

Kunnskapsoppsummering og klassifisering av tiltaket:

De Utrolige Årenes (DUÅs) skole- og barnehageprogram (1. utg.)

DUÅs skole- og barnehageprogram er et universalforebyggende opplæringstiltak for personalgrupper i barnehager, skoler og SFO/AKS. Hovedmålsettingen er målrettet arbeid med barns (3-8 år) psykisk helse, med særlig fokus på forebygging og håndtering av atferdsproblemer.

BAKGRUNN

Artikkelen er en kunnskapsoppsummering om effektene av det norske DUÅs skole- og barnehageprogram som bygger på Webster-Strattons indikerte program IY-TCM (Incredible Years Teacher Classroom Management Training). Den universelle versjonen tilbys personell (maks 25 personer pr. gruppe) som jobber med barn i 3-8 års alder i skoler, fritidsordning og barnehager, hovedsakelig i kommuner som har implementert DUÅs foreldretreningsprogram.

Skole- og barnehageprogrammets mål er en felles faglig plattform for målrettet arbeid med barns psykisk helse og forebygging av atferdsproblemer. Programopplæringen ledes av to gruppeledere over ca. ett skole-/barnehageår, og innbefatter seks opplæringsdager med påfølgende veiledning. Den strukturerte opplæringen fokuserer på emner som: 1) positiv relasjonsbygging, 2) ros, oppmuntring, oppmerksomhet, veiledning 3) belønningssystemer, 4) reduksjon av negativ atferd, 5) naturlige/arrangerte konsekvenser, og 6) emosjonell regulering og problemløsning. Oppfølgingsveiledning gis. Programeier er Regionalt kunnskapssenter for barn og unge Nord (RKBU Nord), UiT Norges arktiske universitet (Tromsø).

METODE

Kunnskapsoppsummeringen bygger på et systematisk litteratursøk i de nasjonale og internasjonale databasene Embase, Medline, Psykinfo, NORART, Cochrane, Cristin, NORA, SCOPUS, SweMed, Bibsys og Google Scholar. Det ble også innhentet informasjon om tiltaket fra tiltakseier.

RESULTATER

Resultatene inkluderer en vurdering av tiltakets beskrivelse, foreliggende effektstudier, forskningsmetodisk kvalitet og implementeringskvalitet. Programmet har et solid teoretisk fundament og det gis en god beskrivelse av programmets innholdskomponenter og opplæringsstrategier. To norske kvasi-eksperimentelle studier basert på lærervurderinger, publisert i fire artikler, ble inkludert i kunnskapsoppsummeringen. Resultatene indikerer at programmet kan redusere problematferd og øke sosial kompetanse hos barn i 3-8 års alderen, samt bidra til positive endringer i elev-lærer relasjoner og foreldres involvering i barnas skolegang. Med utgangspunkt i totalutvalget synes kortids-effektene imidlertid å være små til marginale. Resultatene tyder på at programmet har en større effekt på reduksjon av eksternaliseringsproblemer og elev-lærer konflikter for høyrisikobarn. Innenfor denne gruppen ble det i tillegg funnet en moderat effekt på skolefaglige prestasjoner. Det knytter seg noe usikkerhet til evalueringresultatene pålitelighet (særlig indre validitet). Per dags dato bør ikke programmet omtales som skoleomfattende, da det kun implementeres på 1. – 3. klassetrinn. Det er behov for en implementeringsmanual for skoler og barnehager og bedre kvalitetskontroll.

KONKLUSJON

DUÅs skole- og barnehageprogram klassifiseres på evidensnivå 4 – Tiltak med tilfredsstillende dokumentasjon på effekt.

Innledning

Bakgrunn

Grunnskolelærere står daglig overfor store utfordringer med å tilrettelegge undervisningsinnhold,

egen praksis og læringsmiljøet i klassen slik at det/den i størst mulig grad imøtekommer den enkelte



Mari-Anne Sørli

m.a.sorlie@nubu.no


NUBU Nasjonalt kunnskaps-
utviklingssenter for barn
og unge AS



Kyrre Breivik

kybr@norceresearch.no

RKBU Vest – Regionalt kunns-
kapssenter for barn og unge

 Engelsk sammendrag på
nettsiden

ISSN 2464-2142

Tidsskriftet Ungsinn utgis av Re-
gionalt kunnskapssenter for
barn og unge- nord (RKBU
Nord) ved UiT Norges arktiske
universitet.



elevs skolefaglige og sosial-emosjonelle forutsetninger og behov. Det er en klar sammenheng mellom elevenes skolefaglige prestasjoner, sosiale ferdigheter og skoleatferd (Epstein, Atkins, Cullinan, Kutash & Weaver, 2008). For eksempel vet man at barn som begynner på skolen med svake sosiale ferdigheter og/eller atferdsproblemer er i betydelig risiko for vedvarende dårlig tilpasning, for ikke å få gode venner og for liten deltakelse i klasseromsundervisningen, noe som kan lede til svake skoleprestasjoner både på kort og lengre sikt (eks. Raver & Knitze, 2002). Barn som tidlig viser atferdsproblemer er også i betydelig risiko for en rekke andre negative livsutfall, slik som å droppe ut av skolen, kriminalitet, rusmisbruk, arbeidsledighet, angst og depresjon (Skogen & Torsvik, 2013). Nordahl og kolleger (2009) oppsummerte norske og danske skoleundersøkelser og fant at 7-12% av elever i alderen 10-17 år har et så stort omfang av uønsket atferd at det er rimelig å omtale det som atferdsproblemer. Mange læringsmuligheter og verdifull undervisningstid har gått tapt på grunn av forstyrrende og regelbrytende elevatferd (Anderson & Kincaid, 2005; Walker, Ramsey & Gresham, 2004) og slik atferd synes å være en viktig kilde til stress og utbrenthet hos lærere (Aloe, Shisler, Norris, Nickerson & Rinker, 2014). Et utvalg norske rektorer rangerte da også elever med alvorlige atferdsproblemer som den aller største utfordringen i dagens skole (Sørli & Ogden, 2014).

Flere studier har vist at effektiv klasseledelse (Oliver, Wehby & Reschly, 2011; Simonsen, Fairbanks, Briesch, Myers & Sugai, 2008) samt sosial ferdighetstrening i skole og barnehage (Durlak, Weissberg, Dymnicki, Taylor & Schellinger, 2011; Webster-Stratton m.fl., 2008) kan ha en positiv effekt på barns atferdsproblemer.

I tråd med dette, har flere skolebaserte intervensjoner rettet mot spesifikke barn med eller i høy risiko for å utvikle alvorlige atferdsproblemer (indikerte tiltak) vist seg effektive både m.h.t. redusert eksternalisert og internalisert problematferd og fremming av sosiale ferdigheter (Stoltz, van Londen, Dekovic, Orobio de Castro & Prinzie, 2012). Men også vel integrerte universelle intervensjoner (dvs. primærforebyggende tiltak rettet mot alle barn i en bestemt kontekst eller gruppe) antas å kunne være effektive for høyrisikoelever, selv om evidensen foreløpig er relativt begrenset (Wolpert, Humphrey, Belsky & Deighton, 2013). Gjennomsnittseffekten av universelle skolebaserte intervensjoner på eksternalisert problematferd rapportert som relativt små (Cohens d = cirka ,20) i metaanalyser (Durlak m.fl., 2011; Wilson & Lipsey, 2007). Effektene av skolebaserte intervensjoner for høyrisikoelever (indikerte tiltak) rapporteres gjerne å være noe høyere, og da særlig for yngre barn (Durlak m. fl., 2011; Wilson & Lipsey, 2007). I Norge som i de fleste europeiske land, har skoler og barnehager hittil bare i beskjeden

grad blitt oppmuntret til å benytte slike forhåndsstrukturerte og forskningsstøttede intervensjonsprogram, det være seg både av universell og indikert karakter.

De Utrolige Årenes (DUÅ) skole- og barnehageprogram er imidlertid et sjeldent eksempel på en universell intervensjon med fokus på forebygging av atferdsproblemer og fremming av sosiale ferdigheter hos barn i 3-8 års alderen. Regionalt kunnskaps-senter for barn og unge (RKBU Nord, UiT Norges arktiske universitet) er programtilbyder, i samarbeid med RKBU Vest (NORCE), RKBU Midt-Norge (Norges teknisk-vitenskapelig universitet, NTNU) og RBUP Øst og Sør.

Målsetninger og målgruppe

Den primære målsettingen med DUÅs skole- og barnehageprogram er å fremme prososial atferd og forebygge atferdsproblemer blant yngre barn (3-8 år). Programmet består av 42 timer ferdighetsrettet opplæring for personalgrupper i barnehager, SFO/AKS og for lærere på småskoletrinnet (1-3 klasse) med påfølgende mellomveiledning. Opplæringstilbudet har til hensikt å utvikle en felles faglig plattform m.h.t. til målrettet arbeid for barns psykiske helse og god klasse-/gruppeledelse slik at personalet lettere kan skape positive relasjoner til barna og sikre dem en gunstig læringsposisjon samt forebygge og håndtere atferdsproblemer. På DUÅs hjemmeside er målsettingen nærmere spesifisert som «...å utvikle god klasse- eller gruppeledelse og derved sette ansatte i skolen/SFO/AKS/barnehagen i stand til å skape positive relasjoner til barn, bringe barn i læringsposisjon, forebygge uro og problemer og håndtere atferdsproblemer, der det forekommer» (<http://dua.uit.no/1-skole-og-barnehage/>).

Programmet beskrives videre som skoleomfattende, dvs. at hele personalgruppen i grunnskoler (1-10 trinn) kan tilbys programopplæring, forutsatt tilpasning for henholdsvis 4-7 trinn og 8-10 trinn. Nærmere informasjon om hvilke tilpasninger som er nødvendige for bruk på 4-10 trinn foreligger imidlertid ikke.

Beskrivelse og materiell

DUÅs universalforebyggende skole- og barnehageprogram er en norsk tilpasning av Carolyn Webster-Stratton's IY-TCM program (Incredible Years -Teacher Classroom Management) – en indikert, vel utprøvd og evaluert intervensjon med høyrisikobarn som målgruppe (Webster-Stratton, 2003; Webster-Stratton & Reid, 2002). DUÅs skole- og barnehageprogram gis som et lokalt organisert opplæringstilbud for grupper av småskole-, SFO- eller barnehageansatte (maks 25 deltakere), og ledes av to gruppeledere (primært sertifiserte). Programmet er manual-

basert, dvs. at opplæringen følger en forhåndsstrukturert plan der bestemte temaer fokuseres i en fastlagt rekkefølge. Programopplæringen strekker seg over seks dager á 7 timer i løpet av et ½ - 1 skoleår (ca. 1 dag pr. måned). Til nå har ca. 500 fagpersoner gjennomgått opplæringen og ca. 70 av disse er sertifiserte.

DUÅs skole- og barnehageprogram omfatter en håndbok for gruppeledere, to lærebøker for gruppeledere/veiledere; «*Hvordan fremme sosial og emosjonell kompetanse hos barn*» (Webster-Stratton, 2005) og «*Utrolige lærere*» (Webster-Stratton, 2018), et videosett med 7 DVDer og ulike plakater (f.eks. «Læringspyramiden») og laminerte kort for oppheng/bruk i klasserommet/SFO/barnehagen. Håndboken med tilhørende korte videovignetter er delt inn i seks kapitler, ett for hver innholdskomponent (omtalt som delprogram) som suksessivt gjennomgås på kursdagene:

- 1) relasjonsbygging via proaktive ansatte,
- 2) oppmerksomhet, ros, oppmuntring og veiledning,
- 3) belønnings- og motivasjonssystemer,
- 4) reduksjon av negativ atferd ved bruk av ignorering og avledning,
- 5) reduksjon av negativ atferd ved bruk av naturlige og arrangerte konsekvenser,
- 6) emosjonell regulering, sosiale ferdigheter og problemløsning.

Det gis en detaljert beskrivelse av hva hver komponent og opplæringsdag (omtalt som workshops) skal omhandle, hvordan gruppelederen praktisk skal gjennomføre opplæringen og i hvilken rekkefølge de ulike temaene skal presenteres. For hver komponent medfølger skriftlig materiell til utdeling blant kursdeltakerne. Før utgivelsen av «*Utrolige lærere*» (Webster-Stratton, 2018), inkluderte utdelt materiale et oversatt kapittel om veiledning (coaching) fra boken «*Incredible teachers*» (Webster-Stratton, 2012). Antall vignetter pr. komponent varierer. For eksempel tematiseres og eksemplifiseres «*Betydningen av oppmerksomhet, veiledning og ros fra den voksne*» gjennom 56 vignetter og fire tilleggs-vignetter.

Gjennomføring av tiltaket og metoder som anvendes

I forkant av opplæringen skal det foreligge en spesifisert samarbeidsavtale mellom den aktuelle enhetsleder i skole eller barnehage og enheten som er ansvarlig for opplæringen. Enhetsleder må bl.a. forplikte seg til selv å delta aktivt i programopplæringen. For at de ansatte skal få opplæring i programmet, stilles det ingen andre krav enn at de

må være tilsatte medarbeidere i skole, SFO eller barnehage. I følge programeier er det nå satt et krav om at 80% av deltakerne samtykker i at opplæringen skal gis i deres enhet.

De seks opplæringsdagene følger i hovedsak samme struktur: (1) Introduksjon. Her beskrives det aktuelle temaets (dvs. delprogram/innholdskomponent) relevans og hvilke nøkkeltemaer som skal belyses (f.eks. «Føler ikke de andre barna seg utelatt når bare ett barn får ros?», «Noen barn ser ikke ut til å reagere på ros – hva er grunnen til dette»). For hvert spørsmål er det utarbeidet en tekst som beskriver hva gruppeleder skal vektlegge, (2) Kort introduksjon til den enkelte vignett, (3) Gruppediskusjon, inklusive spørsmål for å sette i gang diskusjonen, (d) Betrachtinger, (e) Rollespill/øvelse og (f) Oppsummering. Den enkelte sesjon avsluttes med en (g) Evaluering der hver deltaker fyller ut et spørreskjema. I forkant av hver sesjon henstilles deltakerne også om til å fylle ut en sjekklister til eget bruk.

Mellom opplæringsdagene gis det øvingsoppgaver og gruppearbeid, i tillegg til fem gruppebaserte veiledninger (gitt av gruppelederne), for å modellere og hjelpe deltakerne med å generalisere de nye ferdighetene, og for å bistå dem i utarbeidelsen av en utviklingsplan (atferdsplan) tilpasset det enkelte klasserom/barnegruppe. På denne måten søker man å gi rom både for kulturell tilpasning og tilpasning til deltakernes behov og ferdighetsnivå (Webster-Stratton, 2012; Webster-Stratton, 2018). Videre gis det oppfølgingsveiledning syv ganger (á 1/2 dag) fordelt over en to-års periode etter fullført programopplæring. Det er utviklet en egen manual som beskriver innholdet i mellom- og oppfølgingsveiledningene. Det understrekes at oppfølgingsveiledningen må planlegges og skreddersys til situasjonen i den enkelte skole/barnehage.

I tråd med tidligere forskning (f.eks. Knight, 2010) benyttes flere samarbeidsorienterte og aktive innlæringsmetoder, deriblant gruppediskusjoner, brainstorming og erfaringsutveksling i smågrupper, selvrefleksjon og eksperimentell læring via strukturerte rollespill. Også praktisk problemløsning og spesifikke ferdighetstreningmetoder benyttes for å lære deltakerne effektiv gruppe-/klasseledelse, selvregulering og stressmestring. I tillegg legges det betydelig vekt på bruk av visualisering og eksemplifisering i form av videovignetter. I noen grad benyttes også tradisjonell kunnskapstegnelse via lesing av utvalgt fagstoff.

Gruppelederne skal i opplæringen basere seg på følgende prinsipper: a) vektlegge samarbeid og relasjonsbygging, b) kursdeltakerne skal sette seg mål og vurdere egen praksis, c) bruke «Læringspyramiden» som et veikart i formidlingen av programinnhold (se Figur 1), d) bygge opp deltakernes tillit og tro på egen mestring, e) legg vekt både på kognisjon,

emosjoner og atferd, f) bruke eksperimentelle og selvrefleksive læringsmetoder og g) kontekstualisere læringsprosessen (Webster-Stratton & Reid, 2002).

Det teoretiske grunnlaget for metoden

Programserien De Utrolige Årene, herunder skole- og barnehageprogrammet, er basert på et multi-teoretisk fundament som blant annet inkluderer anvendt atferdsanalyse, utviklingspsykologi, relasjons- og tilknytningsteorier (f.eks. Ainsworth, 1989; Baer, Wolf & Risley, 1969; Bowlby, 1982), sosial-kognitive læringsteorier, som Pattersons teori om aggresjonsutvikling hos barn og det tvingende samspillet betydning (coercive process) (Patterson, Reid & Dishion, 1992) og Piagets teori om interaktive læringsmetoder og kognitiv utvikling (Piaget & Inhelder, 1962). Programserien støtter seg også på Banduras teorier om modell-læring og opplevd mestringskompetanse (self-efficacy) (Bandura, 1977). DUÅs teorigrunnlag er nærmere beskrevet i flere bøker og fagartikler (se f.eks. Webster-Stratton, 2007;2011; Løndal & Aasheim, 2016).

Den spesifikke tilnærmingen i skole- og barnehageprogrammet er for øvrig bygget rundt konseptet «Læringspyramiden» (Webster-Stratton, 2006). Læringspyramidens «byggesteiner» fremkommer i Figur 1 og visualiserer en modell som legger vekt på å etablere støttende relasjoner med barn ved raus bruk av positive virkemidler i undervisningen (nederste del) i kombinasjon med mer forsiktig bruk av atferdsregulerende virkemidler (øverste del) både for å nå afterdsmessige og skolefaglige mål.

Figur 1. Læringspyramiden



Implementeringsstrategier, kvalitetssikring og kostnader

Regionalt kunnskapssenter for barn og unge Nord (RKBU Nord ved UiT Norges arktiske universitet) har hovedansvaret for opplæring i og implementering av alle DUÅ-programmene i Norge, med RKBU Midt-Norge (Norges Teknisk-naturvitenskapelig universitet, NTNU), RKBU Vest (NORCE) og RBUP Øst og Sør som nære samarbeidspartnere. Det er etablert et sentralt lønnet opplærings- og mentorkorps ved RKBUEne/RBUP, som er ansvarlige for opplæring og sertifisering av lokale/regionale DUÅ-veiledere og gruppeledere i skole- og barnehageprogrammet. DUÅ finansieres av Helsedirektoratet.

Gruppeledere til skole- og barnehageprogrammet rekrutteres lokalt fra PP-tjenesten eller andre pedagogiske støtteteam (f.eks. ambulerende team) i kommuner som fortrinnsvis tidligere har implementert DUÅs foreldretreningsprogram (eget søknads-skjema), hvor satsningen på programmet er politisk/organisatorisk forankret, og hvor det foreligger en skriftlig forpliktende avtale med DUÅ sentralt (RKBU/RBUP). En forutsetning for at opplæring av gruppeledere iverksettes er at kommunen stiller med minst to fagpersoner med minimum fullført bachelorgrad eller minimum 80 studiepoeng i pedagogiske-, helse-, eller sosialfag, samt stillingsprosent på minimum 2 x 30 % i en relevant tjeneste i kommunen. I følge programeier anbefales nå nye enheter om å sende minst tre fagpersoner for å forebygge sårbarhet i tilbudet. Det arrangeres opplæring av nye gruppeveiledere minst én gang i året.

Det er ikke et krav at gruppelederne må være sertifiserte for å kunne tilby opplæring, men det oppfordres til at flest mulig søker å oppnå dette i løpet av en to-års periode etter oppstartsoplæring. Gruppeledere må selv ha gjennomgått opplæring i skole- og barnehageprogrammet for å kunne tilby opplæringen. For å bli sertifiserte må de i tillegg ha ledet minst to komplette opplæringsforløp for skole-, SFO- eller barnehageansatte. Disse må videotapes og sendes inn for vurdering av en sertifisert mentor ved DUÅ sentralt (RKBU/RBUP). De må også gå i jevnlig veiledning hos en mentor. Mentorer er sertifiserte/akkrediterte DUÅ-gruppeledere og kollega-veiledere som har bred erfaring med henholdsvis foreldretreningsprogrammet, skole- og barnehageprogrammet eller barneprogrammet og som har blitt invitert til å ta tilleggsopplæring for å bli mentor innen ett spesifikt program. Mentors veiledning baserer seg på gruppeledernes video-opptak, sjekklister og evalueringer fra gruppeopplæringen av skole-, SFO- og barnehageansatte. Etter en eventuell godkjenning som gruppeleder (sertifisering) oppfordres han/hun til å «fra tid til annen» å sende inn en video til mentor for skriftlig tilbakemelding, dette

for å unngå avdrift fra programmet. I henhold til programutviklers sertifiseringsstandard anbefales det at gruppeledere bør lede en gruppe og gjennomgå en fidelitetsjekk minst én gang i løpet av en 18-måneders periode (<http://www.incredibleyears.com/certification-gl/>).

Frem til 2014 ble skole- og barnehageprogrammet hovedsakelig implementert i enheter som deltok i den første effektevalueringen i Norge, eller i piloteringen i forkant av denne. En forutsetning for at de aktuelle skolene og barnehagene har fått tilgang til programmet har fortrinnsvis vært at kommunen de ligger i først har inngått en skriftlig forpliktende avtale om implementering av DUÅs foreldretreningsprogram. Kommunen har i neste omgang sendt inn en egen søknad om opplæring av veiledere (gruppeledere) og fagpersonell i skole- og barnehageprogrammet. Fagpersonell i ca. 50 av de ca. 140 kommunene som foreløpig har implementert DUÅs foreldretreningsprogram har fått opplæring i skole- og barnehageprogrammet. Videre spredning av skole- og barnehageprogrammet avhenger av evalueringsresultatene, fremtidig ressurstilgang og etterspørsel.

Ifølge programutvikler, er utvikling og sikring av tilstrekkelig implementeringskvalitet (fidelitet) og spredningsstøttende mekanismer innebygd i selve programdesignet i form av det standardiserte opplæringsprogrammet, forskningsstøttede og prinsipstyrte innlæringsmetoder med særlig vekt på samarbeid, gruppeformatet, positiv støtte og praktiske øvelser både under og mellom opplæringsdagene, samt i form av individuelt og kontekstuell tilpasset veiledning over tid (Webster-Stratton, Reinke, Herman & Newcomer, 2011). Antagelsen er altså at de implisitte implementerings- og spredningsprinsippene er tilstrekkelige for å sikre høy implementeringskvalitet både på kort og lang sikt.

Implementeringskvalitet refererer til *hvor godt* svar det er mellom hvordan en intervensjon faktisk ble gjennomført, og hvordan den opprinnelig er beskrevet og testet gjennom forskning, uten brudd med mål, retningslinjer og underliggende teori (Domitrovich m.fl., 2008). Implementeringskvaliteten er viktig, fordi den i stor grad er bestemmende for hvor godt utbytte mottakerne får av intervensjonen. En rekke enkeltstudier og metaanalyser har vist at det oppnås størst effekter når et program implementeres i nær overensstemmelse med den idealtypiske modellen (dvs. høy implementeringskvalitet) (Sørлие, Ogden, Solholm & Olseth, 2010). Dette betyr imidlertid ikke at implementeringen må være perfekt for å oppnå positive resultater. Når programmer implementeres i ordinær praksis (f.eks. vanlige skoler) vil det ofte være behov for noe lokal tilpasning, men uten å miste formålet og kjernen av syne. For å unngå at det går ut over implementeringskvaliteten (troskap mot programmet), er det viktig at de justeringene som foretas ikke

må være store (f.eks. utelate en kjernekomponent) eller tilfeldige, men være godt begrunnede, planlagte og ideelt sett avklart med programutvikler på forhånd.

De skolene og barnehagene som hittil har deltatt i programmet har fått dekket det meste av sine utgifter til materiell og opplæring (manual, videovignetter, lærebøker, workshops, veiledning), inklusive en avgift på drøyt 200 kroner pr. gruppeleder for å bli registrert som gruppelederkandidat hos The Incredible Years i USA og en sertifiseringsavgift på drøyt 2000 kroner pr. gruppeleder. Kommunene har selv dekket lønn til gruppelederne i tillegg til utgifter til forfriskninger på opplæringsdagene og videoutstyr til opptak av opplæringen. De har også betalt kopieringsutgifter og øvrig programmaterieell som leker, flipp-over, klistremerker og opphengsmagneter. Kommunene har likeledes dekket gruppeledernes reise- og oppholdsutgifter i forbindelse med opplæring og veiledning samt stilt med egnede grupperom med opptaksmuligheter og andre nødvendige fasiliteter.

Problemstilling for artikkelen

Formålet med denne artikkelen er å vurdere om DUÅs skole- og barnehageprogram kan antas å være virksomt benyttet i vanlig praksis i Norge.

Metode

Litteratursøk etter studier om tiltaket ble gjennomført i databasene Embase, Medline og Psykinfo, NORRART, Cochrane, Cristin, NORA, SCOPUS, Swe-Med, Cambell, NREPP, NICE, Blueprint og CEBC (søkedato: 23.02.16). I tillegg ble det gjort søk i Google Scholar og Bibsys (søkedatoer: 23.11.17, 02.02.18). For øvrig ble tiltakseier i Norge kontaktet og bedt om å sende inn dokumentasjon som ikke ble fanget opp av søket, som for eksempel manualer, veiledere, upublisert forskningsdokumentasjon, implementeringsstrategier og kvalitetssikringsrutiner. Aktuelle forskere involvert i evalueringen av programmet ble dessuten kontaktet for nærmere avklaring.

To forskere gjennomgikk forskningsdokumentasjonen uavhengig av hverandre og identifiserte den aktuelle litteraturen for nordiske effektstudier, internasjonale kunnskapsoppsummeringer, samt andre norske evalueringer i henhold til kriteriene for klassifisering av tiltak i Ungsinn (Martinussen m.fl., 2016).

Resultater

Resultater fra litteratursøk

Litteratursøkene resulterte i 159 treff hvorav totalt tre artikler ble inkludert i denne kunnskapsoppsummeringen. De inkluderte artiklene formidlet funn fra to norske effektstudier; en barnehagestudie og en skolestudie (Aasheim, Reedtz, Handegård, Martiniussen & Mørch, 2018a; Fossum, Handegård & Drugli, 2017; Kirkhaug, Drugli, Handegård, Lydersen, Åsheim & Fossum, 2016). En ny resultatartikkel fra skolestudien er publisert etter litteratursøket og er også inkludert i vurderingen (Aasheim, Drugli, Reedtz, Handegård & Martinussen, 2018b). De fire resultatartiklene er publisert i internasjonale eller skandinaviske tidsskrift med refereedordning.

De resterende 156 publikasjonene fra litteratursøket ble ekskludert fordi de: a) ikke omhandlet DUÅs universelle skole- og barnehageprogram, b) omhandlet originalprogrammet IY-TMC eller c) omhandlet andre varianter av IY-TMC utprøvd i Wales og Iran. I tillegg ble d) studentarbeider, e) et bokkapittel med preliminare funn fra skolestudien og f) duplikater ekskludert.

Gjennomgang av beskrivelser

Den norske universalforebyggende versjonen av IY TCM-programmet er godt beskrevet i en norskspråklig håndbok for gruppeledere og tre fagbøker (Webster-Stratton 2005; 2012; 2018); herunder teoretisk fundament, hovedhensikt, målgruppe, kjernekomponenter, opplæringskrav og -struktur samt arbeidsmetoder (Tabell 1). Inklusjons- og eksklusjonskriterier (dvs. hvilke skoler/barnehager tiltaket antas å passe/ikke å passe for) er ikke spesifisert. Slike kriterier kan for eksempel avspeile om programmet er like godt egnet for skoler som sliter med mye problematferd som skoler med lite problematferd, om det passer bedre for store skoler enn små skoler eller om det stilles krav til geografisk beliggenhet, økonomiske ressurser o.l. Det gis imidlertid en rekke konkrete eksempler på og råd om hvordan gruppeledere praktisk og fleksibelt bør formidle programets kjernekomponenter til skole- og barnehageansatte, blant annet ved bruk av videovignetter, rollespill, diskusjonstemaer og visuelt bilde-materiale. Hva programmet spesifikt har til hensikt å oppnå (f.eks. redusert forekomst av mer eller mindre

alvorlig problematferd, økt sosial kompetanse, mer bruk av proaktive og positive strategier blant de ansatte etc.) er i mindre grad operasjonalisert. Implementeringen er i noen grad beskrevet, men det mangler en implementerings- og kvalitetssikringsmanual. Hva som må til av støttestrukturer, ressurser og organisering på skole- og barnehagenivå for at programmet skal kunne etableres og vedlikeholdes som et enhetsomfattende intervensjonsprogram, fremkommer ikke i programbeskrivelsen.

Tabell 1. Vurdering av tiltakets beskrivelse.

Dimensjoner i beskrivelsen	Ikke beskrevet	Noe beskrevet	Godt beskrevet
Problembeskrivelse			X
Målgruppe			X
Hovedmål			X
Sekundærmål		X	
Utforming av tiltaket			X
Metoder som benyttes			X
Kjerneelementer/fleksibilitet			X
Utøvere av tiltaket			X
Manual/veileder for utøvere			X
Materiell for mottakere av tiltaket		X	
Undersøkelser som styrker beskrivelsen	X		

Norske effektstudier

Begge de norske studiene har evaluert DUÅs skole- og barnehageprogrammers effekt ved hjelp av et kvasi-eksperimentelt kontrollgruppedesign med pre-post test, hvor rekrutteringen av skoler/barnehager foregikk fortløpende i en fem års periode fra 2009 til 2014 (Tabell 2). Pre-mål ble innhentet på høsten før programstart og post-mål ble innhentet på vårparten samme skole-/barnehageår, ca. tre uker etter at programoppleringen var avsluttet. Det var omlag åtte måneder mellom pre og post målingene. Lærer-rapporterte mål på barns fungering ble innhentet fra et tilfeldig utvalg barn ($n = 7$) for hver barnehage-avdeling/skoleklasse.

Tabell 2. Inkluderte nordiske effektstudier ($K = 2$).

Studie/ Artikkel	Forskningsdesign	Deltakere	Utvalget	Måletidspunkt
Fossum, Handegård & Drugli (2017)	Kontrollgruppedesign med pre-post test	Utvalg fra Norge $N = 1049$ barnehagebarn	92 barnehager (46 i eksperimentgruppen), 49 % jenter. Snittalder 4,4 år	Pretest, posttest, 8-9 mnd. mellom måletidspunktene
Aasheim, Reedtz, Handegård, Martinussen & Mørch (2018a)	Kontrollgruppedesign med pre-post test	Utvalg fra Norge, $N = 1518$ småskolebarn	1 – 3 klassinger fra 43 skoler (21 i eksperimentgruppen), 46,7 % jenter. Snittalder 7,26 år	Pretest, posttest, 8-9 mnd. mellom måletidspunktene
Aasheim, Drugli, Reedtz, Handegård & Martinussen (2018b)	Kontrollgruppedesign med pre-post test	Utvalg fra Norge, $N = 1518$ småskolebarn	1 – 3 klassinger fra 43 skoler (21 i eksperimentgruppen), 46,7 % jenter. Snittalder 7,26 år	Pretest, posttest, 8-9 mnd. mellom måletidspunktene
Kirkhaug, Drugli, Handegård, Lydersen, Åsheim & Fossum (2016)	Kontrollgruppedesign med pre-post test	Subutvalg fra Aasheim m.fl. $N = 83$	Høyrisikobarn som skårer over klinisk grenseverdi på eksternalisering. 34 skoler. 21,7 % jenter	Pretest, posttest, 8-9 mnd. mellom måletidspunktene

I den ene studien undersøkte Fossum og kollegaer (Fossum, Handegård & Drugli, 2017) effekten av programmet i 46 barnehager ($n = 511$ barn) i kommuner som tidligere hadde implementert DUÅs foreldretreningsprogram. Et tilsvarende antall barnehager ($n = 538$ barn) fra kommuner som ikke hadde implementert foreldretreningsprogrammet fungerte som kontrollgruppe. I den andre studien (Aasheim, Reedtz, Handegård, Martinussen & Mørch, 2018a; Aasheim, Drugli, Reedtz, Handegård & Martinussen, 2018b; Kirkhaug, Drugli, Handegård, Lydersen, Åsheim & Fossum, 2016) undersøkte man effektene av DUÅs skole- og barnehageprogram for skolebarn i 1-3 klasse. Rekrutteringen av 21 intervensjonsskoler ($n = 829$ barn) foregikk i samme geografiske områder som barnehagestudien for å få kvalifiserte gruppeledere til programmet. For å kunne delta som intervensjonsskole var det et krav at minst 80 % av lærerne stilte seg positive til innføring av intervensjonen. I tillegg måtte skolen samtykke i å innføre programmet på hele småskoletrinnet (1-3 klasse). Kontrollgruppen, bestående av 23 skoler ($n = 689$ barn), ble som i barnehagestudien hovedsakelig hentet fra kommuner som ikke hadde implementert DUÅs

foreldretreningsprogram. For å minske sjansen for programkontaminering ble det satt som inklusjonskriterium for både intervensjons- og kontrollskolene at de ikke benyttet seg av eller hadde benyttet andre atferdsrettede evidensbaserte intervensjoner i løpet av siste år.

Både barnehagestudien ($n = 106$ barn) og skolestudien ($n = 83$ barn) gjorde tilleggsanalyser hvor en undersøkte om programmet hadde effekt for høyrisikobarn, dvs. barn som skåret over klinisk grenseverdi på atferdsvansker ved pre-test. I skolestudien undersøkte man også om effektene av programmet på foreldreinnvolvering i skolen og lærernes relasjon til foreldrene og elevene varierte med klassetrinn og elevenes kjønn (Aasheim, Drugli, Reedtz, Handegård & Martinussen, 2018b). Noen av resultatene for høyrisikobarna ble for øvrig presentert i en separat artikkel (Kirkhaug, Drugli, Handegård, Lydersen, Åsheim & Fossum, 2016).

Forskningsmetodisk kvalitet

Vurderingen av studienes forskningsmessige kvalitet jf. Ungsinns kriterier, fremgår av Tabell 3 og er nærmere begrunnet i det følgende.

Tabell 3 Vurdering av forskningsmetodisk kvalitet.

Studie	1. Statistiske analyser	2. Måling	3. Indre validitet	4. Tro mot tiltaket	5. Ytre validitet	Gj. snitt
Fossum m.fl. (2017)	3	3	3	2	3	2,8
Aasheim m.fl. (2018a)	3	3	3	2	3	2,8
Aasheim m.fl. (2018b)	3	2	3	2	3	2,6
Kirkhaug m.fl. (2016)	2	3	1,5	2	3	2,3

Note. Ikke rapportert eller undersøkt = 0; Dårlig/utilfredsstillende = 1; Tilfredsstillende = 2; Godt = 3; Svært godt = 4.

Statistiske analyser

I begge studiene ble hovedanalysene utført ved hjelp av flernivåanalyser. Flernivåanalyser er en adekvat statistisk metode for å analysere den typen data som fremkommer i disse studiene. Ved bruk av metoden kan en ta hensyn til avhengigheten i dataene som skyldes at observasjonene (nivå 1) er klustret i individ (nivå 2) som igjen er klustret i barnehagepersonell/lærer (nivå 3) og til slutt i skole/barnehage (nivå 4). En svakhet ved analysene i begge studiene er at det ikke ble tatt hensyn til alle disse nivåene, noe som potensielt kan gi økt risiko for å feilaktig forkaste null-hypotesen (dvs. feilaktig slutte at det er en programeffekt, type I-feil). Resultater fra styrkeberegningsanalyser er ikke oppgitt, men vi anser det som relativt sannsynlig at analysene fra begge studiene har tilstrekkelig styrke til å kunne finne små til moderate effekter, så lenge det fokuseres på totalutvalget. Analysene som ble gjort på høyriskikutvalget synes derimot i begge studiene å ha for liten statistisk styrke til å kunne finne effekter med tilstrekkelig grad av sannsynlighet, med mindre disse er relativt store (type II feil, dvs. feilaktig slutte at det er ingen forskjell mellom intervensjons- og kontrollgruppen, altså ingen programeffekt).

Det blir gjennomgående angitt hvor stort frafallet fra pre- til post-test var, og egne frafallsanalyser ble utført for å undersøke om det var systematiske forskjeller mellom de som svarte på begge måltidspunktene og de som kun svarte på pre-test. Det mangler derimot frafallsanalyser som fokuserer på manglende data på enkeltvariabler blant de som svarte på spørreskjemaet. Det kommer ikke klart frem hvordan manglende data ble håndtert i barnehagestudien. I to av artiklene til skolestudien (Åsheim m.fl., 2018a, b) ble manglende data håndtert ved hjelp av multippel imputasjon. I artikkelen som fokuserer på skolebarn som skåret over klinisk grenseverdi, ble manglende data håndtert mer implisitt ved å la alle barna være med i flernivå-analysene selv om noen kun hadde svart på ett måletidspunkt. Begge disse metodene bygger på antagelsen om «missing at random» (MAR) og gir korrekte resultater dersom manglende data er relatert til andre observerte variabler i analysemodellen.

Målinger

Begge studiene benyttet i hovedsak de samme måleinstrumentene for blant annet å kartlegge det enkelte barns pre-post problematferd, sosiale ferdigheter, skolefaglige prestasjoner og elev-lærer relasjoner. Kartleggingene er kun basert på lærer-/barnehagelærervurderinger. Inkludering av andre informanter/observasjonsdata ville styrket kvaliteten på målingene. Måleinstrumentene er internasjonalt vel utprøvde med påvist tilfredsstillende psykometriske egenskaper.

Reliabilitet/indre konsistens i de norske studiene er angitt i form av Cronbachs alfa (α). Fra hvilke(t) målepunkt reliabilitetsinformasjonen gjelder, fremgår ikke konsekvent.

Problematferd

To mål på problematferd ble gjennomgående benyttet; SESBI-R (Sutter-Eyberg Student Behavior Inventory- Revised, Eyberg & Pincus, 1999) og TRF (Teacher Report Form, Ashenbach & Rescorla, 2001).

SESBI-R består av 38 spørsmål og måler hhv. forekomst (1 = aldri, 7 = alltid) og intensitet (1 = ja, 0 = nei) av mer og mindre alvorlige former for problematferd blant barn i 2-16 års alder. Måleinstrumentet synes å ha tilfredsstillende måleegenskaper også i en norsk kontekst (Kirkhaug m.fl., 2012). I tråd med dette var alfa-verdiene for problem- og/eller intensitetsskalaene høyere enn 0,90 både i barnehagestudien (Fossum m.fl., 2017) og skolestudien (Aasheim m.fl., 2018a).

TRF inngår i multi-informant kartleggingsverktøyet ASEBA (Achenbach System of Empirically Based Assessment) som måler både kompetanse, styrker, tilpasningsevne, atferdsproblemer, emosjonelle problemer og sosiale problemer fra alderen 1½ til over 90. Lærerversjonen (TRF) for vurdering atferdsproblemer hos barn og unge i 6–18 års alder består av 99 spørsmål og inkluderer kartlegging av internalisering (engstelig/deprimert, tilbaketrukket/deprimert, somatiske plager, sosiale problemer), eksterialisering (tankeproblemer, oppmerksomhetsproblemer, regelbrytende atferd, aggressiv atferd), og skolefaglig fungering. En tilsvarende versjon finnes for barn i førskolealder (1-5 ½ år). Hvert ledd vurderes på en skala fra 0 til 2 (0 = ikke sant så vidt jeg vet, 1 = noe eller av og til sant, 2 = veldig eller ofte sant).

Dokumentasjonen av måleegenskapene til den norske versjonen av TRF er blitt vurdert i Psyktestbarn som noe mangelfull ettersom det er behov for mer forskning rundt skalaens begrepsvaliditet, diagnostiske nøyaktighet, test-retest reliabilitet og normer (Kornør & Drugli, 2011). Om en tar utgangspunkt i Chronbachs alfa er det dog støtte for at skalaene har god reliabilitet også under norske forhold. Aasheim m.fl. (2018) rapporterte alfaverdier på over 0,90 for både aggresjon og oppmerksomhets-skalaen. Kirkhaug m.fl. (2016) benyttet totalskalaen for internaliserende problemer, men oppga kun alfaverdiene for de ulike delskalaene (α fra 0,53 til 0,79). Tidligere norsk forskning gir imidlertid støtte til at Chronbachs alfa er tilfredsstillende også for den totale internaliseringskalaen (Kornør & Drugli, 2011). Det bør likevel bemerkes at Chronbachs alfa ikke er særlig egnet for å måle

reliabilitet for slike overordnede/generelle dimensjoner og at bedre alternativer nå eksisterer (Raykov & Zinbarg, 2011; Yang & Green, 2011). I barnehagestudien administrerte Fossum m.fl. (2017) skalaene for internaliserende problemer, aggresjon og oppmerksomhetsvansker, men oppga dessverre ingen mål på indre konsistens. Resultatene fra en annen norsk studie gir dog støtte for at skalaene har god indre konsistens (målt ved Chronbachs alfa) også for barn i denne aldersgruppen (Berg-Nielsen m.fl., 2012).

Sosiale ferdigheter

Sosiale ferdigheter ble i skolestudien (Kirkhaug m.fl., 2016; Aasheim m.fl., 2018) målt med lærerversjonen av SSRS (Social Skills Rating System, Gresham & Elliott, 1990) som består av 30 spørsmål om sosiale ferdigheter og inkluderer ferdighetsdimensjonene samarbeid, selvkontroll (emosjonell regulering) og selvhevdelse. Forekomsten av de ulike sosiale ferdighetene vurderes på en skal fra 1 = aldri til 4 = svært ofte. Det er verd å bemerke at responskalaen er utvidet fra tre til fire kategorier i den norske versjonen (Ogden, 2003). I tillegg inngår en problematferdsskala (27 spørsmål), men ingen av studiene benyttet denne. En norsk valideringsstudie har gitt støtte til at skalaen har tilfredsstillende reliabilitet og validitet også i vår kontekst (Ogden, 2003). I skolestudien varierte den indre konsistensen, uttrykt ved Chronbachs alfa, fra 0,87 til 0,91 for sub-dimensjonene (Aasheim m.fl., 2018a), mens den var 0,94 for den overordnede sumskåren (Kirkhaug m. fl., 2016). I barnehagestudien benyttet Fossum m.fl. (2017) den lærervurderte sosial kompetanseskalaen fra målingsbatteriet SCBE-80 (2 ½-6 år, Social Competence and Behavior Evaluation, LaFreniere & Dumas, 1995). Barna vurderes på 40 spørsmål om sosiale ferdigheter med en seks-punkts skala (1 = aldri, 6 = alltid). Fossum og hans kollegaer oppgir ikke mål på indre konsistens for denne skalaen, men Chronbachs alfa har tidligere vist seg å være meget høy i en norsk studie på et klinisk utvalg (Fossum m.fl., 2007).

Elev-lærer relasjoner

Aasheim m.fl. (2018b) og Kirkhaug m.fl. (2016) inkluderte også elev-lærerrelasjoner som en utfallsvariabel i sin studie av DUÅs skole- og barnehageprogram. De benyttet en kortversjon av The Student-Teacher Relationship Scale (Pianta, 2001) (STRS-SF, 15 spørsmål) som måler to kvalitetsdimensjoner: konflikt og nærhet (Settanni, Longobardi, Sclavo, Fraire & Prino, 2015). Denne faktorstrukturen har fått støtte også under norske forhold (Drugli & Hjemdal, 2012). Lærerne vurderer graden av varme eller disharmoni i sin relasjon til den enkelte elev på en skala fra 1 = passer definitivt ikke til 5 = passer definitivt. Indre konsistens for nærhetsdimensjonen var

$a = 0,81$ og $0,84$ for konfliktdimensjonen (Kirkhaug m.fl., 2016; Aasheim m.fl., 2018b).

Foreldreinvolvering i skolen

I Aasheim m.fl. (2018b) benyttet man dessuten en revidert versjon av måleinstrumentet The Teacher Involvement Questionnaire (INVOLVE-T, Webster-Stratton, 1998) hvor lærere blir bedt om å vurdere foreldres involvering i sine barns skolegang og angi hvor ofte det er kontakt mellom hjem og skole. To av de tre originale subskalaene (Parent Involvement in Education, $a = 0,79$, og Parent Involvement with School/Teacher, $a = 0,76$,) ble slått sammen i analysen av dataene og referert til som Foreldreinvolvering i skolen (Parent Involvement in School, 13 spørsmål). Det ble ikke rapportert reliabilitetsmål på denne sammenslåtte skalaen. I tillegg inngikk subskalaen Lærer – foreldrerelasjon (Teacher Bonding with Parent, 7 spørsmål, $a = 0,61$). Svarene ble gitt på en fem-punkts skala der 0 = ingen involvering og 4 = høy involvering. Aasheim og hennes kollegaer (2018b) peker på at det var relativt lite variasjon i lærersvarene på INVOLVE-T, noe som kan tyde på at måleinstrumentet ikke er tilstrekkelig tilpasset norsk skolemiljø. Den indre konsistensen for subskalaen som måler Lærer-foreldrerelasjon må videre regnes som utilfredsstillende ($a < 0,70$). En annen norsk studie ga videre kun delvis støtte til den forventede faktorstrukturen (Kirkhaug m.fl., 2013), noe som understreker behovet for ytterligere forskning på måleinstrumentets psykometriske egenskaper i en norsk kontekst.

Indre validitet

Begge studiene benyttet et kvasi-eksperimentelt kontrollgruppedesign med pre-post tester. Bruken av kontrollgruppe er fordelaktig ettersom en kan, i alle fall til en viss grad, kontrollere for eventuelle historie-, modnings- og testeffekter. I skolestudien ble kontrollgruppen matchet med intervensjonsgruppen mht. skolestørrelse og geografisk beliggenhet. I begge studiene valgte en ikke å ta med kontrollskoler fra samme kommuner som intervensjonsskolene, dette for å hindre potensielle smitteeffekter. Ettersom skolene ikke ble randomisert til de ulike betingelsene, vil seleksjonseffekter representere en potensiell trussel mot den indre validiteten ved at gruppene kan være ulike på observerte og/eller uobserverte variabler som påvirker utfallsvariablene.

I barnehagestudien var barna i intervensjonsgruppen signifikant yngre enn barna i kontrollgruppen, mens i skolestudien var det flere barn som var regnet som ikke-norske. Det er imidlertid liten grunn til å tro at disse forskjellene har hatt innvirkning på resultatene, da de var enten minimale eller ble statistisk kontrollert for i analysene. Mer bekymringsfullt er det at intervensjonsgruppen

i skolestudien i snitt skåret systematisk høyere på alle mål på problematferd og lavere på alle mål på sosial kompetanse og foreldreinvolvering i skolen ved pre-test. Selv om forskjellen var statistisk signifikant kun for enkelte av utfallsmålene (samarbeid, selvkontroll, foreldreinvolvering i skolen, lærers bånd til foreldre), vil slike initiale forskjeller bety at en ikke kan utelukke at positive effekter av programmet helt eller delvis kan tilskrives regresjon mot gjennomsnittet (regression to the mean). Det er også verd å minne om at intervensjons- og kontrollgruppen skilte seg mht. om skole- og barnehageprogrammet ble innført i kommuner hvor det tidligere var innført DUÅs foreldretreningsprogram eller ei. Det faktum at enkelte av barna i intervensjonsgruppen også var eksponert for foreldretreningsprogrammet, har innvirkning på hvor sikker en kan være på at de positive effektene skyldes skole- og barnehageprogrammet i seg selv. Det bør også bemerkes at det knytter seg noe usikkerhet til den indre konsistensen, da relevante reliabilitetsindikatorer mangler for flere av målene.

I skolestudien var det en betydelig grad av frafall fra pre-test til post-test i intervensjonsgruppen. Det at frafallet var betydelig større i tiltaksgruppen enn i kontrollgruppen er en klar trussel mot den indre validiteten dersom frafallet ikke oppfylte antakelsen til metoden som ble brukt til å håndtere manglende data (MAR, missing at random). Det ble ikke foretatt sensitivitetanalyse for å undersøke hvor stor innflytelse brudd på denne antakelsen kunne ha. En annen trussel mot indre validitet, som gjelder for begge studiene, er at kun lærere er brukt som informanter. Fordi lærerne også er de som innførte tiltaket, kan dette ha ført til at de ga mindre objektive svar enn lærerne i kontrollgruppen – særlig ved post-test.

Som nevnt tidligere, antas den statistiske styrken å være for liten til å undersøke sannsynlige effekter for høyrisikoelever, med mindre disse er relativt store. Lav statistisk styrke er med andre ord en trussel mot studiens indre validitet, ettersom en da står i fare for å ikke finne eller for å trekke konklusjoner om at det ikke er effekter av programmet for denne gruppen elever, selv om reelle effekter kanskje finnes.

Tro mot tiltaket (fidelity)

Grad av troskap mot tiltaket ble verken beskrevet eller eksplisitt målt i noen av publikasjonene. Det er derfor uvisst i hvilken grad programmet faktisk ble implementert. Når slik informasjon om implementeringsgrad/-kvalitet mangler, kan det ikke trekkes

sikre slutninger om programmets effekt (Dix, Slee, Lawson & Keeves, 2012; Domitrovich m.fl., 2008).

Men en del momenter ble inkludert som en kan forvente har fasilitert implementeringen. Barnehage- og skolepersonalet gjennomgikk en relativt omfattende programopplæring over 42 timer. Personalet ble også bedt om å øve på det de hadde lært fra gang til gang og gi tilbakemelding om sine erfaringer på neste opplæringsdag. De mottok veiledning fra kursleder i periodene mellom kursdagene. Som en del av denne opplæringen fikk hver enhet ett gruppesett á 20 bøker av boken "Hvordan fremme sosial og emosjonell kompetanse hos barn". Både personalet og gruppelederne fylte ut sjekklister i opplærings-/implementeringsperioden for å fremme troskap mot tiltaket. En mulig svakhet er at det var relativt kort tid mellom pre-test og post-test (ca. 8-9 mnd.), noe som kan ha gjort det vanskelig å implementere alt de lærte.

Ytre validitet

Ytre validitet omhandler i hvilken grad (de kausale) effektene kan overføres på tvers av personer, kontekster, behandlinger og utfall både innenfor og utenfor eksperimentet (Shadish, Cook & Campbell, 2002). DUÅs skole- og barnehageprogram ble testet innenfor rammen av vanlig skole- og barnehagepraksis i en rekke kommuner i Norge, noe som bidrar til å sannsynliggjøre at resultatene kan generaliseres til hele den norske populasjonen av barnehager og skoler. Det er midlertidig viktig å ha i mente de tidligere nevnte inklusjonskriteriene (f.eks. at de inkluderte skolene/barnehagene ikke skulle ha benyttet andre atferdsrettede evidensbaserte intervensjoner året før studiene startet), ettersom det er mer usikkert om effekten kan generaliseres til de skoler/barnehager som ikke oppfyller disse. Med unntak av analysene foretatt på høyrisikobarn, presenteres resultatene som hovedeffekter og det er uvisst i hvilken grad effektene kan generaliseres på tvers av subgrupper (basert på f.eks. sosioøkonomisk status og etnisitet) innenfor studienes utvalg. Det er p.t. ikke foretatt oppfølgingsstudier, noe som gjør at det er usikkert om programeffektene vedvarer over tid.

Effektstørrelser i barnehagestudien

Resultatene viste signifikant reduksjon av symptomer på eksterneringsproblemer (SESBI – intensitet, SESBI-problem, TRF-aggresjon), internaliseringsproblemer og oppmerksomhetsvansker. I tillegg ga resultatene støtte for at programmet kan føre til en økning i sosial kompetanse (SCBE-T). Effektstørrelsene ble oppgitt ved hjelp av Cohen's *d* som uttrykker hvor stor effekten er i standardavviksenheter (Tabell 4 og 5).

Tabell 4. Mellom-gruppe effektstørrelser (*d*) i barnehagestudien.

		SESBI Intensitet	SESBI Problem	C-TRF Internalisering	C-TRF Aggresjon	C-TRF Oppmerksomhet	SCBE-T Sosial kompetanse
Fossum m.fl. (2017)	<i>Alle</i>	0,14**	0,14**	0,12*	0,15**	0,14**	0,21***
	<i>Høyrisiko</i>	0,47	0,20	0,24	0,34	0,18	0,48*

Note. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, d = Cohens d

Effektene for hele utvalget var relativt små, med de fleste under 0,20. Selv om effektstørrelsene gjennomgående var noe høyere for høyrisiko-gruppen (små til middels), var dette utvalget så lite at signifikante effekter kun ble observert for én av utfallsvariablene (sosial kompetanse). En svakhet ved studien er at det ikke ble foretatt interaksjonsanalyser hvor en undersøkte om programmet var mer virksomt for høyrisikogruppen sammenlignet med resten av utvalget.

andre elever når det gjaldt eksternaliseringsproblemer (Aasheim m.fl., 2018a) og lærer-elev konflikt (Aasheim m.fl., 2018b).

Effektstørrelser skolestudien

Også i skolestudien fant en statistisk signifikante og positive effekter på de fleste utfallsvariablene når hele utvalget inngikk i analysene. Resultatene for totalutvalget viste en signifikant reduksjon av symptomer på eksternaliseringsatferd (SESBI-intensitet og SESBI-problem), oppmerksomhetsvansker, totale vansker og en signifikant økning i barnas samarbeids- og selvkontrollferdigheter samt på et overordnet mål for sosiale ferdigheter (Aasheim m.fl., 2018a). Videre indikerte resultatene positive effekter på elev-lærer relasjonene (nærhet, konflikt) samt i lærerens vurderinger av foreldrenes involvering i barnas skolegang. Effekten på grad av elev-lærer konflikt gjaldt særlig for høyrisikoelever (Aasheim m.fl., 2018b). Tiltaket var ikke assosiert med statistisk signifikante endringer i barnas aggresjon, skolefaglige prestasjoner og selvhevdelse eller på læreres opplevde bånd/tilknytning til foreldrene. Det er verd å merke seg at effektstørrelsene i de fleste tilfellene må regnes som små til marginale, og generelt mindre enn i barnehagestudien (Tabell 5). Et unntak gjaldt for grad av involvering av foreldre i skolen, hvor det var en liten til moderat effekt.

Studien inkluderte ikke internaliseringsproblemer som utfallsvariabel selv om dette ble undersøkt i høyrisikoutvalget (se nedenfor). Også i denne studien ble det foretatt separate undersøkelser på et subutvalg med høyrisikobarn som skåret over 90-persentilen på aggressiv atferd. Analysene av denne gruppen er hovedsakelig presentert i en egen artikkel (Kirkhaug mfl., 2006), men Aasheim og hennes kollegaer fant støtte for at programmet hadde en signifikant større effekt på høyrisiko-gruppen enn

Tabell 5. Mellom-gruppe effektstørrelser i skolestudien.

	Alle Effektstørrelse <i>d</i> Aasheim m.fl. (2018a)	Alle Effektstørrelse <i>d</i> Aasheim m.fl. (2018b)	Høyrisiko Effektstørrelse <i>g</i> Kirkhaug m.fl. (2016)
SESBI Intensitet	0,08**!		0,57
SESBI Problem	0,09*		
TRF Aggresjon	0,08		
TRF Oppmerksomhet	0,08**		
TRF Internaliserende problem			0,25
TRF Akademisk	0,08		0,50*
TRF Total	0,09*		
SSRS Samarbeid	0,17**		
SSRS Selvhedelse	0,11		
SSRS Selvkontroll	0,20*		
SSRS sosiale ferdigheter	0,19*!		0,29
STRS Lærer-elev konflikt		0,15*!	0,65*
STRS Lærer-elev nærhet		0,22*!	0,20
INVOLVE- T Foreldreinvolvering i skolen		0,40*	
INVOLVE –T Lærer – foreldrerelasjon		0,16	

Note: * $p < 0,05$. ** $p < 0,01$, d = Cohens d , g = Hedges d , ! = signifikante interaksjonseffekter: SESBI Intensitet og STRS Lærer-elev konflikt -> større effekter for høyrisikobarna. SSRS sosiale ferdigheter og STRS Lærer-elev nærhet -> større effekter i 2 klasse sammenliknet med 3 klasse.

Effektstørrelsene i Kirkhaug m.fl. (2018a) er uttrykt ved hjelp av Hedges g som er mye lik Cohens d idet begge tar utgangspunkt i standardiserte gjennomsnittsforskjeller. I denne studien fokuserte man på følgende utfallvariabler; eksternaliseringsproblemer (SESBI Intensitet), internaliseringsproblemer (TRF internalisering), skolefaglige ferdigheter (TRF Akademisk) og sosiale ferdigheter (SSRS). I tillegg ble nærhet til og konflikt med lærere inkludert som utfallsvariabler. Selv om kun 2 av 6 (skolefaglige ferdigheter og konflikt med lærere) effektmålene var statistisk signifikante, med den begrensede statistiske styrken i analysene, var effektstørrelsene gjennomgående større enn i hovedartikkelen. Dette gjaldt særlig for de to overnevnte utfallsvariablene, samt eksternaliseringsproblemer hvor effektstørrelsene kan regnes som moderate ($g \geq 0,50$).

Internasjonale kunnskapsoppsummeringer

IY-TCM programmet som DUÅs skole- og barnehageprogram bygger på, har blitt evaluert som et lovende program av en ekspertkomité med tilknytning til justisdepartement i USA (<https://www.blueprintsprograms.org/evaluation-abstract/incredible-years-teacher-classroom-management>). TCM-programmet har videre blitt evaluert med positive effekter både for barn og personale i randomiserte kontrollstudier i Storbritannia (Hutchings, Martin-Forbes, Daley & Williams, 2013) og Irland (McGilloway m.fl., 2010) samt i et noe tilpasset program fra Jamaica (Baker-Henningham, Scott, Jones & Walker, 2012).

Det påpekes imidlertid at de fleste studiene er basert på små utvalg og at ingen har oppfølgingsmålinger eller har målt grad av troskap til programmet.

Implementeringskvalitet

Det foreligger en god beskrivelse av programmets opplærings- og veiledningsstruktur (for gruppeledere og mentorer), men det savnes en egen implementeringsmanual eller overordnet skriftlig plan hvor det også beskrives støttestrukturer og krav til skolene/barnehagene som implementerer programmet (inkl. forventet tids- og ressursbruk, nødvendig lokal forankring/oppslutning etc.), samt forventet fidelitet (lojalitet, troskap til programmet) hos den enkelte lærer/SFO-/barnehageansatt. Det eksisterer p.t. heller ingen lokale eller sentrale kvalitetssikringssystemer for monitorering og tilbakemelding til skolene/barnehagene om deres implementeringsnivå og fremdrift (resultater) over tid. Programformidler opplyser imidlertid at et data-basert system for monitorering av implementeringen av skole- og barnehageprogrammet er under utvikling.

Vurderingen av programmets implementeringssystem anskueliggjøres i Tabell 6.

Tabell 6. Implementeringskvalitet.

Kategori	Ja	Nei	Ikke relevant
Implementeringsstøtte	X		
Kvalifikasjonskrav	X		
Opplæring	X		
Sertifiseringsordninger	X ¹		
Monitorering av fidelity/ etterlevelse		X	
Veiledning	X		
Identifisering av målgrupper		X ²	
Kartleggings- og vedlikeholdsverktøy		X ³	
Strategier for tilpasning	X		
Samlet skåre	∑ 6/9		

Note. 1 = ikke obligatorisk, 2 = ingen inklusjons- eller eksklusjonskriterer, 3 = finnes en sjekkliste for bruk av mentorer. Informasjon om reliabilitet/validering foreligger ikke.

Diskusjon

Hovedmålet med denne artikkelen var å undersøke om De Utrolige Årenes skole- og barnehageprogram er et virksomt tiltak benyttet i normal praksis i Norge, basert på Ungsinns nye kriterier for klassifisering. Resultatene diskuteres på bakgrunn av disse i påfølgende avsnitt.

Beskrivelse og teoretisk rasjonale

Beskrivelse

DUÅs skole- og barnehageprogram som er tenkt benyttet i barnehager og på grunnskolenes småtrinn (1-3 klasse), er godt beskrevet og eksemplifisert i en håndbok for gruppeledere/veiledere og i medfølgende tilleggsmateriell. Hovedmål, målgruppe og hvem som kan utøve programmet fremgår. Det mangler imidlertid en beskrivelse av programmet slik det er tenkt brukt som en skoleomfattende intervensjon. Det mangler også en nærmere spesifisering av hva som er primære og sekundære utfallsmål, samt en tydeliggjøring av for hvilke skoler og barnehager programmet eventuelt ikke anses som egnet eller tilstrekkelig. Det vil styrke beskrivelsen av programmet ytterligere dersom påpekte manglende informasjon ble tilføyd.

Teoretisk rasjonale

Programserien De Utrolige Årene, herunder skole- og barnehageprogrammet, har et vel begrunnet og tydeliggjort multi-teoretisk og evidensbasert fundament. Dette fundamentet og programmets antatte virkningsmekanismer er nærmere beskrevet i de tre fagbøkene som inngår som en del av opplæringsmaterialet, samt i flere engelskspråklige

artikler av den amerikanske utvikleren av originalprogrammet IY-TCM (Webster-Stratton).

Antall effektstudier

Det foreligger to effektstudier av den norske universalforebyggende versjonen av IY-TCM, en barnehagestudie og en skolestudie. Effektene av det kompetanseutviklende programmet for ansatte i skoler og barnehager ble målt på flere antatt relevante barn/eleverelaterede utfallsvariabler 8-9 måneder etter opplæringsstart. Per dags dato foreligger ingen publiserte studier av potensielle effekter for skole-/barnehageansatte, brukerundersøkelser eller studier der langtidseffekter er undersøkt.

Det bør tilføyes at det i flere amerikanske studier med høy forskningskvalitet er påvist en rekke positive effekter av programmets originalmodell (IY-TCM) som retter seg spesifikt mot høyrisiko-barn (f.eks. Herman, Borden, Reinke & Webster-Stratton, 2011; Webster-Stratton, Reid & Stoolmiller, 2008; Williford & Shelton, 2008).

Forskningsdesign og forskningsmetodisk kvalitet

Begge de foreliggende effektstudiene er basert på et ikke-randomisert pre-post kontrollgruppedesign. Dette er ansett som et egnet design for evaluering av intervensjonseffekter, men medfører flere potensielt alvorlige trusler mot resultatenes indre validitet (Shadish, Cook & Campbell, 2002). Selv om den ytre validiteten anses som god, knytter det seg altså noe usikkerhet til de aktuelle studieresultatenes pålitelighet.

Studiene statistiske styrke anses som tilstrekkelig for evaluering av små til moderate hovedeffekter (totalutvalg), men den statistiske styrken antas i begge tilfeller som for liten til å undersøke sannsynlige effekter for mindre subutvalg (eks. høyrisiko elever). Det ble i begge studiene benyttet stort sett vel utprøvde måleinstrumenter med gode psykometriske egenskaper, hvilket er en styrke. Studienes analytiske tilnærming i form av flernivåanalyser der avhengigheten i dataene (nesting) ble tatt høyde for, er høyst relevant (selv om ikke alle aktuelle nivåer ble inkludert). Det er videre en styrke ved begge studier at effektestimater er oppgitt.

Det er imidlertid en potensiell svakhet at kun lærervurderinger ble benyttet. Et multi-informant design, gjerne i kombinasjon med observasjonsdata, kunne ha styrket resultatene. Det at noen av barna kan ha blitt eksponert for DUÅs skole-/barnehageprogram i tillegg til foreldretreningsprogrammet, problematiseres ikke i noen av publikasjonene. Slike «dobbel-dose tilfeller» kan ha påvirket effektstørrelsene og vanskeliggjør vurderingene av hvor

virksomt skole- og barnehageprogrammet (aleine) er. Det er for øvrig en klar svakhet ved begge studier at implementeringskvaliteten (troskap til tiltaket, fidelitet) verken ble målt eller beskrevet. I hvilken grad man faktisk har målt effektene av programmet slik det er forutsatt gjennomført, er derfor usikkert. Med andre ord, vet man ikke sikkert om de observerte effektene kan tilskrives programmet, god/dårlig gjennomføring eller andre uobserverte forhold som kan ha påvirket utfallene (eks. foreldretreningsprogrammet).

Replikasjonsstudier (hhv. i barnehage og skole) med et fortrinnsvis sterkere forskningsdesign (eks. randomisert eksperiment, forsterket kvasi-eksperiment) der også sammenhengen mellom utfall og implementeringskvalitet undersøkes, anses som nødvendig før mer sikre konklusjoner om program-effekt kan trekkes. Dersom programmet i fremtiden skal videreutvikles og formidles som en skole-omfattende intervensjon, bør mulige effekter for eldre barn (4-10 klasse) også undersøkes.

Effekter

Både skole- og barnehagestudien indikerte positive om enn små til marginale endringer i forekomsten av eksternalisert problematferd. Det ble også observert positive effekter på barnas sosiale ferdigheter, elev-lærer relasjoner og foreldres støtte til barnas skolegang. Ingen negative innvirkninger av programmet er rapportert. Gjennomsnittlig effektstørrelse rapportert i meta-analyser av universelle skoleprogrammer er liten til moderat (Durlak m. fl., 2011; Wilson & Lipsey, 2007). Relativt begrensede effektstørrelser i evalueringer av forebyggende skoleintervensjoner må forventes, fordi de aller fleste barn fungerer bra og har det bra. Effektene av DUÅs skole- og barnehageprogram fremstår likevel i all hovedsak som mindre enn de man vanligvis finner i evalueringer av tilsvarende skoleintervensjoner. Selv om effektene må regnes som små i statistisk forstand, er det viktig å være klar over at de likevel kan ha en praktisk betydning, ettersom såpass mange barn blir eksponert for programmet (j.f. forebyggingsparadokset, Rose, 2008; Weare & Nind, 2011). Når det gjelder høyrisikobarns registrerte positive utbytte av DUÅs skole- og barnehageprogram, vurderes studie-resultatene som langt mer usikre p.g.a. svak statistisk styrke. Dette har sannsynligvis medført at det var relativt få statistisk signifikante program-effekter for denne gruppen. Det er likevel viktig å merke seg at effektstørrelsene var gjennomgående noe større enn i totalutvalget, og at det var statistisk signifikant større effekt på redusering av eksternaliseringsproblemer og elev-lærer konflikter for høyrisikobarna. Innenfor denne gruppen ble det i tillegg funnet en statistisk signifikant effekt (moderat effektstørrelse) av programmet på skole-

faglige prestasjoner, noe som ikke ble funnet i totalutvalget. Foreliggende forskningsresultater gir grunnlag for videreutvikling og evaluering av programmet.

Implementeringskvalitet

Dersom implementerings- og kvalitetssikrings-systemer mangler, er det stor sannsynlighet for at det skjer avdrift fra programmet, dvs. at programmet ikke implementeres slik som forutsatt, og at problemer i implementeringsprosessen ikke oppdages og justeres - og dermed fører til redusert, manglende eller i verste fall negativt utbytte for målgruppen (her; barn, ansatte) (eks. Fixsen & Ogden, 2014). I dette tilfellet foreligger det en implementeringsstrategi for å sikre at programformidlerne (gruppeledere, veiledere) har nødvendig kompetanse til å gjennomføre programmet med tilstrekkelig fidelitet. Som tidligere nevnt, mangler det imidlertid en beskrivelse hva som må til av støttestrukturer, ressurser og organisering på skole- og barnehagenivå for at programmet skal kunne etableres og vedlikeholdes over tid. Foreløpig mangler det også et monitorerings- og feedbacksystem for å: a) holde oversikt over om kvaliteten på den praktiske utøvelsen av programmet i skoler/barnehager faktisk er tilstrekkelig, b) støtte implementeringsprosessen i den enkelte enhet, og c) for å kunne gi løpende databasert informasjon om resultatutviklingen i den enkelte skole/barnehage. For å kunne sikre tilfredsstillende implementering over tid, anbefales det derfor å utarbeide en egen implementerings- og spredningsplan, samt at man ferdigstiller det planlagte monitorerings- og feedbacksystemet for programskolene/barnehagene basert på data, fortrinnsvis fra reliable og validerte måleinstrumenter. Det bør videre vurderes om det bør stilles krav om (heller enn å oppfordre til) sertifisering/re-sertifisering og fidelitysjekk av gruppeledere.

Konklusjon

DUÅs skole- og barnehageprogram har et solid teoretisk fundament. Programmet som er tenkt brukt i barnehager og på småskoletrinnet (1-3 klasse), er godt beskrevet og er et av svært få universelt forebyggende program som eksisterer for yngre barn i Norge og Norden. Resultater fra to norske kvasi-eksperimentelle effektstudier tyder på at programmet lar seg innpasse i ordinær praksis, og det kan bidra til noe redusert problematferd, økt sosial kompetanse, bedre elev-lærere relasjoner og økt skolerelatert foreldrestøtte for barn i 3-8 års alderen. Effektene synes imidlertid å være relativt små til marginale. Det er også indikasjoner på at barn i særlig risiko for å utvikle alvorlige og vedvarende atferdsproblemer kan ha bra, og i noen tilfeller særlig bra, utbytte av programmet. Det

knytter seg imidlertid noe usikkerhet til forskningsresultatene validitet (pålitelighet). Manglende informasjon om/analyser av troskap mot tiltaket og lav statistisk styrke for vurdering av effekt for høyrisikobarn utpeker seg som særlige svakheter ved de foreliggende effektstudiene. Ytterligere dokumentasjon fra studier med et mer robust design (fortrinnsvis RCTer med multi-informant vurderinger, lengre tidsramme og gjennomført i utvalg som tidligere ikke har blitt eksponert for andre DUÅ-programmer) er nødvendig før sikre konklusjoner om programeffekt kan trekkes. En beskrivelse og evaluering av programmet som en skoleomfattende intervensjon og utvikling av et løpende monitorerings- og feedbacksystem for å sikre nødvendig implementeringskvalitet over tid tilrådes.

Med ovennevnte forbehold i mente, klassifiseres tiltaket som funksjonelt virksomt på evidensnivå 4: Tiltak med tilfredsstillende dokumentasjon på effekt.

Referanser

- Aasheim, M., Reedt, C., Handegård, B.H., Martinussen, M. & Mørch, W-T. (2018a). Evaluation of the Incredible Years Teacher Classroom Management program in a regular Norwegian school setting. *Scandinavian Journal of Education Research*, published online 29.05.18. <https://www.tandfonline.com/doi/full/10.1080/00313831.2018.1466357>
- Aasheim, M., Drugli, M.-B., Reedt, C., Handegård, B.-H. & Martinussen, M. (2018b). Change in teacher-student relationships and parent involvement after implementation of the Incredible Years Teacher Classroom Management programme in a regular Norwegian school setting. *British Educational Research Journal*, 44, 1064-1083. <https://doi.org/10.1002/berj.3479>
- Ainsworth, M. D. S. (1989). Attachment beyond infancy. *American Psychologist*, 4, 709-716.
- Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B. & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review*, 12, 30-44.
- Anderson, C. M. & Kincaid, D. (2005). Applying behavior analysis to school violence and discipline problems: Schoolwide positive behavior support. *The Behavior Analyst*, 28(1), 49-63.
- Aschenbach, T. M. & Rescorla, L. A. (2001). ASEBA school-age forms and profiles: youth self report. Burlington VT: Library of Congress.
- Baer, D.M., Wolf, M.M. & Risley, T.R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91-97.
- Bandura, A. (1977). Self-efficacy. Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Baker-Henningham, H., Scott, S., Jones, K. & Walker, S. (2012). Reducing child conduct problems and promoting social skills in a middle-income country: cluster randomised controlled trial. *British Journal of Psychiatry*, 201, 101-108.
- Berg-Nielsen, T. S., Solheim, E., Belsky, J. & Wichstrom, L. (2012). Preschoolers' psychosocial problems: In the eyes of the beholder? Adding teacher characteristics as determinants of discrepant parent-teacher reports. *Child Psychiatry & Human Development*, 43, 393-413. <http://dx.doi.org/10.1007/s10578-011-0271-0>
- Bowlby, J. (1982). *Attachment and loss. Vol.1. Attachment* (2nd ed.). New York: Basic Books.
- Dix, K. L., Slee, P. T., Lawson, M. J. & Keeves, J. P. (2012). Implementation quality of whole-school mental health promotion and students' academic performance. *Child and Adolescent Mental Health*, 17, 45-51. doi:10.1111/j.1475-3588.2011.00608.x
- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Haagwood, K., Buckley, J. A., Olin, S., m.fl. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, 1, 6-28.
- Durlak, J.A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D. & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning. A meta-Analysis of school-based universal interventions. *Child Development*, 82, 405-432.
- Epstein, M., Atkins, M., Cullinan, D., Kutash, K. & Weaver, R. (2008). Reducing Behavior Problems in the Elementary School Classroom. IES Practice Guide. NCEE 2008-012. *What Works Clearinghouse*.
- Eyberg, S.M. & Pincus, D. (1999). *Eyeberg Child Behaviour Inventory and Sutter-Eyberg Student Behaviour Inventory - Revised*. Odessa, FL: Psychological Assessment Resources.
- Fixsen, D. L., & Ogden, T. (2014). Implementation research, practice and policy. *Zeitschrift für Psychologie*, IV, 68 pp. ISBN: 978-0-88937-459-1.
- Fossum, S., Handegård, B. H. & Drugli, M. B. (2017). The Incredible Years Teacher Classroom Management Programme in kindergartens: effects of a universal preventive effort. *Journal of Child and Family Studies*, 26, 2215-2223.
- Fossum, S., Mørch, W.-T., Handegård, B. H. & Drugli, M. B. (2007). Childhood disruptive behaviors and family functioning in clinically referred children: Are girls different from boys? *Scandinavian Journal of Psychology*, 48, 375-382.
- Gresham, F. M. & Elliott, S. N. (1990). *Social skills rating system. Manual*. Circle Pines: American Guidance Service.
- Herman, K. C., Borden, L. A., Reinke, W. M. & Webster-Stratton, C. (2011). The impact of the Incredible Years Parent, Child, and Teacher training programs on children's co-occurring internalizing symptoms. *School Psychology Quarterly*, 26, 189-201. doi: 10.1037/a0025228
- Hutchings, J., Martin-Forbes, P., Daley, D. & Williams, M.

- E. (2013). A randomized controlled trial of the impact of a teacher classroom management program on the classroom behavior of children with and without behavior problems. *Journal of School Psychology, 51*, 571–585. doi.org/10.1016/j.jsp.2013.08.001
- Kirkhaug, B., Drugli, M.B., Handegård, B. H., Lydersen, S., Åsheim, M. & Fossum, S. (2016). Does the Incredible Years Teacher Classroom Management Training programme have positive effects for young children exhibiting severe externalizing problems in school?: A quasi-experimental pre-post study. *BMC Psychiatry, 16*, 1-11. doi: 10.1186/s12888-016-1077-1
- Kirkhaug, B., Drugli, M. B., Mørch, W. T. & Handegård, B. H. (2012). Teacher Report of Children's Problem Behavior on The Sutter-Eyberg Student Behavior Inventory—Revised (SESBI-R) in a Norwegian Sample of Preschool and School Children. *Scandinavian Journal of Educational Research, 56*, 139-153. doi:10.1080/00313831.2011.568672
- Knight, J. (2010). *Unmistakable impact: A partnership approach to improving instruction*. New York: Corwin.
- Kornør, H. & Drugli, M. B. (2011). Måleegenskaper ved den norske versjonen av Teacher's Report Form (TRF). *PsykTestBarn, 1*:7.
- LaFreniere, P. J. & Dumas, J. E. (1995). *Social competence and behavior evaluation, Preschool edition (SCBE)*. Los Angeles: Western Psychological Services.
- LaFreniere, P. J., Masataka, N., Butovskaya, M., Chen, Q., Dessen, M. A., Atwanger, K., Schreiner, S., Montiroso, R. & Frigerio, A. (2002). Cross-cultural analysis of Social Competence and Behavior Problems in Preschoolers. *Early Education & Development, 13*, 201-219.
- Løndal, O. & Aasheim, M. (2016). De Utrolige Årenes Skole og barnehageprogram i H. Eng, S. Ertesvåg, I. Frønes & Kjølbi, J. (red.), *Inkludering og utvikling - førebyggende intervensjoner i barnehage og skule* (s. 141-160). Oslo: Gyldendal Akademisk.
- Martinussen, M., Reedtz, C., Eng, H., Neumer, S.-P., Patras, J. & Mørch, W-T. (2016). *Kriterier og prosedyrer for vurdering og klassifisering av tiltak*. 2. utg. Ungsinn. Tidsskrift for virksomme tiltak for barn og unge. Tromsø: Universitetet i Tromsø
- McGilloway, S., Mháille, G., Lodge, A., O`Neill, D., Kelly, P., Leckey, Y. m.fl. (2011). *Positive classrooms: positive children: a randomized controlled trial to investigate the effectiveness of the Incredible Years Teacher Classroom Management Training Programme in an Irish context (short term outcomes)*. Summary report prepared for Archway. Nedlastet 09.09.17 fra <http://www.incredibleyears.com>
- Nordahl, T., Mausethagen, S. & Kostøl, A. K. (2009). *Skoler med liten og stor forekomst av atferdsproblemer. En kvantitativ og kvalitativ analyse av forskjeller og likheter mellom skolene*. Rapportnr. 3/2009. Hamar: Høgskolen i Hedmark.
- Ogden, T. (2003). The validity of teacher ratings of adolescents' social skills. *Scandinavian journal of educational research, 47*, 63-76.
- Oliver, R. M., Wehby, J. & Reschly, D. J. (2011). Teacher classroom management practices: effects on disruptive or aggressive student behavior. *Campbell Systematic Reviews, 4*, 1-55. doi: 10.4073/csr.2011.4
- Patterson, G. R., Reid, J. B. & Dishion, T. J. (1992). *Antisocial boys: A social interactional approach* (Vol. 4). Eugene, OR: Castalia.
- Piaget, J. & Inhelder, B. (1962). *Le développement des quantités physiques chez l'enfant* (2d ed.) Paris: Delachaux & Niesté.
- Pianta, R. C. (2001). *Student-Teacher Relationship Scale: Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- Raver, C. C. & Knitze, J. (2002). Promoting the emotional well-being of children and families. *Policy Paper, 3*, 1-24.
- Raykov, T. & Zinbarg, R. E. (2011). Proportion of general factor variance in a hierarchical multiple-component measuring instrument: A note on a confidence interval estimation procedure. *British Journal of Mathematical and Statistical Psychology, 64*, 193-207. doi:10.1348/000711009X479714
- Rose, G. (2008). *The strategy of preventive medicine*. Oxford: Oxford University Press.
- Settanni, M., Longobardi, C., Sclavo, E., Fraire, M. & Prino, L. E. (2015). Development and psychometric analysis of the student-teacher relationship scale – short form. *Frontiers in Psychology, 6*, 898. doi:10.3389/fpsyg.2015.00898
- Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Simonsen, B., Fairbanks, S., Briesch, A., Myers, D. & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*, 351-380.
- Skogen, J. C. & Torvik, F.A. (2013). *Atferdsforstyrrelser blant barn og unge i Norge. Beregnet forekomst og bruk av hjelpetjenester. Rapport 13:4*. Oslo: Folkehelseinstituttet.
- Stoltz, S., van Londen, M., Dekovic´ M., Orobio de Castro, B. & Prinzie, P. (2012). Effectiveness of individually delivered indicated school-based interventions on externalizing behavior. *International Journal of Behavioral Development, 36*, 381–388.
- Sørli, M. A. & Ogden, T. (2014). Mindre problematferd i grunnskolen? -Lærervurderinger i et 10-års perspektiv. *Norsk pedagogisk tidsskrift, 98*, 190-202.
- Sørli, M-A, Ogden, T., Solholm, R., & Olseth, A. R. (2010). Implementeringskvalitet – om å få tiltak til å virke. En oversikt. *Tidsskrift for Norsk Psykologiforening, 47*, 315-321.

- Walker, H. M., Ramsey, E. & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices*. Belmont, CA: Wadsworth Publishing Company.
- Weare, K. & Nind, M. (2011). Mental health promotion and problem prevention in schools: what does the evidence say? *Health Promotion International*, 26(suppl_1), i29-i69.
- Webster-Stratton, C. (2003). *The Teacher Classroom Management Programme* (Rev. ed.). Washington, DC: Incredible Years Ltd.
- Webster-Stratton, C. (2005). *Hvordan fremme sosial og emosjonell kompetanse hos barn*. Oslo: Gyldendal Norsk Forlag AS.
- Webster-Stratton, C. (2006). *How to promote children's social and emotional competence*. Thousand Oaks, CA: Sage Publication Inc.
- Webster-Stratton, C. (2007). *De utrolige årene. En foreldreveileder*. 2. utgave. Oslo: Gyldendal Norsk Forlag AS.
- Webster-Stratton, C. (2012). *Incredible teachers. Nurturing children's social, emotional, and academic competence*. Seattle, WA: Incredible Years Inc.
- Webster-Stratton, C. (2018). *Utrolige lærere: Hvordan fremme barns sosiale, emosjonelle og faglige ferdigheter i skole og barnehage*. Oslo: Gyldendal.
- Webster-Stratton, C. & Reid, M. J. (2002). *The Incredible Years Classroom Management Teacher Training Program: Content, methods, and process*. Lastet ned 02.10.17 fra: <http://www.incredibleyears.com/wp-content/uploads/teacher-training-program-classroom-mgmt-02.pdf>
- Webster-Stratton, C., Reid, M. J. & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: evaluation of the Incredible Years Teacher and Child Training Programs in high-risk schools. *Journal of Child Psychology and Psychiatry*, 49, 471–488.
- Webster-Stratton, C., Reinke, W. M., Herman, K. C. & Newcomer, L. L. (2011). The Incredible Years Teacher Classroom Management Training: the methods and principles that support fidelity of training delivery. *School Psychology Review*, 40, 509-529.
- Williford, A. P. & Shelton, T. L. (2008). Using mental health consultation to decrease disruptive behaviors in preschoolers: Adapting an empirically supported intervention. *Journal of Child Psychology, Psychiatry and Allied Disciplines*, 49, 191-200. doi:10.1111/j.14697610.2007.01839.x
- Wilson, S. J. & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine*, 33, 130 –143.
- Wolpert, M., Humphrey, N., Belsy, J. & Deighton, J. (2013). Embedding mental health support in schools: learning from the Targeted Mental Health in Schools (TaMHS) national evaluation. *Emotional and Behavioural Difficulties*, 18, 270-283. Doi:10.1080/13632752.2013.819253
- Yang, Y. & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, 29, 377-392. doi:10.1177/0734282911406668
- Åsheim, M., Reedt, C., Handegård, B. H., Martinussen, M. & Mørch, W.-T. (2018). Evaluation of the Incredible Years Teacher Classroom Management Program in a regular Norwegian school setting. *Scandinavian Journal of Educational Research*. doi:10.1080/00313831.2018.1466357

Forfatterne erklærer at de ikke har noen interessekonflikter.

Mottatt 20.08.2018

Godkjent 13.09.2019

Publisert 20.10.2019.

Redaktør: Bente Storm Haugland

