

Måleegenskaper ved den norske versjonen av Brief Problem Monitor (BPM)

Agathe Backer-Grøndahl¹ (agathe@r-bup.no)

Monica Martinussen² (monica.martinussen@uit.no)

PsykTestBarn 2018, 1:2

ISSN 1893-9910

Mottatt dato 9. mars 2018

Publisert dato 25. juni 2018

Artikkelens URL <https://psyktestbarn.r-bup.no/no/artikler/bpm-brief-problem-monitor>

DOI 10.21337/0057

Som alle artikler i PsykTestBarn, kan denne fagfellevurderte artikkelen arkiveres og distribueres fritt for alle slags formål på følgende vilkår: korrekt referanse skal oppgis (se under), ingen kommersiell bruk og ingen bearbeidelse av tekst eller innhold.

Denne artikkelen skal siteres på følgende måte:

Backer-Grøndahl, A. & Martinussen, M. (2018) Måleegenskaper ved den norske versjonen av Brief Problem Monitor (BPM). PsykTestBarn, 1:2. doi:10.21337/0057

¹ Regionsenter for barn og unges psykiske helse, Helseregion Øst og Sør

² RKBU Nord, UiT Norges arktiske universitet

Sammendrag

Beskrivelse. Brief Problem Monitor (BPM) er en kortversjon av kartleggingssystemet Achenbach System of Empirical Based Assessment (ASEBA). Instrumentet er utviklet for at man enkelt skal kunne følge med på barnets/ungdommens utvikling på eksternaliserende problemer, internaliserende problemer og oppmerksomhetsproblemer, inkludert eventuell endring etter en intervensjon. Det finnes tre versjoner; én for foreldre (BPM-P), én for lærere (BPM-T), og én for ungdom (BPM-Y). Det er 19 spørreskjema-ledd (18 i BPM-T), og disse er hentet fra de tilsvarende lengre ASEBA-skjemaene (CBCL, TRF og YSR). BPM består av tre delskalaer (én skala for hvert problemområde), samt en samleskår. BPM er utviklet av Achenbach, McConaughy, Ivanova, and Rescorla (2011). I den norske versjonen har man hentet ut ledd fra CBCL 6-18, TRF og YSR, oversatt av Torunn S. Nøvik og Sonja Heyerdahl og publisert i 1986/88, 1993 og 2002 (Kornør & Jozefiak, 2012).

Litteratursøk. Litteratursøket resulterte i kun én artikkel der det rapporteres resultater på BPM i et norsk utvalg. Artikkelen inkluderer mål på indre konsistens (reliabilitet) og resultater som belyser begrepsvaliditeten basert på to store norske normalutvalg i barne- og ungdomsskolealder.

Psykometri. Resultatene tyder på god reliabilitet for samleskåren i de tre versjonene av BPM, mens det hefter noe mer usikkerhet rundt enkelte av delskalaene, spesielt i BPM-Y. Korrelasjonene mellom delskalaene i BPM og de tilsvarende skalaene i de lengre instrumentene er svært høye og støtter begrepsvaliditeten til BPM.

Konklusjon. BPM kan muligens være et nyttig instrument for å måle atferds- og emosjonsproblemer hos barn og unge, men det trengs flere studier på skandinaviske utvalg for å kunne konkludere om de psykometriske egenskapene. Ettersom hensikten med instrumentet er å følge barns utvikling på problemene over tid eller som respons på en intervensjon, er det ønskelig med studier der man benytter BPM over tid og der deltakerne mottar en intervensjon.

Abstract

Description. Brief Problem Monitor (BPM) is a screening instrument for monitoring children's functioning in different problem areas (externalizing problems, internalizing problems, and attention problems), as well as for assessing responses to therapeutic interventions. It is a short version of the Achenbach System of Empirical Based Assessment (ASEBA) forms for children and youths aged 6–18 years. Separate BPM forms were developed for different informants: parents (BPM-P), teachers (BPM-T), and a self-report form for youths (BPM-Y). The BPM items (19 items; 18 in the BPM-T) were selected from the corresponding longer ASEBA scales (i.e., CBCL, TRF, and YSR). The BPM was developed by Achenbach, et al. (2011). The Norwegian BPM-items were derived from the translations of the CBCL, TRF, and YSR by Torunn S. Nøvik and Sonja Heyerdahl, published in 1986/88, 1993, and 2002, respectively (Kornør & Jozefiak, 2012).

Literature search. The systematic literature search yielded only one study that reported results on the BPM in a Norwegian sample. The study reports analyses of internal consistency (reliability) and validity of the BPM in two large population-based samples of school-age children (4th-10th grades and 1st-7th grades).

Psychometrics. The results suggest good reliability for the total problem scale across reporters, while the reliability coefficients are substantially lower for some of the subscales; in particular,

for the BPM-Y. Correlation coefficients between the subscales and corresponding scales in the longer versions indicate good construct validity.

Conclusion. The BPM may provide useful information regarding behavioral and emotional problems in children and youths. However, more studies on Norwegian or Scandinavian samples are needed to examine the adequacy of the psychometric properties of the Norwegian versions of BPM. Also, as the BPM was developed to monitor children's functioning over a period of time, including their responses to interventions, more studies are needed that focus on - or include - these aspects.

Innledning

Brief Problem Monitor (BPM) er et kartleggingsinstrument som er ment å måle barn og ungdoms (6–18 år) eksternaliserende problemer, internaliserende problemer og oppmerksomhetsproblemer. Instrumentet er utviklet for enkelt å kunne følge med på barnets/ungdommens utvikling på disse problemområdene, inkludert eventuell endring etter intervensjon eller behandling (Achenbach, et al., 2011). BPM er en del av Achenbach System of Empirical Based Assessment (ASEBA) og er primært utviklet for å ha et instrument som krever mindre tid og ressurser enn de allerede etablerte ASEBA-instrumentene. Det finnes tre versjoner; én for utfylling av foreldre til barn/ungdom mellom 6 og 18 år (BPM-P), én for utfylling av lærer til barn mellom 6 og 18 år (BPM-T), og én for utfylling av ungdom mellom 11 og 18 år (BPM-Y). Spørreskjemaene er hentet fra de tilsvarende og lengre versjonene Child Behavior Checklist for Ages 6–18 (CBCL), Teacher Report Form (TRF) og Youth Self-Report (YSR). BPM kan dermed betraktes som en kortversjon av dette kartleggingssystemet.

ASEBA og BPM er utviklet av Thomas Achenbach og hans kollegaer (Achenbach & Edelbrock, 1978; Achenbach, et al., 2011; Achenbach & Rescorla, 2001). CBCL, YSR og TRF er oversatt til norsk av Torunn S. Nøvik og Sonja Heyerdahl, og er publisert i 1986/88, 1993 og 2002 (Kornør & Jozefiak, 2012), og leddene i den norske versjonen av BPM er hentet fra disse oversettelsene. Erasmus University Rotterdam i Nederland eier rettighetene til den norske versjonen av BPM (og de andre ASEBA-instrumentene).

Bakgrunnen for utviklingen av BPM var at Chorpita og kollegaer ville lage en kortere versjon av CBCL og YSR. De ønsket å enklere kunne måle og følge med på barn og ungdoms potensielle endring som følge av et tiltak eller behandling (Chorpita, et al., 2010). Ved å bruke item response theory (IRT) og faktoranalyse på spørreskjemaene fra CBCL og YSR kom de fram til et 12-ledds spørreskjema som dekket eksternaliserende og internaliserende problemer (Brief Problem Checklist; BPC). Achenbach, et al. (2011) ønsket å inkludere et mål på oppmerksomhetsproblemer i tillegg, ettersom slike problemer gjerne blir adressert i intervensjoner. De identifiserte fem spørreskjemaene fra CBCL/YSR som diskriminerte mellom en gruppe barn/ungdom med ADHD og en gruppe uten ADHD. I tillegg inkluderte de et ekstra ledd som har vist seg å lade høyt på en ADHD-faktor (Achenbach, et al., 2011).

BPM består av 19 ledd (18 i BPM-T) formet som utsagn om barnets atferd de siste seks månedene, og disse skal skåres 0 = *stemmer ikke*, 1 = *stemmer delvis* eller 2 = *stemmer veldig bra eller ofte*. Leddene dekker delskalaene eksternaliserende problemer (7 ledd i BPM-P og BPM-Y, 6 ledd i BPM-T), internaliserende problemer (6 ledd) og oppmerksomhetsproblemer (6 ledd). Det kan også kalkuleres en samleskår, basert på alle leddene. Testen tar kun et par minutter å fylle ut, og kan brukes i klinisk praksis og til forskningsformål. Det er verdt å merke seg at BPM først og fremst er egnet til å følge med på utviklingen av problemer og respons på intervensjoner/behandling. Utviklerne oppfordrer til å bruke CBCL, TRF og/eller YSR for mer omfattende og spesifikk kartlegging av problemområder for å bestemme egnet intervensjon eller behandling. Kravene for å administrere og skåre BPM er de samme som for

CBCL; opplæring, samt at man har pedagog-, psykolog- eller legeutdannelse, eller minst treårig høyskoleutdanning innen helse-/sosialfag (Kornør & Jozefiak, 2012).

I tillegg til de psykometriske analysene som ble foretatt i forbindelse med utviklingen av BPC og BPM, har vi funnet tre internasjonale artikler som har undersøkt psykometriske egenskaper ved BPM-P: én studie av den engelske versjonen i et amerikansk/kanadisk utvalg (Piper, Gray, Raber, & Birkett, 2014), én studie av den spanske versjonen i et spansk utvalg (Penelo, de la Osa, Navarro, Domènech, & Ezpeleta, 2017) og én av den tyske versjonen i et tysk utvalg (Rodenacker, Plücker, & Döpfner, 2015)³. Resultatene fra studiene viser blant annet høy indre konsistens (for eksempel Cronbach's alpha $\geq 0,78$) og god kriterievaliditet målt som samsvar med tilsvarende skalaer fra CBCL (for eksempel $r \geq 0,74$) (Penelo, et al., 2017; Piper, et al., 2014) og med foreldres rapport på Strength and Difficulties Questionnaire (SDQ) ($r \geq 0,52$) (Penelo, et al., 2017), en tilfredsstillende tilpasning til den antatte tre-faktorstrukturen og målingsinvarians på tvers av kjønn og alder (Penelo, et al., 2017). Videre tyder resultatene på at BPM-P skårene hadde god prediktiv verdi når det gjelder å skille mellom dem som brukte og ikke brukte psykisk helsetjeneste i det spanske utvalget (OR $\geq 1,12$) (Penelo, et al., 2017).

Metode

Bibliotekarer Mari Elvsåshagen og Sølvi Biedilæ ved Regionsenter for barn og unges psykiske helse, helseregion Øst og Sør, søkte etter dokumentasjon på testens psykometriske egenskaper i databasene, PsycINFO, Medline, Embase, Cochrane Library, Oria (BIBSYS), Norart, SveMed+, PubMed, CRISTin.no, NORA.no, Forskningsdatabasen.dk og Swepub. Søkedato: 28.09-02.10.2017. Søkestrategien er tilgjengelig på <http://www.psyktestbarn.no/CMS/ptb.nsf/pages/bpm-brief-problem-monitor>

Vi inkluderte alle publikasjoner av studier som har undersøkt og rapportert minst ett av følgende i skandinaviske utvalg:

- normdata for testen
- reliabilitet: indre konsistens, test-retest, interrater og endringssensitivitet
- validitet: samsvar med liknende testskårer, samsvar med referansestandard eller annet kriterium, og/eller faktorstruktur

Forfatterne gikk gjennom alle identifiserte publikasjoners sammendrag, og foretok vurderingene uavhengig av hverandre.

Begge forfatterne vurderte normering, validitet og reliabilitet uavhengig av hverandre, ved hjelp av en tilpasset versjon av Test Review Form and Notes for Reviewers (European Federation of Psychologists' Association (EFPA), 2013).

Resultater

Litteratursøk

Søket resulterte i 14 publikasjoner (3 norske og 11 svenske/danske). Det ble kun rapportert resultater fra bruk av BPM i én av disse, og vi baserer oss følgelig kun på denne ene publikasjonen (Richter, 2015) presentert i Tabell 1. De ekskluderte artiklene handlet i hovedsak om andre tema og inkluderte ikke BPM (12 treff/artikler). I kun én artikkel utover Richter (2015) er BPM brukt (Patras et al., 2015), men ettersom det er en protokoll for en randomisert studie er det ikke rapportert resultater i artikkelen. I Richter (2015) er analysene av BPM-P (mors rapportering; $N = 2582$) og BPM-Y (selvrapportering; $N = 1353$) basert på to delutvalg av en studie der CBCL og YSR ble benyttet i et stort populasjonsbasert, tilfeldig, stratifisert utvalg av elever i fjerde til tiende klasse i Sør-Trøndelag ($N = 4082$). Analysen av BPM-T (lærers rapportering; $N = 1353$) er basert på en studie der TRF ble benyttet i et stratifisert tilfeldig utvalg elever i første til syvende klasse i Norge ($N = 1409$). I artikkelen undersøkes og rapporteres reliabilitet (Cronbach's alpha og intraklasse korrelasjonskoeffisient) og begrepsvaliditet (samsvar med liknende testskårer for fullversjonene) for de tre versjonene av BPM (Richter, 2015).

³ Denne foreligger kun på tysk, og omtales ikke av oss i det følgende.

Tabell 1. Inkludert norsk studie

Referanse	Design	Populasjon	N	Mål	Rapporterte egenskaper, verdier
Richter (2015)	Tversnitt	Tilfeldig utvalg av skolebarn fra Sør-Trøndelag. Alder: $M = 10,6$ ($SD = 2,8$).	2582	BPM-Parent, Totalskår, delskalaer (internalisering, eksternalisering og oppmerksomhetsproblemer)	Reliabilitet, validitet
		Delutvalg av barna som inngikk i evaluering av BPM-P. Alder: $M = 13,2$ ($SD = 1,6$)	1150	BPM-Youth, Totalskår, delskalaer (internalisering, eksternalisering og oppmerksomhetsproblemer)	
		Stratifisert populasjonsutvalg fra Norge (et barn fra hver klasse). Alder: $M = 9,4$, ($SD = 2,0$).	1409	BPM-Teacher, Totalskår, delskalaer (internalisering, eksternalisering og oppmerksomhetsproblemer)	

Reliabilitet

Reliabilitetsanalyse ble først brukt for å identifisere de seks spørreskjemaledene i hver skala/hvert problemområdene (eksternaliserende, internaliserende og oppmerksomhet) fra CBCL, TRF og YSR som resulterte i høyest indre konsistens (Cronbach's alpha) (Richter, 2015). En trinnsvis «alpha-if-item-deleted»-prosedyre resulterte i de samme leddene som rapportert av Chorpita, et al. (2010) og Achenbach, et al. (2011)– altså de leddene som inngår i de tre versjonene av den engelske BPM.

I henhold til kriteriene til European Federation of Psychologists' Association (EFPA; 2013) er indre konsistens tilfredsstillende med Cronbach's alpha $< 0,70$, tilfredsstillende med Cronbach's alpha mellom $0,70$ og $0,79$, god med Cronbach's alpha mellom $0,80$ og $0,90$ og utmerket med Cronbach's alpha $> 0,90$. I Richter (2015) varierte Cronbach's alpha for delskalaene mellom $0,67$ (oppmerksomhetsproblemer i BPM-Y) og $0,88$ (oppmerksomhetsproblemer i BPM-T), og var generelt sett noe lavere enn de tilsvarende koeffisientene fra delskalaene i de lengre versjonene CBCL, TFR og YSR (Richter, 2015). To av delskalaene hadde alpha-verdi på henholdsvis $0,67$ (BPM-Y oppmerksomhetsproblemer) og $0,69$ (BPM-Y eksternaliserende problemer). De øvrige syv skalaene hadde alpha-verdier som kan

beskrives som adekvate til gode i henhold til EFPA, selv om reliabiliteten for delskalaen internaliserende problemer i BPM-T ($0,70$) og BPM-P ($0,71$) er akkurat over grenseverdien for det som kan karakteriseres som tilfredsstillende. Cronbach's alpha for samleskåren var $0,90$ for BPM-Y, $0,93$ for BPM-P og $0,96$ for BPM-Y, noe som ifølge kriteriene kan klassifiseres som utmerket. Intraklasse korrelasjon ble også beregnet som mål på indre konsistens, og viste samme mønster som rapportert for Cronbach's alpha (lavere på internaliserende vansker i BPM-P [$0,69$] og BPM-T [$0,68$] enn BPM-Y [$0,79$], og lavere på eksternaliserende vansker og oppmerksomhetsproblemer i BPM-Y [$0,67$; $0,68$] enn i BPM-P [$0,75$; $0,77$] og BPM-T [$0,85$; $0,88$]).

Resultatene tyder på en noe svakere indre konsistens i svarene på BPM-Y for delskalaene eksternaliserende problemer og oppmerksomhetsproblemer enn for internaliserende problemer, og også sammenlignet med svarene på eksternaliserende problemer og oppmerksomhetsproblemer i BPM-P og særlig BPM-T. I BPM-T og BPM-P, derimot, indikerer resultatene svakere indre konsistens i svarene for delskalaen internaliserende problemer enn i svarene på de to andre delskalaene, og også sammenlignet med svarene på internaliserende problemer i BPM-Y (Richter, 2015).

Validitet

Begrepsvaliditet er indikert ved korrelasjonskoeffisienter mellom delskalaene (dvs. eksternaliserende problemer, internaliserende problemer og oppmerksomhetsproblemer) i hver av de tre BPM-versjonene med tilsvarende delskalaer fra CBCL, TRF og YSR, samt mellom delskalaene og underskalaene for eksternaliserende og internaliserende problemer (f.eks. aggresjon, engstelig/deprimert og somatiske plager) i CBCL, TRF og YSR. I tillegg rapporteres regresjonskoeffisienter fra analyser der delskalaene fra de tre BPM-versjonene predikerer de tilsvarende delskalene i CBCL, TRF og YSR (Richter, 2015). Resultatene viser at delskalaene fra de tre BPM-versjonene korrelerte høyt ($r > 0,84$) med de tilsvarende delskalaene fra CBCL, TRF og YSR. Av de tre delskalaene i BPM var det oppmerksomhetsproblemer som var høyest korrelert med tilsvarende delskala fra CBCL, TRF og YSR (Richter, 2015). Delskalaene som måler internaliserende vansker i de tre versjonene av BPM korrelerte 0,87–0,92 med underskalaen «engstelig/deprimert» i de lengre versjonene, noe lavere med underskalaen «tilbaketrukket/deprimert» (0,60–0,73) og som forventet enda lavere (mellom 0,30 og 0,57) med underskalaen «somatiske problemer». Eksternaliserende vansker i BPM korrelerte høyt med underskalaen «Regelbrytende atferd» i TRF (0,82), men noe lavere med den tilsvarende underskalaen i CBCL (0,65) og YSR (0,63).

De standardiserte regresjonskoeffisientene varierte mellom 0,72 (internaliserende problemer; BPM-T/TRF) og 0,93 (oppmerksomhetsproblemer; BPM-P/CBCL og BPM-T/TRF), og viser at delskalaene i BPM forklarer omtrent tre fjerdedeler eller mer av variasjonen i de tilsvarende delskalaene i CBCL, TRF og YSR (Richter, 2015).

Diskusjon og konklusjon

Litteratursøket resulterte i kun én publikasjon som rapporterte resultater om BPM i et norsk eller skandinavisk utvalg. Dette er i seg selv et for lite kunnskapsgrunnlag, og det trengs mer forskning for å kunne konkludere angående de psykometriske egenskapene til den norske versjonen av BPM.

Studien som ligger til grunn for vår vurdering av den norske oversettelsen av BPM hadde som mål å presentere noen foreløpige resultater om reliabilitet og validitet ved bruk av BPM-P, BPM-T og BPM-Y i Norge. De to utvalgene som ble brukt er store og antatt representative for populasjonen.

Reliabiliteten til de fleste delskalaene er tilfredsstillende eller god, med unntak av delskalaene for eksternaliserende problemer og oppmerksomhetsproblemer i BPM-Y (EFPA, 2013). Lavt samsvar mellom ulike rapportører av barn og ungdoms atferd og emosjonelle vansker er mye diskutert i litteraturen (De Los Reyes & Kazdin, 2004). At de to delskalaene har lavere reliabilitet i BPM-Y enn i BPM-P og BPM-T, kan være med på å bidra til lavt samsvar mellom rapportører. Andre årsaker som f.eks. hvilke situasjoner informantene observerer barnet og dermed legger til grunn for sine svar, er imidlertid antatt å være viktigere enn målingsfeil for å forklare diskrepansen man finner mellom ulike rapportører (De Los Reyes, 2011).

Reliabiliteten til samleskårene er utmerket. Dette gjenspeiler sannsynligvis at Cronbach's alpha blir større med økende antall ledd (Pedhazur & Schmelkin, 1991), noe som også kommer fram i sammenligningen av koeffisientene for BPM og de lengre versjonene.

Intraklassekorrelasjonskoeffisienten er et mindre brukt mål på indre konsistens enn Cronbach's alpha og gir mye av den samme informasjonen. Resultatene fulgte mønsteret som diskutert over i forbindelse med Cronbach's alpha (Richter, 2015).

Begrepsvaliditeten er belyst ved å sammenligne kortutgavene med de tilsvarende lengre versjonene. Om man antar at dette er veletablerte instrumenter med godt dokumentert begrepsvaliditet (Achenbach & Edelbrock, 1978; Kornør & Jozefiak, 2012), støtter de høye korrelasjonene antakelsen om at BPM måler de samme begrepene som de opprinnelige testene. Det er ifølge kriteriene (EFPA, 2013) utmerket samsvar mellom delskalaene fra de tre versjonene av BPM og de tilsvarende delskalaene fra CBCL, TRF og YSR (Richter, 2015). Det er imidlertid en stor begrensning at analysene av BPM og de lengre versjonene er basert på de samme dataene (Smith, McCarthy, & Anderson, 2000). Det er ikke

overraskende at det er høyt samsvar mellom kort- og langversjonene når dataene i kortversjonene inngår i langversjonene. I analysene blir dataene i kortversjonen brukt to ganger – de er med på «begge sider av ligningen» – og vil på denne måten bidra til en kunstig høy korrelasjon (Smith, et al., 2000). En bedre måte å undersøke samsvar mellom lang- og kortversjoner av spørreskjemaer, er å administrere de to versjonene i to separate test-sesjoner. Funnet som viser lavt samsvar mellom delskalaen internaliserende problemer i BPM med underskalaen «Somatiske problemer» belyser dette problemet – her er det sannsynligvis lavt samsvar blant annet fordi leddene som måler somatiske plager ikke inngår i BPM. De psykometriske egenskapene ved BPM burde vært undersøkt i et design der man følger utvalget over et begrenset tidsrom og først administrerer CBCL, TRF og YSR, og deretter BPM, samt et annet instrument som måler de samme problemområdene (f.eks. Strengths and Difficulties Questionnaire).

Artikkelen til Richter (2015) omhandler ikke normdata eller gjennomsnittsverdier og varians i kliniske utvalg eller andre utvalg. Det er heller ikke foretatt noen faktoranalyse som kan belyse faktorstrukturen.

Den norske versjonen av BPM kan muligens være nyttig når det kommer til å følge med på barn og ungdoms problemer over tid eller som respons på en intervensjon eller behandling. Det trengs imidlertid flere studier som følger barna over noe tid, og ev. også underveis i behandling eller mens de mottar en intervensjon, før man kan si noe mer sikkert om instrumentet fungerer etter hensikten. Det hefter noe usikkerhet rundt påliteligheten til enkelte av delskalaene for de ulike rapportørene, mens samleskåren ser ut til å gi et sikkert mål på om det foreligger problemer eller ikke. Det er høy korrelasjon mellom de tre versjonene av BPM og de tilsvarende lengre ASEBA-instrumentene. Dette er imidlertid å forvente ettersom analysene av BPM og de øvre ASEBA-instrumentene er basert på samme data. Det bør derfor gjennomføres flere studier på norske eller skandinaviske utvalg der man administrerer BPM i tillegg til andre mål på eksternaliserende problemer, internaliserende problemer og oppmerksomhetsproblemer.

Referanser

- Achenbach, T. M. & Edelbrock, C. S. (1978). The Classification of Child Psychopathology: A Review and Analysis of Empirical Efforts. *Psychological Bulletin*, 85(6), 1275-1301. doi: 10.1037/0033-2909.85.6.1275
- Achenbach, T. M., McConaughy, S., Ivanova, M. & Rescorla, L. (2011). *Manual for the ASEBA brief problem monitor (BPM)*. Burlington, VT: ASEBA.
- Achenbach, T. M. & Rescorla, L. (2001). *Manual for the ASEBA school-age forms & profiles : child behavior checklist for ages 6-18, teacher's report form, youth self-report : an integrated system of multi-informant assessment*. Burlington, Vt: ASEBA.
- Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., Krull, J. L. et al. (2010). Evaluation of the Brief Problem Checklist: Child and Caregiver Interviews to Measure Clinical Progress. *Journal of Consulting and Clinical Psychology*, 78(4), 526-536. doi: 10.1037/a0019602
- De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, 40(1), 1-9.
- De Los Reyes, A. & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment*, 16(3), 330.
- European Federation of Psychologists' Association (EFPA). (2013). EFPA Review model for the description and evaluation of psychological tests: Test review form and notes for reviewers, v 4.2.6: European Federation of Psychologists' Association (EFPA).
- Kornør, H. & Jozefiak, T. (2012). Måleegenskaper ved den norske versjonen av Child Behavior Checklist- versjon 2-3, 4-18, 11/2-5 og 6-18 (CBCL). *PsykTestBarn*, 2(1:3).
- Pedhazur, E. J. & Schmelkin, L. P. (1991). *Measurement, Design, and Analysis : An Integrated Approach*. London: Taylor and Francis.
- Penelo, E., de la Osa, N., Navarro, J. B., Domènech, J. M. & Ezpeleta, L. (2017). The Brief Problem Monitor-Parent form (BPM-P), a short version of the Child Behavior Checklist: Psychometric properties in Spanish 6-to 8-year-old children.

- Psychological Assessment*, 29(11), 1309-1320.
- Piper, B. J., Gray, H. M., Raber, J. & Birkett, M. A. (2014). Reliability and validity of Brief Problem Monitor, an abbreviated form of the Child Behavior Checklist. *Psychiatry and Clinical Neurosciences*, 68(10), 759-767.
- Richter, J. (2015). Preliminary evidence for good psychometric properties of the Norwegian version of the Brief Problems Monitor (BPM). *Nordic Journal of Psychiatry*, 69(3), 174-178.
- Rodenacker, K., Plück, J. & Döpfner, M. (2015). Fragebogen zum Problem-Monitoring für Eltern, Lehrer und Jugendliche—eine deutsche Fassung des Brief Problem Monitor (BPM). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 44, 197-206. doi: 10.1026/1616-3443/a000307
- Smith, G. T., McCarthy, D. M. & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12(1), 102.