

Erfaringer med at undersøge kvalitet nationalt

Forarbejde til national kvalitetsundersøgelse –
leverance 3



DANMARKS
EVALUERINGSINSTITUT

INDHOLD

Erfaringer med at undersøge kvalitet nationalt

1	Sammenfatning	5
<hr/>		
2	Indledning	7
2.1	Formål med leverance 3	8
2.2	Design og metode	8
<hr/>		
3	Norge	9
3.1	Baggrund	9
3.2	Valg af redskab	11
3.3	Observation med ECERS-R, ITERS-R og CIP	11
3.4	Udvælgelse	13
3.5	Data	14
3.6	Sammenhængsanalyser	15
3.7	Erfaringer med at gennemføre undersøgelse	16
3.8	Reaktioner på undersøgelsen	17
3.9	Resultater med særlig betydning for en kommende national kvalitetsundersøgelse	19
<hr/>		
4	Danmark	20
4.1	Baggrund	20
4.2	Valg af redskab	20
4.3	Observation med ECERS-3	20
4.4	Udvælgelse, stikprøvestørrelse og repræsentativitet	22
4.5	Repræsentativitet beregnet for forældre til børnehavebørn	26
4.6	Strukturdata	28
4.7	Sammenhængsanalyser	29

4.8	Erfaringer med at gennemføre undersøgelsen	32
4.9	Reaktioner på undersøgelsen	33
4.10	Resultater med særlig betydning for en kommende national kvalitetsundersøgelse	34
<hr/>		
	Appendiks A – Litteraturliste	35
<hr/>		
	Appendiks B – oversigt over centrale studier i GoBaN	37
<hr/>		

1 Sammenfatning

Delleverance 3 er et beskrivende notat, der giver overblik over relevante danske og norske erfaringer med at gennemføre nationale kvalitetsundersøgelser. I forhold til Norge drejer det sig om erfaringer gjort ifm. undersøgelsen GoBaN, og ift. danske erfaringer sættes fokus på Danmarks Evalueringsinstituts (EVA) undersøgelse af læringsmiljøkvaliteten i kommunale børnehaver fra 2019, som er den eneste nationale undersøgelse af kvalitet i danske dagtilbud. Der er fokus på både de gode og de mindre gode erfaringer.

Nedenstående tabel giver et overblik over ligheder og forskelle mellem den EVA's undersøgelse af læringsmiljøkvalitet i kommunale børnehaver og den norske kvalitetsundersøgelse GoBan.

TABEL 1.1

Overblik over nationale kvalitetsundersøgelser i Danmark og Norge

	Norge	Danmark
Projekttitel	GoBaN (Gode barnehager for barn i Norge)	Læringsmiljø i kommunale børnehaver
Årstal	2012-2019	2019
Hvem udførte projektet?	Forskere fra Høgskolen i Oslo og Akershus (HiOA) ¹ og andre forskningsinstitutioner	Danmarks Evalueringsinstitut
Antal medvirkende dagtilbud	93 dagtilbud med i alt 206 stuer	88 dagtilbud med i alt 165 stuer
Aldersgrupper	0-2,5 år	3-5 år
Måleredskab	2,5-5 år	ECERS-3
Udvælgelse af deltagende institutioner	ECERS-R (Early Childhood Environment Rating Scale, Revised Edition)	Stratificeret tilfældig udvælgelse
	ITERS-R (Infant/Toddler Environment Rating Scale, Revised Edition). Caregiver Interaction Profile (CIP)	Læringsmiljø i kommunale børnehaver
Økonomisk ramme	Lamers kortlægningskema ²	

Kilde: Danmarks Evalueringsinstitut.

¹ Blev i 2018 samlet til det, der kaldes OsloMet.

² Totalsum rapporteret til NFR er 38 millioner (dette inkluderer summen fra NFR og egenfinansiering fra OsloMet) – dertil kommer beløb fra uformelle partnere, samlet set vurderer projektleder Elisabeth Bjørnstad, at totalsummen er knap 50 millioner NOK.

I det følgende fremhæves udvalgte pointer, som med fordel kan tages med i arbejdsgruppens og kvalitetsforums overvejelser omkring en national undersøgelse.

Erfaringer fra Norge:

- Brug af udenlandske redskaber: I Norge har det vist sig at være brugbart at anvende udenlandsk udviklede redskaber, og disse har vist sig også at være brugbare i en norsk kontekst, fx redskabet CIP.
- Brug af opfølgende interview: I ITERS-R indgår et opfølgende interview, som afvikles i dagtilbudet straks efter observationen. Dette har vist sig at være rigtig brugbart og værdifuldt for praksis ifølge forskergruppen. Det samme gælder de efterfølgende konferencer om undersøgelsens resultater.
- Udvikling af nyt redskab: Erfaringerne fra Norge viser, at det er tids- og ressourcekrævende at igangsætte et arbejde med at udvikle et nyt redskab.
- Tidlig inddragelse af faglige organisationer: Erfaringerne fra Norge tyder på, at tidlig inddragelse af sektoren via de faglige organisationer i projektet kan være med til at styrke modtagelsen af en national kvalitetsundersøgelse.

Erfaringer fra Danmark:

- Stikprøvestørrelsen er vigtig i det omfang, at man ønsker at lave sammenhængsanalyser. Stikprøven, som blev udtaget i den danske undersøgelse, viste sig at være for lille til at gennemføre sammenhængsanalyser. I Norge er der fundet enkelte signifikante resultater i de gennemførte sammenhængsanalyser.
- Uddannelse af observatører: Observatørerne var allerede uddannet. Det er vigtigt at være opmærksom på, at det kræver en form for uddannelse at gennemføre pålidelige observationer, hvad enten man vælger et eksisterende observationsredskab eller udvikler et nyt.
- Observatøreffekt: Det var erfaringen både hos observatører og personale, at selve observationen ikke påvirkede den pædagogiske praksis. I begyndelsen af observationstiden var personale bevidste om, at de blev observeret, men denne bevidsthed fortog sig efter ganske kort tid.

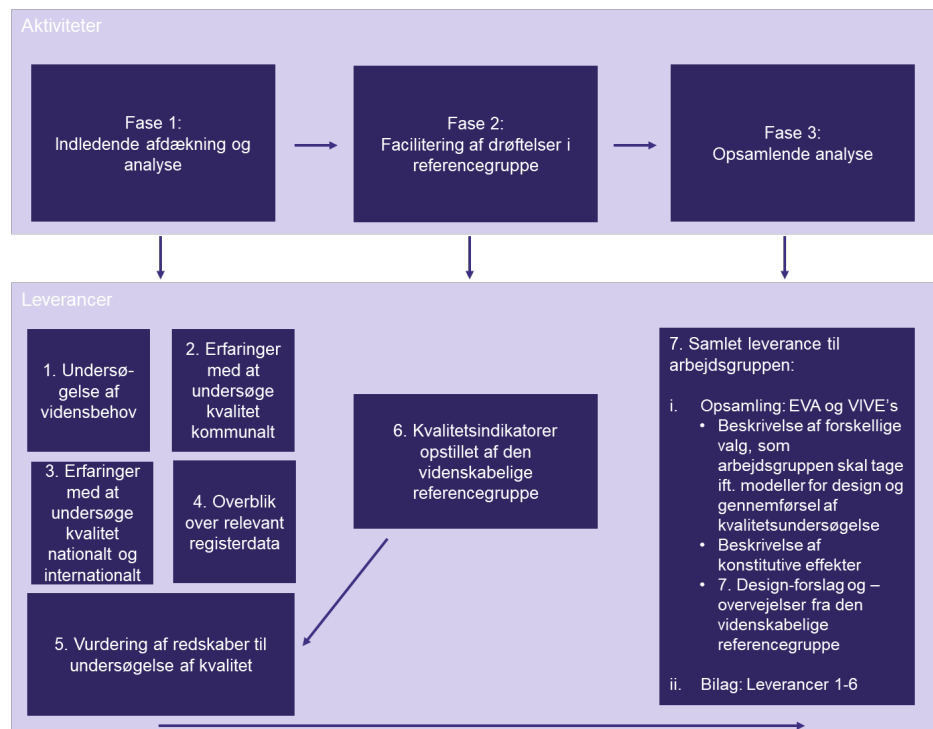
2 Indledning

Det fremgår af Aftale om kommunernes økonomi 2022, at regeringen og KL er enige om løbende at følge udviklingen af kvalitet i dagtilbud. Derfor igangsættes et arbejde med en kvalitetsundersøgelse på dagtilbudsområdet, der tager afsæt i den styrkede pædagogiske læreplan, den danske dagtilbudstradition og relevante data. Som led i kvalitetsundersøgelsen gennemføres et forarbejde.

Formålet med forarbejdet til kvalitetsundersøgelsen er overordnet at sætte retning for det videre arbejde med kvalitetsundersøgelsen. Forarbejdet til kvalitetsstudiet skal skabe et vidensbaseret grundlag for, at der kan tages stilling til kvalitetsundersøgelsens form, indhold og proces for gennemførelse. Som led i forarbejdet udarbejdes seks delleverancer og et opsamlende notat (leverance 7) til arbejdsgruppen. Disse er illustreret i nedenstående figur. Dette notat udgør delleverance 3.

FIGUR 2.1

Overblik over leverancer ifm. forarbejdet



Note: "Arbejdsgruppe" henviser til Arbejdsgruppen for undersøgelse af kvalitet i dagtilbud, som består af KL, BUPL, FOA, Andreas Rasch-Christensen (formand for Partnerskabet for kvalitet i dagtilbud), Finansministeriet og Børne- og Undervisningsministeriet.

2.1 Formål med leverance 3

Leverance 3 er et beskrivende notat, der giver overblik over relevante danske og norske erfaringer med at gennemføre nationale kvalitetsundersøgelser. Der er fokus på både de gode og de mindre gode erfaringer.

Notatet bidrager til forarbejdet med viden om tidligere erfaringer, der er relevante at indtænke i designet af den nationale kvalitetsundersøgelse samt i den efterfølgende tilrettelæggelse af remediering mv.

2.2 Design og metode

Gennem desk research kortlægges og beskrives erfaringer med nationale kvalitetsundersøgelser i Danmark samt Norge. Norge er valgt, da arbejdsgruppen har peget på dette land.

Kortlægningen tager udgangspunkt i rapporter og andre dokumenter samt telefoninterview med projektlederen på projektet i Norge, som har foretaget undersøgelserne, og med en repræsentant fra en faglig organisation.

Gennem interviewene kortlægger vi, *hvilke* nationale kvalitetsundersøgelser der er foretaget i det pågældende land, og gennem interview og dokumentanalyse vil vi for hver kvalitetsundersøgelse beskrive følgende:

- Overordnet design og omfang af undersøgelsen
- Valg af redskab/metode til at vurdere/undersøge kvalitet, herunder fordele og ulemper ved redskabet (konstitutive effekter mv.)
- Om/hvordan undersøgelsen inkluderer indikatorer for hhv. struktur- og resultat kvalitet
- Overordnede resultater og afreporteringsformater
- (Tiltænkt) anvendelse af resultater, herunder på hvilke niveauer
- Eventuelle indsigter i, hvordan undersøgelsen blev modtaget, fx i sektoren eller i praksis (interview med repræsentant for faglig organisation).

Der er gennemført skype-interview med professor Elisabeth Bjørnstad, OsloMet, som var projektleder på GoBan. Det primære formål med interviewet var få besvaret en række opklarende spørgsmål, som er opstået under desk researchen. Derudover er der gennemført interviews med rådgiver Ingri Bjørnevik fra Fagforbundet³, som også sad i den videnskabelig referencegruppe tilknyttet projektet. Formålet med dette interview var at afdække sektorens opfattelse af projektets gennemførelse og resultater.

3 Fagforbundets medlemmer dækker både pædagoger (Barnehagelærere) og pædagogiske assistenter (Fagarbejdere).

3 Norge

I dette kapitel præsenteres viden om den store norske nationale kvalitetsundersøgelse GoBaN (Gode barnehager⁴ for barn i Norge). Vi har indsamlet viden om undersøgelsen via desk research og telefoninterviews med projektleder Elisabeth Bjørnstad og rådgiver Ingri Bjørnevik fra Fagforbundet. I kapitlet beskrives baggrunden for undersøgelsen, overblik over, hvilke studier der er indgår som en del af projektet, og formålet med studierne. Derudover præsenteres erfaringer med at gennemføre undersøgelsen og reaktioner fra omverdenen. Formålet med kapitlet er at udpege udvalgte pointer, som kan tages med i arbejdsgruppens og Kvalitetsforums overvejelser omkring en national undersøgelse.

3.1 Baggrund

GoBaN er det største forskningsprojektet inden for dagtilbudsområdet i Norge og er det første projekt nogensinde i Norge, som undersøger kvaliteten af det pædagogiske læringsmiljø i dagtilbud, og hvilken betydning det har for børnene. GoBaN blev igangsat efter ønske fra Kunnskapsdepartementet⁵, som efterspurgte forskningsbaseret viden inden for dagtilbudsområdet. Norske børn tilbringer samtidig en stor del af deres vågne timer i dagtilbuddene, og derfor oplevede Kunnskapsdepartementet et stort behov for mere forskning i, hvordan psykologiske mekanismer og læringsprocesser påvirkes af de strukturelle faktorer i dagtilbuddene.

Projektet havde til formål at undersøge, hvad der karakteriserer gode dagtilbud, og hvilke faktorer som har indflydelse på børnenes trivsel udvikling og læring. Formålet med projektet var at undersøge sammenhængen mellem forskellige faktorer som struktur, organisering, pædagogiske processer og børnenes trivsel, udvikling og læring ([www. GoBan.no](http://www.GoBan.no); Bjørnstad et al., 2013; Interview med Elisabeth Bjørnstad 10.12.2021). Derudover havde projektet til formål at udvikle et nyt redskab til måling af kvalitet, som var tilpasset norske dagtilbud. Dette formål er dog endnu ikke blevet indfriet. Arbejdet blev påbegyndt, men blev afbrudt, da det blev tydeligt, at det var en større opgave at udvikle et sådant redskab (Interview med Bjørnstad, 10.12.2021).

GoBaN blev igangsat i oktober 2012, mens den første dataindsamling blev gennemført i oktober 2013. Hovedprojektet varede frem til 2019, men der er også efter 2019 blevet gennemført en række forskellige delstudier. På nuværende tidspunkt skal projektgruppen ligeledes i gang med et opfølgende studie, hvor fokus er på sammenhængen mellem kvaliteten af det pædagogiske lærings-

4 På norsk er der ikke et ord for hhv. børnehaver og vuggestuer. Barnehage dækker derfor både over vuggestuer og børnehaver.

5 Kunnskapsdepartementet har ansvaret for dagtilbud, grundskole, kulturskole, ungdomsuddannelser, fagskoleuddannelser, videregående uddannelser samt voksen- og efteruddannelse.

miljø i børnehaven med senere skolepræstationer (Interview med Bjørnestad, 10.12.2021). I hovedprojektet GoBaN har der været særligt fokus på interaktionen mellem pædagogisk personale og barn, men også på interaktionerne mellem børnene. Forskningsprojektet var i sit fokus og metode inspireret af det engelske EPPE-studie (Early Provision of Preschool Education) og det hollandske studie NCKO (Dutch Child Research Consortium, Netherlands).

Projektet var et longitudinelt studie, som fulgte 1.152 børn med forskellig social baggrund, deres hverdag i dagtilbud samt kvaliteten af dagtilbuddene over fire år, fra barnet var omkring to år, til det startede i skole, med mulighed for at følge barnet længere i dets uddannelsesforløb.

GoBaN var et nationalt og internationalt samarbejdsprojekt med forskere fra forskellige forskningsinstitutter Høgskolen i Oslo og Akershus (HiOA)⁶; Norsk institutt for forskning om oppvekst, velferd og aldring (NOVA); Universitetet i Stavanger (UiS); Høgskolen i Vestfold (HiVe); Birkbeck; Nord Universitet; Utrecht University; University of London; og Dutch Child Research Consortium, Netherlands (NCKO).

I dele af dataindsamlingen er der samarbejdet med forskningsprojektet *Blikk for Barn* (Searching for qualities). *Blikk for Barn* adskilte sig fra GoBaN ved kun at undersøge de mindste børn, og børnene blev ikke fulgt over tid. Dvs. fokus var alene på at afdække kvaliteten af dagtilbuddene målrettet de mindste børn. Mere information om forskningsprojektet *Blikk for Barn* i nedenstående boks.

Blikk for Barn

Forskningsprojektet *Blikk for barn* blev gennemført i perioden 2012-2018. Projektet havde til formål at undersøge forskellige aspekter af kvalitet af dagtilbud for børn i alderen 0-3 år i forskellige gruppetyper. Projektet undersøgte derfor mulige sammenhænge mellem de strukturelle og organisatoriske rammer for det processuelle, relationelle og indholdsmæssige element af dagtilbuddene.

Dataindsamling:

I projektet blev der anvendt flere forskellige datakilder, både kvalitative og kvantitative. Der blev gennemført en række etnografiske studier. Derudover blev der indsamlet kvantitative data i form af observationer med måleredskaberne Infant/Toddler Environment Rating Scale, Revised Edition (ITERS-R) og Caregiver Interactive Profile scales (CIP). Endelig er der også gennemført spørgeskemaundersøgelser blandt forældre og pædagogiske ledere. Der deltog i alt 168 ansatte i projektet.

Centrale resultater:

Resultaterne pegede på en række udfordringer på dagtilbudsområdet i Norge. Resultaterne viste, at det pædagogiske personale gennemgående var venlige mod børnene, udviste respekt for deres autonomi og skabte struktur i børnenes hverdag, men samtidig pegede resultaterne også på, at den emotionelle støtte kan forbedres, hvor bl.a. personalets verbale interaktion kan styrkes. Derudover viste målingerne foretaget med ITERS-R, at der var plads til

6 Blev i 2018 samlet til det der kaldes OsloMet.

forbedringer i relation til hygiejne, sikkerhed samt tilstrækkeligt og tilgængeligt materiale til lege- og læringsaktiviteter.

Derudover viste resultaterne, at særligt de store dagtilbudsgrupper er præget af stram logistik, som kan gå ud over det vedvarende samspil mellem barn og pædagogisk personale. Personalet udtrykker også selv, at de strukturelle faktorer i mange sammenhænge gør det vanskeligt at få nok tid til direkte samvær med børnene.

Kilde: Os & Winger, 2019.

3.2 Valg af redskab

Kvaliteten af det pædagogiske læringsmiljø i de norske dagtilbud er undersøgt med en række internationale anerkendte redskaber. Der er gennemført observationer med redskaberne ECERS-R (Early Childhood Environment Rating Scale, Revised Edition) og ITERS-R (Infant/Toddler Environment Rating Scale, Revised Edition). Det var de engelske manualer, som blev anvendt, og ikke norske oversættelser af redskabet. Der blev dog ifølge projektlederen foretaget tilpasninger og konkretiseringer, ift. hvordan man skulle forstå indikatorerne i en norsk kontekst. Projektet fulgte børnene og deres dagtilbud i en række år, og derfor blev der foretaget observationer med begge redskaber. ITERS-R blev anvendt til de første målinger, hvor den udvalgte børnegruppe (født 2011) var under 30 måneder. Da børnene rykkede op i børnehaven, blev der gennemført observationer med ECERS-R, som er målrettet børn over 30 måneder (Bjørnstad et al., 2013).

Derudover blev det hollandske redskab *Caregiver Interaction Profile* (CIP) også anvendt, da redskabet har særligt fokus på interaktion mellem barn og personale. ITERS-R og ECERS-R har også fokus på interaktion, men hvor fokus i disse redskaber er på den samlede personalegruppe, er fokus i CIP på hvert enkelt pædagogisk personale og dennes interaktion med børnene (Bjørnstad et al., 2019). Redskabet er baseret på videoobservationer, som stiller skarpt på interaktionen mellem barn og pædagogisk personale. Der blev kun foretaget videooptagelser i første runde af forskningsprojektet, da det var for kompliceret at foretage opfølgende videooptagelser, når børnene rykkede op i børnehaven, da der ifølge interview med projektlederen kunne være kommet nye børn til, og dermed skulle der indhentes samtykke på ny fra disse børn.

Samlet set blev 1.152 børn (født mellem 2011 og 2012) fulgt i fire år, fra barnet var to år, til de fyldte fem år (longitudinelt studie fra 2013-2019). Børnene var fordelt på 93 dagtilbud med i alt 206 stuer, som senere er blevet fordelt på 205 børnehavestuer. Der er gennemført videoobservationer med redskabet CIP i 110 af stuerne for de mindste børn (vuggestuebørn), hvilket dækker over 110 pædagoger og 58 assistenter.

3.3 Observation med ECERS-R, ITERS-R og CIP

Der er gennemført ITERS-observationer i 93 dagtilbud fordelt på 205 stuer med alt 2.356 børn og 758 medarbejdere. Mens der er gennemført ECERS-R observationer i 95 dagtilbud fordelt på 205 stuer med 3489 børn og 684 medarbejdere (Bjørnstad præsentation, 2019).

3.3.1 ITERS-R og ECERS-R

En observation varede mellem tre og fire timer, hvor en ITERS-R/ECERS-R certificeret observatør var til stede på stuen og efterfølgende foretog interview med dagtilbuddets leder. Alle interaktionsindikatorer skulle observeres inden for de tre-fire timer, men det var muligt at tillægge tid til optælling af fx legetøj og andre materialer. Under observationen iagttog observatøren, hvorvidt over 400 indikatorer for kvalitet var opfyldt. Alle indikatorer forholder sig til barnets behov for et trygt miljø, hvori det kan opbygge positive relationer og blive stimuleret til at lære. Indikatorerne er kategoriseret efter syv delskalaer, som hver er inddelt i 39 punkter i ITERS-R og 43 punkter i ECERS-R. De syv delskalaer er:

- Interaktion
- Sprog og literacy
- Organisationsstruktur
- Læringsaktiviteter
- Plads og indretning
- Rutiner for personlig pleje
- Forældre og personale (ikke medtaget i projektet).

Observatørerne vurderer på baggrund af kvalitetsindikatorerne kvaliteten inden for de seks delskalaer ud fra følgende overordnede spørgsmål:

- Hvordan er læringsmiljøet organiseret?
- Hvordan møder personalet børnene og arbejder med børnefællesskaber?
- Hvordan er personalets tilgang til læring?
- Hvordan viser personalet omsorg, og hvordan varetager de børnenes sundhed og sikkerhed?

CIP

CIP (Caregiver Interaction Profile) er udviklet af dagtilbudsforskere i Holland og er et redskab til at undersøge kvaliteten af interaktionen mellem personalet og børn i dagtilbud. Redskabet er oversat til norsk af Elisabeth Bjørnstad og Ellen Os i samarbejde med NCKO-teamet (The Netherlands Consortium for Research in Child Care), som oprindeligt udviklede redskabet. Redskabet beskriver forskellige interaktionsegenskaber, som antages at have en positiv indvirkning på børns trivsel og udvikling. Egenskaberne er fordelt på to hovedområder: 1) grundlæggende interaktion og 2) læringsmæssig interaktion. Hvert hovedområde består af tre underdimensioner:

Grundlæggende interaktion:

1. Sensitive responsiveness
2. Respect for autonomy
3. Structure and limit setting.

Læringsmæssig interaktion:

1. Verbal communication

2. Developmental stimulation
3. Fostering positive peer interactions.

Redskabet bygger på videooptagelser af interaktionssituationer, som efterfølgende kodes ud fra en standardiseret score. En certificeret observatør stod for optagelserne, mens en anden observatør stod for den efterfølgende scoring. Videooptagelserne varede omkring otte og ti minutter pr. deltagende pædagogisk personale i fire forskellige situationer. De fire situationer var måltid, bleskift, overgange og fri leg. Optagelserne blev herefter kodet ud fra en 7-punkts likert-skala, hvor 1 indikerer interaktion af meget lav kvalitet, mens 7 indikerer meget høj.

For hver af de seks dimensioner fremgår først en generel definition af den pågældende interaktions egenskab efterfulgt af tre korte beskrivelser, som karakteriserer, hvordan egenskaben kan komme til udtryk ved hhv. høj (6, 7), moderat (3, 4, 5) og lav (1, 2) score.

Der er gennemført CIP-optagelser på 111 stuer i 66 dagtilbud, dvs. i godt halvdelen af vuggestuerne fra det samlede forskningsprojekt. Optagelserne blev indsamlet i august 2016, hvor observatørerne var til stede i dagtilbuddet en hel dag fra kl. 08.00-15.00. Det var frivilligt, hvorvidt personalet ønskede at lade sig videooptage, og kun børn, hvis forældre havde givet samtykke, blev videooptaget. I alt har 168 medarbejdere ladet sig videooptage (Interview med Elisabeth Bjørnstad, 10.12.2021).

3.4 Udvalgelse

Dagtilbuddene blev udvalgt ud fra en stratificeret tilfældig udvælgelsesstrategi. Det blev besluttet, at der skulle udvælges dagtilbud, som lå i en overkommelig geografisk afstand til de institutioner, som forskerne var tilknyttet. Der blev derfor udvalgt dagtilbud fra følgende områder: Oslo, Akershus, Vestfold, Rogaland, Nordland og Troms.

Projektgruppen modtog en liste med samtlige børn født i 2011, som var indskrevet i et dagtilbud i Norge. Ud fra denne liste forsøgte man at udtrække en stikprøve, som afspejlede den nationale fordeling baseret på udvalgte kriterier, fx som privat/kommunal og dagtilbuddets størrelse. Ud fra listen rettede forskergruppen direkte henvendelse til dagtilbuddene med henblik på at rekruttere til undersøgelsen. Det var frivilligt, hvorvidt man som dagtilbud ønskede at deltage. Det var ikke et krav, at der skulle indsamles skriftligt samtykke for at foretage ITERS/ECERS-observationen i dagtilbuddene, men der skulle indsamles skriftligt samtykke fra personalet ift. at videooptage og skriftligt samtykke fra forældre til børn født i 2011, som var de børn, som projektet skulle følge fremadrettet (Interview med Elisabeth Bjørnstad, 10.12.2021).

Der var dagtilbud, som ikke ønskede at deltage, og disse udgik derfor fra projektet. I tilfælde, hvor der kun var tale om enkelte medarbejdere, som ikke ville deltage, udgik dagtilbuddet ikke af projektet. I de tilfælde blev der stadig foretaget ITERS-R/ECERS-R/CIP-observationer i dagtilbuddet, men de respektive medarbejdere blev ikke videooptaget med CIP, og deres data, fx medarbejdernes uddannelsesniveaue eller anciennitet, indgik ikke i den efterfølgende analyse. Det samme gjorde sig gældende for de forældre, som takkede nej til videooptagelse. Her blev børnene ikke videooptaget (Interview med Elisabeth Bjørnstad, 10.12.2021). Ift. ITERS/ECERS behandles ingen børn individuelt, og derfor var der ikke behov for samtykke. De børn, hvor forældrene ikke ønskede at deltage i hele projektet, indgik ikke i de efterfølgende analyser med test/kortlægning af barnets udvikling.

I alt indvilligede 1.152 forældre i, at deres barn deltog i projektet, men i de efterfølgende delstudier kan analyserne dog fortsat have været baseret på et mindre antal, da forældrene eksempelvis kan

have sagt ja til, at børnenes sproglige udvikling må testes, men ikke deres sociale udvikling. Det kan også være, at barnet på dagen ikke har haft lyst til at deltage, hvorfor barnet også er udgået af det konkrete delprojekt. Hvis forældre på en konkret stue ikke gav samtykke til videooptagelser med CIP, er disse ikke gennemført på stuen, da det blev vurderet, at det ikke ville være til at styre, om børnene ville gå ind foran kameraet. For personalet kan der godt have været enkelte, som har takket nej, hvor der så alligevel er gennemført videooptagelser på stuen, dog uden at filme den pågældende medarbejder. I det opfølgende studie, som pågår netop nu, er det ca. halvdelen af de deltagende børn fra hovedprojektet, som har sagt ja til at deltage (Interview med Elisabeth Bjørnestad, 10.12.2021).

Ud over de udvalgte dagtilbud var der også en række dagtilbud, som på eget initiativ ønskede at deltage (Interview med Elisabeth Bjørnestad, 10.12.2021).

3.5 Data

I projektet er der indsamlet flere forskellige data med forskellige indsamlingsmetoder. Foruden observation med måleredskaberne ITERS-R, ECER-R og CIP er der gennemført interview med forældrene, når barnet er omkring 30 måneder og fem år samt udsendt spørgeskema til den pædagogiske leder af to omgange og gennemført interviews. I projektet er der både indsamlet strukturdata, procesdata og data om børnenes udvikling.

3.5.1 Forældreinterview og interview med pædagogisk leder

Der er gennemført interviews med forældre til børn på 30 måneder ved projektets opstart i efteråret 2013 og et opfølgende interview, da barnet var fire-fem år. Formålene med interviewene var at kortlægge forhold i hjemmet. I det første interview er der indsamlet baggrundsinformation om bl.a. barnet og demografiske data. Interviewguiden er oversat og tilpasset fra EPPE-studiet. Der er foretaget en del tilpasninger for at gøre spørgsmålene relevante i en norsk kontekst, hvilket har vejet tungere end hensynet til at kunne sammenligne med det engelske studie.

For at indsamle baggrundsinformation om børnehaven er de pædagogiske ledere i dagtilbuddene blevet interviewet. Her har fokus været på generel information om dagtilbuddet, forældreinddragelse, de ansattes kompetencer, stabilitet, arbejdsbetingelser samt pædagogisk profil. Det er igen taget udgangspunkt i den interviewguide, som blev benyttet i EPPE-studie, men igen med visse tilpasninger til en norsk kontekst (Bjørnestad et al., 2013).

Følgende **strukturdata** er indsamlet:

- Andelen af kvalificeret pædagogisk personale
- Normering
- Gruppestørrelse
- Gruppeorganisering (stabile vs. fleksible grupper)
- Ejerform (kommunal vs. privat)
- Pædagogisk leder til stede/ikke til stede på afdelingen.

Derudover er der indsamlet **data om børnenes udvikling** med fokus på deres sproglige udvikling og socioemotional udvikling.

Børnenes sproglige udvikling blev målt med en oversat version af redskabet BAS3 (British Ability Scales, 3 edition), da barnet var hhv. to år og otte måneder samt fem år. Redskabet blev afprøvet før den egentlige dataindsamling, hvor det blev konkluderet, at redskabet fungerede godt i en norsk kontekst. Udgangspunktet for testen er at kortlægge, hvor mange ord barnet kender, når det bliver præsenteret for en række forskellige billeder.

Børnenes socioemotionelle udvikling blev målt ud fra et skema, som den pædagogiske leder fra barnets stue har udfyldt, da barnet var hhv. to år og otte måneder samt fem år. Skemaet bygger på Kari Lamers kortlægningsskema for social adfærd og EPPE studiet. Der er benyttet 30 spørgsmål fra Lamers skala samt 10 spørgsmål fra EPPE-studiet.

3.6 Sammenhængsanalyser

GoBaN-projektet har affødt en række forskellige delprojekter, hvor data fra projektet er blevet brugt til at undersøge forskellige problemstillinger relateret til kvalitet i dagtilbud. Der er foretaget deskriptive analyser, hvor kvaliteten af læringsmiljøet beskrives ud fra ITERS-R- og ECERS-R-scorer samt scorer på delskalaerne og de enkelte punkter. Der er også deskriptive analyser, som beskriver kvaliteten af den processuelle kvalitet på baggrund af CIP, hvor der ses nærmere på CIP-redskabets underdimensioner, og hvilke dimensioner der scores højest på.

Der er også gennemført sammenhængsstudier, som har undersøgt, hvilke strukturelle faktorer som har betydning for kvaliteten af det pædagogiske læringsmiljø i norske dagtilbud. Nogle sammenhængsstudier bruger ECERS-R/ITERS-RS til måling af kvaliteten af det pædagogiske læringsmiljø, andre har udvalgt underskalaer fra redskaberne, mens nogle studier i stedet anvender data fra CIP-optagelser til at afdække kvaliteten af det pædagogiske læringsmiljø. Sammenhængsanalyserne har set på fire aspekter:

- Normering
- Andel pædagogisk uddannet personale
- Organisering (stabile vs. fleksible grupper)
- Ejerskab (privat vs. kommunal).

Et studie har bl.a. fundet, at der ikke ses en sammenhæng mellem andelen af uddannet personale og den overordnede ITERS-R-score, men der ses en positiv signifikant sammenhæng mellem andelen af uddannet personale og enkelte items fra delskalaen *interaktion*, hvor uddannet personale har en positiv betydning for kvaliteten. Studiet finder også en positiv sammenhæng mellem normering og den overordnede ITERS-R-score samt delskalaerne omhandlende interaktion og sprog, hvor færre voksne pr. barn hænger sammen med højere score. Studiet finder også, at stabile grupper scorer signifikant højere end stuer med fleksible grupper. Det gælder både ift. den overordnede ITERS-R-score og for delskalaerne med undtagelse af delskalaen *Rutiner for personlig pleje*. I studiet ses der ingen forskel mellem scoren i private og kommunale dagtilbud (Bjørnstad & Os, 2018).

Derudover er der to studier, som har undersøgt sammenhængen mellem kvaliteten af det pædagogiske læringsmiljø og børnenes udvikling. Studierne har undersøgt:

- Sammenhæng ml. kvaliteten af det pædagogiske læringsmiljø og børns socioemotionelle udvikling.

- Sammenhæng ml. kvaliteten af kvaliteten af det pædagogiske læringsmiljø og børns kognitive udvikling.

Et studie finder bl.a., at der er en svag positiv sammenhæng mellem kvaliteten af det pædagogiske læringsmiljø målt med ITERS-R og børns socioemotionelle udvikling. Her viser resultaterne, at der er en sammenhæng mellem interaktionskvalitet og børnenes empati, når de er tre år, men ingen sammenhæng med andre dimensioner af socioemotionel kompetence. Studiet finder desuden en marginal signifikant sammenhæng mellem interaktionskvalitet og selvkontrol, når børnene er fem år, men dette fund er behæftet med en vis usikkerhed. Der blev ikke fundet sammenhæng mellem interaktionskvalitet og andre socioemotionelle kompetencer (Løkken et al., 2018).

Et andet studie finder ingen sammenhæng ml. ITERS-R-score og børns kognitive udvikling (Eliassen et al., 2018). Forfatterens bud på en årsag til den manglende sammenhæng er bl.a. manglende variation på den uafhængige variabel, dvs. kvaliteten af det pædagogiske læringsmiljø målt med ITERS-R.

Eftersom CIP-skalaen ikke tidligere har været anvendt i Norge, er der som en del af GoBaN-projektet desuden foretaget en validering af redskabets gyldighed i en norsk kontekst ved at sammenholde redskabet med ITERS-R. Resultaterne viser bl.a., at der for nogle delskalaer var stærke korrelationer, mens det ikke var tilfældet for andre. Eksempelvis fandt man ingen korrelation mellem CIP redskabet og delskalaerne *Plads og indretning* og *Rutiner for personlig pleje* i ITERS-R, mens der var den stærkeste korrelation ml. ITERS-R's delskalaer *Sprog og literacy*, *Interaktioner* og *organisationsstruktur* med CIP skalaerne *verbal kommunikation* og *udviklingsmæssig stimulering*. Dette indikerer, at CIP redskabet måler noget andet end ITERS-R, hvilket peger på, at redskabet er relevant at bruge som tilføjelse til målinger med ITERS-R.

Til slut har en række studier anvendt data fra GoBaN-projektet til at udvælge cases til kvalitative dybdegående interviews eller etnografiske studier, som dykker ned i bestemt felt inden for dagtilbudsforskningen, fx sprogarbejde, risikofyldt leg eller hente-bringesituationer, dog stadig med fokus på kvalitet.

I appendiks B ses en oversigt over relevante studier.

3.7 Erfaringer med at gennemføre undersøgelse

Som en del af ITERS-R indgår et opfølgende interview med dagtilbuddet efter gennemført observation. Under interviewet kunne observatøren stille uddybende spørgsmål, fx til ting, som ikke blev observeret under observationstiden. Ifølge Elisabeth Bjørnstad peger tilbagemeldingerne på de opfølgende interviews på, at dagtilbuddene også fandt disse opfølgninger meget interessante. Det opfølgende interview fik dem til at reflektere over deres praksis, og det blev betragtet som et positivt element ved undersøgelsen (Interview med Elisabeth Bjørnstad, 10.12.2021).

Det er ikke undersøgt, hvordan observatører og personalet oplevede hhv. at observere og blive observeret.

Som nævnt er CIP-redskabet oversat i tæt samarbejde med NCKO-temaet, og kriterierne kunne nemt tilpasses den norske kontekst. Iflg. brugen af British Ability Acales (BAS) til måling af børnenes sproglige udvikling uddyber Elisabeth Bjørnstad, at de har oversat ordene og afdækket, hvilke ord som har samme betydning, og hvilke som ikke har. På nuværende tidspunkt pågår et arbejde med

at udarbejde videnskabelige artikler, hvor det undersøges, hvor godt ordene i BAS er tilpasset det norske sprog.

Ifølge Ingri Bjørnevik var der ikke en opfattelse af, at gennemførelsen af CIP-målinger, hvor personalet skulle videofilmes, gav anledning til modstand og kritik. Bjørnevik oplevede ikke, at praksisfeltet havde fundet denne dataindsamlingsmetode problematisk og understregede i den sammenhæng, at deltagelse netop var frivillig, og det nok også var derfor, at der ikke var modstand imod dette. Samtidig var der tillid til, at de enkelte individers videooptagelser alene indgik i en større sammenhæng og ikke blev brugt imod den enkeltes pædagogiske praksis.

3.8 Reaktioner på undersøgelsen

Hovedresultaterne fra GoBaN blev offentliggjort i januar 2019, men ifølge projektleder Elisabeth Bjørnstad var der ikke megen omtale af undersøgelsen i de landsdækkende aviser. Der var dog nogle reaktioner på de sociale medier omkring offentliggørelsestidspunktet. Samme opfattelse ses hos repræsentant fra Fagforbundet, Ingri Bjørnevik. I et interview i januar 2022 uddyber hun, at offentliggørelsen af projektets resultater ikke fik særlig stor opmærksomhed, da offentliggørelsen faldt sammen med ikrafttrædelse af den nye norske rammeplan (læreplan). Fokus blandt Fagforbundet og dets medlemmer var derfor i højere grad rettet mod de nye opgaver, som rammeplanen medførte.

I februar 2019 udgav Utdanningsforbundet en artikel om undersøgelsens resultater med titlen *"Flere barnehagelærere gir bedre barnehager"*. I artiklen præsenteres centrale resultater fra undersøgelsen og en beskrivelse af, hvad disse kan bruges til fremadrettet: *"Prosjektet gir nyttig informasjon. Mye er bra, men vi har absolutt forbedringspotensial på en del områder, særlig på tilbudet til de yngste barna, sier nestleder Hege Valås i Utdanningsforbundet."* ([Flere barnehagelærere gir bedre barnehager \(utdanningsforbundet.no\)](https://www.utdanningsforbundet.no/)).

Ovenstående stemmer overens med opfattelsen blandt projektets leder Elisabeth Bjørnstad, som under et interview med EVA i december 2021 fortæller, at undersøgelsen overordnet set blev modtaget positivt, både blandt praksis og blandt faglige organisationer. Selvom fagforeningen Utdanningsforbundet forholder sig positivt til undersøgelsen i citatet ovenfor, fremhæver Elisabeth Bjørnstad dog, at fagforeningen Utdanningsforbundet ikke kun var positivt indstillet over for undersøgelsen. Ifølge hende har det generelt været den organisation, som har forholdt sig mest skeptisk over for undersøgelsen. Deres skepsis dækkede bredere end den konkrete undersøgelse og var en mere generel kritik af kortlægninger og ratings, hvor holdningen var den, at man ikke kan måle kvalitet på den måde. Det var dog ikke hele fagforeningen, som forholdt sig kritisk til undersøgelsen. Det var især folk fra de centralstyrede poster, mens de lokale medlemmer så mindre kritisk på undersøgelsen. Hege Valås, 2. næstformand i Utdanningsforbundet, fremhævede i et indlæg i 2017 et opmærksomhedspunkt ift. vurderingen af kvalitet i dagtilbud ved brug af måleredskaber: *"Hvilket fotfeste har så slike kvalitetsmanualer her hjemme? De brukes av noen forskere. GoBan-forskningen er et eksempel og Agder-prosjektet er et annet. Dette behøver ikke å være problematisk, så lenge forskerne har den nødvendige ydmykheten i tolkningen av resultatene. Det er viktig at det brede forskningsfeltet bidrar til å sette undersøkelsene inn i en kulturell og sosial kontekst, og at funn speiles i allsidig forskning."* og uddyber ydere: *"La oss lytte til den belgiske utdanningsforskeren Michel Vandenbroeck, når han advarer mot å tro at det finnes en objektiv sannhet om hva som er en god barnehage. Ingen kvalitetsmanualer kan tilpasses en hvilken som helst rammeplan eller læreplan."* ([Utdanningsforbundet.no](https://www.utdanningsforbundet.no/)).

Derudover har Turi Pålerud, politisk rådgiver i Utdanningsforbundet, udtalt følgende om redskaberne CLASS og ITERS i et indlæg på Utdanningsnytt.no: *"En av mine sterkeste innvendinger mot*

både CLASS og ITERS er at begge verktøy ser ut til å hvile på tanken om at kvalitet skapes av pedagogen – og gis til barnet. Pedagogen tilfører barnet kvalitet, ut fra sin intensjon og sitt perspektiv. Hvorvidt det som pedagogen gjør og sier er meningsfullt for barn, er rett og slett lite etterspurt” (Utdanningsnytt.no).

Derudover sætter hun spørsmålsteget ved, hvorvidt kvalitet kan opgøres kvantitativt:
”Bruk av slike manualer beskrives ofte som kvantitativ forskning, som i større grad enn den kvalitative forskningen skal kunne gi holdbar og solid kunnskap. Tja? Manualene er laget for å kunne måle og tallfeste barnehagers kvalitet, og for å kunne gi objektive sammenligninger. Men bak hver eneste indikator ligger det kvalitative valg, som i liten grad begrunnes eller legitimeres. Når vi taper kvalitetsbegrepet for subjektive begrunnelser, står vi igjen med følgende definisjon på barnehagekvalitet: Kvalitet i barnehagen er det man måler med en kvalitetsmanual” (Utdanningsnytt.no).

Forbundet udkom også med en kritisk artikel efter offentliggørelsen (Interview med Elisabeth Bjørnstad, 10.12.2021).

Ifølge projektleder Elisabeth Bjørnstad tog praksisfeltet godt imod undersøgelsen. Forskergruppen var en smule skeptisk, før resultaterne skulle præsenteres, da undersøgelsens resultater også pegede på nogle udfordringer. Men praksis tog godt imod resultaterne, da de netop kunne genkende resultaterne, da de afspejlede den virkelighed, som de stod i. Forskergruppen har også inviteret praksisfeltet ind og været i dialog med dem hele vejen, hvilket muligvis har været medvirkende til den positive modtagelse af undersøgelsens resultater (Interview med Elisabeth Bjørnstad, 10.12.2021). Efterfølgende har forskergruppen afholdt foredrag og workshops for kommuner og børnehaver. Målet har været at gøre det helt tydeligt, hvad undersøgelsen lægger i de enkelte kriterier for kvalitetsvurderingen, men også at gå i dialog med praksis om, hvordan resultaterne bedst kan bruges fremadrettet (EVA, 2020).

Ingri Bjørneviks udtalelser fra interviewet stemmer overens med Elisabeth Bjørnstads beretning om modtagelsen af undersøgelsens igangsættelse og resultater blandt de faglige organisationer og praksisfeltet. Ifølge Ingri Bjørnevik var Fagforbundet overordnet set positivt indstillet over for GoBAN-projektet. De fandt det meget positivt, at de fik mulighed for at indgå i referencegruppen, og at de tidligt i processen blev inddraget i projektet og dets forløb. Ifølge Ingri Bjørnevik var der en stor åbenhed i referencegruppen, hvilket var med til at styrke projektets modtagelse positivt. Der var plads til faglige diskussioner, hvor bl.a. brugen af et amerikansk redskab (ITERS-R) i en norsk kontekst blev drøftet. Her fandt Fagforbundet det positivt, at redskabet i projektet blev tilpasset den norske kontekst. Ifølge Ingri Bjørnevik har projektet været med til at løfte sektoren forskningsmæssigt, og det har været gavnligt at være med fra start. På den måde har fagforbundet kunne besvare spørgsmål og berolige sektoren fra projektets start, fordi Fagforbundet som repræsentant i referencegruppen blev inddraget i projektet.

Selvom undersøgelsen blev modtaget positivt, mener Ingri Bjørnevik, at det ville have styrket undersøgelsen yderligere, hvis der var foretaget en anden inddeling af de forskellige faggrupper i den efterfølgende afrapportering. En del af Fagforbundets medlemmer er pædagogiske assistenter, og der mangler særligt forskning om denne faggruppe, hvorfor det er problematisk, at der ikke er særskilt fokus på denne faggruppe, men at den i stedet sammenlægges med pædagogiske medhjælpere, som ikke varetager samme opgavefunktioner.

Ingri Bjørnevik udtaler yderligere, at projektets resultater også modtaget positivt i sektoren. Resultaterne satte gang i gode faglige diskussioner og gav en del ahaoplevelser, som var med til at udvikle praksis. Hun påpeger dog samtidig, at sektoren ikke har haft det store kendskab til selve projektet. Det er primært Fagforbundet, som efterfølgende præsenterede medlemmerne for centrale

resultater ved forskellige lejligheder. Her blev der heller ikke gået helt tæt på selve projektet, det blev i stedet omtalt som dagtilbudsforskning generelt, og ikke specifikt GoBaN-projektet.

3.9 Resultater med særlig betydning for en kommende national kvalitetsundersøgelse

I det følgende fremhæves udvalgte pointer, som med fordel kan tages med i arbejdsgruppens og kvalitetsforummets overvejelser omkring en national undersøgelse.

Brug af udenlandske redskaber: I Norge har det vist sig at være brugbart at anvende udenlandske redskaber, og disse har vist sig også at være brugbare i en norsk kontekst, fx redskabet CIP og ITERS.

Brug af opfølgende interviews: I ITERS-R indgår et opfølgende interview, som afvikles i dagtilbudet straks efter observationen. Dette har vist sig at være rigtig brugbart og værdifuldt for praksis, som følge af at forskergruppen oplevede det som en god anledning til at reflektere over egen praksis og til at stille uddybende spørgsmål. Det samme gælder de efterfølgende konferencer om undersøgelsens resultater.

Udvikling af nyt redskab: Erfaringerne fra Norge viser, at det er tids- og ressourcekrævende at igangsætte et arbejde med at udvikle et nyt redskab.

Tidlig inddragelse af faglige organisationer: Erfaringerne fra Norge tyder på, at tidlig inddragelse af sektoren via de faglige organisationer i projektet kan være med til at styrke modtagelsen af en national kvalitetsundersøgelse.

4 Danmark

4.1 Baggrund

I 2020 offentliggjorde Danmarks Evalueringsinstitut den første nationale undersøgelse af kvaliteten af det pædagogiske læringsmiljø i kommunale børnehaver. Baggrund for undersøgelsen var, at forskningen peger på, at børns trivsel og læring i de første leveår er af afgørende betydning for både det gode børneliv og børnenes muligheder på sigt. Da tre ud af fire børn i 3-5-årsalderen i Danmark bruger de fleste af deres vågne timer i et dagtilbud, er dagtilbuddene en vigtig arena i forhold til at sikre, at alle børn trives og udvikler sig. Der eksisterede en del viden om, hvad der skaber gode dagtilbud, og hvor børn trives og lærer, men der fandtes ikke en national undersøgelse, som giver systematisk viden om kvaliteten i kommunale dagtilbud på nationalt plan i Danmark.

Undersøgelsen blev lavet i samarbejde med University College Nordjylland og med hjælp fra University College Absalon og rådgivningsvirksomheden Sprogin.

4.2 Valg af redskab

Kvaliteten af det pædagogiske læringsmiljø er undersøgt ved hjælp af det internationale observationsredskab Early Childhood Environment Rating Scale, Third Edition (ECERS-3), som også er blevet brugt til nationale undersøgelser af kvaliteten i dagtilbud i bl.a. England og Norge.

EVA valgte redskabet, da det på daværende tidspunkt var det eneste redskab til måling af proceskvalitet, som var underlagt systematisk kvalitetskontrol, og hvor der var dokumentation for redskabets gennemsigtighed, gyldighed og pålidelighed. Siden undersøgelsen har andre observationsredskaber som fx KID's forbedret deres målekvalitet. For mere information om redskaber og deres kvalitet, se delleverance 5.

4.3 Observation med ECERS-3

Der blev gennemført observationerne på 88 institutioner, hvor 77 af disse inkluderer 2 observationer (stuer) og 11 institutioner blot en enkelt observation. De 88 institutioner er fordelt på 56 kommuner på tværs af alle landets fem regioner. Dataudvælgelsen er foregået på en måde, så der er blevet stratificeret på regionsniveau, og valget om at udvælge 165 stuer fordelt på 88 institutioner er truffet for at sikre repræsentativitet på et nationalt niveau.

En observation strakte sig over tre formiddagstimer (fra kl. 9-12), hvor en ECERS-3-certificeret observatør var til stede på stuen. Under observationen iagttog observatøren, hvorvidt i alt 460 indikatorer for kvalitet var opfyldt. Alle kriterier forholder sig til barnets behov for et trygt miljø, hvori det kan opbygge positive relationer og blive stimuleret til at lære. Indikatorerne er kategoriseret efter seks delskalaer, som hver er inddelt i en række punkter (35). De seks delskalaer er:

- Interaktion
- Sprog og literacy
- Organisationsstruktur
- Læringsaktiviteter
- Plads og indretning
- Rutiner for personlig pleje.

Observatørerne vurderer på baggrund af kvalitetskriterierne kvaliteten inden for de seks delskalaer ud fra følgende overordnede spørgsmål:

- Hvordan er læringsmiljøet organiseret?
- Hvordan møder personalet børnene og arbejder med børnefællesskaber?
- Hvordan er personalets tilgang til læring?
- Hvordan viser personalet omsorg, og hvordan varetager de børnenes sundhed og sikkerhed?

I alt var der 17 observatører, der alle var uddannet på SPROGIN's uddannelse. På ECERS-certificeringskurset kan observatørerne blive certificeret ECERS-3-observatør efter de amerikanske ERS-standarder. Kurset har en varighed af fire dage. Den første dag præsenteres teorien bag ECERS og på kursets anden, tredje og fjerde dag skal de ud og børnemiljøvurdere i tre forskellige børnehaver. Efterfølgende gennemgås de enkelte scoringer. Forudsætningen for at bestå en certificering er, at observatøren har 85 % korrekte svar sammenlignet med en erfaren observatør.

For at sikre inter-rater reliability, dvs. at forskellige observatører bruger redskabet så ensartet som muligt, så tilfældige målefejl minimeres, blev observationerne gennemført således, at hver femte af de 180 observationer blev foretaget med to observatører. Inter-rater reliability beregnes for hver enkelt observatør og fordeles jævnt over dataindsamlingsperioden, så observatøren har maksimalt fire observationer i træk uden en medobservatør. I øvrigt fordeles de 180 observationer ligeligt mellem observatørerne for at minimere eventuel bias for enkelte observatører.

Observatører, som havde gennemført mindre end fem observationer som certificeret observatør eller ikke har gennemført observationer inden for det seneste år, havde en medobservatør på første, tredje og femte observation og herefter minimum hver femte observation, som de øvrige observatører.

4.4 Udvælgelse, stikprøvestørrelse og repræsentativitet

I dette afsnit kortlægges, hvilke beslutninger der er truffet om undersøgelsens design, og hvilket grundlag de er truffet på. I første del beskrives, hvorfor der var to målinger pr. institution frem for en, og i anden del beskrives, hvordan der blev truffet beslutning om en stikprøvestørrelse på 90 institutioner med afsæt i usikkerhedsberegninger. Hertil fremlægger EVA de antagelser, beregningerne beror på. Sidste men ikke mindst fremlægges analyserne af repræsentativitet.

4.4.1 To observationer pr. institution

I undersøgelsen foretages der to ECERS-3-observationer pr. institution, dvs. observationer på to stuer. Det muliggør en variansanalyse af forskelle i kvalitet mellem og inden for institutionerne. De to stuer udvælges tilfældigt for at sikre repræsentativitet.

Der var på tidspunktet for design af undersøgelsen grund til at antage, at to målinger pr. institution var tilstrækkeligt til at afdække varians mellem stuer, da beregninger med afsæt i eksisterende viden viser, at danske daginstitutioner for børn i alderen 3-5 år i gennemsnit har mellem to og tre stuer.⁷

4.4.2 Stikprøvestørrelse og usikkerhedsberegninger

Stikprøven er udtrukket på baggrund af listen over landets daginstitutioner, som den så ud i februar 2019 på Børne- og Undervisningsministeriets Dagtilbudsportal (Børne- og Undervisningsministeriet, 2019). EVA udtrak en stikprøve på 91 kommunale daginstitutioner, hvoraf de 88 har deltaget i undersøgelsen. Størrelsen på stikprøven er fastlagt ud fra en afvejning mellem usikkerhedsberegninger og marginale omkostninger.

Beregning af stikprøvestørrelse og usikkerhed: antagelser og tilgang

Den optimale stikprøvestørrelse afhænger af, hvor stor en usikkerhed vi vil acceptere på vores estimater. Usikkerheden på estimaterne afhænger af, hvor meget målinger varierer fra gennemsnittet (standardafvigelsen), og hvor meget målinger inden for den samme institution korrelerer med hinanden (intraklasse-korrelationen, ICC).

Standardafvigelsen er den gennemsnitlige afvigelse fra gennemsnittet. Den er et samlet mål for variationen og indeholder altså både den variation, som findes mellem institutioner og mellem stuer inden for samme institution. Tidligere studier fra andre lande, der har udført ECERS-observationer, viser standardafvigelser på mellem 0,3 og 1,06 (Barnett, Tarr, Lamy, & Frede, 2001; Early, Sideris, Neitzel, LaForett, & Nehler, 2018). Vi antager en høj standardafvigelse på 2,0 for på den måde at være på den sikre side.

⁷ Udregningerne er foretaget på dagtilbudsregisterdata fra Danmark Statistik. Data fra 2017 indeholder opgørelser over antal indskrevne børn i kommunale og selvejende institutioner i 70 kommuner. Beregninger viser, at der i gennemsnit er 50 børn pr. institution. Glavind & Pade (2015) finder desuden, at den gennemsnitlige stuestørrelse i danske børnehaver er 21,5 børn. Kombinerer vi de to datakilder, fås et gennemsnit på 2,34 stuer pr. institution.

ICC måler forskellen eller ligheden mellem målinger foretaget inden for samme institution. Hvis ICC er høj, betyder det, at stuernes kvalitetsmålinger ligner hinanden meget. Det vil sige, at man ikke får meget ny information ved at fortage observationer på en ekstra stue. Hvis ICC generelt er høj i danske institutioner, er det ikke nødvendigt at foretage så mange observationer pr. institution for at få et retvisende billede af institutionens kvalitet. Omvendt skal der i så fald bruges flere målinger fra forskellige institutioner for at skabe et retvisende nationalt estimat. EVA antager her, at ICC er relativt lav på baggrund af erfaringer fra andre ECERS-undersøgelser i dansk kontekst.

Usikkerheden ved forskellige stikprøvestørrelser er udregnet ved statistikprogrammet Stata's powerberegningsmodul. I beregningerne varieres stikprøvestørrelse, standardafvigelse og ICC for at give indblik i variationens størrelse. Antallet af observationer pr. institution holdes konstant på 2, ligesom signifikansniveauet er sat til 0,05.

Usikkerhed ved 90 institutioner

Usikkerhedsberegningerne viser, at EVA med en stikprøve på 90 institutioner kan forvente en usikkerhed på +/- 0,44, når man antager høj standardafvigelse (2,0) og lav ICC (0,1). Det betyder fx, at det nationale gennemsnit med 95 % sikkerhed befinder sig mellem 3,06 og 3,94, hvis der måles et gennemsnit på 3,5.

Beregningerne viser, at den marginale fordel ved at tilføje flere institutioner falder med antallet af institutioner. Hvis man øger stikprøvestørrelsen fra 50 til 70, reducerer man usikkerheden med 0,09, hvis man øger fra 70 til 90, reducerer man usikkerheden med 0,06, og øger man stikprøven fra 90 til 110, reducerer man usikkerheden med 0,04 (ved en standardafvigelse på 2,0 og ICC på 0,1). Vi vælger at udtrække 91 institutioner, da de marginale omkostninger ved ekstra institutioner er høje, og de marginale fordele efter 91 institutioner er forholdsvis lave. På grund af afrundinger under stratificering af stikprøven ender den endelige stikprøve på 91 institutioner.

Tabel 4.1 viser, hvordan den beregnede usikkerhed ser ud, når man bruger den faktiske standardafvigelse og intraklasse-korrelation (ICC), som er resultatet af dataindsamlingen. Det ses af tabellen, at den overordnede usikkerhed på målingsresultaterne er på +/- 0,21 score på ECERS-3-skalaen. For de enkelte temaer spænder usikkerheden fra +/- 0,13 til +/- 0,36.

TABEL 4.1

Beregnet usikkerhed fordelt på temaer

	Standardafvigelse	ICC	Usikkerhed
Overordnet	0,83	0,39	+/- 0,21
Plads og indretning	0,85	0,39	+/- 0,21
Rutiner for personlig pleje	0,92	0,42	+/- 0,23
Sprog og literacy	1,02	0,22	+/- 0,24
Læringsaktiviteter	0,55	0,29	+/- 0,13
Interaktion	1,43	0,46	+/- 0,36
Organisationsstruktur	1,45	0,30	+/- 0,35

Kilde: Danmarks Evalueringsinstitut.

Note: Tabellen viser resultaterne af en beregning foretaget i Stata's powerberegningsmodul med de kendte standardafvigelser og intraklasse-korrelationer.

Frafald

Der har været et lille frafald i undersøgelsen, idet tre af de udtrukne institutioner er udgået af undersøgelsen. Den ene udgik, fordi institutionen ved nærmere kontakt viste sig ikke at være et kommunalt dagtilbud, der opererer under dagtilbudsloven. De to andre institutioner udgik, fordi der var helt særlige omstændigheder i børnegruppen, som udelukkede observationer. Dermed består det endelige datasæt af observationer udført i 88 institutioner, hvilket svarer til, at 97 % af de udtrukne institutioner har deltaget i undersøgelsen. I 11 af de 88 institutioner er der udført observationer på én stue, da der kun var en stue at observere på. I de resterende 77 institutioner er der foretaget observationer på to stuer. Der er dermed i alt udført 165 observationer i 88 institutioner i 56 kommuner. Dermed lever stikprøven op til de krav, som EVA på forhånd havde sat for at opnå repræsentativitet.

Stratificering af stikprøven

Stikprøven er udtrukket ved stratificeret tilfældig udvælgelse. På grund af afrundinger i stratificeringen er der udtrukket 91 institutioner i stedet for 90. EVA valgte at stratificere stikprøven på regions- og kommunistørrelse, da EVA ønskede en stikprøve med et bredt udsnit af landets kommuner. Forud for valget om at stratificere udtrak EVA tre teststikprøver ved simpel tilfældig udtrækning. Teststikprøverne viste en overvægt af institutioner i store kommuner og i Jylland. Desuden kom de udtrukne institutioner fra et sted mellem 21 og 23 forskellige kommuner. Ved at stratificere stikprøven på kommunistørrelse og region endte EVA i stedet med en stikprøve med institutioner fra 56 forskellige kommuner repræsentativt fordelt over landets regioner og kommunistørrelse.

Kommunale daginstitutioner for børn i alderen 3-5 år fordeler sig regionalt og kommunalt som vist i tabel 4.2 nedenfor. Kommunistørrelsen er bestemt ud fra befolkningsantallet i kommunerne i 2018. Kommunerne er på baggrund af befolkningsantallet inddelt i fire grupper.

TABEL 4.2

Geografisk placering af kommunale daginstitutioner for 3-5-årige

Kommunistørrelse	Hovedstaden	Midtjylland	Nordjylland	Sjælland	Syddanmark	Total
Færre end 25.000 indbyggere	11 %	4 %	2 %	6 %	3 %	6 %
25.001-50.000 indbyggere	36 %	20 %	38 %	39 %	22 %	30 %
50.001-75.000 indbyggere	23 %	14 %	17 %	22 %	31 %	21 %
Flere end 75.000 indbyggere	31 %	63 %	43 %	33 %	44 %	43 %
Total	100 %	100 %	100 %	100 %	100 %	100 %

Kilde: EVA's beregninger på baggrund af tal fra Danmarks Statistik.

På baggrund af tallene i ovenstående tabel er sandsynligheden for, at en institution befinder sig i hver af de 20 strata, udregnet. Sandsynlighederne er herefter brugt til at udregne, hvor mange institutioner der skal udtrækkes til stikprøven fra hver strata. Af tabel 4.2 ses det, at 43 % af de udtrukne institutioner kommer fra en stor kommune med flere end 75.000 indbyggere, 22 % er beliggende i en kommune med 50.001-75.000 indbyggere, mens 29 % af institutionerne er beliggende i en kommune med 25.001-50.000 indbyggere, og endelig ligger 7 % af de udtrukne institutioner i små kommuner med færre end 25.000 indbyggere. Institutionerne er udtrukket tilfældigt fra hver strata.

TABEL 4.3

Geografisk placering af stikprøvens kommunale daginstitutioner

Kommunestørrelse	Hovedstaden	Midtjylland	Nordjylland	Sjælland	Syddanmark	Total
Færre end 25.000 indbyggere	12 %	5 %	0 %	8 %	5 %	7 %
25.001-50.000 indbyggere	35 %	18 %	36 %	38 %	21 %	29 %
50.001-75.000 indbyggere	23 %	14 %	18 %	23 %	32 %	22 %
Flere end 75.000 indbyggere	31 %	64 %	45 %	31 %	42 %	43 %
Total	100 %	100 %	100 %	100 %	100 %	100 %

Kilde: EVA's beregninger på baggrund af tal fra Danmarks Statistik.

4.4.3 Repræsentativitet beregnet for børnehavebørnene

Undersøgelsens design er valgt for at sikre en repræsentativ stikprøve. Med en repræsentativitetsanalyse kan EVA undersøge, om det er lykkedes at udvælge en stikprøve af kommunale børnehaver, hvori børn og personale er sammenlignelige med børn og personale i samtlige kommunale børnehaver i Danmark.

Tabel 4.4 viser forskellen mellem de børn, der var til stede under observationerne på de 165 børnehavestuer, og de børn i alderen 3-5 år, der er indskrevet i samtlige kommunale daginstitutioner i Danmark. Tabellen viser, at der på de observerede børnehavestuer var 50,9 % drenge til stede, mens der generelt er indskrevet 51,8 % drenge i alderen 3-5 år i landets kommunale daginstitutioner. Denne forskel er ikke signifikant. Tabellen viser derudover, at der er en signifikant forskel på andelen af børn med etnisk dansk baggrund i de observerede børnehavestuer i forhold til alle børnehavebørn i kommunale daginstitutioner. I de observerede børnehavestuer er andelen 87,8 %, hvilket er 2,4 procentpoint højere end for den brede gruppe af børnehavebørn.

Når man sammenligner gennemsnitsalderen og aldersfordelingen på de observerede børnehavestuer med alle kommunale børnehaver, ser man ligeledes, at der er signifikant forskel. Gennemsnitsalderen for børn i de observerede børnehavestuer er 3,3 år, mens den er knap 3,5 år i alle kommunale børnehaver. Der er altså i gennemsnit knap 0,2 års forskel mellem de to grupper.

TABEL 4.4

Baggrundsoplysninger for børnehavebørnene

	Observerede børnehaver	Alle kommunale børnehaver i Danmark
Andel drenge	50,9 %	51,8 %
Etnisk dansk baggrund	87,8 %	85,4 %
Gennemsnitsalder	3,3 år	3,5 år
Andel 2-årige	21,1 %	17,5 %
Andel 3-årige	36,7 %	33,2 %
Andel 4-årige	32,8 %	34,1 %
Andel 5-årige	9,0 %	14,3 %
Andel 6-årige	0,4 %	0,9 %

Kilde: Danmarks Evalueringsinstitut og Danmarks Statistik.

Note: Tal med **fed** indikerer, at der er tale om signifikante forskelle mellem de to grupper på 5 %-signifikansniveau. De observerede børnehavestuer er blevet observeret fra april til september 2019. Data fra alle kommunale børnehaver er taget som et øjeblikbillede af de børn og det pædagogiske personale, som var tilknyttet institutionerne pr. 1. juni 2018. På den måde opnår vi to grupper, som er så sammenlignelige som muligt.

4.5 Repræsentativitet beregnet for forældre til børnehavebørn

Tabel 4.5 viser forskellen mellem forældre til børn på de observerede børnehavestuer og forældre til børn indskrevet i samtlige kommunale børnehaver. Tabellen viser, at der ikke er store forskelle mellem forældregrupperne til børnene på de observerede børnehavestuer og forældregrupperne til børn indskrevet i samtlige kommunale børnehaver. For både mødrene og fædrene er der, ligesom for børnene, forskel på andelen, som har etnisk dansk baggrund. I gennemsnit er mødrene og fædrene 2,8 procentpoint oftere etnisk danske i de observerede børnehavestuer i forhold til i alle kommunale børnehaver. Derudover er fædrene til børn på de observerede børnehavestuer oftere i beskæftigelse, idet 91,5 % af dem er i beskæftigelse mod 90,3 % af fædrene til børn i alle kommunale børnehaver.

TABEL 4.5

Baggrundsoplysninger for forældre til børnehavebørn

	Observerede børnehaver	Alle kommunale børnehaver i Danmark
Baggrundsoplysninger for mødre		
Alder	34,3 år	34,3 år
Etnisk dansk baggrund	81,3 %	78,5 %
Indkomst	328.000 kr.	338.600 kr.
I beskæftigelse	74,7 %	75,6 %
Ufaglært	11,5 %	10,9 %

	Observerede børnehaver	Alle kommunale børnehaver i Danmark
Faglært/gymnasial uddannelse	35,2 %	35,7 %
Kort videregående uddannelse	4,9 %	5,1 %
Mellemlang videregående uddannelse	29,4 %	28,9 %
Lang videregående uddannelse	18,9 %	19,5 %
Baggrundsoplysninger for fædre		
Alder	36,9 år	37,0 år
Etnisk dansk baggrund	83,0 %	80,8 %
Indkomst	454.100 kr.	472.200 kr.
I beskæftigelse	91,5 %	90,3 %
Ufaglært	10,5 %	11,1 %
Faglært/gymnasial uddannelse	48,9 %	46,7 %
Kort videregående uddannelse	6,7 %	7,6 %
Mellemlang videregående uddannelse	15,9 %	15,3 %
Lang videregående uddannelse	18,0 %	19,3 %

Kilde: Danmarks Evalueringsinstitut og Danmarks Statistik.

Note: Tal med **fed** indikerer, at der er tale om signifikante forskelle mellem de to grupper på 5 %-signifikansniveau. Undersøgelsen er foretaget fra april til september 2019. Data for forældre til børn i alle daginstitutioner dækker forældre til børn, som var indskrevet i en kommunal daginstitution pr. 1. juni 2018. På den måde opnår vi et tidsnedslag, der er så sammenligneligt som muligt med det tilsvarende for vores stikprøve.

4.5.1 Repræsentativitet beregnet for pædagogisk personale

Tabel C.3 viser forskellen mellem det pædagogiske personale, der var til stede på de observerede børnehavestuer, og det pædagogiske personale, der er ansat i samtlige kommunale daginstitutioner. For flere baggrundsoplysninger er der ikke signifikant forskel mellem personalet fra vores stikprøve og fra samtlige kommunale daginstitutioner i Danmark, fx alder, anciennitet og andel med pædagoguddannelse.

For andre baggrundsoplysninger er der en signifikant, men lille forskel. Der er således en signifikant forskel på andelen af kvindeligt pædagogisk personale til stede i de observerede børnehavestuer sammenlignet med andelen i samtlige kommunale børnehaver. I vores stikprøve er andelen på 82,1 %, mens den er 89,0 % i samtlige kommunale børnehaver. Derudover er der forskel på andelen af personale, som er etnisk danske. I de børnehavestuer, der er foretaget observationer på, er andelen 96,2 %, mens den er 91,6 % i samtlige kommunale børnehaver. Endelig er der forskel på den andel af det pædagogiske personale, som har en pædagogisk grunduddannelse eller en assistentuddannelse, hvor det er 3,5 % i vores stikprøve, men 6,3 % i samtlige kommunale børnehaver.

TABEL 4.6

Baggrundsoplysninger for pædagogisk personale

	Observerede børnehaver	Samtlige kommunale børnehaver i Danmark
Alder	41,6 år	40,7 år
Andel kvinder	82,1 %	89,0 %
Etnisk dansk baggrund	96,2 %	91,6 %
Anciennitet (fra år 2009-2018)	7,4 år	7,8 år
Deltagelse i pædagogisk VEU	39,4 %	42,5 %
PAU-/PGU-uddannet	3,5 %	6,3 %
Pædagoguddannet	55,6 %	51,2 %
Pædagogik eller anden LVU	0,7 %	0,7 %

Kilde: Danmarks Evalueringsinstitut og Danmarks Statistik.

Note: VEU – Voksen- og efteruddannelse; LVU – Lang videregående uddannelse. Tal med **fed** indikerer, at der er tale om signifikante forskelle mellem de to grupper på 5 %-signifikansniveau. Undersøgelsen er foretaget fra april til september 2019. Data for pædagogisk personale i alle kommunale daginstitutioner dækker over pædagogisk personale, som var ansat i en kommunal daginstitution pr. 1. juni 2018. På den måde opnår vi et tidsnedslag, der er så sammenligneligt som muligt med det tilsvarende for vores stikprøve.

4.6 Strukturdata

I forbindelse med undersøgelsen blev der indsamlet en række forskellig strukturdata. Data om *ledelsesstrukturen* og *ledesspændet* i institutionerne blev indhentet via interview med lederne, hvor de har skullet svare på, hvilken type ledelsesstruktur de arbejder under, ligesom de har skullet angive antallet af medarbejdere, som de er ansvarlige for. Desuden er data for *normeringer* også indsamlet via observationer og dataindsamling hos institutionerne, som har leveret cpr-oplysninger på samtlige ansatte og samtlige indskrevne børn.

Data om det pædagogiske personales *uddannelsesbaggrund* og deres *anciennitet* blev indhentet fra Danmarks Statistiks registre. Fra Danmarks Statistik er ligeledes indsamlet information om børnene, deres forældre, søskende og yderligere oplysninger om personalet:

- Børn: køn, alder, etnicitet, familieforhold (bor med to eller en forælder), sociale foranstaltninger, antal søskende
- Forældre: alder, etnicitet, uddannelsesbaggrund, indkomst
- Personale: alder, køn, uddannelsesoplysninger, anciennitet.

4.7 Sammenhængsanalyser

Da projektet om kvaliteten af det pædagogiske læringsmiljø i børnehaver blev planlagt og designet, og mens data blev indsamlet, var det tanken at gennemføre en række sammenhængsanalyser. De skulle afdække, hvilke faktorer der påvirker læringsmiljøkvalitet og dermed forklarer den variation, der er i kvaliteten i de danske børnehaver. Derfor indsamlede EVA data i 2019 på en måde, der sikrede så meget information som muligt om både personalet og børnene i de daginstitutioner, hvor observationerne har fundet sted. I starten af 2020 gennemførte EVA sammenhængsanalyserne. Sammenhængsanalyserne viste sig ikke at være entydige i deres resultater. Derudover fremkom en lang række af 'nulfund', som virker både ulogiske og potentielt skadelige for debatten om, hvad der fører til god kvalitet i en dansk børnehave. Af disse årsager traf EVA en beslutning om ikke at lægge resultaterne fra analysen frem, men EVA ønsker at være åbne omkring de metodiske overvejelser gjort undervejs, hvordan EVA har udført analyserne, og tanker om, hvorfor resultaterne ikke umiddelbart vil bidrage positivt til debatten. Resultaterne fremlægges her alene til brug for arbejdsgruppens overvejelser om design af en national kvalitetsundersøgelse.

I sammenhængsanalyserne så EVA på fem aspekter:

- Normeringen (som observeret på observationstidspunktet)
- Andel pædagogisk uddannet personale på stuen (som observeret på observationstidspunktet)
- Personalets gennemsnitlige anciennitet på stuen
- Ledelsesstrukturen(-typen) i institutionen
- Ledesspændet i institutionen.

EVA ønskede at se nærmere på, hvordan disse fem aspekter ved de forskellige stuer og institutioner hænger sammen med deres kvalitetsmåling foretaget med ECERS-3. De fem aspekter er de såkaldte interessevariable, og i en analysesammenhæng betragtes de i udgangspunktet som uafhængige af kvaliteten i institutionen eller på stuen, dvs. forventningen er, at det er de fem aspekter, som påvirker kvaliteten og ikke omvendt. Som mål for kvaliteten medtog EVA de samlede scorer, de seks overordnede temaer og hvert af de 35 undertemaer i ECERS-3, som de afhængige variable, som antages at blive påvirket af de fem interessevariable. Der er dermed udført 42 regressionsanalyser for hver af de fem interessevariable.

EVA samarbejdede med eksterne metodeeksperter, ift. hvilke regressionsmodeller der vil være bedst til at analysere sammenhænge, ud fra hvilke mekanismer der er på spil. Nedenstående tabel viser disse overvejelser fordelt på de fem aspekter:

TABEL 4.7

Metodiske overvejelser

Aspekt	Hypoteser/overvejelser	Mekanismer
Normering Andel uddannede Anciennitet	Lavere kvalitet fører til højere normering/andel uddannede/højere anciennitet (omvendt kausalitet)	Det vil være naturligt, at man hæver normeringen (eller de andre aspekter) på stuer eller institutioner (eller i kommuner), hvor der er udfordringer med lav kvalitet i læringsmiljøet. Kan betragtes som en art 'social' normering.
	Højere normering/andel uddannede/anciennitet fører til højere kvalitet, alt andet lige	Det virker logisk at antage, at en given stue under de nuværende omstændigheder vil kunne øge kvaliteten af sit læringsmiljø, hvis vi sætter ekstra/mere erfarent personale/flere uddannede ind på stuen (antaget, at der ikke i forvejen er for mange personer)
	Inden for den enkelte institution er det let at ændre normeringen/andel uddannede/anciennitet på dagen	De daglige ledere i børnehaverne kan let fra dag til dag se, hvordan børnene og personalet trives. Lederen kan derfor relativt nemt justere antallet af voksne på stuen i tilfælde af sygdom, perioder med uro osv. Alle aspekter kan let påvirkes på den enkelte stue, men normeringen er dog nok lettest at ændre hurtigt vha. vikarer og deltidsansatte medhjælpere.
Ledelsesstrukturen	Ledelsesstrukturen er forholdsvis fastbestemt, så den ændres ikke så let. Der er ikke umiddelbart en klar hypotese om, hvilken struktur der skulle være bedst for kvaliteten, men man kunne forvente at strukturer, hvor ledelsen er tæt på praksis, har en positiv indflydelse. Aspektet er desuden ens for flere stuer på samme institution	-
Ledelsesspændet	Større ledelsesspænd fører til lavere kvalitet	Lederen har mindre føling med, hvordan det går i de enkelte stuer, og dermed har de sværere ved at reagere, hvis noget går skidt.

Den overordnede grund til, at EVA ikke gik videre med at præsentere resultaterne, er, at EVA ikke mente, at det ville bidrage positivt eller videnskabeligt til den debat, der kørte om normeringer, uddannelse af pædagoger og kvaliteten i dagtilbud generelt. Tabel 5.2 nedenfor viser resultaterne fra EVA's analyser af sammenhængene mellem en række strukturelle faktorer og kvaliteten i dagtilbud. Som det fremgår af tabellen nedenfor, er de fleste resultater insignifikante (signifikante resultater er markeret med ***). Et resultat med et minustegn foran i tabellen indikerer, at eksempelvis parameteret en højere normering umiddelbart hænger sammen med en lavere score i ECERS. Det betyder, at hvis der er flere børn pr. voksne i institutionen, så scorer institutionen lavere i målingen. Det omvendte gør sig gældende for parametre uden et minustegn. For ledelsesstrukturen er variabelen konstrueret, så institutionerne har et 1-tal, hvis de har område-/klyngeledelse, hvorfor resultaterne umiddelbart peger på, at denne ledelsesform hænger sammen med højere kvalitet. For ledelsesspænd, andel uddannede og anciennitet er fortolkningen sådan, at et større ledelsesspænd, en højere andel eller anciennitet hænger sammen med enten højere eller lavere kvalitet, afhængigt af om resultatet er hhv. positivt eller negativt. Som man kan se i tabellen, så er blot et af resultaterne signifikante (ledelsesstruktur), hvilket blot kan skyldes, at der er lavet 35 tests. Der kan derfor være tale om en type-1-fejl.

TABEL 4.8

Regressionsresultater

	Normering	Ledelsesstruktur	Ledelsesspænd	Andel uddannede	Anciennitet
Plads og indretning	-0.079	0.229	-0.001	-0.452	0.092
Rutiner for personlig pleje	0.045	0.092	-0.004	-0.231	0.101
Sprog og literacy	-0.071	0.517***	-0.002	0.546	0.054
Læringsaktiviteter	0.012	0.125	-0.003	0.144	0.061
Interaktion	0.015	0.276	0.009	-0.001	0.039
Organisationsstruktur	0.007	0.367	0.010	-0.403	0.008
Samlet score	-0.012	0.268	0.001	-0.066	0.059

Kilde: Danmarks Evalueringsinstitut, 2020.

Grundlæggende kan der være tale om flere årsager til, at analysen ikke finder flere meningsfulde signifikante resultater. Herunder beskrives et par af de mest sandsynlige:

For få observationer og/eller for lille variation: Det er oplagt, at det kan være et problem for analysen, at der blot er 165 observationer at arbejde med, da det kan være svært i forhold til at kunne observere statistiske sammenhænge med en vis sikkerhed. Det kræver ydermere en vis mængde variation i både de afhængige variable (kvalitetsmål) og uafhængige variable (aspekterne), og her opstår et andet problem: Fx er det et oplagt problem, at der mellem stuer i samme institution ikke nødvendigvis er forskel på andelen af uddannet personale på de to stuer (det er almindeligt forekommende, at begge stuer har 2/3 eller 2/4 personaler, som er uddannet). Desuden er der selvfølgelig en vis spredning på kvalitetsmålene, men i mange tilfælde ligger mange stuer relativt tæt, når man betragter de rent numeriske værdier. Vi ser på mange af temaerne, at langt de fleste stuer befinder sig i et interval på 1-2 point på ECERS-skalaen.

Omvendt kausalitet: Det er derudover en mulighed, at der kan være tale om såkaldt omvendt kausalitet. Det er fx oplagt, at man inden for en institution skeler til, hvordan det går på de forskellige stuer, når disse skal bemandedes. Derudover er det også muligt, at man kommunalt skeler til, hvordan det går i de forskellige institutioner, når man beslutter sig for den institutionelle normering. Endelig er der mulighed for, at der er en forskel i normeringen på tværs af kommuner, som også gør det svært at få et klart indtryk af, hvad det isoleret set betyder for kvaliteten, at man hæver normeringen, andel uddannede eller erfaringen på den enkelte stue. Det vil trække de estimerede sammenhænge i den modsatte retning af, hvad vi ellers ville forvente.

Der er ingen sammenhæng: Endelig skal det selvfølgelig nævnes, at det er oplagt, at der kan være flere af de testede sammenhænge, som simpelthen bare ikke findes.

4.8 Erfaringer med at gennemføre undersøgelsen

Efter observationerne i dagtilbud har observatørerne foretaget interviews med det pædagogiske personale om deres oplevelser med at blive observeret, mens EVA har interviewet en række observatører fra undersøgelsen om det samme. Interviewene giver indblik i både observatører og pædagogisk personales oplevelser med at deltage i undersøgelsen. På baggrund af interviewene er der fem centrale erfaringer om måling af kvalitet i dagtilbud.

4.8.1 En gruppe af observatører oplever, at de kan være fluen på væggen under observationen

Ved brug af måleredskabet ECERS-3 er hensigten, at observatøren skal indtage rollen som fluen på væggen. Et dominerende perspektiv er, at observatørerne oplever, at det i høj grad lykkes for dem at være fluen på væggen under observationerne. De er klar over, at det er idealet og er derfor opmærksomme på, hvordan de skal agere, for at det kan lykkes.

En gruppe observatører begrundet deres oplevelse af at lykkes med, at børnene og det pædagogisk personales opførsel ligner det, man kunne forvente, hvis ikke der var en observatør. Andre henviser til, at personalet i deres interview giver udtryk for, at de glemmer observatørens tilstedeværelse eller er ikke generet af den.

4.8.2 En gruppe af observatører oplever, at personalet vil i dialog om deres pædagogik

Når der i ECERS-3 er et ideal om at være fluen på væggen, er der dermed også et fokus på, at observatørerne skal have så få samtaler med børn og personale som muligt under observationen. Ved observatørernes besøg i børnehaverne har nogle observatører oplevet, at personalet har haft et behov for at forklare eller fortælle om deres pædagogik og gerne vil i dialog med observatørerne. Det kan for eksempel være, at personalet vil gøre opmærksom på, hvorfor et barn agerer, som det gør i en situationen, eller hvilken pædagogik børnehaven arbejder med. På samme måde har flere observatører også oplevet, at personalet og ledelsen ved observationens afslutning har efterspurgt feedback eller refleksioner fra observatøren. Det er observatørernes oplevelse, at det meste af personalet og ledelsen er klar over præmissen om, at resultaterne skulle bruges samlet i forbindelse med undersøgelsen, men at de er naturligt nysgerrige og professionelt interesserede i en tilbagemelding på deres egen praksis.

4.8.3 Personalet har en positiv oplevelse, men en gruppe oplever utryghed ved observationen

Når det pædagogiske personale bliver spurgt til deres generelle oplevelse med at blive observeret, er de overordnet positive. En gruppe beskriver det enten som en 'fin' eller 'okay' oplevelse, og en gruppe lægger vægt på, at de er vant til eller før har prøvet at blive observeret. En gruppe giver udtryk for, at det er underligt eller ubehageligt at være i situationen, fx fordi de føler sig kigget over skulderen eller overvåget. De peger dog på, at følelsen af ubehag gik over i løbet af observationen.

4.8.4 En gruppe personaler oplever, at observatørens tilstedeværelse kun har indflydelse på dem i begyndelsen

En gruppe af personalet beskriver, at de ikke har været bevidste om observatørens tilstedeværelse, mens andre oplever, at de har haft en opmærksomhed omkring det. Et dominerende perspektiv er, at deres bevidsthed omkring det var mest tydelig i begyndelsen og derefter fortog sig i løbet af dagen, fordi deres fokus på den normale praksis fik dem til at glemme observatøren. En gruppe af personalet oplever dermed heller ikke, at observatørens tilstedeværelse har indflydelse på deres egen eller kollegers adfærd. De giver udtryk for, at deres praksis med børnene er, som den plejer. En gruppe af personalet peger dog på, at deres bevidstheden om observatørerne ved begyndelsen kan have haft en indvirkning på deres væremåde, fordi de var bevidste om observationen.

Observatører og personale oplever generelt ikke, at observationen påvirker børnenes adfærd

Overordnet havde hverken observatører og personale en oplevelse af, at børnene reagerede særligt på observatørens besøg. De beskriver, at børnenes væremåde var, som den plejer, og at de ikke virkede til at være interesserede i observatøren. En gruppe af personalet var dertil overraskede over, at børnene ikke reagerede mere, end det var tilfældet. Der peges på, at nogle børn kan være følsomme, hvis der sker nyt, men at personalet ikke havde en oplevelse af dette under observationen. En gruppe af observatørerne pegede på det samme og gav blandt andet udtryk for, at deres tilstedeværelse så ud til kun at fylde lidt for børnene.

Personalet håndterer de reaktioner, der kommer fra børnene

Selvom mange børn ikke reagerede på besøget, oplevede både observatører og personale alligevel forskellige reaktionsmønstre hos nogle børn i løbet af dagen. De børn, der reagerede, gjorde det enten ved at være nysgerrige og stille spørgsmål eller være tilbageholdende. Hos den første gruppe reagerede børnene især ved at stille spørgsmål som "hvad skal du lave her?", "hvad skriver du?" eller "hvad hedder du?" til observatørerne. Den anden gruppe børn kunne virke utrygge eller betydede, som en pædagog beskriver det i interviewet, og tage afstand til observatøren.

Observatørerne beskriver, at personalet håndterede børnenes reaktioner ved at forklare dem om observatøren, hvis børnene var opsøgende eller reagerede på tilstedeværelsen. Personalet fortæller også enkelte gange om, at de forsøgte at tage hånd om børnene, hvis de reagerede på besøget. I nogle børnehaver var der en fælles introduktion af observatørerne til samling, hvor observatørerne i nogle tilfælde også selv fortalte, hvorfor de var til stede i børnehaven.

4.9 Reaktioner på undersøgelsen

Undersøgelsen blev offentliggjort i januar 2020. Der var flere reaktioner på rapporten både i tv og diverse aviser. Nedenfor gengives reaktionen fra FOA og BUPL.

FOA udgav bl.a. artiklen "Rystende rapport om kvalitet i børnehaver". I artiklen blev der lagt vægt på, at resultaterne fra undersøgelsen var bekymrende særligt for børn i udsatte positioner. Daværende formand for Pædagogisk Sektor i FOA Kim Henriksen udtaler blandt andet, at "undersøgelsen viser, at der mangler tydeligt mål og retning for den pædagogiske indsats, som virker tilfældig, hvis ikke helt fraværende," siger formand for Pædagogisk Sektor i FOA Kim Henriksen. Han siger videre, at undersøgelsen viser, at man ikke forlade sig på, at den fulde indfasning af minimumsnormeringer vil ændre markant på forholdene, som belyses i rapporten. For selv om tilstrækkeligt personale er en forudsætning for at lykkes med det pædagogiske arbejde, handler kvalitet også om andet og mere end det.

BUPL var meget kritiske over for undersøgelsen og offentliggjorde flere artikler i samarbejder med forskere om deres kritik. Neden for er gengivet citater fra artiklen ”5 forskere dumper kvalitetsmåling: Hvor er børneperspektivet?” (<https://bupl.dk/artikel/5-forskere-dumper-kvalitetsmaaling-hvor-er-boerneperspektivet/>; downloadet 22.11.2021).

No Emil Kampmann er lektor og ph.d. på Københavns Professionshøjskole og har forsket i kvalitet i dagtilbud. For ham gør ECERS-3 op med den danske pædagogiske tradition og kan derfor ikke bruges til at måle kvaliteten i danske dagtilbud. ”At mange timers interaktion, børneliv og børnesamvær skal koges ned til en score, der hedder 4,5 eller 2,3, giver ikke mening. Det er voldsomt problematisk, fordi det forandrer blikket på pædagogikken fra at være noget, der er komplekse processer, til noget, der er entydige resultater.”

Jan Thorhauge Frederiksen, lektor i pædagogik på Københavns Universitet, siger: ”ECERS-3 bruges til vurdering af læringsmiljøer. Redskabet måler, hvordan rammerne fungerer, men det siger intet om børns udvikling og den pædagogiske kvalitet.”

Et ofte fremført, jf. ovenfor, kritikpunkt af ECERS-3 er det manglende børneperspektiv i metoden. Flere af forskerne mener, at ECERS-3 har for meget fokus på, at kvaliteten skal kunne måles ud fra x antal faktorer uden at overveje børneperspektivet. Det er dybt problematisk, mener ph.d. og lektor ved Roskilde Universitet Annegrethe Ahrenkiel: ”Det er et problem, at man opfatter kvalitet som noget selvindlysende, som noget objektivt, der venter på at blive målt og sammenlignet. Som om der er universelle og entydige normer for, hvad kvalitet er. Og at eksperter derfor kan komme udefra og måle kvalitet. I dagtilbudsloven står der, at der skal tages udgangspunkt i et børneperspektiv. Hvor er det henne i ECERS-3?” siger hun.

Forsker Pernille Juhl, ph.d. og lektor ved Center for Daginstitutionsforskning på Roskilde Universitet udtaler: ”ECERS-3 kan ikke fortælle noget om kvalitet fra et børneperspektiv. Uden det kan vi ikke blive kloge på, hvad det, vi voksne definerer som kvalitet, betyder for børnene.”

4.10 Resultater med særlig betydning for en kommende national kvalitetsundersøgelse

I det følgende fremhæves udvalgte pointer, som med fordel kan tages med i arbejdsgruppens og kvalitetsforums overvejelser omkring en national undersøgelse.

Stikprøvestørrelsen er vigtig i det omfang, at man ønsker at lave sammenhængsanalyser. Stikprøven, som blev udtaget i denne undersøgelse, viste sig at være for lille til at gennemføre sammenhængsanalyser. I Norge er der fundet enkelte signifikante resultater i de gennemførte sammenhængsanalyser. Det norske projekt bestod stikprøven af 93 dagtilbud fordelt på 205 stuer.

Uddannelse af observatører: Observatørerne var allerede uddannet. Det er vigtigt at være opmærksom på, at det kræver en form for uddannelse at gennemføre pålidelige observationer, hvad enten man vælger et eksisterende observationsredskab eller udvikler et nyt.

Appendiks A – Litteraturliste

Alvestad, M., Gjems, L., Myrvang, E., Berger Storli, J., Espedal Tungland, I.B., Lønning Velde, K., & Bjørnstad, E. (2019). *Kvalitet i barnehagen. Rapport fra dybdestudien i det longitudinelle forskningsprosjektet Gode barnehager for barn i Norge (GoBaN)*. Rapport nr. 85, Universitetet i Stavanger.

Bjørnstad, E., Gulbrandsen, L., Johansson, J.-E., & Os, E. (2013). *Metodiske idealkrav og nødvendige tilpasninger: Foreløpig tilstandsrapport fra prosjektet 'Better provision for Norway's children in ECEC: A study of Children's well-being and Development in ECEC.'* Oslo: Høgskolen i Oslo og Akershus (HiOA).

Bjørnstad, Elisabeth, & Ellen Os. (2018). Quality in Norwegian childcare for toddlers using ITERS-R. *European Early Childhood Education Research Journal*. Vol 26 (1), p.111-127. DOI: 10.1080/1350293X.2018.1412051 .

Bjørnstad, E., Broekhuizen, M., Os, E., & Baustad, AG. (2019). Interaction quality in Norwegian ECEC for toddlers measured with Caregiver Interaction Profile Scales (CIP). *Scandinavian Journal of Educational Research*, DOI:<https://www.tandfonline.com/doi/full/10.1080/00313831.2019.1639813>.

Baustad, A.G., & Bjørnstad, E. (2020). Everyday interaction between staff and children aged 1-5 in Norwegian ECEC. *Early Years*.
<https://www.tandfonline.com/doi/pdf/10.1080/09575146.2020.1819207?needAccess=true>.

Eliassen, Erik, Henrik Daae Zachrisson, & Edward Melhuish. (2018). Is cognitive development at three years of age associated with ECEC quality in Norway? *European Early Childhood Education Research Journal*. Vol.26 (1), p. 97-110. <https://www.tandfonline.com/doi/full/10.1080/1350293X.2018.1412050>.

Hansen, Joakim Evensen, & Marit Alvestad. (2018). Educational language practices described by preschool teachers in Norwegian kindergartens. *European early childhood education research journal*. Vol.26 (1), p.128-141.

Johannessen, K., Bjørnstad, E. Nilsen, A.D.O., Ylvisåker, E., Nornes-Nymark, M., Engesæter, M., Pedersen, L., & E. Aadland. (2020). The association between preschool quality and objectively assessed physical activity among Norwegian preschoolers. *Journal of Research in Arts and Sport Education*. 4(2), 7-25. <https://jased.net/index.php/jased/article/view/2485>.

Hansen, Joakim Evensen (2020): Characteristics of educational language practices observed in four toddler groups in Norwegian ECEC, *International Journal of Early Years Education*. 28.4, 382-396.

Kleppe, Rasmus (2018). "Characteristics of staff–child interaction in 1–3-year-olds' risky play in early childhood education and care." *Early Child Development and Care* 188.10: 1487-1501.

Kleppe, R. (2018). Affordances for 1- to 3-year-olds' risky play in Early Childhood Education and Care. *Journal of Early Childhood Research*, 16(3), 258–275.

Løkken, Ingrid M., Martine L. Broekhuizen, Jacqueline Barnes, Thomas Moser, & Elisabeth Bjørnstad (2018a). Interaction Quality and Children's Social-Emotional Competence in Norwegian ECEC. *Journal of Early Childhood Education Research*, 7(2). p. 338-361. <https://jecer.org/wp-content/uploads/2018/12/Lokken-Broekhuizen-Barnes-Moser-Bjornstad-issue7-2.pdf>.

Løkken, Ingrid Midteide., Elisabeth Bjørnstad, Martine Broekhuizen, & Thomas Moser. (2018b). The relationship between structural factors and interaction quality in Norwegian ECEC for toddlers. *International Journal of Child Care and Education Policy*, 12(1), 9.

Lønning Velde, Kjersti (2016). *Foreldrenes opplevelse av hentesituasjonen i barnehagen knyttet til kvalitet*. Masteroppgave, 06-2016. Fakultet for utdanningsvitenskap og humaniora, Universitetet i Stavanger.

Os, E., & Winger, N. (2019). Blikk for barn. *Nordisk Barnehageforskning*, 18
<https://doi.org/10.7577/nbf.3419>.

Rove Nilsen, T. (2021). Pedagogical intentions or practical considerations when facilitating children's play? Teachers' beliefs about the availability of play materials in the indoor ECEC environment. *International Journal of Child Care and Education Policy*, 15(1), 1.16.

Sukke Pettersen, Solveig (2014). *Små barn i testsituasjoner – en fenomenologisk-hermeneutisk studie om små barns trivsel*. Masteroppgave, 05-2014. Fakultet for utdanningsvitenskap og humaniora, Universitetet i Stavanger.

Utdanningsforbundet. *Finnes Universell barnehagekvalitet?* Lokaliseret 12.01.2022 på <https://www.utdanningsforbundet.no/nyheter/2017/finnes-universell-barnehagekvalitet/>.

Utdanningsnytt. *Hvad er kvalitet i barnehagen?* Lokaliseret 12.01.2022 på <https://www.utdanningsnytt.no/barnehagelaerer-fagartikkel/hva-er-kvalitet-i-barnehagen/114046>.

Appendiks B – oversigt over centrale studier i GoBaN

TABEL B.1

Centrale studier fra GoBaN

Titel	Formål/Tema	Data fra GoBaN	Kvantitativt vs. kvalitativt studie
Bjørnestad & Os, 2018	Sammenhæng ml. strukturvariable og kvaliteten af det pædagogiske læringsmiljø.	Strukturdata (normering, personalets uddannelse, gruppeorganisering, ejerform) Observationsdata (ITERS-R).	Kvantitativt
Hansen & Alvstad, 2018	Sprogarbejde i dagtilbud af høj kvalitet.	Strategisk Caseudvælgelse (høj og lav score på Sprogskala)	Kvalitativt
Lønning, 2016	Hentesituationen og kvalitet i dagtilbuddet uddybes dybdegående med udgangspunkt i GoBaN-projektets ITERS-r-score på delskalaen 5,88.	Caseudvælgelse – frivillig deltagelse.	Kvalitativt
Sukke, 2014	Undersøgelse af børns trivsel i testsituationer	Caseudvælgelse – frivillig deltagelse	Kvalitativt
Kleppe, 2017	Karakteristika ved dagtilbuds personales interaktion med børn i alderen 1-3 år i risikofyldt leg	Strategisk caseudvælgelse (høj og lav score på interaktionskala + yderligere kriterier)	Kvalitativt
Hansen, 2018	Karakteristika ved sproglæringsmiljøer og uddannelsesmæssige praksisser i vuggestuer.	Strategisk caseudvælgelse (score på ml. 4-7 på sprogskalaen)	Kvalitativt
Kleppe, 2018	Mulighed for risikofyldt leg blandt 1-3 årige i dagtilbud, herunder det fysiske miljø og udstyr, samt hvordan det relaterer sig til kvalitet.	Strategisk caseudvælgelse (høj og lav ITERS-R-score + yderligere kriterier)	Kvalitativt
Nielsen, 2021	Dybdegående indblik i børnehavers praksis med legetøj. Fokuserer på skalaen læringsaktiviteter, hvor Norge har scoren minimal kvalitet (3.2 ECERS-R), særligt grundet ikke	Strategisk caseudvælgelse (høj og lav ECERS-R-score på læringsaktivitetsskalaen)	Kvalitativt

Titel	Formål/Tema	Data fra GoBaN	Kvantitativt vs. kvalitativt studie
	tilstrækkelig adgang til legetøj.		
Baustad & Bjørnestad, 2020	Proceskvalitet: personalets interaktion med børnene i dagtilbud.	Observationsdata (ITERS-R, ECERS-3 + CIP).	Kvantitativt
Alvestad et al., 2019	Hvad karakteriserer børnehaver med pædagogiske læringsmiljøer af høj kvalitet? Opfølgning på ECERS-R resultater	Strategisk udvælgelse. (dagtilbud med høj score)	Kvalitativt.
Bjørnestad et al., 2019	Sammenhængsstudie: Betydning af persons uddannelse for proceskvalitet + validering af CIP-skalaen i en norsk kontekst	Observationsdata (ITERS-R + CIP) Strukturdata (personalets uddannelse)	Kvantitativt
Løkken et al., 2018a	Sammenhæng ml. kvalitet (fokus på interaktion) og børns socioemotionelle udvikling	Observationsdata (ITERS-R, udvalgte punkter) Betydning for børnene: Competence in Preschool scale (LSCIP)	Kvantitativt
Eliassen et al., 2018	Sammenhæng ml. kvaliteten af det pædagogiske læringsmiljø og børns kognitive udvikling	Observationsdata (ITERS-R) Betydning for børnene: To deltests fra "British Ability Scales 3"	Kvantitativt
Løkken et al., 2018b	Sammenhæng ml. to strukturelle faktorer og interaktionskvalitet. De to strukturelle faktorer er 1) organiseringsform (stabile eller fleksible grupper) og 2) normering	Observationsdata (ITERS-R, fokus på underskalaerne Interaktion, Sprog og Literacy og organisationsstruktur)	Kvantitativt

Erfaringer med at undersøge kvalitet nationalt

© 2022 Danmarks Evalueringsinstitut

Citat med kildeangivelse er tilladt

Publikationen er kun udgivet i elektronisk form på: www.eva.dk

ISBN (www) 978-87-7182-609-8

Danmarks Evalueringsinstitut (EVA) gør uddannelse og dagtilbud bedre. Vi leverer viden, der bruges på alle niveauer – fra institutioner og skoler til kommuner og ministerier.



**DANMARKS
EVALUERINGSINSTITUT**

T 35 55 01 01
E eva@eva.dk
H www.eva.dk