

KAPITTEL 3

Mestrer jeg trenerrollen?

Validitet og reliabilitet i en skala til måling av mestringstro i et utvalg norske breddefotballtrenere

Yngvar Ommundsen^{1,2*}, Bård Erlend Solstad¹,
Rune Høigaard² & Bente Wold³

¹Norges idrettshøgskole, Seksjon for coaching og psykologi; ²Universitetet i Agder, Institutt for folkehelse, idrett og ernæring; ³Universitetet i Bergen, HEMIL-senteret, Det psykologiske fakultet

Sammendrag: Formålet med denne studien var å teste dimensjons-, konvergens-, diskriminant- og kriterievaliditet for en norsk versjon av Feltz, Chase, Moritz og Sullivan (1999) – *Coaching Efficacy Scale*. Invarians over tid og mellom grupper ble også inkludert med tanke på prospektiv bruk av skalaen i form av lengdesnittstudier og effektstudier. Utvalget omfattet 281 norske breddefotballtrenere i tre norske fotballkretser. Disse deltok i datainnsamlinger før (T1) og etter sesong (T2) i den norske delen (Solstad, 2017) av det internasjonale PAPA-prosjektet (Quested et al., 2013). Innledende eksplorativ faktoranalyse i *Mplus* støttet en forventet 4-faktors mestringstromodell som reflekterte henholdsvis: Motivasjon, Kampstrategi, Teknikk og Karakterbygging. To utsagn viste høy residualvarians og ble eliminert. En oppfølgende analyse basert på 22 utsagn gav tilfredsstillende tilpasningsindekser (S-B $\chi^2 = [df = 149, n = 222] = 312.901, p < .001$; RMSEA = .070; 90% konfidensintervall (CI) = .059–.081); CFI = .94; SRMR = .025). Ved parselleringsteknikk (Kline, 2011; Little, Rhemtulla, Gibson, & Schoemann, 2013) ble 22 utsagn konvertert til elleve parseller som basis for sjekk av faktorvaliditet og invarians over tid og mellom grupper ved hjelp av latent variabelanalyse. De elleve parsellene gav svært gode tilpasningsindekser; (S-B $\chi^2 = [df = 29, n = 222] = 38.622, p < .011$ (ns.); RMSEA = .039; 90% (CI) = .000–.068); CFI = .99; SRMR = .023). I tillegg fant vi entydig støtte for sterk skaleringsinvarians over tid (fotballsesongen; T1–T2) for intervensjons- og kontrollgruppe samlet. Invarians splittet på intervensjons-/kontrollgruppe kunne derimot ikke påvises. Vi observerte også god konvergent og diskriminerende validitet. Kriterievaliditet ble også bekreftet. I sum synes skalaen å være velegnet til undersøkelse av tverrsnitt- og lengdesnittstudier av treneres mestringstro i barne- og ungdomsidrett. Skalaen trenger ytterligere validering før bruk i intervensjonsstudier.

Nøkkelord: treneres mestringstro, skalavalidering, breddefotball

*Korresponderende forfatter: Yngvar Ommundsen, Norges idrettshøgskole, Seksjon for coaching og psykologi, Postboks 4014 Ullevål stadion, 0806 Oslo. Epost: yngvar.ommundsen@nih.no

Sitering av denne artikkelen: Ommundsen, Y., Solstad, B.E., Høigaard, R. & Wold, B. (2018). Mestrer jeg trenerrollen? Validitet og reliabilitet i en skala til måling av mestringstro i et utvalg norske breddefotballtrenere. I T. Haugen & R. Høigaard (red). *Trender i idrettspsykologisk forskning i Skandinavia* (Kap. 3, s. 47–72). Oslo: Cappelen Damm Akademisk. DOI: <https://doi.org/10.23865/noasp.39.ch3>
Lisens CC-BY 4.0

Abstract: The purpose of this study was to examine factorial, convergent, discriminant, and criteria validity of a multi-dimensional scale measuring coaching efficacy (Feltz, Chase, Moritz, & Sullivan, 1999). The sample consisted of 281 Norwegian grassroots football coaches taking part in two data collections (across the football season) within the Norwegian arm of the international PAPA project (intervention group, $n = 193$) and control group, $n = 88$). Latent variable modeling supported a 4-factor model reflecting Motivation efficacy, Strategy efficacy, Technique efficacy and Character efficacy. An introductory ESEM analysis of the 24-item coaching efficacy scale yielded an acceptable solution after omitting two items. A parceling approach of the twenty-two items were then converted into eleven parcels to examine factorial validity and time and group invariance. The four-factor model based on parcels obtained good fit indices: (S-B $\chi^2 = [df = 29, n = 222] = 38.622, p < .011$ (ns.); RMSEA = .039; 90% (CI = .000–.068); CFI = .99; SRMR = .023). Moreover, we obtained strong support for time invariance for a merged solution of the two groups of coaches. Invariance across time for the intervention and control group separately was supported at time 1, but not at time 2. Satisfactorily convergent, discriminant, and criteria validity were also supported. The scale needs further validation before use in intervention studies.

Keywords: coaching efficacy, scale validation, grassroots football

Introduksjon

Mestringstro (efficacy) har vært definert som en persons tro på egne ferdigheter i å planlegge og gjennomføre ulike typer handlinger for å nå mål (Bandura, 1997). Mestringstro genereres via sosial læring – gjennom individets interaksjon med seg selv og sine sosiale og fysiske omgivelser (Bandura, 1997). Forskning har vist at mestringstro bidrar til økt motivasjon, høyere målsettinger, økt anstrengelse og dermed bedre resultater (Benight & Bandura, 2004; Feltz, Short, & Sullivan, 2008). Teorien om mestringstro har blitt applisert spesifikt på ulike individuelle roller, områder og grupper (dvs. studenters mestringstro, mestringstro blant lærere og kollektiv mestringstro; Bandura, 1997; Feltz et al., 2008).

Feltz et al. (1999) utviklet en skala til måling av mestringstro i rollen som trener. Den er integrert i en modell hvor innholdet i mestringstro spesifiseres, samt forløpere og konsekvenser. Skalaen måler fire dimensjoner: Motivasjon, Kampstrategi, Teknikk og Karakterbygging. «Motivasjon» inkluderer tillit til at man kan motivere og influere på psykologiske ferdigheter og tilstander blant egne utøvere. «Kampstrategi» er knyttet til strategisk tenkning og valg, og innebærer tiltro til at man i konkurranser evner å lede egne utøvere til gode prestasjoner. «Teknikk» er troen på at en har tiltro til egne ferdigheter som trener, og det er knyttet til å instruere, diagnostisere og veilede samt demonstrere gode ferdighetsbilder i egen idrett. Og «Karakterbygging» er å

ha tiltro til egen evne til å få utøverne til å tenke og handle i tråd med gode etiske retningslinjer (f.eks. fair play) når de utøver idretten sin.

Trenere som har sterk tiltro til det å kunne motivere, foreta taktiske valg, formidle tekniske ferdigheter samt formidle betydningen av god sportsånd, vil merke påvirkning på egen motivasjon og atferd, noe som i sin tur påvirker utøvernes tankemønstre, følelser og atferd (Feltz et al., 1999; Mageau & Vallerand, 2003). Forskning har vist at trenere som opplever høy mestringstro i rollen som trenere, synes å være mer tilbøyelige til å lede på en måte som er til beste for utøverne (Feltz et al., 1999; Sullivan, Paquette, Holt, & Bloom, 2012; Sullivan & Kent, 2003). Det er for eksempel vist at lederskap preget av mestringstro evner å skape indre motivasjon blant utøverne og forebygge frafallet i organisert idrett (Hollembeak & Amorose, 2005; Pelletier, Fortier, Vallerand, & Briere, 2001).

En skala om treneres mestringstro - psykometriske kvaliteter

I de senere årene har skalaen for måling av mestringstro for trenere gjennomgått utstrakt psykometrisk testing (for en oversikt, se Myers, Chase, Pierce, & Martin, 2011). Tidlige tester av skalaens modelltilpasning viste i flere tilfeller uakseptable verdier; eksempelvis .87–.89 for CFI og .08–.09 for RMSEA. Spesielt så man dette i studier blant idrettsutøvere innen amerikansk idrett på videregående/gymnasnivå (Feltz et al., 1999). Samtidig viser senere valideringsforsøk på tilsvarende utvalg noe bedre resultater (for en oversikt, se Myers, Feltz, & Wolfe, 2008; Myers et al., 2011).

Mestringstro i trenerrollen har fått overraskende lite oppmerksomhet blant trenere i barne- og ungdomsidretten (Feltz, Hepler, Roman & Paiement, 2009; Myers et al., 2008). Feltz et al. (2009) undersøkte psykometriske egenskaper ved skalaen på et utvalg amerikanske ungdomsidrettstrenere, hvor de fant blandet støtte for skalaens faktorvaliditet ved bruk av latent variabelanalyse (bekreftende faktoranalyse, CFA). En mer nylig kanadisk studie innen barne- og ungdomsidretten undersøkte betydningen av idretts-/veiledningskonteksten og trenernes formelle utdanning for deres mestringstro, og hvordan en slik forventning var forbundet med deres opplevelse av hvordan de veiledet og ledet utøverne sine (Sullivan et al., 2012). Resultatene viste at formell trenerutdanning

er av betydning for deres tro på det å motivere, foreta taktiske valg, formidle tekniske ferdigheter samt formidle betydningen av god sportsånd. I den studien målte man imidlertid mestringsstro overordnet, ikke på de respektive fire nevnte dimensjonene. Dermed får man liten innsikt i hvordan trenerutdanning eksplisitt påvirker for eksempel tiltro til å motivere versus tiltro til ferdighet i å foreta taktiske valg i møte med utøverne.

Mestringsstroskalaen for trenere behøver ytterligere validering

Horn (2002) har pekt på at utvidet validering av skalaen innenfor rammen av barne- og ungdomsidrett åpner opp for en mer inngående undersøkelse av psykologiske faktorer og prosesser som kan påvirke idrettsdeltakelsen, sett både fra utøvernes ståsted og fra trenerne selv. Uavhengig av nylige forsøk på bruk av skalaen og testing av dens psykometriske egenskaper i andre kulturelle kontekster (Malete, Chow, & Feltz, 2013) synes det viktig å validere skalaen til bruk her i landet. Det finnes i det hele tatt lite validering og bruk av skalaen i en europeisk idrettskontekst, og formålet med denne studien er derfor å undersøke den psykometriske kvaliteten ved den norske versjonen av Feltz et al. (1999) – *Coaching Efficacy Scale*.

Innen psykologisk forskning pågår det en debatt vedrørende metodikk i forbindelse med språklig og kulturell oversettelse av denne type skalaer. Dette er problemstillinger som ofte underkommuniseres blant forskere (Segall, Lonner, & Berry, 1998). I denne studien har vi forsøkt å vise transparens med hensyn til slike problemstillinger, og vi har fulgt anbefalte anvisninger i forbindelse med oversettelse og psykometrisk testing av skalaen i en norsk idrettskontekst (Byrne, 2012). Det innbefatter bl.a. stringent oversettelse og re-oversettelse fra engelsk til norsk (Epstein, Santo, & Guillemin, 2015), samt bruk av ulike analyseformer for å undersøke validitet og reliabilitet.

Invarians over tid og mellom grupper

Undersøkelse av psykometrisk kvalitet innebærer også hvorvidt skalaen egner seg ved prospektive analyser og analyse av intervensjonseffekter.

Følgelig tok vi et steg lenger enn hva tidligere forskning har gjort ved å teste den norske versjonen med hensyn til invarians over tid, samt invarians mellom to grupper (en intervensjonsgruppe og en kontrollgruppe) blant trenere som inngikk i den norske delen av det internasjonale PAPA-prosjektet (se utvalgs karakteristika for nærmere omtale).

Invarians i et måleinstrument er avgjørende ved test av endring i skårer over tid og/eller mellom grupper. Gjennomsnittsverdiene til latente faktorer kan ikke sammenliknes dersom måleinstrumentene ikke oppfyller kravene til faktorinvarians (Little et al., 2013; Sass, 2011). Tidligere studier har undersøkt i hvilken grad utdanningsprogrammer for trenere påvirker deres mestringstro (Campbell & Sullivan, 2005; Maleté & Feltz, 2000), men så vidt oss bekjent testet ikke disse eller andre studier mestringstrodimensjonene for invarians over tid og/eller mellom grupper. Oppnår man invarians, kan man være rimelig sikker på at enhver påvist forskjell i skår på skalaen over tid og/eller mellom grupper er reell, og ikke en artefakt av manglende invarians. Rent overordnet innebærer invarians at utsagnene (observert nivå) som er underliggende for de fire latente dimensjonene i mestringstro, er homogene og stabile over tid og/eller mellom grupper (Geiser, 2013; se statistisk analyse for ytterligere detaljer). Bare under slike betingelser kan man ha tiltro til resultatenes validitet (for en oversikt, se Geiser, 2013).

Kriterievaliditet

Kriterievaliditet ble sjekket ved å undersøke de fire dimensjonene av mestringstro opp mot trenernes positive emosjonelle tilstand i forbindelse med trenergjerningen. Rasjonalet for dette er basert på tidligere forskning, som har vist at eksperimentelt induisert økt mestringstro relatert til gjennomføring av fysisk aktivitet er assosiert med en økning i positive følelser etter å ha trent (Jerome et al., 2002; McAuley, Talbot, & Martinez, 1999). I og med at selvrapportert vitalitet i dagliglivet kan tilsløre forholdet mellom mestringstro i rollen som trener og trenerens positive emosjonelle tilstand i trenerrollen, kontrollerte vi for vitalitet i dagliglivet.

Metode

Deltakere

Utvalget bestod av 281 norske trenere innen breddefotball fra tre norske fotballkretser (menn $n = 258$, kvinner $n = 23$). Gjennomsnittsalder var 42 år ($SD = 6,1$ år); aldersintervall fra 16 til 60 år. De var hovedsakelig etnisk norske (94%). Trenerne hadde ulike roller i eget fotballag; hovedtrenere ($n = 44$), assistenttrenere ($n = 86$), trenere med delt ansvar ($n = 71$) og 19 trenere som rapporterte andre varianter av ansvarsfordeling, samt 61 trenere som ikke oppgav noe vedrørende egen rollefunksjon i laget. Gjennomsnittlig erfaring som trener var 7 år ($SD = 4,9$ år); og gjennomsnittlig antall år som trener for nåværende lag var 4,3 år ($SD = 2,1$ år). Når det gjaldt formell trenerutdanning, rapporterte 48% av totalutvalget en viss formell trenerutdanning (grovt rangert som lavt, medium og høyt nivå). 52% svarte ikke på dette spørsmålet. Utvalget ble randomisert delt inn i en intervensjonsgruppe ($n = 193$) og en kontrollgruppe ($n = 88$). Datainnsamlingen og utvalget var integrert i den norske delen av det internasjonale *Promoting Adolescent Physical Activity*-prosjektet (PAPA) (Duda et al., 2013). PAPA er et intervensjonsprosjekt med mål om å stimulere til økt motiverende lederskap blant trenere i den organiserte breddefotballen blant barn og unge (Quested et al., 2013; Solstad, Van Hoyer, & Ommundsen, 2015). Påvirkning av treneres mestringsforventning var ikke et formalisert formål med det internasjonale PAPA-prosjektet (Quested et al., 2013). I den norske delen av datainnsamlingen ble imidlertid treneres mestringstro i trenerrollen lagt til i spørreskjemaet med tanke på evaluering av måleinstrumentet på norsk, samt med tanke på en eventuell endring av slik mestringstro som følge av intervensjonen. Et teoretisk rasjonale for dette er dokumentert i tidligere forskning ved at kontekstuell stimulering kan gi næring til et mestringsfokuseret målperspektiv med positive implikasjoner på egen tiltro til å mestre ulike oppgaver i rollen som trener (Duda & Nicholls, 1992; Jourdan, Bandura, & Banfield, 1991). En validering av instrumentet i en norsk idrettskontekst åpner også opp for bruk i eventuelle framtidige studier som ønsker å undersøke treneres mestringstro i rollen som trener.

Framgangsmåte

Før rekruttering av fotballklubber til prosjektet ble datainnsamlingen blant trenere og spillere godkjent av NSD (Norsk senter for forskningsdata). Trenere i disse klubbene, som hadde lag i breddefotballen i aldersintervallet 11–14 år, ble invitert til å delta. Begge datainnsamlingene (T1 og T2) ble gjennomført etter forhåndsavtale med trenerne før eller etter en trening med laget, eller på et annet passende tidspunkt for dem. Trenerne ble skriftlig og muntlig orientert om hva deltakelse i prosjektet innebar, og informert om at de når som helst kunne trekke seg. Datainnsamlingen foregikk ved at trenerne fylte ut standardiserte spørreskjemaer.

Måleinstrumenter

Demografiske bakgrunnsdata

Det ble innhentet data for alder, kjønn, etnisitet og fotballbakgrunn (trener og som spiller), og med spesifisert informasjon om trenergjerningen. Dette inkluderte trenererfaring (nåværende og tidligere lag), antall år som trener, type lag, formell trenerutdanning (kurs, sertifisering), samt ansvarsfunksjon i eget lag.

Treneres mestringstro

En norsk versjon av Coaching Efficacy Scale (Feltz, et al., 1999) ble benyttet for å måle treneres mestringstro. Skalaen består av 24 utsagn og skiller mellom fire ulike kategorier av mestringstro: Motivasjon (dvs. tro på egen evne til å motivere egne spillere), Kampstrategi (dvs. tro på egen evne til å foreta taktiske valg), Teknikk (dvs. tro på egen evne til utvikle tekniske ferdigheter) og Karakterbygging (dvs. tro på egen evne til å formidle betydningen av god sportsånd). Utsagnene fra den originale versjonen ble modifisert med sikte på å knytte mestringstro spesifikt til trenerrollen i breddefotballen for barn og unge. Således ble de fire aspektene ved mestringstro knyttet spesifikt til deres forventning om å mestre dette overfor spillere i eget fotballag. Det innebar at «*utøvere*» ble erstattet med «*spillere*» i alle utsagnene, og inngangen til utsagnene var følgende: *Alle trenere er ulike med hensyn til hva de føler at de gjør bra eller dårlig sammen med laget*

sitt. Sett ring rundt det tallet som representerer hvor stor tro du har på dine evner med tanke på de ulike ferdighetene som står under, fulgt opp av: Hvor stor tro har du på at du kan. På linje med Feltz et al. (1999) ble alle utsagnene besvart på en responsskala fra 0 (*ingen tro*) til 9 (*svært stor tro*). Da instrumentet ble brukt første gang i en norsk idrettskontekst, gjennomgikk det en engelsk–norsk oversettelse og re-oversettelse ved hjelp av kyndige tospråklige personer med innsikt i tematikken. Målet med oversettelsesprosedyren var å reflektere meningsinnholdet i den engelske versjonen (kilden), snarere enn å foreta en bokstavelig oversettelse av ordbruken i kildeversjonen (Epstein et al., 2015; Su & Parham, 2002).

Positiv affektiv tilstand

En kortversjon (seks utsagn) av *The Positive and Negative Affect Schedule* (PANAS) ble benyttet (Watson, Clark, & Tellegen, 1988). Skalaen ble tilpasset fotballkonteksten via inngangssetningen: *Sett en ring rundt det svaret som passer best med hvor ofte de følgende følelsene har preget deg når du har trent dette laget i løpet av den siste måneden*, etterfulgt av: *I løpet av den siste måneden mens jeg har trent dette laget, har jeg stort sett følt meg:* (a) glad, (b) lykkelig, (c) oppstemt, (d) begeistret/frydefull, (e) entusiastisk, og (f) stolt. En syvpunkts responsskala fra 1 (*sjelden/aldri*) til 7 (*alltid*) ble benyttet.

Vitalitet

Vi anvendte en fempunkts utsagnsskala basert på arbeidet til Ryan og Frederick (1997), som måler vitalitet i dagliglivet. Inngangen til skalaen var: *Under står noen utsagn relatert til hverdagslivet ditt (alt du vanligvis gjør) og ikke bare til fotballgjerningen din. Sett en ring rundt svarene som passer best med hvordan du generelt har kjent deg den siste måneden*, etterfulgt av: *Den siste måneden har jeg stort sett:* (a) følt meg opplagt, (b) vært i godt humør, (c) gledet meg til hver dag, (d) følt meg kvikk og våken, og (e) hatt masse energi.

Statistiske analyser

I tråd med anbefalinger (Kline, 2011) anvendte vi eksplorativ faktoranalyse (ESEM; Mplus versjon 8.0) i en innledende fase på den opprinnelige

24-punkts utsagnskalaen med mål om avklaring av beste tilpasning (en faktor versus 2-, 3- eller 4-faktorløsning), etterfulgt av sjekk av faktorladninger samt residualvarians. Vi anvendte deretter bekreftende faktoranalyse (CFA) med estimering av maksimal sannsynlighet (robust maximum likelihood estimation; MLR-estimator) på en utsagnsløsning med tilfredsstillende estimater avdekket i ESEM-analysen.

Parsellering

Vi tok i bruk parsellering ved måling av faktorvaliditet og invarians. Dette innebærer at hvert enkelt utsagn deles opp i ikke-overlappende enheter, eller sett av utsagn (Kline, 2011). Det skjer via en totrinnsprosess (total disaggregering; Bagozzi & Edwards, 1998) hvor utsagnene først aggregeres til parseller som hver består av to eller flere utsagn. Deretter anvendes de aggregerte utsagnene som manifeste indikatorer for et latent begrep i en flerdimensjonal målemodell (delvis disaggregering; Bagozzi & Edwards, 1998). I motsetning til å operere med enkelt-utsagn som manifeste variabler, har bruk av parseller som manifeste variabler følgende fortrinn: (a) færre parametere som skal estimeres, (b) lavere indikator-til-deltaker-ratio, (c) mindre sannsynlighet for korrelerte restledd (residualer), og (d) doble faktorladninger samt færre kilder som kan forårsake utvalgsfeil (for en oversikt, se Little et al., 2013). Parsellering bidrar til å redusere kompleksiteten i målemodeller og strukturmodeller som inkluderer flere latente begreper; modeller som ellers kan vise seg for komplekse til at de lar seg estimere. Dette krever imidlertid at man innen hver parsell opererer med endimensjonale begrep (Kline, 2011, Little et al., 2013). Av den grunn brukte vi som nevnt først ESEM-analyse for å bekrefte grunnstrukturen (4-faktorløsning) i mestringstroskalaen. Little et al. (2013) har argumentert for en tilnærming ved opprettelse av parseller hvor disse forsøkes balansert (dvs. *the balancing approach*). Hovedhensikten er å få en mest mulig balansert representasjon av de ulike utsagnenes styrke til å representere det latente begrepet. Operativt gjøres det ved at man lager par av utsagn som har henholdsvis høyest og lavest faktorladning, nest høyest og nest lavest osv. (Little et al., 2013).

Modelltilpasningen for de fire dimensjonene ved mestringstro – Motivasjon, Kampstrategi, Teknikk og Karakterbygging – ble undersøkt ved hjelp av en kji-kvadrattest (Satorra–Bentler chi-square test; S-B χ^2 difference test; for en oversikt, se Satorra & Bentler, 2001), sammenliknende tilpasningsindeks (CFI, Bentler, 1990), standardisert kvadratrot-residual (SRMR; Bentler, 1995), tilnærmet kvadratrot-basert feilestimat (RMSEA; Browne & Cudek, 1993), med tilhørende 90% konfidensintervall rundt RMSEA. Generelt reflekterer RMSEA verdier nært opp til .06 (Hu & Bentler, 1999) akseptabel tilpasning, og verdier lik eller mindre enn .05 god tilpasning (Browne & Cudek, 1993). CFI-verdier lik .95 eller over, og SRMR verdier mindre enn eller lik .08 speiler en tilstrekkelig modelltilpasning til data (Hu & Bentler, 1999; Marsh, Wen, & Hau, 2004).

Invarians over tid og/eller mellom grupper

Prospektive analyser (endringsanalyser) av mestringstro betinger invarians over tid av de målingene som gjøres. Invarians innebærer sammenlikning av en serie nestete modeller, hvor man suksessivt pålegger dem flere restriksjoner (Dimitrov, 2010). *Konfigurativ invarians* reflekterer at det er konsistens over tid og/eller mellom grupper i *antallet* indikatorer tilhørende et latent begrep, samt at korrespondansen mellom indikator og latent begrep er lik over tid. *Metrisk invarians* speiler hvorvidt ladningene som tilhører de respektive indikatorene for et latent begrep, er like over tid og/eller mellom grupper. *Skaleringsinvarians* speiler hvorvidt innslags-/avskjæringspunktet (intercept) for de respektive indikatorene er likt over tid og/eller mellom grupper. Serien av nestete modeller med suksessivt flere restriksjoner blir så sammenliknet for tilpasning. Endringer i tilpasningsindekser over de tre nivåene konfigurativ, metrisk og skaleringsinvarians anses akseptabel og indikerer invarians ved ikke-signifikant endring i kji-kvadrat, samt $\Delta\text{CFI} \leq .01$ og $\Delta\text{RMSEA} \leq .015$) ved å sammenlikne tilpasningsindeksene mellom de ulike invariansnivåene (for en oversikt, se Chen, 2007; Cheung & Rensvold, 2002).

Konvergent og diskriminant validitet samt kriterievaliditet

Vi testet konvergent validitet gjennom å sjekke størrelse og signifikansnivå for koeffisientene på de respektive observerte parsellene under de fire latente begrepene for mestringsstro (Cole, 1987). Signifikante verdier på de parselleringsbaserte koeffisientene indikerer god konvergent validitet. I tillegg så vi også på størrelsen på koeffisientene (dvs. faktorladningene). Vi testet diskriminantvaliditet ved inspeksjon av korrelasjoner mellom de fire ulike dimensjonene av mestringsforventning. Korrelasjonsstørrelse $< .85$ er tegn på tilstrekkelig diskriminerende validitet (Kline, 2011). Test av diskriminerende validitet ble supplert ved hjelp av en mer sofistikert tilnærming (Fornell & Larcker, 1981; se resultatdel for en mer utfyllende beskrivelse). Kriterievaliditet ble testet ved hjelp av partiell korrelasjonsanalyse av enkle sumskårer i SPSS.

Reliabilitet

Reliabilitetskår av skalaen ble estimert ved hjelp av sammensatt reliabilitet (Composite reliability; Raikovs Rho) i *Mplus*. Fordelen med denne er at den ikke betinger at alle utsagnene bidrar likt til faktorvarians. Samtidig tar Raikovs Rho høyde for korrelert feilvarians (Yang & Green, 2010).

Manglende data

Av 281 trenere fylte 222 ut spørreskjemaet ved T₁ og 204 ved T₂. Vi fikk data fra 145 både ved T₁ og ved T₂, mens 77 trenere fylte ut spørreskjemaet kun ved T₁ og 59 trenere kun ved T₂. Imidlertid aksepterte alle trenerne i utvalget vår invitasjon om å delta både ved T₁ og ved T₂, og det var ingen som trakk seg fra studien. Ingenting tyder på at manglende data (dvs. *missing data analysis*) skyldes annet enn organisasjonsmessige forhold knyttet til datainnsamlingene, som f.eks. lang reiseavstand for de som hadde ansvaret for datainnsamlingen, samt at trenere i flere tilfeller ikke så seg i stand til å tilpasse endrede tidspunkter initiert

av de som samlet inn data, eller at de selv – grunnet uforutsette faktorer – ikke kunne møte til avtalt tid for datainnsamling med laget sitt. Av den grunn anså vi data som manglende ved tilfeldighet (*missing at random*; MAR), og vi anvendte derfor analysen *Full Information Maximum Likelihood* (FIML). Den betraktes som overlegen andre tradisjonelle teknikker ved at den tar i bruk alle datapunkter og genererer fullverdig informasjon for tilpasning til en statistisk modell basert på alle tilgjengelige data (Little et al., 2013). FIML betraktes dessuten som effektiv med hensyn til å redusere skjevhet grunnet ikke-tilfeldig oppstått manglende data (*missing not at random*; MNAR), og den evner å gi reliable analyse-resultater selv ved en sats på 60% manglende data (for en oversikt, se Dong & Peng, 2013).

Resultater

Beskrivende statistikk

Gjennomsnittsverdier, standardavvik og estimer for sammensatt reliabilitet for de fire dimensjonene av mestringstro, positive følelsetilstander og vitalitet i dagliglivet er gjengitt i Tabell 1.

Tabell 1. Beskrivende statistikk, reliabilitetsestimater og estimer vedrørende diskriminerende validitet (AVE).

	Gj.snitt	SD	composite RHO#	AVE skår
MT: Motivasjon	7.14	0.92	.92	0.64
MT: Karakterbygging	7.73	0.93	.90	0.70
MT: Kampstrategi	6.30	1.20	.90	0.58
MT: Teknikk	6.81	1.17	.92	0.67
Positive følelser	5.66	0.79	.88	*
Vitalitet	3.89	0.63	.92	*

Note. *ikke relevant; #utsagnsbasert RHO; MT = mestringstro.

Sammensatt reliabilitet var tilfredsstillende med Raikovs Rho-koeffisienter i størrelsesorden fra .90 til .92. Trenerne rapporterte relativt høy mestringstro, høyest for Karakterbygging og lavest for Kampstrategi.

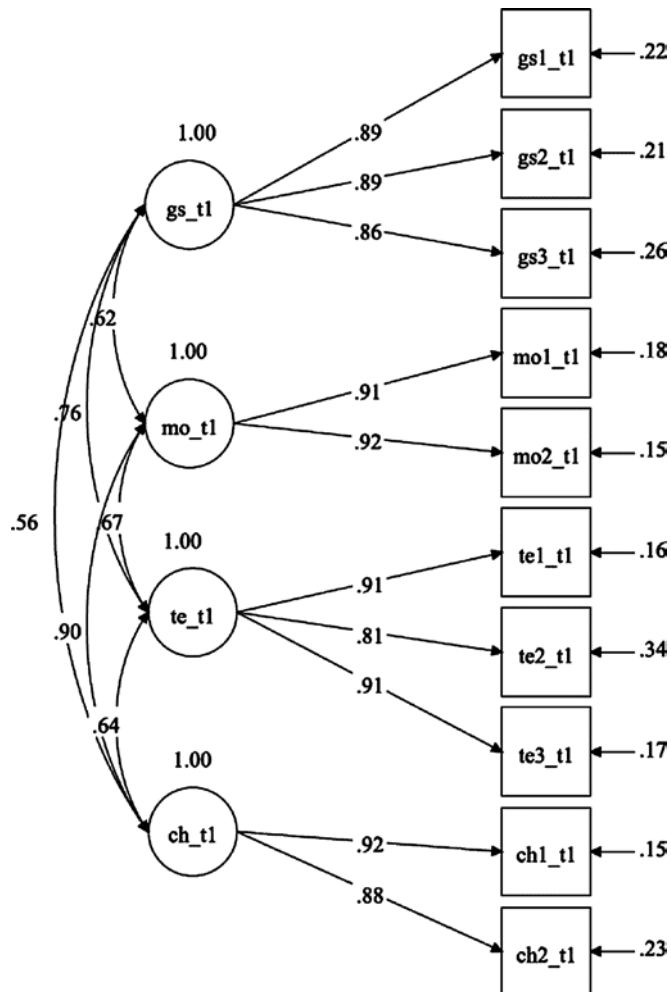
En enveis variansanalyse i SPSS med oppfølgende Tukey-b post-hoc-test viste signifikant høyere skår på teknisk mestringstro blant trenere med middels ($M = 7.04$) eller høyere nivå av formell trenerutdanning ($M = 7.08$), sammenliknet med gruppen lav ($M = 6.05$). Ingen forskjeller kunne påvises for de andre dimensjonene. En t-test for uavhengige utvalg ble brukt for å undersøke forskjeller i mestringstro mellom trenere som oppgav ulike roller i eget lag (dvs. hovedtrener og likt ansvar versus assistenttrener). Resultatene viste at trenere med hovedansvar eller delt ansvar skåret signifikant høyere på mestringstro relatert til *Teknikk* ($t = 4.40, p < .001$), *Motivasjon* ($t = 2.34, p < .01$) og *Kampstrategi* ($t = 2.71, p < .01$) enn assistenttrenerne (alle disse analysene ble foretatt i SPSS i en innledende fase basert på enkle summerte skårer på de fire dimensjonene av mestringstro).

Den 24-utsagnsbaserte 4-faktormodellen som bygget på ESEM-analyse, gav akseptabel tilpasning til data, men to utsagn – (a) *finne ut av motstanderlagets sterke sider i kamp*, og (b) *forberede spillerne mentalt på kamp* – ble tatt ut grunnet høy residual varians (se vedlegg, og se diskusjonsdel for legitimering). En modifisert ESEM-versjon basert på de gjenværende 22 utsagnene gav klart akseptabel tilpasning til data: (S-B $\chi^2 = [df = 149, n = 222] = 312.901, p < .001$; RMSEA = .070; 90% (CI = 0.059–0.081); CFI = .94; SRMR = .0025). GEOMIN-roterte faktorladninger var høye og signifikante; i størrelsesorden fra .46 til .89 på de respektive forventende utsagnene.

En CFA-versjon basert på den samme 22-utsagnsversjonen gav en akseptabel tilpasning til data: (S-B $\chi^2 = [df = 203, n = 222] = 458.425, p < .001$; RMSEA = .080; 90% (CI = .066–.084); CFI = .91; SRMR = .066). Standardiserte faktorladninger var høye og signifikante, i størrelsesorden fra .70 til .89, og residualvariens var overveiende lav. Resultatene fra de innledende ESEM-analysene, og CFA-analysene, styrker konklusjonen om en *samlet sett* klart akseptabel tilpasning for 22-utsagnsmodellen, og kriteriet for å gå videre med en parselleringstilnærming synes å være til stede (Kline, 2011).

De elleve parsellene tilhørende henholdsvis dimensjonene *mestringstro knyttet til å: (a) motivere, (b) foreta taktiske valg, (c) utvikle tekniske ferdigheter, samt (d) formidle betydningen av god sportsånd*, basert på 22-utsagnsskalaen, gav svært gode tilpasningsindekser: (S-B $\chi^2 = [df = 29, n = 222] = 38.622, p < .011$ (ns.); RMSEA = .039; 90% (CI = .000–.068);

CFI = .99; SRMR = .023). Standardiserte faktorladninger for de ulike observerte skårene (parsellene) for den enkelte dimensjon ladet signifikant, sterkt og entydig på de respektive fire latente dimensjonene uten indikasjoner på bedre modelltilpasning ved tillegg av kryssladninger. Dette indikerer god begreps- eller dimensjonsvaliditet for skalaene. Det er også en indikator på god konvergerende validitet (Cole, 1987). Som vist i Figur 1 er de standardiserte vektene i størrelsesorden fra .81 til .96.



Figur 1. 22-utsagnsskala for mestringsstro - faktorladninger, residualer og inter-korrelasjoner mellom dimensjoner.

Note. gs = Kampstrategi; mo = Motivasjon; te = Teknikk; ch = Karakterbygging.

Videre viste resultatene at minst 50% av variansen (69–93%) for hver av de ti parsellbaserte observerte indikatorene forklares av modellen. Restleddene (residualene) er svært små – og størrelsesmessig jevnt fordelt. Sammen med tilpasningsindeksene og de standardiserte faktorladningene indikerer disse funnene at den teoretiske målemodellen har god tilpasning til data, og at de elleve parsellene dermed uttrykker god dimensjonsvaliditet og konvergerende validitet for de fire latente begrepene: *Motivasjon*, *Kampstrategi*, *Teknikk* og *Karakterbygging* (Anderson & Gerbing, 1988; Cole, 1987; Kline, 2011).

Reliabilitetsestimater vises både via reliabilitetskårer og sammensatt reliabilitet (Raikovs Rho) i Tabell 1. Som også vist i Tabell 1 er estimatene for Raikovs Rho på mellom .90 og .92, med noe lavere verdier for reliabilitetskårene. I sum er det støtte for god reliabilitet i de fire dimensjonene av mestringstro.

Invarians over tid og gruppe

Ifølge Tabell 2 viste resultatene entydig støtte til *Konfigurativ invarians*, *Metrisk invarians* og *Skaleringsinvarians* fra T1 til T2 for intervensjons- og kontrollgruppen samlet. Ingen endring i kji-kvadratverdier ble observert, CFI-verdiene forholdt seg høye, og RMSEA-verdiene forholdt seg lave, med kun marginale justeringer fra Konfigurativ invarians til Skaleringsinvarians ($\Delta CFI \leq .01$, $\Delta RMSEA \leq .015$).

Samlet indikerer dette invarians ved sammenlikning av de respektive nestete modellene (Chen, 2007; Cheung og Rensvold, 2002), og således støtte til *Skaleringsinvarians*. Det åpner dermed opp for at skalaen kan brukes til sammenlikning av gjennomsnittsverdier på skårer over tid (Dimitrov, 2010). Vi testet også gruppeinvarians (intervensjon/kontrollgruppe) på henholdsvis T1 og T2 (data ikke vist). I sum fant vi støtte for Skaleringsinvarians på T1 ($\Delta CFI \leq .01$, $\Delta RMSEA \leq .015$), men ikke på T2, ifølge anbefalte retningslinjer, ved at CFI-verdien (.11) oversteg disse retningslinjene allerede i basismodellen (Chen, 2007; Cheung & Rensvold, 2002).

Tabell 2. Invariansestimater over tid.

Modell	X ²	df	Modell sammenlikning	X ² diff	df endring	CFI	ΔCFI	RMSEA (90% CI)	Δ RMSEA	SRMR
Grunnmodell T1	38.622	29				.99		.04 (.00-.07)		.02
M0 Konfigurativ	232.688	132				.97		.05 (.04-.06)		.04
M1Metrisk	245.940	138	M1-M0	13.26 n.s.	6	.97	.000	.05 (.04-.06)	.00	.06
M2 Skalering	254.362	144	M2-M1	8.42 n.s.	6	.97	.000	.06 (.04-.06)	.00	.06

Note. X² = kji-kvadrattest (robust maximum likelihood estimering MLR); CFI = sammenliknende tilpasningsindeks; RMSEA = tilnærmet kvadratrot basert feilestimat; SRMR = standardisert kvadratrotresidual; Grunnmodell (ingen invarians pålagt), M0 = likhet i antall faktorladninger per latent begrep; M1 = invariant størrelse på faktorladninger; M2 = invariant innslags-/avskjæringspunkt for de respektive indikatorene (intercept). Ikke-signifikant Δ i kji-kvadrat mellom modeller så vel som ΔCFI > .01, and Δ RMSEA > .015 indikerer invarians mellom respektive nestete modeller (Cheung & Rensvold, 2002; Chen, 2007).

Diskriminantvaliditet

Som vist i Figur 1 er korrelasjonene mellom de fire latente dimensjonene i størrelsesorden fra .56 til .90. Med unntak av korrelasjonen mellom mestringstro vedrørende *Motivasjon* og *Karakterbygging* ($r = .90$) er det ingen av de andre dimensjonene som overstiger grensekriteriet ($< .85$) for tilfredsstillende diskriminerende validitet (Kline, 2011). Fornell og Larcker (1981) beskrev imidlertid en mer stringent test for avklaring av dette. For at et begrepspar skal ha diskriminerende validitet, må gjennomsnittlig ekstrahert varians (AVE) for hver dimensjon i parene være større enn kvadratet av faktorkoeffisienten for de respektive dimensjonene. Kvadratet av faktorkoeffisienten for de respektive dimensjonene var i størrelsesorden fra .33 til .77. Bare kvadratet av koeffisienten til dimensjonene mestringsforventning om *Motivasjon* og *Karakterbygging* (.77) var større enn AVE for hver av de fire dimensjonene i Tabell 1 (område .64–.70). Oppsummert, selv ved bruk av denne mer sofistikerte tilnærmingen til måling av diskriminerende validitet, finner vi støtte for diskriminerende validitet mellom alle dimensjonene så nær som mellom mestringstro med henblikk på *Motivasjon* og *Karakterbygging*.

Kriterievaliditet

Pearsons korrelasjonsanalyse viste en positiv sammenheng mellom positiv affekt og alle de fire dimensjonene av mestringstro: *Motivasjon* ($r = .49$, ($r = .41$), $p < .001$), *Kampstrategi* ($r = .34$, ($r = .31$), $p < .001$), *Teknikk* ($r = .31$, ($r = .24$), $p < .001$), og *Karakterbygging* ($r = .46$, ($r = .38$), $p < .001$). Partiell korrelasjonsanalyse med kontroll for vitalitet i dagliglivet endret lite på korrelasjonsmønsteret (partielle korrelasjoner i parentes). Summert står de positive sammenhengene seg mellom alle de fire dimensjonene av mestringstro og positiv affekt i rollen som trener, også ved kontroll for egenrapportert følt vitalitet i dagliglivet blant trenerne.

Diskusjon

Formålet med denne studien var å undersøke måleegenskapene til den norske versjonen av skalaen til måling av mestringstro i rollen som trener

(Feltz et al., 1999; 2009; Myers et al., 2008; Sullivan et al., 2012). Vi benyttet ESEM, etterfulgt av flere CFA-er på et utvalg norske trenere i organisert breddefotball. Vi testet måleegenskaper i form av dimensjons-, konvergens-, diskriminant- og kriterievaliditet for skalaen. I tillegg undersøkte vi skalaens reliabilitet.

ESEM-analysene viste høye residualer på utsagn 2 og 3. Byrne (2012) og Kline (2011) legitimerer modifiseringer av en teoretisk basert skala ved oversettelse dersom det kan gis et rasjonale for modifiseringene. I vårt tilfelle grunngis dette på et primært empirisk grunnlag gjennom høye residualer på de to utsagnene. For øvrig er vi av den oppfatning at utsagnene (a) finne ut av motstanderlagets sterke sider i kamp, og (b) forberede spillerne mentalt på kamp, med stor sannsynlighet oppfattes som relativt marginale blant de rolleforventninger som skal mestres for en trener med spillere i aldersklassen 11–14 år i norsk breddefotball. Slike utsagn har større relevans om man tenker seg at kampresultater og prestasjonsfokus står sentralt blant trenere i dette segmentet. Lite tyder på det, snarere synes det å være motsatt. En nylig studie viste at trenere i norsk breddefotball fokuserer på å ivareta mestring og sosial inklusjon i trenergjerningen (for en oversikt, se Solstad, 2017).

De innledende ESEM-analysene som er basert på 22 utsagn, viste god tilpasning til data: (S-B $\chi^2 = [df = 149, n = 222] = 312.901, p < .001$; RMSEA = .070; 90% (CI = .059-.081); CFI = .94; SRMR = .025). Selv om tilpasningen ble noe moderert i CFA-modellen, ble den betydelig bedre ved bruk av parseller. Det er vanskelig å sammenlikne våre funn med tidligere relevante studier (se for eksempel Feltz et al., 2009), da det er forskjeller i analysetilnærminger, responsskalering og i hva som oppgis av tilpasningsindekser. Men uavhengig av dette gir våre funn støtte til at bruk av skalaen basert på parsellering synes fruktbar ved studier av treneres mestringstro. Vi fant også at tilpasningen av modellen til data var invariant over tid. Det betyr at skalaen også er egnet ved måling av endringer i mestringstro blant trenere over tid. Invarians testet mellom grupper (dvs. intervensjons og kontrollgruppe) ble imidlertid ikke oppnådd. Det betyr at skalaen i vårt tilfelle ikke kan brukes til å undersøke eventuelle effekter av tiltak hvor man prøver å påvirke treneres mestringstro.

Analysene støtter samlet opp under uttalte fordeler ved bruk av parcelleringsteknikk, ved at denne skaper mer robuste forutsetninger ved tilpasning av en målemodell til data. Little (2013) peker i den forbindelse på følgende: «*Parceling reduces both the sampling variability of the selected sample and the amount of incorrectness of my model in the population. The benefits of reducing the likelihood of correlated residuals and dual-factor loadings are both aspects of how parcels reduce sampling variability or population misfit*» (s. 24–25). Vi vil anbefale at andre som måtte ta i bruk skalaen i en eventuelt framtidig undersøkelse, sørger for å sjekke mønsteret i faktorladningene med tanke på balansert oppsett av parseller, da mønsteret i faktorladningene rimeligvis vil endre seg i studier basert på andre utvalg (Kline, 2011).

I tråd med anbefalinger (Byrne, 2012) undersøkte vi også konvergerende og diskriminerende validitet på de fire mestringstrodimensjonene. Estimaten var gjennomgående gode, og bedre enn hva som er funnet tidligere innenfor tilsvarende trenerutvalg (Feltz et al., 2009). Et klart unntak var manglende diskriminerende validitet mellom mestringstrodimensjonene: *Motivasjon* og *Karakterbygging* ($r = .90$). Feltz et al. (2009) har argumentert for at man ikke nødvendigvis vil finne at ulike aspekter ved dimensjonsvaliditeten i en skala for trenere i barne- og ungdomsidrett evner å skille like klart som blant trenere i mer elitepreget voksenidrett. Det er mulig at det å motivere til å bygge laget blant spillerne veves inn i det å bygge karakter og lagånd i moralsk forstand. Dermed vil trenerne lettere kunne komme til å anse motivering til å opptre på en respektfull måte overfor andre lags spillere og det å motivere til å bygge laget som to sider av samme sak.

Diskriminerende validitet var også marginal (.76) mellom mestingstrodimensjonene *Kampstrategi* og *Teknikk*. Én grunn til dette kan være at opplæringskulturen i norsk barne- og ungdomsfotball i stor grad vektlegger treningsformer knyttet til ulike varianter av spill, hvor teknikk og taktikk flyter over i hverandre i trenernes veiledning.

Med hensyn til kriterievaliditet tilknyttet treneres mestingstro, har to tidligere studier undersøkt dette innenfor rammen av barne- og ungdomsidrett. Feltz et al. (2009) og Sullivan et al. (2012) undersøkte mestringsdimensjonene i relasjon til spiller- og trenererfaring, samt lederskaps erfaring. Man fant at disse gav støtte til kriterievaliditeten i

studien. Slike funn kan man dra paralleller til i vår studie, hvor vi finner at en høy skår for mestringstro er positivt forbundet med positive følelsetilstander i trenerrollen. Vitalitetsnivå i dagliglivet forkludrer i marginal grad disse funnene og underbygger at slike sammenhenger ikke er spuriøse. Funnene er i tråd med Banduras sosial-kognitive teori (Bandura, 1997) og empiriske funn (Hu, Motl, McAuley, & Konopack, 2007) som viser at mestringstro utløser glede og andre positive følelsetilstander under utøvelse av idrett. Man finner tilsvarende sammenhenger innen konteksten skole, hvor mestringstro i rollen som lærer (motivere elever, samarbeide med kolleger, håndtere utfordringer osv.) er relatert til tilfredshet i læregjeringen (Skaalvik & Skaalvik, 2010). Vi analyserte også forholdet mellom mestringstro og nivå av formell trenerutdanning, og dette viste at trenere med middels eller høyt nivå av trenerutdanning skåret høyere på mestringstrodimensjonen *Teknikk* enn de med lite formell utdanning. Resultatet er i tråd med funn blant trenere i Singapore, som viste paralleller blant henholdsvis sertifiserte og ikke-sertifiserte trenere (Lee, Malet, & Feltz, 2002). Erfaring og kompetanse er gjerne sammenvevd, og Kavussanu og medarbeidere (Kavussanu et al., 2008) fant at antall år med erfaring som trener predikerte mestringstro med henblikk på å utvikle tekniske ferdigheter. Ut fra premisset om at trenere med mer erfaring og formell kompetanse påtar seg mer ansvar, viser jo også våre analyser at mestringstroen gjennomgående er høyere blant trenere som tar hovedansvar for laget, eller som deler dette med en annen person.

Styrker og begrensninger

Styrker ved denne studien inkluderer grundig analyse av dimensjonsvaliditet, diskriminerende validitet, konvergerende validitet samt kriterievaliditet ved hjelp av flere psykometriske indikatorer. Analyse av invarians over tid bidrar dessuten med kunnskap om muligheten for å ta i bruk instrumentet i måling av endringer over tid i treneres mestringstro (Byrne, 2012). Denne studien legger til grunn bruk av samme utvalg ved ESEM-analyse, innledende CFA-analyse og påfølgende bruk av parsellering. Det vil kunne være innsigelser mot en slik tilnærming,

da studier ved bruk av CFA ofte gjør bruk av et utvalg i ESEM-fasen, for så å bruke et nytt utvalg i CFA-fasen. Dette var det ikke mulighet til i vårt tilfelle, og det kan ses som en begrensning ved studien. For øvrig ble to utsagn fra den opprinnelige 24-utsagnsskalaen tatt bort (Feltz et al., 1999, 2009). Dette gjorde vi ved hjelp av ESEM-analyser i en eksplorativ fase, mens vi ikke foretok ytterligere «inngrep» i skalaen i den bekreftende fasen (CFA). Vi mener dermed at vi har et akseptabelt grunnlag for bruk av parsellering som endelig analysestrategi. For øvrig har vi empirisk og teoretisk søkt å grunngi at de to utsagnene (a) *finne ut av motstanderlagets sterke sider i kamp*, og (b) *forberede spillerne mentalt på kamp*, ble tatt ut i ESEM-analysen.

Feltz og medarbeidere (Feltz et al., 2009) anvendte en responskala med fem kategorier basert på anbefalinger fra Myers og medarbeidere (Myers, Wolfe, & Feltz, 2005). Disse observerte at bare en prosent av trenerne tok i bruk de laveste responskategoriene mellom 0 og 4. Gjennomsnittsverdier over de 22 utsagnene i vårt materiale lå mellom 6.7 og 7.8, og kun rundt 3% av trenerne tok i bruk responskategoriene mellom 0 og 4. Dette kan tyde på at man i nye studier med skalaen bør kunne ta i bruk en responskala med syv kategorier (Zimet, Dahlem, Zimet, & Farley, 1988).

Konklusjon

Studien viser at en 22-utsagnsbasert, firedimensjonal målemodell av treneres mestringstro i rollen som trener i breddefotballen gir god tilpasning til data. Parselleringstilnærmingen gav betydelig støtte til god dimensjonsvaliditet for de fire parsellbaserte dimensjonene. Invarians over tid ble bekreftet for intervensjons- og kontrollgruppen samlet. Invarians mellom intervensjons- og kontrollgruppe over tid ble ikke oppnådd. Vi observerte god konvergerende og diskriminerende validitet. Kriterievaliditet ble også bekreftet. I sum synes skalaen å være velegnet til undersøkelse av tverrsnitts- og lengdesnittstudier av mestringstro som trener i barne- og ungdomsidrett. Framtidige studier må bekrefte invarians mellom undergrupper i et materiale. Dette kan for eksempel være grupper som har fått, eventuelt ikke fått et tiltak, før disse sammenliknes på relevante utfallsvariabler.

Referanser

- Anderson, J.C. & Gerbing, D.W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411–423.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bagozzi, R.P. & Edwards, J.R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, 1, 45–87.
- Benight, C.C. & Bandura, A. (2004). Social cognitive theory of posttraumatic recovery: the role of perceived self-efficacy. *Behaviour Research and Therapy*, 42, 1129–1148.
- Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246.
- Bentler, P.M. (1995). *EQS structural equations programs manual*. Encino, CA: Multivariate Software, Inc.
- Browne, M.W. & Cudeck, R. (1993). Alternative ways of assessing model fit. I: K.A. Bollen & J.S. Long (red.), *Testing structural equation models* (s. 136–162). Thousand Oaks, CA: Sage.
- Byrne, B.M. (2012). *Structural equation modeling with Mplus. Basic concepts, applications, and programming*. NY: Routledge.
- Campbell, T. & Sullivan, P.J. (2005). The effect of a standardized coaching education program on the efficacy of novice coaches. *Avante*, 11, 38–45.
- Cheung, G.W. & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255.
- Chen, F.F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 464–504.
- Cole, D.A. (1987). Utility of confirmatory factor analysis in test validation research. *Consulting and Clinical Psychology*, 55, 584–594.
- Dimitrov, D.M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43, 121–149.
- Dong, Y. & Peng, C.-Y.J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2, 1–17.
- Duda, J.L. & Nicholls, J.G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology*, 84, 290–299.
- Epstein, J., Santo, R.M., & Guillemain, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out consensus. *Journal of Clinical Epidemiology*, 68, 435–441.

- Feltz, D.L., Chase, M.A., Moritz, S.E., & Sullivan, P.J. (1999). Development of the multi-dimensional coaching efficacy scale. *Journal of Educational Psychology, 91*, 765-776.
- Feltz, D.L., Short, S.E., & Sullivan, P.J. (2008). *Self-efficacy in sport*. Champaign, IL: Human Kinetics.
- Feltz, D.L., Hepler, T.J., Roman, N., & Paiement, C.A. (2009). Coaching efficacy and youth sport coaches. *The Sport Psychologist, 23*, 24-41.
- Fornell, C., Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39-50.
- Geiser, C. (2013). *Data analysis with Mplus*. New York, The Guilford Press.
- Hollebeak, J. & Amorose, A.J. (2005). Perceived coaching behaviors and college athletes' intrinsic motivation: A test of self-determination theory. *Journal of Applied Sport Psychology, 17*, 20-36.
- Horn, T.S. (2002). Coaching effectiveness in the sport domain. I T.S. Horn (red.), *Advances sport psychology* (s. 309-354). Champaign, IL: Human Kinetics.
- Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.
- Hu, L., Motl, R.W., McAuley, E., & Konopack, J.F. (2007). Effects of self-efficacy on physical activity enjoyment in college-aged women. *International Journal of Behavioural Medicine, 14*, 92-96.
- Jerome, G.J., Marquez, D.X., McAuley, E., Canaclisova, S., Snook, E., & Vickers, M. (2002). Self-efficacy effects on feeling states in women. *International Journal of Behavioral Medicine, 9*, 139-154.
- Jourden, F.J., Bandura, A., & Banfield, J.T. (1991). The impact of conceptions of ability on self-regulatory factors and motor skill acquisition. *Journal of Sport & Exercise Psychology, 8*, 213-226.
- Kavussanu, M., Boardley, I.D., Jutkiewicz, N., Vincent, S., & Ring, C. (2008). Coaching efficacy and coaching effectiveness: examining their predictors and comparing coaches' and athletes' reports. *The Sport Psychologist, 22*, 383-404.
- Kline, R.B. (2011). *Principles and practice of structural equation modeling*. NY: The Guilford Press.
- Lee, K.S., Malet, L., & Feltz, D.L. (2002). The strength of coaching efficacy between certified and noncertified Singapore coaches. *International Journal of Applied Sport Sciences, 14*, 55-67.
- Little, T.D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.

- Little, T.D., Jorgensen, T.D., Lang, K.M., Whitney G., & Moore, E. (2014). On the joys of missing data. *Journal of Pediatric Psychology*, 39, 151–162.
- Little, T.D., Rhemtulla, M., Gibson, K., & Schoemann, A.M. (2013). Why the item versus parcels controversy needn't be one. *Psychological Methods*, 18, 285–300.
- Malete, L. & Feltz, D. (2000). The effect of a coaching education program on coaching efficacy. *The Sport psychologist*, 14, 410–417.
- Malete, L., Chow, G.M., & Feltz, D. (2013). Influence of coaching efficacy and coaching competency on athlete-level moral variables in Botswana youth soccer. *Journal of Applied Social Psychology*, 43, 2107–2119.
- Mageau, G.A., & Vallerand, R.J. (2003). The coach-athlete relationship: A motivational model. *Journal of Sport Sciences*, 21, 883–904.
- McAuley, E., Talbot, H.M., & Martinez, S. (1999). Manipulating self-efficacy in the exercise environment in women: Influences on affective responses. *Health Psychology*, 18, 288–294.
- Marsh, H.W., Wen, Z., & Hau, K.-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, 9, 275–300.
- Myers, N.D., Feltz, D.L., & Wolfe, E.W. (2008). A confirmatory study of rating scale category effectiveness for the coaching efficacy scale. *Research Quarterly for Exercise and Sport*, 79, 300–311.
- Myers, N.D., Chase, M.A., Pierce, S.W., & Martin, E. (2011). Coaching efficacy and exploratory structural equation modeling: a substantive-methodological synergy. *Journal of Sport & Exercise Psychology*, 33, 779–806.
- Pelletier, L.G., Fortier, M.S., Vallerand, R.J., & Briere, N.M. (2001). Associations among perceived autonomy support, forms of self-regulation, and persistence: A prospective study. *Motivation and Emotion*, 25, 279–306.
- Quested, E., Ntoumanis, N., Viladrich, C., Haug, E., Ommundsen, Y., Van Hoye, A., & Duda, J.L. (2013). Intentions to drop-out of youth soccer: a test of the basic needs theory among European youth from five countries. *International Journal of Sport and Exercise Psychology*, 11, 395–407.
- Ryan, R.M. & Frederick, C.M. (1997). On energy, personality and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality*, 65, 529–565.
- Satorra, A. & Bentler, P.M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507–512.
- Segal, M.H., Lonner, W.J., & Berry, J.W. (1998). Cross-cultural psychology as a scholarly discipline. *American Psychologist*, 53, 1101–1110.
- Skaalvik, E. & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, 26, 1059–1069.

- Solstad, B.E. (2017). *Towards a better understanding of the dynamics of sports coaching at the youth level: The coach's perspective. A study of the Youth Football Coaches participation in the Norwegian Arm of the promoting Adolescent Physical Activity (PAPA) project*. Oslo: Norges idrettshøgskole. Doktorgradsavhandling.
- Solstad, B.E., Van Høye, A., & Ommundsen, Y. (2015). Social-contextual and intrapersonal antecedents of coaches' basic need satisfaction: The intervening variable effect of providing autonomy-supportive coaching. *Psychology of Sport and Exercise*, 20, 84–93.
- Su, C.T. & Parham, L.D. (2002). Case report: Generating a valid questionnaire translation for cross-cultural use. *American Journal of Occupational Therapy*, 56, 581–585.
- Sullivan, P. & Kent, A. (2003). Coaching efficacy as a predictor of leadership style in intercollegiate athletics. *Journal of Applied Sport Psychology*, 15, 1–11.
- Sullivan, P., Paquette, K.J., Holt, N.L., & Bloom, G.A. (2012). The relation of coaching context and coach education to coaching efficacy and perceived leadership behaviors in youth sport. *The Sport Psychologist*, 26, 122–134.
- Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070.
- Wiersma, L.D. & Sherman, C.P. (2005). Volunteer youth sport coaches' perspectives of coaching education/certification and parental codes of conduct. *Research Quarterly for Exercise and Sport*, 76, 324–338.
- Yang, Y. & Green, S.B. (2010). A note on structural equation modeling estimates of reliability. *Structural Equation Modeling: A Multidisciplinary Journal*, 17, 66–81.
- Zimet, G.D., Dahlem, N.W., Zimet, S.G., & Farley, G.K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52, 30–41.

Vedlegg. Treneres mestringsforventning, norsk versjon.

Hvor stor tro har du på at du kan:

1. (M) bidra til at spillere bevarer troen på seg selv
 2. (KS#) finne ut av motstanderlagets sterke sider i kamp
 3. (M#) forberede spillerne mentalt på kamp
 4. (KS) anvende ulik taktikk i kamp
 5. (K) oppmuntre til gode moralske holdninger
 6. (M) bygge et positivt selvbilde hos spillerne
 7. (T) demonstrere ferdigheter/øvelser innen fotball
 8. (KS) endre eget spill / egen taktikk slik at det/den tilpasses ulike kampsituasjoner
 9. (KS) finne ut av motstanderes svakheter i kampsituasjoner
 10. (M) motivere spillerne
 11. (KS) ta viktige avgjørelser under kamp
 12. (M) utvikle/bygge samhold i laget
 13. (K) fremme en fair play-holdning hos spillerne
 14. (T) veilede spillerne når det gjelder teknikk/ ferdigheter
 15. (M) utvikle spillernes tro på seg selv
 16. (T) utvikle spillernes talent
 17. (KS) fremme spillernes/lagets sterke sider under kamp
 18. (T) identifisere spillernes talent
 19. (K) fremme god sportsånd
 20. (T) oppdage tekniske/ferdighetsmessige feil
 21. (KS) tilpasse kampstrategier med utgangspunkt i lagets ferdighetsnivå
 22. (T) lære bort ferdigheter i fotball
 23. (M) bygge lagets tro på seg selv
 24. (K) fremme en holdning om å respektere hverandre
-

Note. #utsagn 2 og 3 ble tatt ut grunnet høy residual varians.

M = Motivasjon; KS = Kampstrategi, T = Teknikk; K = Karakterbygging.