/10.1186/s12887-020-2000-y>

Vaezghasemi, M., Eurenus, E., Ivarsson, A., Richter Sundberg, L., Silfverdal, S. A., & Lindkvist, M. (2022). The Ages and Stages Questionnaire: Social-Emotional—What Is the Optimal Cut-Off for 3-Year-Olds in the Swedish Setting? *Frontiers in Pediatrics*, 10. <https://www.frontiersin.org/article/10.3389/fped.2022.756239>

Velikonja, T., Edbrooke-Childs, J., Calderon, A., Slead, M., Brown, A., & Deighton, J. (2017). The psychometric properties of the Ages & Stages Questionnaires for ages 2-2.5: A systematic review. *Child: Care, Health and Development*, 43(1), 1–17. <https://doi.org/10.1111/cch.12397>

## Perceived Stress Scale (PSS)

Nordin & Nordin (2013) kommer fram till att PSS-10 har två tydligt separata faktorer, men ignorerar detta med hänvisning till att det inte finns någon teoretisk grund för multidimensionaliteten, och använder ändå alla items som en unidimensionell skala. Oklart varför man gör en faktoranalys om man tänker ignorera resultatet. Detta är den enda svenska psykometriska artikeln som hittats.

En internationell översiktsartikel (Lee, 2012) konstaterar att både PSS-10 och -14 verkar ha bättre modelpassning med två faktorer än en, samtidigt som tvåfaktormodellen förklarar för lite variation i data. PSS-4 fungerar psykometriskt sämre än de längre versionerna.

RISE har analyserat data från PSS-14 och PSS-10 och båda uppvisar problem med multidimensionalitet. Varken PSS-14, PSS-10 eller PSS-4 bör användas som att de vore unidimensionella (summapoäng kan ej användas). Däremot går det att göra en enkel anpassning av PSS-14 genom att enbart använda de 7 frågor som fokuserar på svårigheter. Dessa utgör en unidimensionell skala som fungerar betydligt bättre psykometriskt. Svarskategorierna fungerar väl, items passar Rasch-modellen acceptabelt, residuallkorrelationer är acceptabla, targeting och reliabilitet är god och inga problem med DIF mellan kön. PSS-10 kan på samma vis kortas ner till de 6 negativa frågorna, men uppvisar sämre reliabilitet än PSS-14 nedkortad till 7 frågor.

RISE vill rikta ett tack till Alexander Rozental och David Forsström för tillgängliggörande av data för analys av PSS-14. Öppna data för PSS-10 har hämtats från <https://doi.org/10.6084/m9.figshare.13042010>

PSS är fritt att använda för icke-kommersiella syften, se <https://www.cmu.edu/dietrich/psychology/stress-immunity-disease-lab/scales/revised-pss-request-reply-for-all-requests.pdf>

Lee, E.-H. (2012). Review of the Psychometric Evidence of the Perceived Stress Scale. *Asian Nursing Research*, 6(4), 121–127. <https://doi.org/10.1016/j.anr.2012.08.004>

Nordin, M., & Nordin, S. (2013). Psychometric evaluation and normative data of the Swedish version of the 10-item perceived stress scale. *Scandinavian Journal of Psychology*, 54(6), 502–507. <https://doi.org/10.1111/sjop.12071>

## Children and parenting strategies

Kjellgren et al (2013) skriver: “The questionnaire used (Janson et al., 2011) has been modified from the Parent Child Conflict Tactics Scale and has previously been used in Swedish research on child physical abuse. The questionnaire consists of 24 items covering socio-demographic information, roles within the family, parental style, experiences of childhood physical abuse and attitudes to parenting.”

Dock finns inga psykometriska egenskaper redovisade i rapporten av Janson et al. (2011). I stället hänvisas till en rapport av Edfeldt (1985), som ej har kunnat hittas online.

Edfeldt, Å. W. (1985). "Allt vad I gören—": Slutrapport från SUSA-projektet om aga och våld i hemmet. Stockholm, Proprius.

Janson, S., Jernbro, C., & Långberg, B. (2011). *Kroppslig bestraffning och annan kränkning av barn i Sverige*. Stiftelsen Allmänna Barnhuset.

Kjellgren, C., Svedin, C. G., & Nilsson, D. (2013). Child Physical Abuse—Experiences of Combined Treatment for Children and their Parents: A Pilot Study. *Child Care in Practice*, 19(3), 275–290. <https://doi.org/10.1080/13575279.2013.785934>

## Parent Expectation Survey (PES)

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Består av 20 frågor som besvaras på en skala 0-10 med etiketter på ändpunkterna samt i mitten. Ursprungsartikeln (Reece, 1992) innehåller ingen psykometrisk analys.

Uppföljningsartikeln av Reece och Harkless (1998) innehåller inte heller någon psykometrisk analys. Pålsson (2020) har översatt PES till svenska, men hänvisar bara till de tidigare nämnda artiklarna gällande psykometri.

Hooge (2010) använder Parenting Knowledge Scale (PKS), som beskrivs som "a 7-item self-report, scale designed by Baby and You program managers through an adaptation of the Parent Expectations Survey (Reece, 1992)." Men här saknas också psykometrisk analys eller referens till någon sådan.

Hooge, S. L. (2010). *Effects of a Brief Parenting Education Program on Parenting Knowledge and Social Support in Mothers of Infants*. x, 117 leaves : ill.; 30 cm.

<https://doi.org/10.11575/PRISM/12590>

Pålsson, P. (2020). *Antenatal preparation for the early parenthood period* [Lund University, Faculty of Medicine].

[https://portal.research.lu.se/portal/files/76896497/Petra\\_P\\_lsson\\_web.pdf](https://portal.research.lu.se/portal/files/76896497/Petra_P_lsson_web.pdf)

Reece, S. M. (1992). The Parent Expectations Survey: A Measure of Perceived Self-Efficacy. *Clinical Nursing Research*, 1(4), 336–346. <https://doi.org/10.1177/105477389200100404>

Reece, S. M., & Harkless, G. (1998). Self-Efficacy, Stress, and Parental Adaptation: Applications to the Care of Childbearing Families. *Journal of Family Nursing*, 4(2), 198–215.

<https://doi.org/10.1177/107484079800400206>



## The Disruptive Behavior Rating Scale–Parent Version (DBRS-PV)

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Erford et al. (2017) analyserar 5 faktorer med 7 frågor vardera, 4 svars-kategorier. CFA med olämpliga estimatorn ML och antagandet att det inte finns någon korrelation mellan residualer. Man skriver också att man inte tittar på modification indices, trots att man får dålig model fit. Oklara resonemang. Man ignorerar helt detta enorma problem med dimensionalitet och går vidare med icke-psykometriska analyser.

Erford, B. T., Bardhoshi, G., Johnson, E. K., & Sherman, M. F. (2017). Disruptive Behavior Rating Scale-2nd Edition-Parent Version (DBRS-II-P): A Study of Reliability and Validity. *Counseling Outcome Research and Evaluation*, 8(2), 77–90.

<https://doi.org/10.1080/21501378.2017.1339185>

## Parental Monitoring

Stattin och Kerr (2000) beskriver nio frågor med fem svarsalternativ, som besvaras av både barn och förälder, angående föräldrarnas kännedom om barnets aktiviteter, kontakter och var hen befinner sig. Cronbach's alpha och test-retest reliabilitet är de enda egenskaper som rapporteras.

Stattin et al. (2011) hänvisar till ovanstående artikel och använder i stort samma frågor. Cronbach's alpha och korrelation mellan barnets och föräldrarnas skattning ( $r = 0.34-0.40$ ).

Stattin, H., & Kerr, M. (2000). Parental Monitoring: A Reinterpretation. *Child Development*, 71(4), 1072–1085. <https://doi.org/10.1111/1467-8624.00210>

Stattin, H., Persson, S., Burk, W., & Kerr, M. (2011). Adolescents' Perceptions of the Democratic Functioning in their Families. *European Psychologist*, 16. <https://doi.org/10.1027/1016-9040/a000039>

## Parent Practices Interview (PPI)

En stor svensk studie på flera föräldraskapsstödsprogram (Högström et al., 2017) beskriver användning av två delskalor från PPI: "harsh treatment" (oklart antal frågor), och "rewarding the child" (oklart antal frågor), med hänvisning till en artikel av Webster-Stratton (1998). Artikeln som refereras nämner inte "Parents' Practice Interview" någonstans. Däremot beskrivs en enkät som används som intervju-underlag, som anpassats från Oregon Social Learning Centers "Discipline Questionnaire", och innehåller bl.a. två delar som benämns som



"harsh style" (7 frågor) och "positive style" (1 fråga). Det är dock oklart om detta är samma, eftersom Högström et al. (2017) inte beskriver antalet frågor, men ger exempel på två frågor gällande "rewarding the child", och eftersom antalet svarskategorier skiljer sig. Webster-Stratton (1998) beskriver 3 svarskategorier för frågorna inom "harsh style" och 5 svarskategorier för "positive style", medan Högström et al. (2017) har 7 svarskategorier för båda.

Varken Högström et al. (2017) eller Webster-Stratton (1998) redovisar några psykometriska egenskaper utöver Cronbach's alpha.

En tidigare svensk artikel av Högström et al. (2015) benämner PPI som "Parent Practices Interview" och ger fler källhänvisningar (Baydar et al., 2003; Reid et al., 2007; Webster-Stratton et al., 2001):

Baydar et al. (2003) benämner PPI explicit och kallar den för "Parenting Practices Interview" (OBS tredje varianten hittills) och gör en EFA med 19 items som resulterar i tre delskalor. Delskalan "supportive parenting" har 4 frågor, och "harsh/negative parenting" har 5 frågor. Man redovisar inte faktorladdningar eller residualkorrelationer. Cronbach's alpha är låg för delskalorna, bara "harsh/negative parenting" är över 0.70. Test-retest korrelation redovisas också, men inga andra psykometriska egenskaper än vad som nämnts i detta stycke.

Webster-Stratton et al. (2001) nämner en Parent Daily Discipline Interview (DDI), men det står inget i artikeln om PPI (i någon av formerna som hittills hittats).

Reid et al. (2007) använder de två delskalorna från PPI som nämns ovan ("harsh" och "supportive" parenting), och redovisar enbart Cronbach's alpha (som är under 0.70 för båda).

En svensk studie av Kling et al. (2010) beskriver PPI som att den har 80 items, med hänvisning till samma källartiklar som nämns ovan.

En norsk studie (Drugli et al. 2010) hänvisar till ännu en källartikel angående PPI:

Webster-Stratton et al. (2004) beskriver de två delskalorna som tidigare nämnts, men denna gång med 7 items för "harsh parenting" och 5 items för "supportive parenting". Ingen psykometrisk analys. Det märkliga med att Drugli et al. (2010) hänvisar till denna artikel är att Drugli menar att det är 14 items för "harsh parenting" och 15 för "supportive parenting".

Baydar, N., Reid, M. J., & Webster-Stratton, C. (2003). The Role of Mental Health Factors and Program Engagement in the Effectiveness of a Preventive Parenting Program for Head Start Mothers. *Child Development*, 74(5), 1433–1453. <https://doi.org/10.1111/1467-8624.00616>

Drugli, M. B., Larsson, B., Fossum, S., & Mørch, W.-T. (2010). Five- to six-year outcome and its prediction for children with ODD/CD treated with parent training. *Journal of Child Psychology and Psychiatry*, 51(5), 559–566. <https://doi.org/10.1111/j.1469-7610.2009.02178.x>

Högström, J., Enebrink, P., Melin, B., & Ghaderi, A. (2015). Eighteen-Month Follow-Up of Internet-Based Parent Management Training for Children with Conduct Problems and the

Relation of Homework Compliance to Outcome. *Child Psychiatry & Human Development*, 46(4), 577–588. <https://doi.org/10.1007/s10578-014-0498-7>

Högström, J., Olofsson, V., Özdemir, M., Enebrink, P., & Stattin, H. (2017). Two-Year Findings from a National Effectiveness Trial: Effectiveness of Behavioral and Non-Behavioral Parenting Programs. *Journal of Abnormal Child Psychology*, 45(3), 527–542. <https://doi.org/10.1007/s10802-016-0178-0>

Kling, Å., Forster, M., Sundell, K., & Melin, L. (2010). A Randomized Controlled Effectiveness Trial of Parent Management Training With Varying Degrees of Therapist Support. *Behavior Therapy*, 41(4), 530–542. <https://doi.org/10.1016/j.beth.2010.02.004>

Reid, M. J., Webster-Stratton, C., & Hammond, M. (2007). Enhancing a Classroom Social Competence and Problem-Solving Curriculum by Offering Parent Training to Families of Moderate- to High-Risk Elementary School Children. *Journal of Clinical Child & Adolescent Psychology*, 36(4), 605–620. <https://doi.org/10.1080/15374410701662741>

Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, 66(5), 715–730. <https://doi.org/10.1037//0022-006x.66.5.715>

Webster-Stratton, C., Reid, J., & Hammond, M. (2001). Social Skills and Problem-solving Training for Children with Early-onset Conduct Problems: Who Benefits? *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(7), 943–952. <https://doi.org/10.1017/S0021963001007776>

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating Children With Early-Onset Conduct Problems: Intervention Outcomes for Parent, Child, and Teacher Training. *Journal of Clinical Child & Adolescent Psychology*, 33(1), 105–124. [https://doi.org/10.1207/S15374424JCCP3301\\_11](https://doi.org/10.1207/S15374424JCCP3301_11)

## Caregiver Strain Questionnaire (CGSQ)

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Brannan et al. (1997) redogör för 21 frågor med 5 svarskategorier, där de 10 första frågorna avser mäta "objective strain", och övriga frågor ska mäta "subjective strain". EFA visar på tre faktorer ("subjective strain" delas i "internalized" och "externalized") och problem med crossloadings och hög korrelation mellan faktorer. Inför CFA skapas nio indikatorer genom att slå ihop items (1-4 items per indikator) och ta bort ett item. Detta innebär såklart att problem med crossloadings (och ev. residuallkorrelationer) "försvinner", och model fit ser bättre ut för 3-faktormodellen man rekommenderar. Men det är en märklig mätmodell som bortser från stora mängder variation i data.

Khanna et al. (2012) gör CFA med ML och utgår från samma typ av item-sammanslagning som Brannan et al. (1997) gjorde. Om vi bortser från att ML är en olämplig estimator för den

sammanslagna data som analyseras så har man ändå stora problem med mätmodellen och behöver lägga in residualkorrelationer för att nå acceptabel model fit.

Utöver problemen med dimensionalitet som nämns ovan så saknas analyser av svarskategorier, targeting, mätosäkerheter, och invarians/DIF.

Brannan, A. M., Heflinger, C. A., & Bickman, L. (1997). The Caregiver Strain Questionnaire: Measuring the Impact on the Family of Living with a Child with Serious Emotional Disturbance. *Journal of Emotional and Behavioral Disorders*, 5(4), 212–222.

<https://doi.org/10.1177/106342669700500404>

Khanna, R., Madhavan, S. S., Smith, M. J., Tworek, C., Patrick, J. H., & Becker-Cottrill, B. (2012). Psychometric properties of the Caregiver Strain Questionnaire (CGSQ) among caregivers of children with autism. *Autism*, 16(2), 179–199. <https://doi.org/10.1177/1362361311406143>

## SOFIA-studien

Frogner et al. (2022) har tagit fram fem frågor för att mäta "social skills", och skriver "These scales have not previously been used in research, but they were chosen because they display adequate face validity and internal consistency reliability." Artikeln innehåller inte heller någon psykometrisk analys.

Klingzell et al. (2016) mäter uppförandeproblem med kriterier för diagnoser i DSM-5 (ODD och CD). Man hänvisar till "Child Problematic Traits Inventory (CPTI)", och källartikeln med titeln "A New Measure to Assess Psychopathic Personality in Children", av Colins et al. (2014).

Colins et al. (2014) skriver att "The aim with the CPTI was to develop a research instrument that enables longitudinal studies that test developmental theories and stability of psychopathic traits and psychopathic personality across different developmental phases." Det förefaller långt från ambitionen med att följa upp föräldraskapsstöd, men vi tittar kortfattat närmare på delskalor och psykometri.

Colins et al. (2014) gör EFA med robust WLS (bra när det är 4 skalsteg) för att undersöka den förväntade 3-faktor-strukturen. Sedan görs CFA med goda model fit indices. Man konstaterar problem med data ("inconsistent categorical values", oklart vad det innebär) när jämförelser ska göras mellan kön- och åldersgrupper, vilket beskrivs som en anledning till att man inte kan göra invarians-test. Om något ger det intrycket att invarians inte råder, men man gör separata CFA för barnets kön och åldersgrupper och får övergripande model fit indices som är nästan identiska (bra). Dock redovisas ej faktorladdningarna för dessa separata CFA. Man har inte testat invarians för kön för den som skattat barnet, utan bara för barnets kön. Det finns ingen analys av svarskategorier, frågornas svårighetsgrad jämfört med respondenternas förmåga, och inte heller mätosäkerheter.

Colins, O. F., Andershed, H., Frogner, L., Lopez-Romero, L., Veen, V., & Andershed, A.-K. (2014). A New Measure to Assess Psychopathic Personality in Children: The Child Problematic Traits Inventory. *Journal of Psychopathology and Behavioral Assessment*, 36(1), 4–21. <https://doi.org/10.1007/s10862-013-9385-y>

Frogner, L., Hellfeldt, K., Ångström, A.-K., Andershed, A.-K., Källström, Å., Fanti, K. A., & Andershed, H. (2022). Stability and Change in Early Social Skills Development in Relation to Early School Performance: A Longitudinal Study of a Swedish Cohort. *Early Education and Development*, 33(1), 17–37.

Klingzell, I., Fanti, K. A., Colins, O. F., Frogner, L., Andershed, A.-K., & Andershed, H. (2016). Early Childhood Trajectories of Conduct Problems and Callous-Unemotional Traits: The Role of Fearlessness and Psychopathic Personality Dimensions. *Child Psychiatry & Human Development*, 47(2), 236–247. <https://doi.org/10.1007/s10578-015-0560-0>

## **Social Competence Scale–Parent (P-Comp)**

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Omnämns i Kling et al. (2010) med hänvisning till Gouley et al. (2008) som beskriver 12 items med 5 svars kategorier. De gör CFA med ML-estimator och finner unidimensionalitet, men beskriver problem med invarians mellan normalpopulation och högriskpopulation (dock bara N = 90 i högriskgruppen). Påtaglig variation i fit indices mellan mättillfällen indikerar också problem med invarians över tid, men även här är det små sampel, så det bör tolkas med försiktighet. I övrigt saknas redovisning av psykometriska egenskaper.

Gouley, K. K., Brotman, L. M., Huang, K.-Y., & Shrout, P. E. (2008). Construct Validation of the Social Competence Scale in Preschool-age Children. *Social Development*, 17(2), 380–398. <https://doi.org/10.1111/j.1467-9507.2007.00430.x>

Kling, Å., Forster, M., Sundell, K., & Melin, L. (2010). A Randomized Controlled Effectiveness Trial of Parent Management Training With Varying Degrees of Therapist Support. *Behavior Therapy*, 41(4), 530–542. <https://doi.org/10.1016/j.beth.2010.02.004>

## **Swanson, Nolan and Pelham Questionnaire (SNAP-IV)**

RISE har inte hittat någon publicerad svensk psykometrisk studie.

26 frågor, varav 9+9 fokuserar på ADHD och 8 frågor om trotssyndrom, samtliga med 4 svars kategorier.

Bussing et al. (2008) skriver i metodsektionen att "CFA was performed using the method of random parceling, a method that reduces a large number of indicators to yield a better model fit", vilket gör resten av artikeln i princip värdelös. Att slå ihop items utan att först analysera items och data utan bearbetning, och helt utan resonemang om varför man gör åtgärden (förutom att artificiellt få bättre model fit) är inte ett rimligt förfarande.

Hall et al. (2020) fokuserar på ADHD-frågorna (18 items) och använder ESEM (explorativ strukturell ekvationsmodellering) med WLSMV-estimator för både föräldra- och lärarskattade versionen. Föräldradata visar acceptabel model fit för de två ADHD-delskalorna (med undantag för items 10 och 12 med faktorladdningar på 0.34 och 0.36). Lärarskattningen visar liknande resultat. Invarians testas över tid separat för lärar- och föräldraskattade versionen och verkar fungera acceptabelt. Däremot föreligger inte invarians mellan lärar- och föräldraskattningen, så dessa bör ej jämföras.

Bussing, R., Fernandez, M., Harwood, M., Hou, W., Garvan, C. W., Eyberg, S. M., & Swanson, J. M. (2008). Parent and Teacher SNAP-IV Ratings of Attention Deficit Hyperactivity Disorder Symptoms: Psychometric Properties and Normative Ratings From a School District Sample. *Assessment*, 15(3), 317–328. <https://doi.org/10.1177/1073191107313888>

Hall, C. L., Guo, B., Valentine, A. Z., Groom, M. J., Daley, D., Sayal, K., & Hollis, C. (2020). The Validity of the SNAP-IV in Children Displaying ADHD Symptoms. *Assessment*, 27(6), 1258–1271. <https://doi.org/10.1177/1073191119842255>

## Center for Epidemiologic Studies-Depression scale (CES-D)

RISE har inte hittat någon publicerad svensk psykometrisk studie.

Olsson & von Knorring (1997) använder CTT för att analysera psykometriska egenskaper hos den svenska ungdomsversionen, CES-DC. Faktoranalys visar på 4 dimensioner i data både för pojkar och flickor. Det framkommer ingen information om faktorladdningar, estimator, residualkorrelation, korrelation mellan faktorer, eller invarians.

Covic et al. (2007) gör Rasch-analys och hittar problem med svarskategorierna (minskar från 4 till 3) och reducerar skalan från 20 till 13 items för att den ska fungera acceptabelt.

Stansbury et al. (2006) gör också Rasch-analys och finner att en 16-itemversion har högre reliabilitet än originalets 20 item. Man tar bort de 4 positivt formulerade frågorna. Även här konstateras problem med svarskategorierna.

Covic, T., Pallant, J. F., Conaghan, P. G., & Tennant, A. (2007). A longitudinal evaluation of the Center for Epidemiologic Studies-Depression scale (CES-D) in a Rheumatoid Arthritis Population using Rasch Analysis. *Health and Quality of Life Outcomes*, 5(1), 41. <https://doi.org/10.1186/1477-7525-5-41>

Olsson, G., & von Knorring, A.-L. (1997). Depression among Swedish adolescents measured by the self rating scale Center for Epidemiology Studies—Depression Child (CES-DC). *European Child & Adolescent Psychiatry*, 6(2), 81–87. <https://doi.org/10.1007/BF00566670>

Stansbury, J. P., Ried, L. D., & Velozo, C. A. (2006). Unidimensionality and Bandwidth in the Center for Epidemiologic Studies Depression (CES-D) Scale. *Journal of Personality Assessment*, 86(1), 10–22. [https://doi.org/10.1207/s15327752jpa8601\\_03](https://doi.org/10.1207/s15327752jpa8601_03)

## Depression Anxiety Stress Scale (DASS)

Ursprungsversionen av DASS (Lovibond & Lovibond, 1995) har tre delskalor med 14 items vardera, som besvaras på en 4-gradig skala. CFA görs med ML (ej optimalt med 4 svars-kategorier) och enbart  $\chi^2$  samt AGFI rapporteras angående model fit, där AGFI når 0.76 för 3-faktormodellen, vilket indikerar brister. Crossloadings nämns, men oklart hur stora de är och vad/om man gör något åt dem. Residualkorrelationer också oklart. En bifaktormodell testas också.

RISE har inte hittat någon publicerad svensk psykometrisk studie som analyserar DASS-42, däremot för DASS-21:

Alfonsson et al. (2017) testar en serie CFA-modeller men beskriver ej vilken estimator som används. Både 3-faktormodellen och en bifaktormodell med tre domänfaktorer visar acceptabel fit. Man identifierar dock signifikanta residualkorrelationer mellan tre par av items tillhörande delskalorna stress och depression, men klargör ej storleken på dessa utan lägger bara till dem i modellen för att få bättre model fit. Det innebär att det finns signifikanta bias-effekter i indexvärden för stress och depression, men att vi inte vet hur mycket detta påverkar. Crossloadings förekommer också för tre items (2, 5, och 8), men i slutänden väljer man att inte göra något åt det. Varken golv- eller takeffekter redovisas.

Shea et al. (2009) gör en Rasch-analys av kortversionen av DASS (21 items) med ett sampel från normalpopulation. Analysen visar behov av att ta bort ett item från varje delskala p.g.a. misfit (item 2, 5 och 11). Targeting är problematisk för samtliga delskalor, med mycket tydliga golveffekter, främst för ångest och depression (35 respektive 40% har 0 poäng). Svarkategorierna fungerar ok överlag. Inga problem med invarians/DIF för kön, ålder eller utbildningsnivå. Man visar att det inte är lämpligt att kombinera samtliga items/delskalor till ett övergripande indexvärde.

Alfonsson, S., Wallin, E., & Maathz, P. (2017). Factor structure and validity of the Depression, Anxiety and Stress Scale-21 in Swedish translation. *Journal of Psychiatric and Mental Health Nursing*, 24(2–3), 154–162. <https://doi.org/10.1111/jpm.12363>

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)

Shea, T. L., Tennant, A., & Pallant, J. F. (2009). Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC Psychiatry*, 9(1), 21. <https://doi.org/10.1186/1471-244X-9-21>

## Andra svenska enkäter

Persson et al. (2004) listar flera enkäter som skulle kunna vara relevanta, och menar att de har finns goda psykometriska egenskaper, men rapporterar bara Cronbach's alpha och test-retest i artikeln. Hänvisningar ges också till tre äldre artiklar:

Stattin & Kerr (2000) rapporterar bara Cronbach's alpha och test-retest.

Kerr et al. (1999) rapporterar bara Cronbach's alpha och test-retest.

Kerr & Stattin (2003) rapporterar bara Cronbach's alpha.

Eftersom ingen artikel som innehåller en grundläggande analys av dimensionalitet har hittats kommer inte ovan nämnda artiklar eller de nämnda enkäterna att beskrivas eller refereras. Det är möjligt att de fungerar väl, men i brist på adekvata analyser kan vi inte göra det antagandet.

ESTER är ett svenskt verktyg för risk- och behovsbedömning för barn och unga (Andershed et al., 2010; Andershed & Andershed, 2010) och innehåller 19 risk-/skyddsfaktorer som grupperas i fyra kategorier. Ingen artikel med psykometrisk analys har hittats.

Andershed, H., & Andershed, A.-K. (2010). Risk-need assessment for youth with or at risk for conduct problems: Introducing the assessment system ESTER. *Procedia Social and Behavioral Sciences*, 7.

Andershed, H., Fredriksson, J., Engelholm, K., Ahlberg, R., Berggren, S., & Andershed, A.-K. (2010). Initial test of a new risk-need assessment instrument for youths with or at risk for conduct problems: ESTER-assessment. *Procedia Social and Behavioral Sciences*, 5.

Andershed, A.-K., & Andershed, H. (2016). Improving evidence-based social work practice with youths exhibiting conduct problems through structured assessment. *European Journal of Social Work*, 19(6), 887–900. <https://doi.org/10.1080/13691457.2015.1043242>

Bond, H. (2013). Further Testing of the Inter-rater Reliability of ESTER-assessment – a Risk-need Assessment Instrument for Youths with or at Risk for Conduct Problems. *American Journal of Applied Psychology*, 2(2), 16. <https://doi.org/10.11648/j.ajap.20130202.11>



Kerr, M., Stattin, H., & Trost, K. (1999). To know you is to trust you: Parents' trust is rooted in child disclosure of information. *Journal of Adolescence*, 22(6), 737–752.

<https://doi.org/10.1006/jado.1999.0266>

Kerr, M., & Stattin, H. (2003). Parenting of adolescents: Action or reaction. In C. Crouter & A. Booth (Eds.), *Children's influence on family dynamics: The neglected side of family relationships* (pp. 121–151). Lawrence Erlbaum Associates, Inc.

Nordqvist, J., & Öhrman, S. (2010). *ESTER-screening och dess konvergenta validitet gentemot tidigare screeninginstrument till unga med eller i riskzonen för normbrytande beteende* [Örebro Universitet]. <http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-14469>

Persson, S., Stattin, H., & Kerr, M. (2004). Adolescents' conceptions of family democracy: Does their own behaviour play a role? *European Journal of Developmental Psychology*, 1(4), 317–330.

<https://doi.org/10.1080/17405620444000238>

Stattin, H., & Kerr, M. (2000). Parental Monitoring: A Reinterpretation. *Child Development*, 71(4), 1072–1085. <https://doi.org/10.1111/1467-8624.00210>

Stattin, H., Persson, S., Burk, W., & Kerr, M. (2011). Adolescents' Perceptions of the Democratic Functioning in their Families. *European Psychologist*, 16. <https://doi.org/10.1027/1016-9040/a000039>



## Psykometriska analyser

RISE har tagit del av anonymiserade data från enkäter som svenska regioner bidragit med. Den ena kommer från arbetet med utökade hembesök (även känt som Rinkebymodellen eller Växa Tryggt), där skattningar görs av den som utför insatsen. Den andra är bestående av 10 frågor angående barns beteende och välbefinnande som är inspirerade av etablerade skattningsskalor och modifierade för föräldraskattning (i stället för självskattning) av barnets mående. Den tredje består av två frågor till blivande föräldrar. Resultatet av analyserna sammanfattas nedan. Slutligen finns en analys av enkätdata från en tidigare rapport från MFoF, som genomfördes av SOM-institutet, gällande tilltro till föräldraförmåga.

### Barns välbefinnande

Den föräldraskattade enkäten har 10 frågor. Alla frågor delar samma fem svarskategorier, och dessa fungerar bra. Dock finns problem med residuallkorrelationer, vilket gör att tre frågor behöver tas bort. Ingen problematisk DIF föreligger för barnets kön eller ålder, och inte heller för föräldrarnas kön eller interaktionen mellan barnets och föräldrarnas kön. Targeting och reliabilitet är acceptabel, drygt 80% av respondenterna befinner sig i området med minst 0.7 i reliabilitet. Dataunderlaget i analysen omfattar barn i åldersspannet 6-17 år. Frågor och svarskategorier finns i slutet på denna bilaga. Denna enkät är att betrakta som en kort och bred enkät som främst lämpar sig för användning på gruppnivå.

### Utökade hembesök

Enkäten kopplad till Utökade hembesök fokuserar på upplevelsen som den som levererar insatsen har av sin förmåga och relation till föräldrarna som tar del av insatsen. De 12 frågorna har antingen 7 eller 10 svarskategorier. Analysen visar att det är för många svarskategorier för att respondenterna med reliabilitet ska kunna skilja på dem. När svarskategori 1-3 slås samman för samtliga frågor, och frågorna därmed i stället får 5 eller 8 svarskategorier, fungerar de bättre. Fyra par av frågor har mycket stor samvariation, vilket är problematiskt om man vill kunna återge enkäten med ett indexvärde. Det innebär också potentiellt att en fråga i varje par kan tas bort utan att tappa särskilt mycket information, beroende på hur resultatet av enkäten ska användas.

### Frågor till blivande föräldrar

Två frågor ställs till blivande föräldrar, båda med en svarsskala från 1-7 med etiketter enbart på ändpunkterna ("Mycket liten" till "Mycket stor"):

- Vad tror du om din förmåga att medverka till att din förlossning blir en positiv upplevelse?
- Vad tror du om din förmåga att bli en bra förälder?

Eftersom detta är två frågor som inte antas mäta samma underliggande konstrukt har ingen psykometrisk analys gjorts (dessutom är två frågor för litet underlag). Att använda enstaka frågor för att genomföra en övergripande mätning kan vara en effektiv strategi som del i ett frågebatteri. Det kräver dock att frågorna först har kalibrerats, d.v.s. att deras mättekniska egenskaper har fastställts genom analyser med fler frågor som avser mäta samma förmåga/upplevelse. En sådan kalibrering medför bl.a. att ta fram parametrar som hur svarskategorierna fungerar, vilka mätosäkerheter som finns runt svarskategoriernas trösklar, samt hur väl frågorna passar målgruppen. Det är extra viktigt när avsikten är att ta fram en enstaka fråga för att mäta en övergripande förmåga/upplevelse att fastställa att frågan verkligen har en bra passning till målgruppens förmåga ("targeting").

Det är tydligt att svarsdata är mycket skeva och ej normalfördelade. Detta indikerar att frågorna och svarsalternativen inte matchar målgruppens förmåga optimalt. Det finns en stark takeffekt för båda frågor, där frågan om förlossning har 35% respondenter i högsta svarskategorin medan frågan om föräldraförmåga har hela 61.2% som svarar i högsta svarskategorin. De lägsta 2-4 svarskategorierna används knappt alls.

## **SOM-institutet & MFoF – tilltro till föräldraförmåga**

RISE har fått ta del av anonymiserad data från SOM-institutet, från en enkätundersökning vars resultat redovisas i rapporten "Att vara förälder idag -- föräldrars upplevelser av föräldrarollen och utmaningar i föräldraskapet" (2020), som Myndigheten för Familjerätt och Föräldraskapsstöd (MFoF) gjort i samarbete med SOM-institutet. Enkäten består av fyra frågor och samma sju svarskategorier används för samtliga frågor. Svarskategorierna har endast etiketter på ändarna, "Instämmer inte alls" till "Instämmer helt och hållet". Analysen visar att de tre lägsta svarskategorierna bör slås samman till en för samtliga frågor. Fråga 1 och 3 har för kraftigt korrelerande residualer, och båda har dessutom undermålig item fit. Fråga 2 fungerar bra, och fråga 4 fungerar acceptabelt. Invarians mellan föräldrarnas kön råder efter att svarskategorierna justerats. Reliabiliteten är dock mycket låg, och når inte upp till gränsvärdet 0.7 i någon del av skalan.

I SOM-institutets rapport skriver man att "Föräldrarna i studien skattar genomgående sin egen föräldraförmåga högt". Det är ett kanske väl starkt påstående att göra utan att ha gjort en analys av targeting och reliabilitet. Föräldrar skattar högt på frågorna, men vi saknar kalibrering av frågorna för att bedöma om det motsvarar en hög föräldraförmåga. Utifrån befintliga frågor ser vi att det finns en takeffekt och för låg reliabilitet för att urskilja mer än i bästa fall två grupperingar i data, även när vi låter alla fyra frågor vara med i analysen.

Det är tveksamt om de få och breda frågorna kan anses representera och beskriva vidden av "tilltro till föräldraförmåga". En kan också undra över om påståendefrågor avsedda att mäta attityden till sin egen förmåga (self-efficacy) är ett optimalt sätt att fånga in "föräldraförmåga", särskilt när frågorna är så brett formulerade. Utan mera beskrivande och detaljerade frågor blir det ett mycket stort tolkningsutrymme för respondenterna. Således är det svårt att veta vad de upplever sig svara på och hur svaren ska tolkas. Kopplingen till den faktiska



föräldraförmågan och hur den används i vardagen är också oklar. Hög tilltro till den egna förmågan medför inte nödvändigtvis att förmågan eller tillämpningen i praktiken är god.

Myndigheten för Familjerätt och Föräldraskapsstöd, & SOM-institutet. (2020). *Att vara förälder idag - föräldrars upplevelser av föräldrarollen och utmaningar i föräldraskapet*. Myndigheten för familjerätt och föräldraskapsstöd. <https://mfof.se/foraldraskapsstod/publikationer-om-foraldraskapsstod/mfofs-publikationer/att-vara-foralder-idag---grundrapport.html>

## Detaljerad information om psykometriska analyser

Mera utförliga redogörelser för de psykometriska analyser RISE har genomfört i samband med denna rapport kommer göras tillgängliga under 2022 på följande webbplats:

<https://doi.org/10.17605/OSF.IO/Y3NJ4>

## Omvandlingstabeller – ordinaldata till intervalldata

När svarsdata på nedanstående enkäter summeras behöver det lägsta svarsalternativet få värdet 0, så att den lägsta summerade råpoängen också blir 0. Därefter kan man gå in i tabellen nedan och ersätta den summerade poängen med respektive intervallpoäng, som är skalad till att gå från 0 till 100.

Värdet i kolumnen för mätosäkerhet indikerar hur stor osäkerhet det är runt intervallpoängen. Exempelvis innebär råpoäng 10 en intervallpoäng på 44,6 och mätosäkerhet på 9,1 - alltså ligger mätvärdet mellan 35,5 och 53,7.

Det medför att för att vi ska kunna säga att två mätvärden med säkerhet är olika, exempelvis när vi följer en individ över tid, behöver förändringen överstiga mätosäkerheten. En individ som hade 10 råpoäng (44,6 intervallpoäng) vid första mätningen behöver således uppnå 15 råpoäng (54,1 intervallpoäng) vid eftermätningen.

### BBQ

Summerad råpoäng	Intervallpoäng	Mätosäkerhet (SE*2)
0	0,0	40,8
1	13,7	22,7
2	22,0	16,4
3	27,0	13,7
4	30,7	12,1
5	33,7	11,1
6	36,3	10,4
7	38,6	9,9
8	40,7	9,5
9	42,7	9,3
10	44,6	9,1
11	46,4	9,1
12	48,3	9,1
13	50,2	9,2
14	52,1	9,3
15	54,1	9,6
16	56,2	9,9
17	58,5	10,3
18	61,0	10,9
19	63,9	11,7
20	67,2	12,8
21	71,4	14,4
22	76,9	17,1
23	85,7	23,4
24	100,0	41,2

**SWEMWBS**

Summerad råpoäng	Intervallpoäng	Mätosäkerhet (SE*2)
0	0,0	30,2
1	10,5	17,3
2	17,3	13,0
3	21,8	11,3
4	25,3	10,3
5	28,4	9,7
6	31,1	9,2
7	33,6	8,9
8	36,0	8,6
9	38,2	8,4
10	40,3	8,2
11	42,3	8,1
12	44,3	8,0
13	46,3	8,0
14	48,3	8,0
15	50,3	8,1
16	52,3	8,2
17	54,4	8,4
18	56,7	8,6
19	59,0	8,9
20	61,6	9,2
21	64,3	9,5
22	67,2	9,9
23	70,3	10,3
24	73,7	10,8
25	77,6	11,7
26	82,3	13,3
27	89,3	17,5
28	100,0	30,3

**PSS-7**

Summerad råpoäng	Intervallpoäng	Mätosäkerhet (SE*2)
0	0,0	33,2
1	12,0	19,5
2	20,0	14,8
3	25,3	12,7
4	29,4	11,5
5	32,9	10,6
6	35,9	10,0
7	38,6	9,6
8	41,1	9,2
9	43,5	8,9
10	45,6	8,7
11	47,7	8,5
12	49,7	8,3
13	51,7	8,2
14	53,5	8,1
15	55,4	8,0
16	57,2	8,0
17	59,0	8,0
18	60,9	8,1
19	62,7	8,2
20	64,7	8,4
21	66,7	8,6
22	68,9	9,0
23	71,4	9,5
24	74,2	10,3
25	77,5	11,5
26	81,9	13,5
27	88,8	18,4
28	100,0	32,5

## Barns välbefinnande

Summerad råpoäng	Intervallpoäng	Mätosäkerhet (SE*2)
0	0.0	32.6
1	11.3	18.6
2	18.5	13.9
3	23.2	11.8
4	26.7	10.6
5	29.7	9.8
6	32.3	9.3
7	34.7	8.9
8	36.9	8.6
9	38.9	8.4
10	40.9	8.3
11	42.9	8.2
12	44.8	8.2
13	46.7	8.2
14	48.7	8.2
15	50.6	8.3
16	52.6	8.4
17	54.6	8.5
18	56.8	8.7
19	59.0	8.9
20	61.3	9.2
21	63.8	9.5
22	66.4	9.9
23	69.4	10.4
24	72.6	11.1
25	76.5	12.2
26	81.3	14.1
27	88.6	18.7
28	100.0	32.6

Frågor om barns välbefinnande återges i tabellen nedan. Svarskategorierna ersätts med siffror 0-4 där "Nästan hela tiden" motsvarar 4, och "Aldrig" motsvarar 0. För frågor med tecknet \* ska svarens siffervärden vara de omvända (Aldrig = 4, Nästan hela tiden = 0). Totalpoängen kan bli mellan 0 och 28, och används sedan med omvandlingstabellen ovan.

		Nästan hela tiden	Ofta	Då och då	Sällan	Aldrig
1.	Barnet verkar spänt, ängsligt eller nervöst*					
2.	Barnet har besvärats av värk, smärta eller andra kroppsliga problem*					
3.	Barnet är nöjd med de saker som hen gjort					
4.	Barnet har haft svårigheter att somna in eller att sova en hel natt*					
5.	Barnet har visat värme eller ömhet för någon (t.ex. gett någon en kram eller sagt något kärleksfullt till någon)					
6.	Barnet har klarat av det mesta som hen behövt göra					
7.	Barnet har varit lättretlig (lätt irriterad) tillsammans med andra*					