

**MASTEROPPGAVE**  
**Læring i komplekse systemer**  
**Juni 2018**

Effekten av lyttetrening for etablering av intraverbale responser

Line Børtveit

**Fakultet for helsefag**  
**Institutt for atferdsvitenskap**

**OsloMet – storbyuniversitetet**

## Forord

Jeg ønsker først og fremst takke deltaker i prosjektet beskrevet i artikkel 2. Takk for at du gladelig stilte opp og svarte på spørsmålene mine, gang på gang. Takk for at jeg fikk bli litt kjent med deg og ta del av ditt univers. Jeg vil også takke deltakers foreldre for at jeg fikk denne muligheten. Jeg retter også en stor takk til deltakers støttepedagog for gode innspill, tips og råd underveis i prosjektet, hjelp med registreringer og SMSer om turdager, takk!

Videre vil jeg også rette en takk til Astri Valmo ved Senter for Tidlig Intervensjon som satte meg i kontakt med deltaker og var behjelpelig med kartlegging og informasjon om deltaker.

Videre må jeg få takke Oda Marie Vister for all hjelp ved gjennomføring av begge artiklene. Takk for at du har stilt opp og hjulpet meg langt over det jeg kan gjengjelde deg til din masteroppgave. Takk for effektive og ikke fult så effektive skrive dager og takk for at du har holdt ut med alle mine spørsmål underveis.

Jeg vil også rette en stor takk til Svein Eikeseth, Inger- Karin Almås og Børge Strømgren for all veiledning og hjelp jeg har fått underveis. Inger- Karin må takkes uvurderlig hjelp ved gjennomføring av prosjektet i artikkel 2, takk for svar på utallige spørsmål og mailkorrespondanser døgnet rundt. Børge Strømgren skal ha mye av æren for artikkel 1. Takk for tålmodig veiledning og for at jeg fikk muligheten til å skrive en slik type artikkel. Det har vært veldig lærerikt og krevende, men har også gitt mersmak. Takk for all hjelp med databehandling, tips til relevant litteratur og tilbakemeldinger på utkast. Hovedveileder Svein Eikeseth vil jeg takke for all hjelp og gode innspill ved planlegging, gjennomføring og evaluering av prosjektet i artikkel 2. Jeg vil takke deg for konstruktive og grundige tilbakemeldinger på utkast til begge artiklene og dessuten oppmuntrende ord underveis.

Takk til familie og venner, og ikke minst Egil og Vigdis Røtterud for at dere har bidratt til at hverdagen min har «gått opp» gjennom hele utdanningsløpet.

## Innhold

<b>Forord</b> .....	ii
<b>Oversikt figurer, tabeller og appendiks</b> .....	v
<b>Sammendrag</b> .....	vi
<b>Abstract</b> .....	vii

### Artikkel 1

<b>En systematisk review av effekten av lyttertrening for etablering av intraverbale responser</b> .....	8
Sammendrag .....	9
Introduksjon.....	10
Lyttertrening .....	10
Evidensbasert praksis.....	11
Vurdering av effekt .....	13
Tidligere reviews .....	13
Formål .....	14
Metode .....	15
Utvelgelse av studier.....	15
Kvalitetsvurdering.....	17
Vurdering av evidens .....	18
Vurdering av effekt ved effektstørrelse og metaanalyse.....	20
Resultater .....	22
Vurderte og utvalgte studier.....	22
Kvalitetsvurdering.....	22
Vurdering av evidens .....	23
Vurdering av effekt ved effektstørrelse og metaanalyse.....	24
Diskusjon .....	24
Litteratursøket .....	25
Deskriptiv vurdering .....	26
Metodologisk kvalitet .....	26
Visuell analyse .....	29
Vurdering av evidens .....	29
Vurdering av effektstørrelse og metaanalyse.....	31
Visuell analyse og effektstørrelse .....	32

Konklusjon .....	33
Referanser .....	34
Tabell 1-4.....	39
Figur 1-3 .....	44
Appendiks A-B .....	47

## Artikkel 2

<b>En single-subjekt-design studie av effekten av lyttertrenting for emergens av intraverbale responser hos en fem år gammel gutt med autisme .....</b>	<b>53</b>
Sammendrag .....	54
Introduksjon.....	55
Metode .....	65
Deltaker.....	65
Kartlegging .....	65
Setting og materiell .....	66
Avhengig variabel og reliabilitet.....	68
Prosedyre.....	69
Prosedyreintegritet .....	72
Design .....	73
Resultat .....	73
Trening .....	73
Emergente intraverbaler .....	74
Emergente bidireksjonale intraverbaler .....	75
ABLLS-r .....	75
Diskusjon.....	75
Konklusjon.....	85
Referanser .....	86
Tabell 1- 4.....	92
Figur 1.....	96
Appendiks A-C .....	97

## Oversikt figurer, tabeller og appendiks

### Artikkel 1

<b>Tabell 1:</b> Artikler vurdert for inklusjon basert på fulltekst	<b>39</b>
<b>Tabell 2:</b> Deskriptive karakteristikk ved inkluderte studier	<b>40</b>
<b>Tabell 3:</b> Total skåring av CECs kvalitetsindikatorer og effekt	<b>41</b>
<b>Tabell 4:</b> Resultat metaanalyse	<b>42</b>
<b>Figur 1:</b> Flytdiagram litteratursøk i databaser	<b>43</b>
<b>Figur 2:</b> Skårene for de inkluderte studiene på CECs kvalitetsindikatorer	<b>44</b>
<b>Figur 3:</b> Forest plott, effektstørrelse og konfidensintervall	<b>45</b>
<b>Appendiks A:</b> CECs metodologiske kvalitetsindikatorer og kriterier for skåring	<b>47</b>
<b>Appendiks B:</b> CECs klassifisering av evidens	<b>51</b>

### Artikkel 2

<b>Tabell 1:</b> Språklige ferdigheter	<b>92</b>
<b>Tabell 2:</b> Når-spørsmålene, målresponser og bildestimuli som ble anvendt.	<b>93</b>
<b>Tabell 3:</b> Spørsmål brukt ved testing for emergente bidireksjonale intraverbaler	<b>94</b>
<b>Tabell 4:</b> Resultater ved test for emergente bidireksjonale intraverbaler	<b>95</b>
<b>Figur 1:</b> Emergente intraverbale responser	<b>96</b>
<b>Appendiks A:</b> Prosedyreintegritet lyttetrening	<b>97</b>
<b>Appendiks B:</b> Prosedyreintegritet tacttrening	<b>98</b>
<b>Appendiks C:</b> Prosedyreintegritet testing	<b>99</b>

## Sammendrag

Lyttertrening er foreslått som en mulig effektiv metode for etablering av intraverbale responser. Innenfor anvendt atferdsanalyse legges det vekt på at tiltak som gjennomføres skal ha empirisk støtte for effekt og at målpersoner skal ha tilgang på de mest effektive tiltakene som finnes tilgjengelig. For å kunne vurdere effekt av tiltak er empiriske studier av tiltakene nødvendige. Denne masteroppgaven består av to artikler der effekten av lyttertrening for etablering av intraverbaler undersøkes.

Artikkel 1 er en systematisk review av tidligere gjennomførte studier som undersøker effekten av lyttertrening på intraverbal respondering. I artikkelen identifiseres slike studier og den metodologiske kvaliteten vurderes. Det blir gjennomført visuelle analyser av effekt, og gjort vurdering av i hvilken grad lyttertrening kan klassifiseres som evidensbasert praksis. Videre blir data hentet ut fra studiene og det blir estimert effektstørrelser. Effektstørrelsene danner utgangspunkt for en metaanalyse. Resultatene viste at den metodologiske kvaliteten var relativt god. Metoden ble klassifisert som mulig evidensbasert. Effekten av tiltaket var noe varierende og metaanalysen viste at variasjonen var for stor til at den kombinerte effektstørrelsen kan sies å være meningsfull. Det er behov for flere studier for å kunne trekke ytterligere konklusjoner om eventuell effekt av metoden.

Artikkel 2 er en singel-subjekt studie av effekten av lyttertrening for emergens av intraverbale responser hos en fem år gammel gutt med autismediagnose. Det ble gjennomført lyttertrening med ni ulike «når-spørsmål» fordelt til tre sett. For det første settet var det nødvendig med tacttrening for å oppnå mestring under testbetingelsen, for sett 2 og 3 ble mestring oppnådd umiddelbart etter lyttertrening. Det ble videre også testet for emergens av bidireksjonale intraverbaler, resultatene viste at dette oppstod for fire av de ni spørsmålene det ble trent på.

## Abstract

Listener training is proposed as a potentially effective method for establishing intraverbal behavior. Applied behavior emphasizes the importance of empirically supported interventions and that individuals receiving interventions should have access to the most effective interventions available. To assess the effects of interventions, empirical studies of the interventions are necessary. This master thesis consists of two articles that explore the effect of listener training on intraverbal behavior.

Article 1 is a systematic review of studies examining the effect of listener training on intraverbal responding. Studies were identified and subsequently an evaluation of the methodological quality was performed. A visual analysis was conducted and to what extent listener training could be classified as evidence-based practice was assessed. Furthermore, data were retrieved from the studies, and effect sizes were estimated. The effect sizes were the basis for a meta-analysis. The results indicated that the overall methodological quality was high. Listener training was classified as a potentially evidence-based practice. The effect of the intervention showed variability and the meta-analysis demonstrated that the variation was too large to make the combined effect size meaningful. Further research is recommended for further conclusions about the effects of the method.

Article 2 is a single-subject design study of the effect of listener training on the emergence of intraverbal responding in a five-year-old boy with autism. Listener training was conducted on nine “when-questions” in three sets. In the first set, tact training was necessary to achieve the mastery criterion in the test condition. In set 2 and 3, mastery was achieved after listener training. A test for emergence of bidirectional intraverbals was also conducted. The results indicated that bidirectionality was achieved for four of the nine when-questions.

**En systematisk review av effekten av lyttertrening for etablering  
av intraverbale responser**

Line Børtveit

OsloMet- storbyuniversitetet



### Sammendrag

Gjennom et systematisk litteratursøk ble det identifisert studier der effekten av lyttertrening på intraverbal respondering studeres ( $n=8$ ). Det ble gjennomført en deskriptiv vurdering. Deretter ble *Council for Exceptional Children* (CEC) standarder for evidensbasert praksis anvendt for å vurdere den metodologiske kvaliteten til de inkluderte studiene. Effekten av lyttertrening ble ved en visuell analyse undersøkt i studiene som hadde tilstrekkelig metodologisk kvalitet. Kvalitet og effekt var utgangspunktet for å vurdere om lyttertrening kunne klassifiseres som evidensbasert praksis basert på retningslinjene i CEC. Resultatene viste at seks av de åtte inkluderte studiene hadde tilstrekkelig metodologisk kvalitet for å inkluderes i vurderingen av evidens. Fire studier viste positive effekter, de resterende hadde nøytral eller blandet effekt. Lyttertrening for etablering av intraverbal respondering ble klassifisert som *mulig evidensbasert praksis*. Det ble videre gjennomført en vurdering av effekt basert på effektstørrelse (Cohen's  $d$ ) for single-subjekt design og en metaanalyse. Metaanalysen viste en stor kombinert effektstørrelse,  $d=1.21$ , for de fem studiene som kunne inkluderes i analysen, estimatet hadde et stort konfidensintervall CI 95% [-1.63, 4.06] og  $I^2$  (85%) indikerte at det var for stor variasjon mellom studiene til at den kombinerte effektstørrelsen kunne vurderes å være meningsfull. Videre forskning på området diskuteres som nødvendig.

*Nøkkelord:* systematisk review, singel-subjekt design, council for exceptional children standarder, lyttertrening, metaanalyse, effektstørrelse.

En systematisk review av effekten av lyttertrening for etablering av intraverbale responser

### **Lyttertrening**

*Lyttertrening* innebærer ofte forsterkning for utvelgelse av visuelle stimuli ved at målpersonen skal peke eller ta på objekter/bilder i nærvær av verbale foranledigende stimuli. En enkel diskriminasjon er etablert når for eksempel barnet peker på nese (respons) når trener sier «pek på nese» (*verbal diskriminativ stimulus* ( $S^D$ )). En kondisjonale diskriminasjon er etablert når barnet peker på en dukke når dokken ligger på bordet sammen med andre ulike objekter og trener sier «pek på dukke». I en kondisjonale diskriminasjon er den verbale foranledigende stimulus en *kondisjonale stimulus* som etablerer ulike visuelle stimuli som diskriminative stimuli ( $S^D$ ) og *Sdelta* ( $S\Delta$ ) for seleksjonsresponser. Den kondisjonale stimulus kan ha ulike grader av kompleksitet og kan være sammensatt av flere stimuli. For eksempel krever spørsmålet «pek på den røde bilen» at målpersonen kan diskriminere mellom ulike farger og skille «bil» fra andre objekter når seleksjonsresponser skal fremvises. Å kunne diskriminere mellom komplekse verbale stimuli beskrives som en forutsetning for kompleks verbal atferd, som for eksempel intraverbaler (Eikeseth & Smith, 2013).

Intraverbal atferd beskrives av Skinner (1957) som verbal atferd som foranlediges av verbal atferd og der konsekvensene er såkalte generaliserte betingede forsterkere. Ved intraverbal atferd er det ikke punkt-til-punkt overensstemmelse mellom de verbale foranledigende stimuli og snakkerens respons. Eksempler på intraverbale responser kan være å respondere «eple» på spørsmål om å nevne en frukt, svare på oppgaver som «hva er to ganger to» eller å oppgi refrenget etter å ha hørt første vers av en sang. I følge Sundberg (2014) er intraverbal atferd avgjørende for å kunne holde sammenhengende samtaler med andre, Skinner beskriver samtaler som *smal talk* som stort sett intraverbal atferd. En kan derfor argumentere for at etableringa av et intraverbalt repertoar er av stor betydning for å kunne kommunisere og samtale med andre.

For de fleste barn er det sjeldent nødvendig med spesielle tiltak for utvikling av et funksjonelt intraverbalt repertoar, derimot kan tiltak være nødvendig for noen, for eksempel personer med autisme (Sundberg & Sundberg, 2011). Fordi disse ferdighetene kan være krevende å etablere argumenterer Eikeseth og Smith (2013) for at det kan være hensiktsmessig å trene inn disse ferdighetene via lyttertrening. På den måten behøver ikke målpersonen å både trene på diskriminasjon av komplekse verbale stimuli og samtidig produsere komplekse verbale responser, da responsen under lyttertrening er ikke-vokal. De diskuteres også at en annen fordel med lyttertrening er at treningen er mulig å gjennomføre selv for individer med begrenset vokalt repertoar. Intraverbale responser som oppstår som følge av lyttetrening vil være emergente intraverbaler, da det under lyttertrening ikke trenes direkte på disse responsene, men utelukkende trenes på utvalgsresponser i form av peking eller lignende.

### **Evidensbasert praksis**

Det er en målsetning innenfor anvendt atferdsanalyse at de tiltak som gjennomføres skal ha empirisk støtte for effekt og utgjøre en reell forskjell for personen hvis atferd tiltaket gjennomføres ovenfor (Baer, Wolf, & Risley, 1968). Personer som mottar atferdsanalytiske tiltak har rett på det mest effektive tiltaket som finnes tilgjengelig (Van Houten et al., 1988). Når det gjelder etablering av intraverbal atferd er det usikkerhet knyttet til hvilke tiltak som er mest effektive og treningsprosedyrer baseres ofte på prosedyrer som har vist seg effektive ved etablering av andre typer verbale operanter (Daar, Negrelli, & Dixon, 2015).

Slocum et al. (2014) argumenterer for at tiltak innenfor atferdsanalyse skal være databaserte og justeringer og endringer skal gjøres på bakgrunn av data. I mange anvendte settinger er det ifølge dem et gap mellom det som utledes fra forskning og det som implementeres i praksis, og de argumenterer for å ta i bruk *evidensbasert praksis* som et rammeverk for å ta gode avgjørelser når det gjelder valg av tiltak. For å kunne ta slike

avgjørelser er det nødvendig at det blir gjennomført vurderinger av kvaliteten på tilgjengelige tiltak for å avgjøre hva som er best tilgjengelige praksis, slik at gapet mellom forskning og praksis kan reduseres (Slocum et al., 2014). Innenfor andre fagområder, som for eksempel medisin, er gullstandarden en systematisk og statistisk gjennomgang av flere randomiserte kontrollerte studier (RCT) for å avgjøre hvilke tiltak som bør velges. Innenfor atferdsanalyse foretrekkes det å gjennomføre studier med en single-subjekt eksperimentelt design der det blir foretatt gjentatte målinger av samme atferd hos samme individer flere ganger, for å avdekke funksjonelle forhold mellom tiltak og endring i atferd (Smith, 2013). I følge Horner et al. (2005) vil studier med singel-subjekt design kunne spille en viktig rolle i utviklingen av evidensbasert praksis innenfor spesialpedagogikkfeltet. Det er derfor nødvendig å utarbeide metoder for å kunne vurdere kvalitet og effekt av tiltak som studeres ved bruk av single-subjekt design. Gode metoder for vurdering og kvalitet vil gjøre det mulig å sammenligne ulike tiltak på måter som tilsvarer lignende vurderinger av RCT-studier. Det er også fordelaktig om det er mulig å vurdere effekt og kvalitet på tvers av forskningsdesign, slik at alle relevante studier kan inkluderes i reviews og metaanalyser uavhengig av design.

*Council for Exceptional Children (CEC)* har utviklet en systematisk tilnærming for å (a) vurdere den metodologiske kvaliteten av individuelle studier gjennom å undersøke hvorvidt hver studie oppfyller en rekke kvalitetsindikatorer og (b) vurdere hvorvidt tiltak som klassifiseres som evidensbaserte på bakgrunn av vurderingene av metodologisk kvalitet og vurderinger av effekt (Cook et al., 2015). Denne tilnærmingen har en rekke styrker. For det første ble standarden ved utarbeiding systematisk gjennomgått av 23 fagpersoner med gjennomsnittlig mer enn 14 års erfaring fra fagfeltet gjennom en *Delphi-prosedyre*. Resultatet var at det ble oppnådd 98% enighet om de ulike kvalitetsindikatorerne og kravene for å vurdere en prosedyre som evidensbasert. Videre har kvalitetsindikatorerne gjennom tester vist seg å kunne gi passende grad av mellomobservatørenighet, noe som innebærer at de er

utformet slik at subjektivitet kanskje i noe mindre grad påvirker resultatene. Det er også en styrke at det er mulig å sammenligne både gruppe- og single-subjekt studier når evidens skal vurderes (Cook et al., 2015).

### **Vurdering av effekt**

For å vurdere effekt av tiltak blir det ved gruppestudier (RCT) ofte anvendt estimater av effektstørrelse, noe som gjør det mulig å sammenligne effekter fra studier med ulike design og utfallsmål. Det finnes flere ulike mål på effektstørrelse som anvendes og er anerkjente innenfor gruppestudietradisjonen (Hedges, Pustejovsky, & Shadish, 2012). Shadish, Hedges, Horner, og Odom (2015) beskriver at det er et problem at data fra single-subjekt studier ofte utelates fra reviews eller at resultater derfra holdes atskilt fra resultater fra gruppestudier. Dette kan ifølge dem forklares med at de som gjennomfører reviews kan være usikre på hvordan resultater fra single-subjekt design skal tolkes, kombineres og sammenlignes, og hvordan det kan estimeres en kombinert effektstørrelse for flere singel-subjekt studier. Hedges et al. (2012) og Hedges, Pustejovsky og Shadish (2013) presenterer en modell for estimering av effektstørrelse for singel-subjekt design som er direkte sammenlignbar med Cohen's  $d$ . Cohen's  $d$  beregnes ofte ved sammenligninger av gruppestudier. Denne tilnærmingen kan være hensiktsmessig fordi den gjør det mulig å uttrykke effekt på samme måte på tvers av eksperimentelle design og legger dermed til rette for større grad av inkludering av resultater fra single-subjekt studier i reviews og metaanalyser.

### **Tidligere reviews**

Det ser ikke ut til at det tidligere er gjennomført systematiske reviews med vurderinger av metodologisk kvalitet og sammenligninger av effekt for studier der effekten av lyttetrening for etablering av intraverbale responser vurderes. Det ser heller ikke ut til at det er gjennomført slike vurderinger av andre metoder for å etablere intraverbale. Derimot er det gjennomført noen litteraturgjennomganger. Cihon (2007) gir i sin litteraturgjennomgang en

oversikt over ulike metoder hadde blitt anvendt for etablering av intraverbale responser. Hun fokuserer først og fremst på *Precision Teaching*, men vurderer også metoder som *Peer mediated interventions*, overføring av stimuluskontroll, bruk av videomodell, *Conversations skills training*, *Discrete trial training* og *Direct instruction*. Styrker og begrensninger ved de ulike metodene blir vurdert. Cihon beskriver de fleste av metodene som effektive, men beskriver ikke bakgrunnen for denne vurderingen.

Axe (2008) gir i sin litteraturgjennomgang beskrivelser av en rekke studier gjort for å etablere intraverbale responser. Fokuset i denne gjennomgangen var etablering av kondisjonale diskriminasjoner. Han beskriver at den mest hensiktsmessige metoden for etablering av en kondisjonal diskriminasjon vil være ved å fokusere på etableringa av to ulike repertoarer: (1) å etablere stimuluskontroll slik at responser kommer under kontroll av ulike verbale stimuli og (2) trening på å fremvise bestemte verbale responser i nærvær av bestemte verbale stimuli. Axe gir, på samme måte som Cihon (2007) ingen systematiske vurderinger av ulike metoders effekt basert på sammenligninger av data. Det er også verdt å bemerke at begge disse litteraturgjennomgangene er gjennomført før publisering av de fleste studier der lyttertrening gjennomføres for etablering av intraverbaler.

Aguirre, Valentino, og LeBlanc (2016) gjør en gjennomgang av publiserte studier fra 2005 til 2015 på etableringa av intraverbaler. De beskriver hvordan de publiserte studiene fordeles på tre hovedområder: studier med direkte trening på intraverbaler (19 studier), emergens av en verbal operant etter trening på en annen (31 studier) og problemløsning (3 studier). Som med de andre litteraturgjennomgangene inkluderer Aguirre et al. ingen systematiske vurderinger av metodenes effekt eller studienes metodologiske kvalitet.

## **Formål**

Med denne reviewartikkelen var formålet å foreta en systematisk gjennomgang av studier som undersøker effekten av lyttertrening på etableringen av intraverbal atferd.

Formålet var å (a) identifisere studienes deskriptive variabler, (b) vurdere studienes metodologiske kvalitet basert på CECs kvalitetsindikatorer, (c) vurdere om lyttertrening kan klassifiseres som en evidensbasert praksis for etablering av intraverbale responser basert på CECs definisjoner og (d) sammenligne effekten av lyttertrening i gjennomførte studier ved estimering av effektstørrelse og deretter foreta en metaanalyse.

## Metode

### Utvelgelse av studier

Studiene ble identifisert på to måter, søk i databaser og ved gjennomgang av referanselister.

**Inklusjonskriterier.** Det ble i forkant av litteratursøket satt fem kriterier for å inkludere studiene: (a) studiene måtte være publisert og fagfellevurdert i et vitenskapelig tidsskrift, (b) artiklene måtte være empiriske. På bakgrunn av dette ble reviewartikler og andre ikke-eksperimentelle studier ekskludert. Også studier med begrensninger i eksperimentelt design, for eksempel der det utelukkende forelå pre-og posttestdata eller ble benyttet AB-design over en deltaker ble ekskludert. Videre måtte (c) studiene som ble inkludert være publisert på engelsk, norsk, svensk eller dansk, (d) studiene måtte ha minst en avhengig variabel som var direkte observerbare intraverbale responser og (e) den uavhengige variabelen måtte være lyttertrening.

**Søkestrategi.** Litteratursøket ble startet 8.03.2018 og avsluttet 9.03.2018. Databasene *Oria*, *Educational Resources Information Center (ERIC)*, *Pubmed Central*, *PsycINFO*, *ProQuest* og *Scopus* ble brukt. I alle databasene ble søk gjennomført med søkeordene «intraverbal», «listener» og «training», kombinert med den boolske operatoren «AND». Ved bruk av de ulike databasenes søkefunksjoner ble treff som ikke var i tråd med inklusjonskriteriene ekskludert, for eksempel var det i alle databaser mulighet for å begrense funnene til at de skulle være publiserte. De gjenstående funnene ble deretter lagt inn i Zotero,

et verktøy for å organisere referanser, der ble duplikater fra de ulike databasene fjernet. De organiserte funnene i Zotero ble videre eksportert til Microsoft Word og satt i tabeller der de ulike inklusjonskriteriene også var beskrevet, for å gjøre gjennomgangen lettere. Forfatter og en medstudent gikk deretter gjennom tabellene i Microsoft Word og skåret uavhengig av hverandre om artiklene skulle inkluderes eller ekskluderes. Funnene ble i første omgang vurdert basert på tittel og sammendrag i henhold til inklusjonskriteriene. De artiklene som etter vurdering av tittel og sammendrag forelå som relevante ble deretter mer detaljert undersøkt ved at forfatter og medstudent gjennomgikk fulltekstversjoner av artiklene.

Enigheten mellom forfatter og medstudent når det gjaldt inklusjon og eksklusjon av artiklene funnet ved søk i databasene var 99,73% (utregnet ved at antall enige ble dividert på antall enige pluss antall uenige, multiplisert med 100%). Forfatter hadde inkludert en studie som medstudenten ikke hadde inkludert (Miguel, Petursdottir, & Carr, 2005). Årsaken til at artikkelen ble ekskludert av medstudenten var at det i tittelen ble skrevet *reseptiv diskriminasjonstrening* og ikke *lyttertrenting*, og artikkelen ble dermed ekskludert basert på tittel. Ved gjennomlesning av metodedelene var forfatter og medstudent enige i at begrepene var synonyme og at artikkelen burde inkluderes i reviewartikkelen.

Videre gjennomførte forfatter og medstudent referanselistene på de artikler som ble funnet relevante ved søket i databasene. Duplikater av studier som allerede var vurdert ble fjernet ved bruk av Zotero. De resterende artiklene ble deretter eksportert til Microsoft Word og vurdert i henhold til samme inklusjonskriterier som studiene fra databasesøkene, først basert på tittel og sammendrag og deretter fulltekst.

Enigheten mellom forfatter og medstudent når det gjaldt inklusjon og eksklusjon av artiklene funnet ved gjennomgang av referanselister var 100%.



## Kvalitetsvurdering

**Deskriptiv vurdering.** En deskriptiv vurdering av karakteristikk ved deltakere, setting og tiltaket ble gjennomført for å beskrive innhold og gi en oversikt over variabler som mulig kunne påvirke resultatene i de ulike studiene. Seks variabler ble beskrevet: (a) deltaker, (b) alder, (c) kjønn, (d) tilstand, (e) hvilken type intraverbaler som ble målt og (f) karakteristikk ved settingen der tiltaket ble gjennomført.

**Metodologisk kvalitetsvurdering.** Forfatter og medstudent vurderte den metodologiske kvaliteten av de utvalgte artiklene ved bruk av en vurderingsprotokoll (Lane, Common, Royer, & Muller, 2014) for Microsoft Excel som var basert på kvalitetsindikatorer beskrevet av CEC (Cook et al., 2015). Forfatter og medstudent skåret de samme artiklene uavhengig av hverandre og vurderte de åtte kvalitetsindikatorer med underliggende komponenter, totalt var det 22 komponentene som var gjeldende for single-subjekt design. Se Appendix A for detaljert beskrivelse av kvalitetsindikatorer og kriterier for skåring. Det ble skåret om den aktuelle studie (1) oppfylte eller (0) ikke oppfylte hver av komponentene i kvalitetsindikatorer. Forfatter og medstudenten hadde i forkant av skåringen sammen skåret tre artikler ved bruk av vurderingsprotokollen til 100% enighet for å i noe større grad sikre likere vurderinger. Under denne treningen ble tre artikler som tidligere var vurdert av Common et al. (2017) ved bruk av CECs kvalitetsindikatorer anvendt slik at eventuell uenighet kunne løses ved å kontrollere hvordan artiklene tidligere var skåret.

Enigheten mellom de to observatørene ved skåring av kvalitetsindikatorer ble beregnet ved at antallet kvalitetsindikatorer komponenter der det var skåret likt ble dividert på det totale antallet skårer, multiplisert med 100%. Enigheten mellom forfatter og medstudent var 94,43%. Det ble også beregnet kappa,  $k=0.98$ . Uenigheten var ved en komponent for en av de åtte inkluderte studiene (komponent 2,1 for Miguel et al. (2005)). Uenigheten ble løst ved at komponenten ble diskutert med forfatter og medstudentens biveileder til enighet.

## Vurdering av evidens

For å avgjøre om lyttertrening kunne defineres som en evidensbasert praksis for etablering av intraverbaler ble (Cook et al., 2015) prosedyrer for vurdering av evidens anvendt. Cook et al. beskriver fem kategorier med kriterier for klassifisering av evidens (a) evidensbasert praksis, (b) mulig evidensbasert praksis, (c) blandet evidens, (d) ikke tilstrekkelig evidens og (e) negative effekter, se Appendix B for en kort beskrivelse av kriteriene for de ulike kategoriene gjeldende for singel-subjekt design, se Cook et al. for detaljert beskrivelse.

Studier der det var brukt et eksperimentelt design: single-subjekt (ABA, multiple baseline, changing criterion eller alternating treatment) eller RCT og som på bakgrunn av kvalitetsindikatorerne ble vurdert å ha tilstrekkelig metodologisk kvalitet kunne ifølge Cook et al. (2015) vurderes. I vurderingsprotokollen (Lane et al., 2014) ble det basert på skåring av komponentene automatisk kalkulert en absolutt og en sammenlagt skåre for hver studie. Den absolutte skåren viste til antall indikatorerne der alle komponentene ble skåret som oppfylt. Den sammenlagte skåren viste til summen av skårene hver indikator hadde. For eksempel om alle komponentene ble skåret som oppfylt fikk indikatoren en skåre på 1, om halvparten av komponentene var oppfylt fikk indikatoren en skåre på 0.5. Det var den sammenlagte skåren som ble lagt til grunn for vurderingen av metodologisk kvalitet, den måtte være på minst 6.4 (80%) for at studien ble vurdert å ha tilstrekkelig metodologisk kvalitet (Lane, Kalberg & Shepcaro, 2009) og kunne inkluderes i vurderinger av evidens.

Videre måtte det også vurderes om studiene som oppfylte kravene til metodologisk kvalitet viste positiv, negativ eller nøytral/blandet effekt av lyttertrening for intraverbal respondering. Bare studiene som hadde tilstrekkelig metodologisk kvalitet ble vurdert for effekt. (Cook et al., 2015) beskriver at vurderinger av effekt ved singel-subjekt studier gjøres ved visuelle analyser av grafene i studiene.

**Visuell analyse.** Det ble undersøkt om det så ut til å være et funksjonelt forhold mellom den uavhengige og den avhengige variabelen, og hvilken effekt den uavhengige variabelen så ut til å ha. Studier som ikke oppfylte kvalitetsindikator-komponent 6.5 ble automatisk ekskludert fra vurdering av evidens, da det var nødvendig med minst tre målinger der eventuell effekt av tiltaket kunne demonstreres, tre faseskifter. Det måtte altså være minst tre caser i studien, for eksempel tre deltakere, atferder, settinger eller grupper med deltakere.

Gjennom visuell analyse av de ulike studienes grafer undersøkte forfatter og medstudenten data på tvers av og innen ulike faser i forhold til endringer i nivå, trender, variabilitet, umiddelbarhet av endringer ved faseskifter og overlapp av data på tvers av faser. Forfatter og medstudent skåret uavhengig av hverandre om hver studie skulle klassifiseres å ha positiv, negativ eller nøytral/blandet effekt.

For at en studie kunne sies å ha positiv effekt måtte det være vist et funksjonelt forhold mellom tiltaket og meningsfull endring (av sosial eller praktisk verdi) i minst 75% av casene. Det kunne ikke være skadelige effekter for noen av casene, for eksempel om det så ut til å være et funksjonelt forhold mellom lyttertrening og reduksjon i verbale responser sammenlignet med baselinemålingen.

For at en studie skulle bli vurdert til å ha negativ effekt måtte det vises et funksjonelt forhold mellom tiltaket og skadelige effekter ved minst 75% av casene. Skadelige effekter ville vært om det var økning i atferd som ønskes redusert, eller reduksjon i atferd det ønskes mer av.

Nøytral eller blandet effekt ble tilskrevet studier det det ikke kunne konkluderes med positive eller negative effekter av tiltaket. For eksempel om det var positiv effekt for 50% av deltakerne, men ingen effekt eller negative effekter for de resterende deltakerne.

Enigheten mellom de forfatter og medstudent når det gjaldt de visuelle analysene var 100%, utregnet ved at antallet studier der forfatter og medstudent var enige om klassifisering av effekt ble dividert på antallet studier totalt, og deretter multiplisert med 100%.

### **Vurdering av effekt ved effektstørrelse og metaanalyse**

**Prosedyre for datainnhenting.** For å hente ut data fra studiene ble programvaren Webplot Digitizer (Rohatgi, 2015) brukt for de studiene som hadde data fremstilt i grafer der det ved visuell inspeksjon var problematisk å nøyaktig vurdere hvilke verdier hvert datapunkt hadde. Dataene ble ved bruk av programvaren konstruert digitalt basert på bilder av grafene, fra programvaren kunne dataene sendes til og lagres i Microsoft Excelformat. I de tilfeller der det ikke var usikkerhet knyttet til hvilke verdier de ulike datapunktene hadde ble verdiene ført direkte inn i en fil i Microsoft Excel. Bruk av lignende verktøy for å hente ut data er vurdert å være en reliabel fremgangsmåte og de uthentede dataene blir nær identiske rådataskårene (Shadish et al., 2009).

**Effektstørrelse.** Cohen's  $d$ , mellom-case standardisert forskjell i gjennomsnitt (*standardized mean difference*) ble beregnet for å beskrive den gjennomsnittlige effekten av tiltaket i hver enkelt studie. Beregningene ble gjort med bakgrunn i Hedges et al. (2012, 2013) som beskriver hvordan Cohen's  $d$  kan beregnes for studier med en singel-subjekt design. For å kunne beregne effektstørrelse på studiene var det nødvendig at de (a) hadde data fra minst tre deltakere og (b) hadde en tilbaketrekkingsdesign, for eksempel ABAB, eller et multiple baseline/probe design over deltakere (Hedges et al., 2013; Valentine, Tanner-Smith, Pustejovsky, & Lau, 2016).

Effektstørrelsen ble estimert ved bruk av appen *SCDHLM* 0.3.1 (Pustejovsky, 2016) for *RStudio* (RTeam, 2015) et program for statistiske beregninger. Appen har to ulike metoder for å gjøre estimater: *moment estimation* og *restricted maximum likelihood* (REML). I denne studien ble REML valgt fordi den ble beskrevet som den mest fleksible modellen (Valentine

et al., 2016). Det ble under estimering benyttet både *fixed* og *random effects* for *level* mellom baseline og tiltaksfasene. Level ble valgt da det ikke var trend i baseline og fordi dette er en enklere variant enn for eksempel lineær regresjon. Det anbefales bruk av en så enkel (*parsimonious*) modell som mulig og at samme modell brukes for alle studiene som inkluderes i metaanalysen.

Forfatter fikk praktisk hjelp av biveileder til å estimere effektstørrelsen.

**Metaanalyse.** Basert på effektstørrelsesestimaterne fra studiene der dette kunne beregnes ble det gjennomført en metaanalyse. Metaanalysen ble gjennomført ved bruk av Meta-Essentials (Suurmond, Rhee, & Hak, 2017) som er et verktøy for å gjøre metaanalyser og det ble brukt en «random effects» modell, som anbefales brukt. Meta-Essentials har syv ulike regneark som lastes ned, deriblant et for Cohen's *d*- effektstørrelse. I Meta-Essentials ble antall deltakere lagt inn for å beregne hvor stor andel av det totale utvalget hver enkelt studie utgjør, *vekt*. Videre ble det også lagt inn antall datapunkter for hele datasettet med singel-subjekt design, da dette ble brukt på samme måte som antall deltakere blir brukt ved estimeringer basert på gruppestudier (Valentine et al., 2016). Etter at effektstørrelse og *standard error* fra studiene man inkluderer i analysen var lagt inn ble det automatisk estimeres blant annet en kombinert effektstørrelse og dataene kunne fremstilles i en *forest plot* som inkluderte effektstørrelse med konfidensintervaller og vekt for hver inkludert studie, i tillegg til kombinert effektstørrelse med konfidensintervall og prediksjonsintervall for hele settet.

Forfatter fikk praktisk hjelp av biveileder til bruk av Meta-Essentials.

## Resultater

### Vurderte og utvalgte studier

Tabell 1 viser en oversikt over artikler som ble vurdert basert på fulltekst ved søk i databasene og gjennomgang av referanselister og årsaker til inklusjon eller eksklusjon av disse.

Figur 1 viser en oversikt over antallet studier funnet under litteratursøket i databasene og antallet studier derfra ekskludert og inkludert i denne reviewartikkelen. Det ble totalt identifisert 1163 unike studier, 1157 av disse som et resultat av søk i databasene. Av disse studiene ble åtte studier vurdert å oppfylle inklusjonskriteriene basert på gjennomlesning av fulltekst. Videre ble seks mulig aktuelle studier funnet ved gjennomgang av referanselistene på de relevante studiene. Fulltekst ble lest for disse seks studiene og de ble alle ekskludert på bakgrunn av inklusjonskriteriene.

### Kvalitetsvurdering

**Deskriptiv vurdering.** Åtte studier oppfylte inklusjonskriteriene ved litteratursøket, i studiene var det totalt 23 deltakere der effekten av lyttertrening på intraverbal respondering ble vurdert. Se tabell 3 for en oversikt over deskriptive karakteristikk ved studiene og deltakerne. Gjennomsnittlig alder på deltakerne var 5.52 år, varierende fra 3 til 15 år. De fleste deltakerne var jenter ( $n=13$ , 56.53%). Både studier gjennomført med normalfungerende barn og barn med diagnoser ble inkludert, de fleste deltakerne var normalfungerende ( $n= 13$ , 56.53%), ti deltakere hadde en autismediagnose (43.47%), og to deltakere hadde i tillegg til autismediagnosen også andre diagnoser. Det ble undersøkt ulike typer intraverbal respondering: kategorisering ( $n=6$ ), hva-spørsmål ( $n=9$ ), hvem-spørsmål ( $n=1$ ) og fremmedspråk ( $n=7$ ). For de fleste deltakerne ble lyttertrening gjennomført på et treningsrom eller en egen treningsplass på skole eller i barnehage ( $n=17$ ) for resterende deltakere var

lyttertrening gjennomført hjemme hos målpersonen. En studie rapporterte ikke setting for tiltaket.

**Metodologisk vurdering.** Figur 2 viser en oversikt over hvordan de åtte inkluderte studiene oppfylte kravene til CEC ulike kvalitetsindikatorer. Totalt ble 144 av de 176 mulige komponentene skåret som oppfylt på tvers av de åtte studiene, noe som utgjorde 81.81% av indikatorene. Gjennomsnittlig ble 81.97% komponenter skåret som oppfylt pr. studie, med en range fra 60% (Kodak & Paden, 2015) til 100% (Miguel et al., 2005; Petursdottir, Carr, Lechago, & Almason, 2008). Gjennomsnittlig absolutt skåre var 6.1 varierende fra 3 til 8. Gjennomsnittlig sammenlagt skåre var 6.82 varierende fra 4.8 til 8.

Tabell 3 viser en oversikt over absolutt og sammenlagt skåre av CECs kvalitetsindikatorer for de inkluderte studiene.

### **Vurdering av evidens**

To studier oppfylte alle CECs kvalitetsindikatorer (Miguel et al., 2005; Petursdottir, Carr, et al., 2008) og fire studier oppfylte minst 80% av indikatorene basert på sammenlagt skåre (Petursdottir & Haflidadottir, 2009; Petursdottir, Olafsdottir, & Aradottir, 2008; Smith et al., 2016; Vallinger-Brown & Rosales, 2014) disse studiene ble dermed vurdert å ha tilstrekkelig metodologisk kvalitet og ble inkludert i vurderingen av evidens. For fire av studiene (Petursdottir & Haflidadottir, 2009; Petursdottir, Olafsdottir, et al., 2008; Smith et al., 2016; Vallinger-Brown & Rosales, 2014) ble det etter visuell inspeksjon konkludert med positive effekter av lyttertrening for intraverbale responser, de to resterende studiene (Miguel et al., 2005; Petursdottir, Carr, et al., 2008) hadde nøytrale eller blandede effekter. Tabell 3 viser oversikt over prosentandel kvalitetsindikatorer oppfylt og effektvurdering basert på den visuelle analysen, for de av studiene der dette kunne vurderes.

Lyttertrening som metode for etablering av intraverbale responser blir i vurderingsprotokollen (Lane et al., 2014) definert som en *mulig evidensbasert praksis*, da det

er fire studier med tilstrekkelig metodologisk kvalitet som viser positiv effekt, ingen studier som viser negativ eller skadelig effekt og en 2:1 rate mellom studier som viser positiv effekt og studier som viser nøytral eller blandet effekt.

### Vurdering av effekt ved effektstørrelse og metaanalyse

**Effektstørrelse.** Figur 3 viser effektstørrelse og konfidensintervall (95%) for inkluderte studier og kombinert effektstørrelse, konfidensintervall og prediksjonsintervall for studiene samlet. Det ble ikke beregnet effektstørrelse for tre av studiene (Kodak & Paden, 2015; Petursdottir & Haflidadottir, 2009; Petursdottir, Olafsdottir, et al., 2008) da disse utelukkende hadde data fra to deltakere, noe som førte til at det ikke var mulig å beregne Cohen's  $d$  (Hedges et al., 2013). Fra studien til Miguel et al. (2005) er effektstørrelse bare beregnet for et av de to settene barna trente på, da appen *SCDHLM* 0.3.1 (Pustejovsky, 2016) for RStudio (RTeam, 2015) ikke kunne estimere effektstørrelse basert på data fra det ene settet (settet der det ble trent på ulike verktøy «tools»), selv ikke når data fra dette settet ble kombinert med data fra det andre settet.

**Metaanalyse.** Fem studier ble inkludert i metaanalysen. For alle studiene som ble inkludert i analysen vises effektstørrelse, konfidensintervaller og vekt i Figur 3. Nederst på Figur 3 viser også den kombinerte effektstørrelsen og konfidensintervallet i tillegg til prediksjonsintervallet (grønn linje). Tabell 4 viser verdier for kombinert effektstørrelse ( $d=1.21$ ), konfidensintervall (95 % CI = [-1.63, 4.06]), to-halet  $p$ -verdi ( $p = 0.24$ ), prediksjonsintervall [-2.92, 5.38] og heterogenitet, eller proporsjon av observert variabilitet i effekt mellom studiene ( $I^2 = 85\%$ ).

### Diskusjon

Formålet med denne reviewartikkelen var å identifisere og evaluere kvaliteten på studier som var gjennomført der effekten av lyttetrening på intraverbale respondering ble studert, undersøke om lyttetrening kunne anses å være en evidensbasert praksis og å



undersøke effekten av lyttertrening på intraverbal respondering ved estimering av effektstørrelse og gjennomføring av en metaanalyse. Resultatene viste at det var relativt høy metodologisk kvalitet på gjennomførte studier. Det kunne konkluderes med at lyttertrening kan anses å være en mulig evidensbasert praksis. Resultatet fra estimatene av effektstørrelse og metaanalysen viste at det var for stor variasjon og for få gjennomførte studier for å få en meningsfull kombinert effektstørrelse.

### **Litteratursøket**

Litteratursøket i databasene og gjennomgang av referanselister førte til at en stor mengde artikler ble undersøkt for inklusjon. Gjennomgang av tidligere litteraturgjennomganger viste at stort sett de samme studiene ble inkludert i denne reviewartikkelen som var inkludert i tidligere gjennomganger, med unntak av artiklene som ikke var publiserte da litteraturgjennomgangene ble skrevet.

Det kan beskrives som en styrke at det var relativt høy enighet mellom forfatter og medstudenten som gjennomførte litteratursøket. Likevel er det en mulighet for at begge observatørene har oversett studier. Med et såpass omfattende søkeresultat, med flere hundre studier å undersøke var det en risiko for at artikler ble oversett om det ikke tydelig fremgikk av tittel at det var gjennomført lyttertrening for etablering av intraverbaler. Det er også en fare for at studier der det var gjennomført flere ulike tiltak i tillegg til lyttertrening kan ha blitt oversett.

Det er selvsagt også muligheter for at relevante studier ble ekskludert på bakgrunn av inklusjonskriteriene. For eksempel kan kravet om at artiklene skulle være publiserte i fagfelleverderte tidsskrifter har ført til at en rekke studier har blitt ekskluderte. Upubliserte artikler, avhandlinger, master- og bacheloroppgaver på temaet kunne muligens har ført til et bedre og bredere utgangspunkt for vurderingene. Ved å utelukkende inkludere publiserte

artikler vil det også være en fare for at studier der det ikke er funnet effekt av tiltaket ikke inkluderes fordi de ikke er publiserte (publikasjonsbias) (Simes, 1986).

Det var stort overlapp mellom treffene i de ulike databasene. Dette kan muligens indikere at søkeordene som ble valgt var presise nok til at relevante studier forekom blant søkeresultatene. Det kan også være et tegn på at relevante studier tidlig ble identifisert og at det muligens ikke ville ført til flere funn om litteratursøket hadde blitt utvidet til ytterligere databaser.

### **Deskriptiv vurdering**

Den deskriptive vurderingen av studiene viste at det var relativt liten variasjon i de gjennomførte studiene. Alle var gjennomført med barn, de fleste i en skole- eller barnehagesetting. Fordelingen av kjønn var ca. halvparten jenter og halvparten gutter, noe som kan antas å være representativt for befolkningen. Det var heller ikke så stor variasjon i hvilken type intraverbaler det ble gjennomført trening på. Dette kan være et mulig argument for at studiene var noenlunde sammenlignbare og at for eksempel estimering av en kombinert effektstørrelse vil kunne forsvares.

Ca. halvparten av studiene var gjennomført med normalfungerende barn, resten av studiene var gjennomført for barn med autismediagnose. Dessverre var datagrunnlaget for lite til at det var mulig å gjennomføre statistiske analyser av om det var forskjell i effekt av lyttertrening for barna med og uten autismediagnose.

### **Metodologisk kvalitet**

Resultatene fra den metodologiske kvalitetsvurderingen basert på CECs kvalitetsindikatorer viste at de inkluderte studiene gjennomsnittlig hadde relativt høy metodologisk kvalitet, med over 80% av komponenten skåret som oppfylt. Bare to studier hadde under 80% av kvalitetsindikatorer oppfylt, hvorav den ene hadde en skåre på 78,7%.

Enigheten mellom forfatter og medstudent ved skåring av kvalitetsindikatorerne var høy, da det utelukkende var uenighet på skåring av en komponent for en av de åtte studiene. Dette er en mulig styrke ved reviewartikkelen da dette kan indikere at skåringene ikke utelukkende er basert på forfatters subjektive vurderinger, men at det er sikret og målt en viss grad av objektivitet.

Den kvalitetsindikatoren som i minst grad ble skåret som oppfylt var indikator 5.0 som omhandlet målinger av prosedyreintegritet. Bare tre av studiene rapporterte om målinger av dette (Miguel et al., 2005; Petursdottir, Carr, et al., 2008; Vallinger-Brown & Rosales, 2014). Manglende målinger av prosedyreintegritet førte også til lavere skårer på indikator 6.0, da 6.1 og 6.7 ikke kunne skåres som oppfylt om det ikke prosedyreintegritet var rapportert. Wolery (2011) beskriver at rapportering av prosedyreintegritet er avgjørende for at det skal kunne trekkes konklusjoner om effekten av en prosedyre. Uten objektive og jevnlig målinger av prosedyreintegritet kan det ikke konkluderes med funksjonelle forhold mellom intervensjon og effekt, da det er knyttet usikkerhet til hva som faktisk er gjennomført. Det er derfor en svakhet ved studiene som undersøkes i denne reviewartikkelen at under halvparten av dem rapporterer om målinger av prosedyreintegritet.

Videre må det diskuteres som en styrke at alle studiene rapporterer om målinger av mellomobservatørsenighet (komponent 7.5). Selv om tilstedeværelse av en ekstra observatør ikke sikrer at lyttertrening ble gjennomført i tråd med prosedyrebeskrivelsene kan det tenkes at det i noe større grad er trolig at dette er tilfellet enn om eksperimentator hadde gjennomført tiltaket helt alene. For eksempel krever ofte mellomobservatørmålinger at prosedyren gjennomføres i bestemte rekkefølger slik at registrering kan planlegges og gjennomføres effektivt mens trening eller testing pågår.

Når det gjelder bruk av CECs retningslinjer må det videre poengteres at de fleste studiene som er inkludert i denne reviewartikkelen er publisert før eller i 2014 ( $n=6$ ). Det er

derfor noe problematisk å stille absolutte krav til hva som skal rapporteres, da forfatterne ikke hadde tilgang til CECs retningslinjer ved publisering av sine resultater. Likevel viser resultatene at studiene stort sett oppfyller 80% kriteriet (Lane et al., 2009) for å kunne inkluderes i vurderingen av evidens.

Det kan videre diskuteres i hvilken grad det å møte 80% av kvalitetsindikatorkomponentene bør defineres som tilstrekkelig metodologisk kvalitet. Cook et al. (2015) beskriver at for at en studie skal defineres å ha tilstrekkelig metodologisk kvalitet må alle kvalitetsindikatorne skåres som oppfylt. I denne reviewartikkelen ville dette innebære at utelukkende to studier (Miguel et al., 2005; Petursdottir, Carr, et al., 2008) kunne inkluderes i vurderingen av evidens. Det ble vurdert som hensiktsmessig å ta utgangspunkt i Lane et al. (2009) som beskriver bruk av et 80% kriterium basert på sammenlagt skåre. Ved bruk av absolutt skåring risikerer man at studier som mangler en eller få komponenter for en indikator utelukkes for videre analyser. Ved å ta utgangspunkt i sammenlagt skåre vil hver komponent som skåres som oppfylt gjenspeiles i totalskåren. Vurderingsprotokollen (Lane et al., 2014) som ble avendt tok også utgangspunkt i 80% kriteriet og gjorde det mulig å beregne grad av evidens basert på studiene som skåret mer 80% (6.4) eller mer som sammenlagt skåre. Det kan selvsagt diskuteres i hvilken grad et 80% kriterium er hensiktsmessig for å tilskrive tilstrekkelig metodologisk kvalitet.

Det er videre også relevant å diskutere at de ulike komponentene er av ulik betydning. For eksempel argumenterer Lane et al. (2009) for at indre validitet er av vesentlig større betydning ved vurdering av evidens enn for eksempel komponenten som omhandler rapportering av sosial validitet. Sosial validitet er selvsagt viktig, men dette kan i mange tilfeller vurderes uten at det er nødvendig med rapporteringer om det i forskningsartikkelen. Det hadde derfor vært fordelaktig om det var mulig å gradere den totale skåren basert på innhold i de ulike indikatorene. Når det gjelder resultatene i denne reviewartikkelen kan det

på dette området diskuteres som problematisk at tre av studiene (Petursdottir & Haflidadottir, 2009; Petursdottir, Olafsdottir, et al., 2008; Smith et al., 2016) ikke rapporterte målinger av prosedyreintegritet (indikator 5.0), men likevel fikk en sammenlagt skåre på over 80% og kunne inkluderes i analysen av evidens. Indikator 5.0 burde muligens vurderes som så viktig for vurdering av funksjonelle forhold mellom intervensjon og målatferd at manglende skårer her burde utelukke studier for videre analyser.

### **Visuell analyse**

Den visuelle analysen som ble gjennomført for de seks studiene som hadde tilstrekkelig metodologisk kvalitet viste blandet effekt av lyttertrening som intervensjon. Det var flest studier som viste positive effekter ( $n=4$ ) og to studier som viste nøytral/blandet eller ingen effekt av lyttertrening. Ingen studier viste negative effekter. Det kan diskuteres som en fordel at den visuelle analysen ble gjennomført av to observatører som skåret uavhengig av hverandre, og at de var 100% enige, da visuelle analyser kan gi noe større rom for subjektive vurderinger enn for eksempel utregning av effektstørrelse.

### **Vurdering av evidens**

Seks av studiene kunne inkluderes i vurderingen av lyttertrening som evidensbasert praksis for etablering av intraverbaler. En slik evaluering var ansett å være nyttig å gjennomføre da det ikke så ut til å være gjort slike evalueringer tidligere og det er knyttet usikkerhet til valg av metode i praktiske settinger for etablering av intraverbaler (Daar et al., 2015). Som beskrevet tidligere er det av stor betydning å få samlet og vurdert tidligere forskning for å kunne ta informerte avgjørelser når det gjelder metodevalg (Slocum et al., 2014).

På bakgrunn av de seks inkluderte studiene ble konklusjonen at lyttertrening for etablering av intraverbale responser kan defineres som en mulig evidensbasert praksis. Dette indikerer at det er nødvendig med ytterligere forskning. For at en metode skal anses å være

evidensbasert basert på single-subjekt studier alene stilles det krav til at metoden er studert i minst fem ulike studier, som totalt sett har mer enn 20 deltakere. I denne reviewartikkelen ble seks studier med totalt 18 deltakere inkludert i analysen. Dette innebærer at det utelukkende var nødvendig med tre flere deltakere med positive effekter i en av studiene, eller ytterligere en studie med positive effekter for tre deltakere, for at lyttertrening som metode skulle kunne anses å være evidensbasert praksis for etablering av intraverbaler.

Det er ikke usannsynlig at det er gjennomført flere studier, med flere deltakere. Cook et al. (2015) uttrykker at det ikke er nødvendig at studier er publiserte i fagfelleverderte tidsskrift for at CECs retningslinjer skal kunne anvendes. En kan dermed tenke seg at om det i denne reviewartikkelen også var inkludert upubliserte studier som for eksempel bachelor og masteroppgaver kunne resultert i andre konklusjoner. Det er en svakhet ved denne reviewartikkelen at utvalget ved inklusjon ble begrenset til publisert og fagfelleverderte artikler. Likevel kan det tenkes at ved å sette dette kravet er studiene som ble inkludert allerede til en viss grad kontrollert for metodologisk kvalitet. Videre ville det også vært utfordringer knyttet til tilgjengeligheten av upubliserte artikler og det ville vært krevende å få oversikt over i hvilken grad flest mulig relevante studier ble identifisert ved litteratursøket. Cook et al. informerer om at studier som skal inkluderes må være offentlig tilgjengelige, noe som vil begrense muligheter for å inkludere upublisert materiale.

Det er som skrevet tidligere en utfordring at tre av studiene som ble inkludert i vurderingen av evidens ikke har rapportere målinger av prosedyreintegritet. For videre utvikling av standarder for vurdering av evidens burde det muligens stilles strenge krav til denne typen rapportering. Når prosedyreintegritet ikke er målt er det problematisk å tilskrive observerte effekter metoden, og dermed blir konklusjoner når det gjelder evidens også problematiske da det er usikkerhet knyttet til i hvilken grad resultatene er sammenlignbare (Wolery, 2011).

### Vurdering av effektstørrelse og metaanalyse

For fem av studiene var det mulig å beregne effektstørrelse, tre studier ble ekskludert på bakgrunn av for få deltakere. To av studien hadde signifikante effektstørrelser som ble vurdert å være store eller svært store basert på Cohen (1988) gjengitt av Ellis (2009). Den kombinerte effektstørrelsen var ikke meningsfull da variasjonen var stor ( $I^2=0.85$ ) og konfidensintervallet overlappet 0.

Det kan diskuteres som en styrke ved de gjennomført estimeringene og analysene at det ble benyttet apper (Pustejovsky, 2016) og ferdigprogrammerte filer (Suurmond et al., 2017) for å gjøre beregningene. Da dette trolig bidro til redusert risiko for å gjøre feilkalkuleringer. Det bør også trekkes frem som en fordel at forfatter fikk hjelp av biveileder til å gjennomføre disse analysene.

Det var noen utfordringer med estimering av effektstørrelse. For en av studiene (Miguel et al., 2005) var det ikke mulig å beregne effektstørrelse for alle dataene i studien. Et av settene de tre deltakerne hadde trent på måtte ekskluderes. Dette kan trolig skyldes at det var for lite variasjon i dataene for dette settet til at det var mulig å gjøre estimater i Rstudio (RTeam, 2015). Det kan tenkes at resultatene både for denne studien og den kombinert effektstørrelsen ville vært forskjellig om alle dataene hadde blitt inkludert.

Videre var det også problematisk at tre av de åtte studiene måtte ekskluderes fra analysene av effekt på bakgrunn av for få (færre enn tre) deltakere. Petursdottir og Haflidadottir (2009) hadde to deltakere som hver hadde trent på to ulike typer intraverbaler. Det innebar at det var mulig å gjøre fire sammenligninger mellom baselinemålinger og målinger gjort etter tiltaket. Petursdottir, Olafsdottir, et al. (2008) hadde to deltakere som hver hadde trent på fire ulike typer intraverbaler (sett og språk) noe som tilsvarte totalt åtte sammenligninger mellom baseline og målinger gjort etter tiltaket (åtte faseskifter). Denne studien var også den av alle studiene i denne reviewartikkelene som hadde flest muligheter for

demonstrasjon av effekt. Det kan derfor anses problematisk å utelukke disse to studiene på bakgrunn av få deltakere når det kan argumenteres for sterke demonstrasjoner av effekt.

Den store variabiliteten ( $I^2=85\%$ ) mellom studiene kan indikere at populasjonene som inngår ikke er sammenlignbare og at det ville vært nyttig å gjennomføre analyser av subgrupper/ moderatorer. I denne reviewartikkelen var det for få inkluderte studier til at slike analyser ble vurdert å være nyttige. For eksempel hadde det vært interessant å undersøke om deltakere med autismediagnose hadde mer eller mindre nytte av lyttertrening enn de andre deltakerne. Den kombinerte effektstørrelsen blir på bakgrunn av den store variasjonen vurdert å ikke være meningsfull.

Metaanalysen viste et relativt bredt prediksjonsintervall. Dette innebærer at det er usikkerhet knyttet til hvilken nytte andre individer vil ha av lyttertrening ved etableringa av intraverbaler. Prediksjonsintervallet vil kunne reduseres ved gjennomføring av flere studier med flere deltakere som kan inkluderes i analysen.

### **Visuell analyse og effektstørrelse**

Av de seks studiene som ble inkludert i den visuelle analysen var det fire studier det ble estimert en effektstørrelse for. To av disse studiene ble det konkludert med nøytral/blandet effekt ved den visuelle analysen, disse studiene var også de som hadde lavest effektstørrelsesestimater (0.02 og 0.24) som heller ikke var signifikante. For de andre to studiene ble det konkludert med positive effekter, disse studiene hadde også store effektstørrelsesestimater (7.07 og 1.31). Det så ut til at resultatene fra den visuelle analysen stemte godt overens med estimatene av effektstørrelse. Dette innebærer at både den visuelle analysen og estimering av effektstørrelse kan diskuteres å være gode metoder å vurdere effekt for single-subjekt studier og at konklusjonene fra de ulike metodene kan bli noenlunde like. Vurderingen av evidens vil også kunne diskuteres som styrket da konklusjonene i forhold til



effekt var lik ved begge metodene for beregning, selv om det i denne reviewartikkelen var den visuelle analysen som ble lagt til grunn for evidensvurderingen.

### **Konklusjon**

Målet med denne reviewartikkelen var å gjøre en systematisk gjennomgang og vurdering av studier som undersøker effekten av lyttertrening ved etablering av intraverbale responser. Åtte studier ble identifisert ved litteratursøket. Studiene holdt generelt høy metodologisk standard. Visuelle analyser viste at de fleste studiene viste positive effekter for intraverbal respondering etter lyttertrening. Lyttertrening ble vurdert som en mulig evidensbasert praksis. Det ble også foretatt estimeringer av effektstørrelse, der det for to av studiene ble vist stor og veldig stor signifikant effekt. Metaanalysen viste at estimatet av kombinerte effektstørrelsen for studiene ikke var meningsfull, da variasjonen var stor. Konklusjonen av denne reviewartikkelen er at det er nødvendig å gjennomføre flere studier som holder god metodologisk kvalitet for å kunne trekke ytterligere konklusjoner om effekten av lyttertrening for etablering av intraverbale responser.

**Referanser**

- Aguirre, A. A., Valentino, A. L., & LeBlanc, L. A. (2016). Empirical investigations of the intraverbal: 2005-2015. *Analysis of Verbal Behavior*, 32(2), 139-153.  
doi:10.1007/s40616-016-0064-4
- Axe, J. B. (2008). Conditional discrimination in the intraverbal relation: A review and recommendations for future research. *Analysis of Verbal Behavior*, 24, 159-174.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1(1), 91-97.  
doi:10.1901/jaba.1968.1-91
- Beitchman, J. H., Brownlie, E. B., & Bao, L. (2014). Age 31 mental health outcomes of childhood language and speech disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(10), 1102-1110.e1108. doi:10.1016/j.jaac.2014.07.006
- Bishop, D. V. M. (2014). *Uncommon understanding: development and disorders of language comprehension in children* (Classic ed. ed.). London: Psychology Press.
- Botting, N., & Conti-Ramsden, G. (2000). Social and behavioural difficulties in children with language impairment. *Child Language Teaching and Therapy*, 16(2), 105-120.  
doi:10.1177/026565900001600201
- Cihon, T. M. (2007). A Review of Training Intraverbal Repertoires: Can Precision Teaching Help? *Analysis of Verbal Behavior*, 23, 123-133.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd. In: Hillsdale, NJ: Erlbaum.
- Common, E. A., Lane, K. L., Royer, D. J., Muller, K., Pustejovsky, J. E., Johnson, A. H., & Johl, L. E. (2017). Functional assessment-based interventions for students with or at-risk for high-incidence disabilities: Field testing single-case synthesis methods. *Remedial and Special Education*, 38(6), 331-352. doi:10.1177/0741932517693320

- Cook, B. G., Buysse, V., Klingner, J., Landrum, T. J., McWilliam, R. A., Tankersley, M., & Test, D. W. (2015). CEC's standards for classifying the evidence base of practices in special education.(Council of Exceptional Children)(Report). *36*(4), 220-234.  
doi:10.1177/0741932514557271
- Daar, J. H., Negrelli, S., & Dixon, M. R. (2015). Derived emergence of WH question–answers in children with autism. *Research in Autism Spectrum Disorders, 19*, 59-71.  
doi:<https://doi.org/10.1016/j.rasd.2015.06.004>
- Eikeseth, S., & Smith, D. P. (2013). An analysis of verbal stimulus control in intraverbal behavior: implications for practice and applied research. *The Analysis of Verbal Behavior, 29*(1), 125. doi:10.1007/BF03393130
- Ellis, P. D. (2009). Thresholds for interpreting effect sizes. Hentet fra  
[http://www.polyu.edu.hk/mm/effectsizafaqs/thresholds\\_for\\_interpreting\\_effect\\_sizes2.html](http://www.polyu.edu.hk/mm/effectsizafaqs/thresholds_for_interpreting_effect_sizes2.html)
- Flensburg-Madsen, T., & Mortensen, E. L. (2017). Associations of early developmental milestones with adult intelligence. *Child development*. doi:10.1111/cdev.12760
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods, 3*(3), 224-239.
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods, 4*(4), 324-341. doi:10.1002/jrsm.1086
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165-179. doi:10.1177/001440290507100203

- \*Kodak, T., & Paden, A. R. (2015). A comparison of intraverbal and listener training for children with autism spectrum disorder. *Analysis of Verbal Behavior, 31*(1), 137-144. doi:10.1007/s40616-015-0033-3
- Lane, K. L., Common, E. A., Royer, D. J., & Muller, K. (2014). *Group comparison and single-case research design quality indicator matrix using Council for Exceptional Children 2014 standards*. <http://www.ci3t.org/practice>. Retrieved from <http://www.ci3t.org/practice>
- Lane, K. L., Kalberg, J. R., & Shepcaro, J. C. (2009). An examination of the evidence base for function-based interventions for students with emotional and/or behavioral disorders attending middle and high schools. *Exceptional Children, 75*(3), 321-340. doi:10.1177/001440290907500304
- \*Miguel, C. F., Petursdottir, A. I., & Carr, J. E. (2005). The effects of multiple-tact and receptive-discrimination training on the acquisition of intraverbal behavior. *The Analysis of Verbal Behavior, 21*(1), 27-41.
- \*Petursdottir, A. I., Carr, J. E., Lechago, S. A., & Almason, S. M. (2008). An evaluation of intraverbal training and listener training for teaching categorization skills. *Journal of Applied Behavior Analysis, 41*(1), 53-68. doi:10.1901/jaba.2008.41-53
- \*Petursdottir, A. I., & Hafliadottir, R. D. (2009). A comparison of four strategies for teaching a small foreign-language vocabulary. *Journal of Applied Behavior Analysis, 42*(3), 685-690. doi:10.1901/jaba.2009.42-685
- \*Petursdottir, A. I., Lepper, T., & Peterson, S. (2014). Effects of collateral response requirements and exemplar training on listener training outcomes in children. *The Psychological Record, 64*(4), 703-717. doi:10.1007/s40732-014-0051-x

- \*Petursdottir, A. I., Olafsdottir, A. R., & Aradottir, B. (2008). The effects of tact and listener training on the emergence of bidirectional intraverbal relations. *Journal of Applied Behavior Analysis, 41*(3), 411-415.
- Pustejovsky, J. E. (2016). scdhlm: A web-based calculator for between-case standardized mean differences (Version 0.3. 1)[Web application]. Hentet fra <https://jepusto.shinyapps.io/scdhlm/>.
- Rohatgi, A. (2015). WebPlotDigitizer (Version 4.0). Austin, Texas, USA. Hentet fra <http://arohatgi.info/WebPlotDigitizer>
- Shadish, W. R., Brasil, I., Illingworth, D., White, K., Galindo, R., Nagler, E., & Rindskopf, D. (2009). Using UnGraph to extract data from image files: Verification of reliability and validity. *Behavior Research Methods, 41*(1), 177-183. doi:10.3758/BRM.41.1.177
- Shadish, W. R., Hedges, L. V., Horner, R. H., & Odom, S. L. (2015). The role of between-case effect size in conducting, interpreting, and summarizing single-case research. *National Center for Education Research*.
- Silva, P. A., Williams, S., & McGee, R. (1987). A longitudinal study of children with developmental language delay at age three: Later intelligence, reading and behaviour problems. *Developmental Medicine & Child Neurology, 29*(5), 630-640. doi:10.1111/j.1469-8749.1987.tb08505.x
- Simes, R. J. (1986). Publication bias: the case for an international registry of clinical trials. *Journal of clinical oncology, 4*(10), 1529-1541.
- Slocum, T., Detrich, R., Wilczynski, S., Spencer, T., Lewis, T., & Wolfe, K. (2014). The Evidence-Based Practice of Applied Behavior Analysis. *The Behavior Analyst, 37*(1), 41-56. doi:10.1007/s40614-014-0005-2
- \*Smith, D. P., Eikeseth, S., Fletcher, S. E., Montebelli, L., Smith, H. R., & Taylor, J. C. (2016). Emergent intraverbal forms may occur as a result of listener training for

- children with autism. *Analysis of Verbal Behavior*, 32(1), 27-37. doi:10.1007/s40616-016-0057-3
- Smith, T. (2013). What is evidence-based behavior analysis? *The Behavior Analyst*, 36(1), 7. doi:10.1007/BF03392290
- Sundberg, M.L. (2014). Verbal behavior. Cooper, J. O., Heron, T. E. & Heward, W. L.(red.), *Applied behavior analysis (2<sup>nd</sup> ed.)*.(s. 536-557). Upper Saddle River, NJ: Pearson.
- Sundberg, M. L., & Sundberg, C. A. (2011). Intraverbal behavior and verbal conditional discriminations in typically developing children and children with autism. *The Analysis of Verbal Behavior*, 27(1), 23.
- Suurmond, R., Rhee, H., & Hak, T. (2017). Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods*, 8(4), 537-553. doi:10.1002/jrsm.1260
- RTeam. (2015). RStudio: integrated development for R. *RStudio, Inc.*, Boston, MA, URL <http://www.rstudio.com>.
- Valentine, J., Tanner-Smith, E., Pustejovsky, J., & Lau, T. (2016). Between-case standardized mean difference effect sizes for single-case designs: a primer and tutorial using the scdhlms web application. In: The Campbell Collaboration.
- \*Vallinger-Brown, M., & Rosales, R. (2014). An investigation of stimulus pairing and listener training to establish emergent intraverbals in children with autism. *Analysis of Verbal Behavior*, 30(2), 148-159. doi:10.1007/s40616-014-0014-y
- Van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *Journal of Applied Behavior Analysis*, 21(4), 381-384. doi:10.1901/jaba.1988.21-381
- Wolery, M. (2011). Intervention research: The importance of fidelity measurement. *Topics in Early Childhood Special Education*, 31(3), 155-157. doi:10.1177/0271121411408621

Tabell 1  
 Artikler vurdert for inklusjon basert på fulltekst

Artikkel	Årsak til eksklusjon	Inkludert
<b>Fra søk i databaser</b>		
Grannan, L., & Rehfeldt, R. (2012). Emergent intraverbal responses via tact and match-to-sample instruction. <i>Journal of Applied Behavior Analysis</i> , 45(3), 601–5. <a href="https://doi.org/10.1901/jaba.2012.45-601">https://doi.org/10.1901/jaba.2012.45-601</a>	Ikke lyttertrening	
Keintz, K. S., Miguel, C. F., Kao, B., & Finn, H. E. (2011). Using conditional discrimination training to produce emergent relations between coins and their values in children with autism. <i>Journal of Applied Behavior Analysis</i> , 44(4), 909–913.	Pre- og posttestdata	
Kodak, T., & Paden, A. R. (2015). A Comparison of Intraverbal and Listener Training for Children with Autism Spectrum Disorder. <i>Analysis of Verbal Behavior</i> , 31(1), 137–144. <a href="https://doi.org/10.1007/s40616-015-0033-3">https://doi.org/10.1007/s40616-015-0033-3</a>		Ja
Lechago, S. A., Carr, J. E., Kisamore, A. N., & Grow, L. L. (2015). The Effects of Multiple Exemplar Instruction on the Relation between Listener and Intraverbal Categorization Repertoires. <i>Analysis of Verbal Behavior</i> , 31(1), 76–95. <a href="https://doi.org/10.1007/s40616-015-0027-1">https://doi.org/10.1007/s40616-015-0027-1</a>	Ikke bare lyttertrening	
Miguel, C. F., Petursdóttir, A. I., & Carr, J. E. (2005). The effects of multiple-tact and receptive-discrimination training on the acquisition of intraverbal behavior. <i>The Analysis of Verbal Behavior</i> , 21, 27–41.		Ja
Petursdóttir, A. I., Carr, J. E., Lechago, S. A., & Almason, S. M. (2008). An evaluation of intraverbal training and listener training for teaching categorization skills. <i>Journal of Applied Behavior Analysis</i> , 41(1), 53–68. <a href="https://doi.org/10.1901/jaba.2008.41-53">https://doi.org/10.1901/jaba.2008.41-53</a>		Ja
Petursdóttir, A. I., & Hafliðadóttir, R. D. (2009). A comparison of four strategies for teaching a small foreign-language vocabulary. <i>Journal of Applied Behavior Analysis</i> , 42(3), 685–690. <a href="https://doi.org/10.1901/jaba.2009.42-685">https://doi.org/10.1901/jaba.2009.42-685</a>		Ja
Petursdóttir, A., Lepper, T., & Peterson, S. (2014). Effects of Collateral Response Requirements and Exemplar Training on Listener Training Outcomes in Children. <i>The Psychological Record</i> , 64(4), 703–717. <a href="https://doi.org/10.1007/s40732-014-0051-x">https://doi.org/10.1007/s40732-014-0051-x</a>		Ja
Petursdóttir, A. I., Ólafsdóttir, A. R., & Aradóttir, B. (2008). The effects of tact and listener training on the emergence of bidirectional intraverbal relations. <i>Journal of Applied Behavior Analysis</i> , 41(3), 411–415. <a href="https://doi.org/10.1901/jaba.2008.41-411">https://doi.org/10.1901/jaba.2008.41-411</a>		Ja

Polson, D. A. D., Grabavac, D. M., & Parsons, J. A. (1997). Intraverbal stimulus-response reversibility: Fluency, familiarity effects, and implications for stimulus equivalence. <i>The Analysis of verbal behavior</i> , 14, 19–40.	Ikke lyttertrening	
Shillingsburg, M. A., Frampton, S. E., Cleveland, S. A., & Cariveau, T. (2017). A clinical application of procedures to promote the emergence of untrained intraverbal relations with children with autism. <i>Learning and Motivation</i> . <a href="https://doi.org/10.1016/j.lmot.2017.02.003">https://doi.org/10.1016/j.lmot.2017.02.003</a>	AB-design	
Smith, D. P., Eikeseth, S., Fletcher, S. E., Montebelli, L., Smith, H. R., & Taylor, J. C. (2016). Emergent intraverbal forms may occur as a result of listener training for children with autism. <i>Analysis of Verbal Behavior</i> , 32(1), 27–37. <a href="https://doi.org/10.1007/s40616-016-0057-3">https://doi.org/10.1007/s40616-016-0057-3</a>		Ja
Vallinger - Brown, M., & Rosales, R. (2015). An investigation of stimulus pairing and listener training to establish emergent intraverbals in children with autism. <i>Analysis of Verbal Behavior</i> , 30(2), 148–159. <a href="https://doi.org/10.1007/s40616-014-0014-y">https://doi.org/10.1007/s40616-014-0014-y</a>		Ja

---

#### Fra referanselister

Sprinkle, E. C., & Miguel, C. F. (2012). The effects of listener and speaker training on emergent relations in children with autism. <i>The Analysis of Verbal Behavior</i> , 28(1), 111. doi:10.1007/BF03393111	Ikke intraverbaler	
Ingvarsson, E. T., & Hollobaugh, T. (2011). A comparison of prompting tactics to establish intraverbals in children with autism. <i>Journal of Applied Behavior Analysis</i> , 44(3), 659-664. doi:10.1901/jaba.2011.44-659	Ikke lyttertrening	
Partington, J. W., & Bailey, J. S. (1993). Teaching intraverbal behavior to preschool children. <i>The Analysis of Verbal Behavior</i> , 11, 9.	Ikke lyttertrening	
Miller, M. A., Cuvo, A. J., & Borakove, L. S. (1977). Teaching naming of coin values—comprehension before production versus production alone. <i>Journal of Applied Behavior Analysis</i> , 10(4), 735-736. doi:10.1901/jaba.1977.10-735	Ikke intraverbaler	
Watkins, C. L., Pack-Teixeira, L., & Howard, J. S. (1989). Teaching intraverbal behavior to severely retarded children. <i>The Analysis of Verbal Behavior</i> , 7, 69.	Ikke lyttertrening	
Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in young children with autism. <i>Behavioral Interventions</i> , 18(4), 245-266. doi:10.1002/bin.142	Ikke intraverbaler	

**Note.** Tabellen viser studier identifisert ved litteratursøk i databasene og ved søk i referanselister som ikke ble ekskludert ved gjennomlesning av tittel og sammendrag, årsaker til eksklusjon basert på lesing av fulltekst og hvilke studier som ble inkludert i denne artikkelen.



Tabell 2

*Deskriptive karakteristikk ved inkluderte studier*

Studie	Deltaker	Alder	Kjønn	Tilstand	Intraverbaler	Setting
Kodak og Paden (2015)	Rosemary	4	J	DBD-NOS og autisme	Hvem-spørsmål	
	Oscar	3	G	PDD-NOS og autisme	Hva-spørsmål	
Miguel, Petursdottir og Carr (2005)	Brandon	3-5	G	NF	Kategorisering	TR
	Tina	3-5	J	NF	Kategorisering	TR
	Jane	3-5	J	NF	Kategorisering	TR
Petursdottir et al. (2008)	Erika	3	J	NF	Kategorisering	TR
	Chanele	3	J	NF	Kategorisering	TR
	Sanjay	3	G	NF	Kategorisering	TR
Petursdottir og Hafliadottir (2009)	Dianna	5	J	NF	Fremmedspråk	TR
	Kamilla	5	J	NF	Fremmedspråk	TR
Petursdottir, Lepper og Peterson (2014)	Maia	4	J	NF	Fremmedspråk	TR
	Lauren	4	J	NF	Fremmedspråk	TR
	Jennifer	4	J	NF	Fremmedspråk	TR
Petursdottir, Ólafsdóttir og Araðóttir (2008)	Lena	5	J	NF	Fremmedspråk	TR
	Karen	5	J	NF	Fremmedspråk	TR
Smith et al. (2016)	Robbie	6	G	Autisme	Hva-spørsmål	Hjem
	Sarah	5	J	Autisme	Hva-spørsmål	Hjem
	Freddie	13	G	Autisme	Hva-spørsmål	Hjem
	Edward	12	G	Autisme	Hva-spørsmål	TR
	Charlie	15	G	Autisme	Hva-spørsmål	Hjem
Vallinger-Brown og Rosales (2014)	Carl	4	G	Autisme	Hva-spørsmål	TR
	Aron	5	G	Autisme	Hva-spørsmål	TR
	John	7	G	Autisme	Hva-spørsmål	TR

*Note.* Alder er oppgitt i antall år. G= gutt; J= jente. DBD-NOS= disruptive behavior disorder- not otherwise specified; PDD-NOS= pervasive developmental disorder- not otherwise specified; NF= normalfungerende. TR= treningsrom/plass på skole eller i barnehage.

Tabell 3  
Total skåring av CECs kvalitetsindikatorer og effekt

	Kodak og Paden (2015)	Miguel, Petursdóttir og Carr (2005)	Petursdóttir et al. (2008)	Petursdóttir og Hafliðadóttir (2009)	Petursdóttir, Lepper og Peterson (2014)	Petursdóttir, Ólafsdóttir og Araóttir (2008)	Smith et al. (2016)	Vallinger-Brown og Rosales (2014)
1.0 Kontekst og setting	0	1	1	1	1	1	1	1
2.0 Deltakere	0.5	1	1	1	1	1	1	0.5
3.0 Trener	1	1	1	1	1	1	1	1
4.0 Beskrivelse av tiltaket	1	1	1	1	1	1	1	1
5.0 Prosedyreintegritet	0	1	1	0	0	0	0	1
6.0 Indre validitet	0.5	1	1	0.67	0.5	0.67	0.67	1
7.0 Avhengig variabel	0.8	1	1	1	0.8	1	1	1
8.0 Dataanalyser	1	1	1	1	1	1	1	1
IOA (%)	100 %	95,4 %	100 %	100 %	100 %	100 %	100 %	100 %
Absolutt skåre	3.0	8.0	8.0	6.0	5.0	6.0	6.0	7.0
Sammenlagt skåre	4.8	8.0	8.0	6.67	6.30	6.67	6.67	7.50
Prosent KI oppfylt	60%	100 %	100 %	83.3%	78.7 %	83.3 %	83.3%	93.7 %
Deltakere (n)	2	3	3	2	3	2	5	3
Effekt (visuell analyse)	IV	N/B	N/B	+	IV	+	+	+

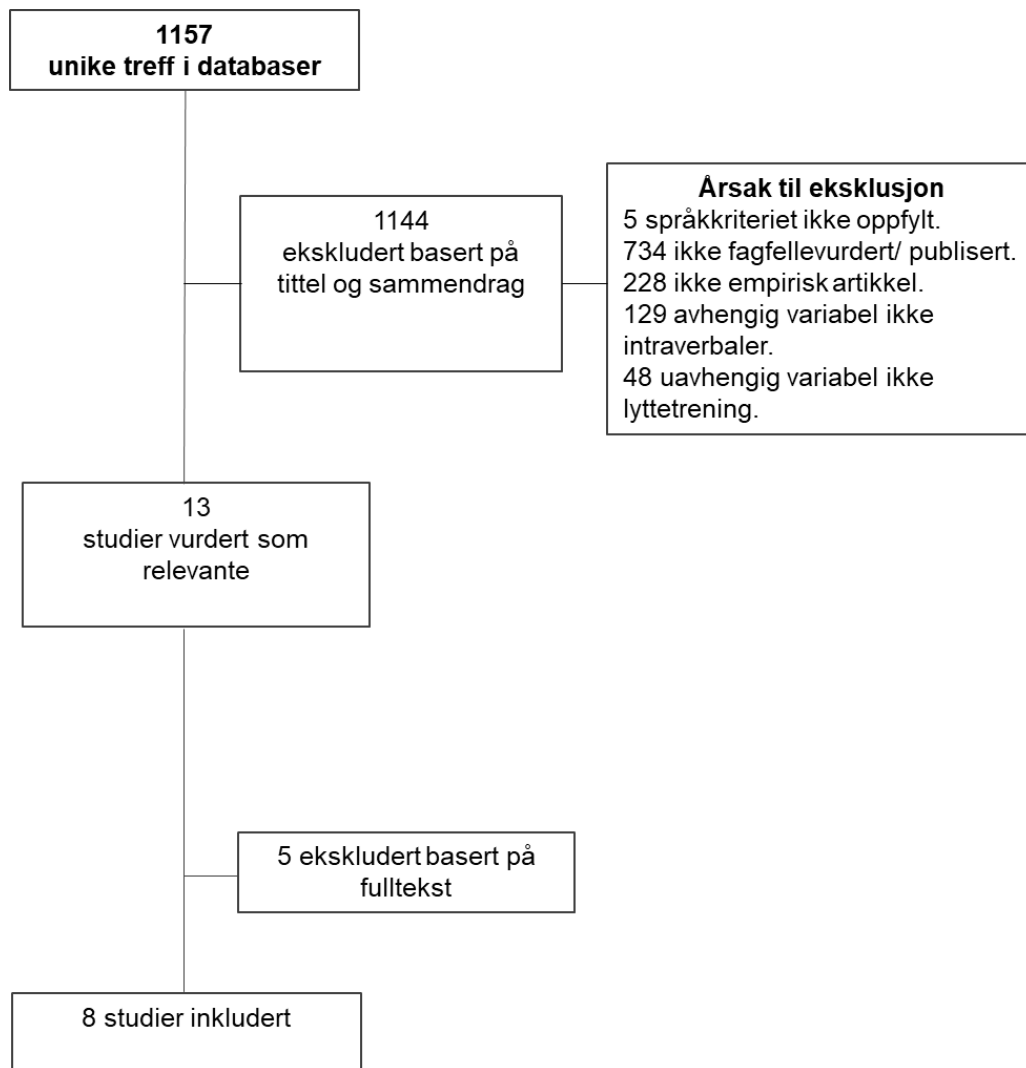
*Note.* IOA= mellomobservatørsenighet ved skåring av kvalitetsindikatorene. Absolutt skåre viser til antallet indikatorer der alle komponenter ble skåret som oppfylt. Sammenlagt skåre viser til summen av andelen komponenter som ble skåret som oppfylt pr. indikator. KI= kvalitetsindikator, her oppgitt i prosent av 8, basert på sammenlagt skåre. IV= ikke vurdert; N/B= nøytral eller blandet effekt; += positiv effekt.

Tabell 4

*Resultat metaanalyse*

Effektstørrelse	1,21
CI nedre grense	-1,63
CI øvre grense	4,06
p-verdi, to halet	0,237
PI nedre grense	-2,92
PI øvre grense	5,38
I <sup>2</sup>	0,85

*Note.* Kombinert effektstørrelse for de inkluderte studiene; CI= øvre og nedre grense for konfidensintervall (95%); PI= øvre og nedre grense for prediksjonsintervallet; I<sup>2</sup>= variabilitet mellom studiene.

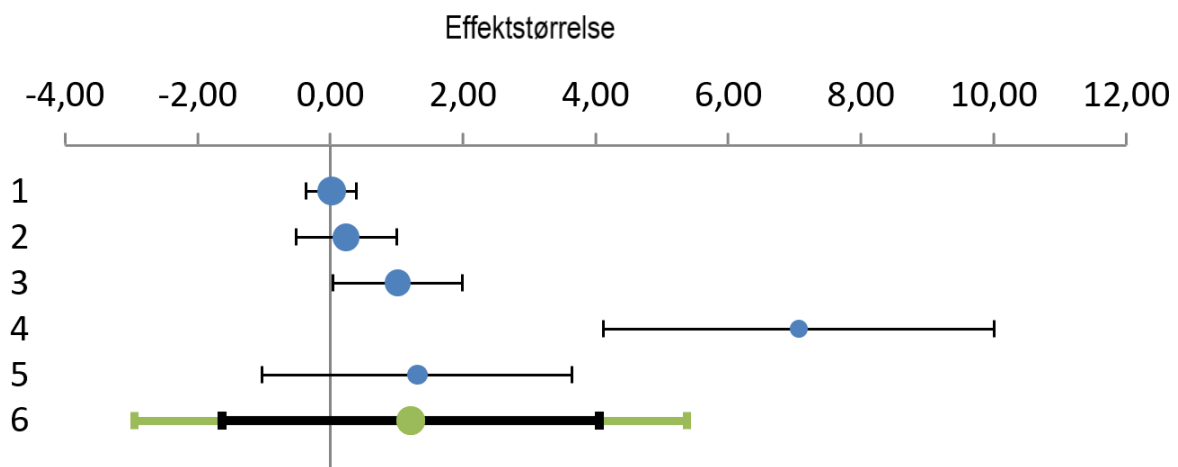


*Figur 1.* Det totale antallet artikler funnet ved litteratursøket i databasene og antallet artikler derfra ekskludert og inkludert i denne artikkelen, med oversikt over årsaker til eksklusjon.

8.3	IR	IR	IR	IR	IR	IR	IR	IR
8.2	1	1	1	1	1	1	1	1
8.1	IR	IR	IR	IR	IR	IR	IR	IR
7.6	IR	IR	IR	IR	IR	IR	IR	IR
7.5	1	1	1	1	1	1	1	1
7.4	0	1	1	1	0	1	1	1
7.3	1	1	1	1	1	1	1	1
7.2	1	1	1	1	1	1	1	1
7.1	1	1	1	1	1	1	1	1
6.9	IR	IR	IR	IR	IR	IR	IR	IR
6.8	IR	IR	IR	IR	IR	IR	IR	IR
6.7	0	1	1	0	0	0	0	1
6.6	1	1	1	1	0	1	1	1
6.5	0	1	1	1	1	1	1	1
6.4	IR	IR	IR	IR	IR	IR	IR	IR
6.3	1	1	1	1	1	1	1	1
6.2	1	1	1	1	1	1	1	1
6.1	0	1	1	0	0	0	0	1
5.3	0	1	1	0	0	0	0	1
5.2	0	1	1	0	0	0	0	1
5.1	0	1	1	0	0	0	0	1
4.2	1	1	1	1	1	1	1	1
4.1	1	1	1	1	1	1	1	1
3.2	1	1	1	1	1	1	1	1
3.1	1	1	1	1	1	1	1	1
2.2	0	1	1	1	1	1	1	0
2.1	1	1	1	1	1	1	1	1
1.1	0	1	1	1	1	1	1	1
	Kodak & Paden (2015)	Miguel, Pettursdottir & Carr (2005)	Pettursdottir et al. (2008)	Pettursdottir & Haflicadottir (2009)	Pettursdottir, Lepper & Peterson (2014)	Pettursdottir, Ólafsdóttir & Araóttir (2008)	Smith et al. (2016)	Vallinger-Brown & Rosales (2014)

**Figur 2.** Skårene for de inkluderte studiene på CECs kvalitetsindikatorer. IR= ikke relevant. 1.1 beskrivelse av setting; 2.1 beskrivelser av deltaker; 2.2 bakgrunn for diagnose; 3.1 trener; 3.2 nødvendig opplæring; 4.1 prosedyrebeskrivelse; 4.2 beskrivelse av materiale; 5.1 prosedyreintegritet; 5.2 beskrivelse av hyppighet av målinger; 5.3 målinger i alle faser; 6.1 systematisk manipulasjon av uavhengig variabel; 6.2 beskrivelser av baseline; 6.3 intervensjon ikke tilgjengelig under baseline; 6.5 tre demonstrasjoner av eksperimentell effekt; 6.6 tre målinger i baseline; 6.7 kontroll for trusler mot indre validitet; 7.1 sosial validitet; 7.2 beskrivelser av avhengig variabel; 7.3 effekt rapporteres for alle mål; 7.4 gjentatte målinger, minst tre i alle faser; 7.5 adekvate IOA mål; 8.2 graf som tydelig presenterer resultatene.

# Studie	Effektstørrelse	95% CI nedre grense	95% CI øvre grense	Vekt
1 Miguel, Petursdottir & Carr 2005*	0,02	-0,36	0,40	27,5 %
2 Petursdottir et al 2008	0,24	-0,52	1,00	25,2 %
3 Petursdottir, Lepper & Peterson 2014	1,02**	0,04	1,99	23,5 %
4 Smith et al 2016	7,07**	4,12	10,01	10,4 %
5 Vallinger-Brown & Rosales 2014	1,31	-1,03	3,65	13,3 %
6 Kombinert effektstørrelse	1,21	-1,63	4,06	



*Figur 3.* Forest plot med effektstørrelse og konfidensintervall for de fem studiene der dette ble kalkulert. Nr. 6 viser kombinert effektstørrelse og konfidensintervall. Den grønne streken representerer prediksjonsintervallet. Tabellen viser en oversikt over effektstørrelse og grensene for konfidensintervall for hver enkelt studie.

Nummereringen i tabellen viser til nummereringen i forest plot. Vekt viser til hvor stor del av utvalget den enkelte studie representerer. \* Fra Miguel, Petursdottir & Carr (2005) ble effektstørrelse bare beregnet for et av de to settene barna trente på (musical instruments). \*\* $p < 0.05$ .

## **Appendiks A**

### **Council for Exceptional Childrens kvalitetsindikatorer og kriterier for skåring**

#### **1.0 Kontekst og setting**

Den første kvalitetsindikatoren var kontekst og setting. For å oppfylle dette kriteriet måtte artiklene tilstrekkelig detaljert beskrive trekk ved kontekst eller setting. Det ble vurdert som tilstrekkelig detaljert om det var beskrevet hvor studien ble gjennomført, som for eksempel «et eget rom i barnehagen», om det ikke ble vurdert som nødvendig med en helt spesiell setting for å gjennomføre tiltaket.

#### **2.0 Deltakere**

Den andre kvalitetsindikatoren bestod av to komponenter. For å oppfylle 2.1 måtte det i artikkelen gis tilstrekkelig informasjon om deltakerne. Dette kriteriet ble skåret som oppfylt, om det var beskrevet minst kjønn, alder og om deltakerne hadde en diagnose eller var normalfungerende. For å møte 2.2 måtte det inkluderes hvilke metoder eller tester som var anvendt for å sette en diagnose i de tilfeller der dette var aktuelt. I studier med normalfungerende deltakere ble 2.2 skåret som oppfylt uten at det ble oppgitt noe grunnlag for vurderingen av funksjonsnivå.

#### **3.0 Trener**

Den tredje kvalitetsindikatoren omhandlet den som gjennomførte tiltaket. Denne indikatoren bestod av to komponenter. For å oppfylle 3.1 måtte det i studien beskrives hvilken rolle den som gjennomførte tiltaket hadde, for eksempel om det var foreldre, en lærer eller eksperimentator. For å oppfylle 3.2 måtte det oppgis hvilken trening den som gjennomfører tiltaket hadde på forhånd, for eksempel hvilken utdanning eller andre erfaringer som var nødvendig for å gjennomføre tiltaket. Dette ble også skåret som oppfylt om det var forfatterne av artikkelen som gjennomførte tiltaket.

#### **4.0 Beskrivelser av tiltaket**

Den fjerde kvalitetsindikatoren omhandlet beskrivelser av selve tiltaket, denne bestod av to komponenter. For at 4.1 skulle være i henhold til standard måtte det i studien gis en detaljert beskrivelse av de prosedyrer som var anvendt. Dette kriteriet ble skåret som oppfylt om beskrivelsene av prosedyren var så detaljert at en replikasjon virket mulig basert på artikkelen alene. For å oppfylle 4.2 måtte nødvendig materiale, og hvor dette eventuelt kunne være tilgjengelig beskrives, 4.2 ble også skåret som oppfylt om det av artikkelen framgikk at det ikke var nødvendig med noe spesielt materiale.

### **5.0 Prosedyreintegritet**

Den femte kvalitetsindikatoren som ble vurdert var prosedyreintegritet. Denne bestod av tre komponenter. For å møte 5.1 måtte det ha vært gjennomført målinger og rapporteringer av prosedyreintegritet. For å møte 5.2 måtte målingene av prosedyreintegritet være gjennomført ved bruk av direkte og reliable målinger. For å møte 5.3 måtte det være gjennomført og rapportert om prosedyreintegritetsmålinger jevnlig og for alle deltakere, settinger og trenere.

### **6.0 Indre validitet**

Den sjettede kvalitetsindikatoren omhandlet indre validitet og bestod av ni komponenter, men bare seks av disse var aktuelle for single-subjekt design og dermed vurdert. For å møte 6.1 måtte forskeren kontrollere og systematisk manipulere den uavhengige variabelen, dette kunne bare vurderes på studier der det også var gjennomført målinger av prosedyreintegritet (kvalitetsindikator 5.0). For studier der det ikke var rapportert prosedyreintegritet ble det skåret 0 på denne komponenten. For å oppfylle 6.2 var det nødvendig av studien hadde en beskrivelse av en baseline- eller kontrollbetingelse. For å møte 6.3 måtte det fremgå i hvilken grad deltaker under kontrollbetingelsen eller baseline kunne tenkes ha tilgang til tiltaket, dette ble vurdert som oppfylt om det var liten sannsynlighet at dette kunne være tilfelle, uten at dette nødvendigvis var eksplisitt beskrevet i artikkelen. 6.4 omhandlet gruppestudier og ble



derfor ikke vurdert. Ved tre eller flere muligheter for demonstrasjoner av effekt, ved minst tre ulike tilfeller ble 6.5 vurdert som oppfylt, dette innebar at studier som bare hadde en pretest og posttest, eller et multiple baseline design over to personer ville skåre 0 på denne komponenten. Det var ikke nødvendig at det ble fremvist en faktisk effekt, det var utelukkende muligheten for demonstrasjoner av effekt som ble vurdert. For å møte 6.6 måtte baselinemålingene inneholde minst tre målinger, men også studier med færre baselinemålinger kunne godtas om det forekom null responser i baseline og det fremgikk som meget usannsynlig at det skulle skje en forbedring uten en intervensjon. For å oppfylle 6.7 måtte studien ha et design som gjorde det mulig å utelukke vanlige trusler mot indre validitet, eksempler på slike design var multiple baseline over minst tre personer/atferder/settinger eller ABAB-design. 6.7 ble i likhet med 6.1 alltid skåret som ikke-oppfylt om det i studiene ikke var gjort målinger av prosedyreintegritet (indikator 5.0). 6.8 og 6.9 var bare aktuelle for gruppestudier og ble derfor ikke vurdert.

## **7.0 Avhengig variabel**

Den syvende kvalitetsindikatoren som ble vurdert var den avhengige variabelen, denne bestod av seks komponenter. For å oppfylle 7.1 måtte resultatene av tiltaket være vurdert som sosialt valide, ved at de var teoretisk eller empirisk knyttet til økt livskvalitet eller viktige for utvikling og læring. Komponenten kunne vurderes som oppfylt uavhengig av om det eksplisitt ble beskrevet i den aktuelle artikkelen om målene ble vurdert å være sosialt valide. For eksempel ville alle studier som omhandlet språktrening være vurdert som sosialt valide, fordi språkferdigheter anses viktig for både livskvalitet, utvikling og læring (Beitchman, Brownlie, & Bao, 2014; Bishop, 2014; Botting & Conti-Ramsden, 2000; Flensburg-Madsen & Mortensen, 2017; Silva, Williams, & McGee, 1987). For å møte 7.2 måtte studien inneholde en tydelig operasjonalisering av den avhengige variabelen. For å møte 7.3 måtte alle resultater beskrives, ikke bare de resultater som viste en positiv eller ønsket effekt, dette ble vurdert som

oppfylt om resultater for alle deltakere i de ulike fasene ble rapportert, for eksempel ved at de ble framstilt i en graf. For å møte 7.4 var det krav om at det måtte være foretatt minst tre målinger i hver fase, med unntak av baselinefaser med liten sannsynlighet for forbedring uten intervensjon. For å møte 7.5 måtte studien ha målt og rapportert mellomobservatørsenighet med resultater på over 80%. 7.6 var ikke aktuell ved singel-subjekt design og ble derfor ikke vurdert.

## **8.0 Dataanalyser**

Den siste og åttende kvalitetsindikatoren bestod av tre komponenter, der bare 8.2 var aktuell for singel-subjekt design. For at 8.2 skulle skåres som oppfylt måtte studien inneholde en graf som tydelig viste utfallsdata over alle faser som gjorde det mulig å vurdere effekten av tiltaket. Grafen skulle være tydelig nok til å kunne gjøre vurderinger av gjennomsnitt, nivåer, trender, overlapp, og datamønstre på tvers av og innen faser.

## **Appendiks B**

### **Council for Exceptional Childrens klassifisering av evidens**

Se Cook et al., 2015 for detaljert beskrivelse av de ulike kriterene for de fem klassifiseringene av evidens. Videre følger en kort beskrivelse av kriteriene for å oppfylle de ulike klassene gjeldende for single-subjekt design.

#### **1. Evidensbasert praksis**

For at en metode skal klassifiseres som evidensbasert praksis må metoden være støttet av minst fem singel-subjekt studier med tilstrekkelig metodologisk kvalitet som viser positive effekter. Totalt må det være mer enn 20 deltakere inkludert i studiene. Alternativt tre single-subjekt studier med tilstrekkelig metodologisk kvalitet og positive effekter for minst ti deltakere og to gruppestudier med minst 60 deltakere, positive effekter og tilstrekkelig metodologisk kvalitet. Det kan ikke forekomme noen studier med tilstrekkelig metodologisk kvalitet som viser negative effekter. Det totale forholdet (alle design) mellom studier som viser positiv effekt og studier som viser nøytral/blandet effekt må være minst 3:1.

#### **2. Mulig evidensbasert praksis**

For at en metode skal klassifiseres som mulig evidensbasert praksis må det forekomme to til fire single-subjekt studier med tilstrekkelig metodologisk kvalitet som viser positive effekter eller en kombinasjon av single-subjekt og gruppestudier med god kvalitet og positive effekter, men ikke tilstrekkelig antall studier og/eller deltakere for å klassifiseres som evidensbasert praksis. Det kan ikke forekomme noen studier med tilstrekkelig metodologisk kvalitet som viser negative effekter. Det totale forholdet (alle design) mellom studier som viser positiv effekt og studier som viser nøytral/blandet effekt må være minst 2:1.

#### **3. Blandet evidens**

Blandet evidens tilskrives metoder der kravene til evidensbasert praksis eller mulig evidensbasert praksis er oppfylt, men forholdet mellom studier med tilstrekkelig metodologisk

kvalitet som viser positive effekter mot studier som viser nøytrale eller blandede effekter er lavere enn 2:1. Det konkluderes også om blandet evidens om finnes det studier med tilstrekkelig metodologisk kvalitet som viser negative effekter, men antallet slike studier er lavere enn antallet som viser positive effekter.

#### **4. Ikke tilstrekkelig evidens**

Ikke tilstrekkelig evidens tilskrives metoder som ikke oppfyller kriteriene for de andre klassifiseringene. Mer forskning er nødvendig.

#### **5. Negative effekter**

Negative effekter tilskrives metoder der mer enn en studie (alle design) med tilstrekkelig metodologisk kvalitet viser negative effekter og antallet studier som viser negative effekter er høyere enn antallet studier som viser positive effekter.

**En single-subjekt-design studie av effekten av lyttertrening for emergens av  
intraverbale responser hos en fem år gammel gutt med autisme**

Line Børtveit

OsloMet- storbyuniversitetet

### Sammendrag

I denne studien ble effekten av lyttertrening for etablering av intraverbale responser studert. Deltaker var en fem år gammel gutt med autismediagnose. Det ble ved lyttertrening trent på ni ulike når-spørsmål, som «når bruker du tannkrem» som deltaker svarte på ved å peke på bilder. I forkant var det kartlagt at deltaker kunne tacte bildestimuliene som ble anvendt under trening. Det ble testet om det oppstod emergente intraverbale responser, ved at samme spørsmål som ble brukt under trening ble presentert, men deltaker måtte respondere vokalt og ikke ved å peke på bildestimuli. Det ble også testet om det oppstod emergente bidireksjonale intraverbaler, ved at spørsmålene ble stilt på en annen måte, som «hva trenger du når du pusser tenner». Resultatene viste at lyttertrening umiddelbart og stabilt førte til emergente intraverbale responser for to av de tre settene det ble trent på (sett 2 og 3). For sett 1 var det nødvendig med tacttrening for å oppnå mestring. Etter tacttrening var det 100% mestring ved alle gjennomførte tester for dette settet. Det oppstod emergente bidireksjonale intraverbaler for to av de tre settene, men var ingen endring for sett 3 ved post-test. Lyttertrening var en mulig effektiv metode for etablering av intraverbale responser for denne deltakeren.

*Nøkkelord:* intraverbaler, lyttertrening, når-spørsmål, verbal atferd, reseptiv diskriminasjonstrening, emergens, bidireksjonalitet.

En single-subjekt-design studie av effekten av lyttertrening for emergens av intraverbale responser hos en fem år gammel gutt med autisme

Autisme er i ICD-10 (World Health Organization, 1992) definert som en gjennomgripende utviklingsforstyrrelse som blant annet gir avvik i mellommenneskelig sosialt samspill og kommunikasjonsmønstre. Disse avvikene vil prege personens fungering i alle situasjoner selv om styrken på avvikene vil være individuelle. Nedsatt eller forsinket utvikling av språk kan ifølge Silva, Williams, og McGee (1987) være en prediktor for at barnet senere vil ha lavere intelligens enn gjennomsnittet for aldersgruppen, dårligere leseferdigheter og dessuten økt risiko for utvikling av atferdsproblemer. Flensburg-Madsen og Mortensen (2017) fant i en longitudinal studie signifikante sammenhenger mellom forsinket språk ved treårs alder og lavere IQ skårer ved 20-34 års alder. Beitchman, Brownlie, og Bao (2014) fant en sammenheng mellom utfordringer med språk og psykiske vansker under barne- og ungdomsårene. Videre fant Botting og Conti-Ramsden (2000) sammenhenger mellom språkproblemer og sosiale utfordringer.

For barn med autisme har intensive opplæringsprogrammer basert på anvendt atferdsanalyse vist seg å være effektive og det argumenteres for at slike programmer bør velges fremfor andre når opplæring skal gjennomføres (Eldevik et al., 2009) (Reichow & Wolery, 2009). Språkopplæring innenfor anvendt atferdsanalyse baseres ofte på en funksjonell tilnærming til språk (Barbera, 2007; Luce, Maurice, & Green, 1996; Løvaas, 2003; McEachin & Leaf, 1999; Sundberg & Partington, 1998). Skinner (1957) forklarer en funksjonell tilnærming til språk og kommunikasjon med utgangspunkt i atferd i relasjon til det miljøet atferden forekommer i og hvilke konsekvenser som etterfølger atferden. Skinner beskriver språk eller kommunikasjon som *verbal atferd* og benytter samme analyseenhet ved analyserer språk som ved analyser av annen atferd: foranledigende eller *diskriminative stimuli* ( $S^D$ ), responser og konsekvenser. Med dette utgangspunktet beskriver han ulike verbale

operanter som er ulike tilfeller av verbal atferd som skilles basert på hvilken funksjon de har i forhold til miljøet, altså hvilke stimuli som foranlediger og etterfølger atferden. Noen av de grunnleggende verbale operantene Skinner beskriver er *mand*, *tact* og *intraverbaler*.

*Mand* er en verbal operant der responsen forsterkes av spesifikke forsterkere som er direkte knyttet til individets tilstand av deprivasjon eller aversiv stimulering. Disse tilstandene og hvordan de påvirker atferd beskrives av Laraway, Snyckerski, Michael, og Poling (2003) som motivasjonelle operasjoner (MO). Mands er verbale operanter der forsterkeren direkte kan knyttes til MO. Den foranledigende stimulusen for en mand er en ikke-verbal stimulus. Et eksempel kan være at barnet er tørst (MO) og ser et glass med vann som står utenfor rekkevidde ( $S^D$ ) sier «vann» og deretter får vannglasset av mor. I følge Skinner (1957) består store deler av småbarns kommunikasjon av mands. Det er diskutert i hvilken grad barns første verbale responser er mands. Horne og Lowe (1996) beskriver at barn først lærer å navngi eller tacte objekter før de fremviser mands, da det er problematisk for barnet å kunne be om å få noe om han ikke har lært seg hva det han ber om kalles.

*Tact* er en verbal operant der responser forsterkes av generaliserte betingede forsterkere. Den foranledigende stimulusen er et bestemt objekt eller en hendelse som forekommer i snakkers omgivelser (Skinner, 1957). Et eksempel kan være om barnet ser en hund og sier «hund», noe som etterfølges av konsekvensen at mor sier «ja, du har rett, det er en hund». Når det forekommer såkalt navngiving av objekter eller hendelser vil dette kunne tolkes som tilfeller av *tact*.

*Intraverbal atferd* er verbal atferd der responser forsterkes av generaliserte betingede forsterkere. Den foranledigende stimulusen er verbal og det er ikke punkt-til-punkt overensstemmelse mellom den foranledigende verbale stimulusen og responsen (Skinner, 1957). Intraverbal atferd gjør det mulig å snakke om objekter og hendelser som ikke fysisk er tilstede, noe som ikke vil kunne være tilfelle ved *mand* eller *tact*. Skinner beskriver en rekke



eksempler på intraverbal atferd, for eksempel når eleven svarer «fire» etter at læreren har spurt «hva er to ganger to» eller når eleven sier «Paris» etter spørsmål om hva som er hovedstaden i Frankrike. Han beskriver videre at også det å kunne svare på spørsmål som «hvordan har du det» kan defineres som intraverbal atferd om svaret «bra» forekommer uten at individet egentlig kjenner etter. Andre eksempler på intraverbal atferd kan være å kunne synge refrenget etter å ha hørt første vers i en sang eller å kunne oppgi bokstavene i alfabetet. Normalfungerende voksne har gjerne flere hundretusen relasjoner mellom verbale  $S^D$ , responser og generaliserte betingede forsterkere, som utgjør deres intraverbale repertoar (Sundberg & Michael, 2001). Skinner forklarer at samtaler og «smaltalk» som oftest er intraverbal atferd og Sundberg (2014) beskriver at det å kunne holde sammenhengende samtaler med andre ikke lar seg gjøre uten et intraverbalt repertoar.

For barn med autisme er ofte utviklingen av et intraverbalt repertoar forsinket, mangelfull eller ikke-eksisterende, dette kan være tilfelle selv om barnet har et stort repertoar av tact-relasjoner og fremviser lytteratferd (reseptivt språk) (Sundberg & Michael, 2001). For eksempel kan barnet kanskje peke på et bilde av en hund når det hører ordet «hund», si «hund» når det ser en hund eller et bilde av en hund, og si «hund» som en mand for å få komme bort til og klappe på en forbipasserende hund, uten å kunne besvare spørsmålet «hvilket dyr sier voff». I følge Sundberg og Michael vil de fleste normalfungerende barn nærmest indirekte gjennom kontakt med det verbale miljøet etablere et intraverbalt repertoar. Problemet for mange barn med autisme er at sosial interaksjon med andre er ikke nødvendigvis forsterkende i seg selv og kontakt med det verbale miljøets generaliserte betingede forsterkere vil dermed ikke være tilstrekkelig for å etablere et intraverbalt repertoar. Det kan dermed være nødvendig å igangsette tiltak for å etablere intraverbal atferd hos barn med autisme (Sundberg & Sundberg, 2011).

At de tiltak som gjennomføres skal være effektive og utgjøre en reell forskjell for personen intervensjonen gjennomføres for er en del av forutsetningene for anvendt atferdsanalyse (Baer, Wolf, & Risley, 1968). Van Houten et al. (1988) beskriver det som en rettighet at de som mottar atferdsanalytiske tiltak skal motta de mest effektive tiltakene som finnes tilgjengelig. Fordi det, som beskrevet tidligere er knyttet en rekke utfordringer til forsinket språk (Beitchman et al., 2014; Botting & Conti-Ramsden, 2000; Flensburg-Madsen & Mortensen, 2017; Silva et al., 1987), og fordi det for barn med autisme ofte ses problemer med intraverbal respondering er det nødvendig å gjennomføre studier for å finne de mest effektive tiltakene for denne gruppen.

Daar, Negrelli, og Dixon (2015) beskriver at fordi det i Skinners analyser av verbal atferd ikke fremgår hvilken forsterkningshistorie som er nødvendig for etablering av et intraverbalt repertoar har mange atferdsanalytikere basert seg på samme fremgangsmåte som brukes for å etablere andre verbale operanter. Dette innebærer presentasjon av en stimulus etterfulgt av en prompt for en spesifikk respons. Tact- trening kan for eksempel se slik ut: barnet vises et bilde av en hund, og trener spør «hvilket dyr er dette» og prompter deretter barnet gjennom å selv si «hund». Ifølge Daar et al. kan noen grunnleggende intraverbale responser gjerne læres ved bruk av prompt på denne måte, men for å kunne holde vanlige samtaler er det sjelden tilstrekkelig med en slik tilnærming. Intraverbale responser i samtaler innebærer ifølge dem ofte komplekse diskriminative stimuli og responser som ikke har blitt direkte trent. For eksempel vil svaret på «hva spiste du til middag i går» variere avhengig av hva du faktisk spiste dagen før. Det er dermed vanskelig å trene på dette ved bruk av prompt uten å måtte gjennomføre trening hver dag fordi det korrekte svaret på spørsmålet stadig endrer seg. Man kan tenke seg mange tilfeller der intraverbale responser vil være responser i former/ topografier personen aldri tidligere har fremvist etter den aktuelle  $S^D$ . Dette kan være

om personen i går prøvde en helt ny matrett til middag og dermed tidligere aldri har respondert med å oppgi denne matretten som svar på «hva spiste du til middag i går?».

Eikeseth og Smith (2013) beskriver at intraverbal atferd er under kompleks diskriminativ kontroll og dette er en mulig forklaring på at intraverbal atferd er så krevende å etablere. De beskriver tre ulike typer stimuluskontroll som vurderes som relevante for etablering av et funksjonelt intraverbalt repertoar: *diskriminerte operanter*, *kondisjonal diskriminasjon* og *sammensatte (compound) stimuli*. Diskriminerte operanter innebærer at en operant kommer under kontroll av foranledigende stimuli ( $S^D$ ), for eksempel at barnet sier hund i nærvær av en hund, men ikke i nærvær av andre dyr eller i fravær av hund. For intraverbal atferd vil dette være når den verbale responsen forekommer i nærvær av en bestemt verbal  $S^D$ , som når barnet sier «gå!» etter å ha hørt utsagnet «klar, ferdig..?», men ikke i nærvær av andre verbale stimuli, som for eksempel «A-B-..?».

Kondisjonal diskriminasjon er når den diskriminerte operanten kommer under kontekstuell kontroll, dette innebærer at stimuli i omgivelsene påvirker om et bestemt stimuli får funksjon som en  $S^D$  eller ikke. For intraverbal atferd vil dette være verbale utsagn som påvirker effekten av andre verbale utsagn. Eikeseth og Smith (2013) nevner eksempler som «Om du heter Charlie, si alfabetet». Det verbale utsagnet «si alfabetet» vil bare fungere som en  $S^D$  som signaliserer forsterkning for individer som faktisk heter Charlie, for andre vil utsagnet ikke fungere som en  $S^D$ , men heller som en *Sdelta* ( $S\Delta$ ) som er en stimulus som signaliserer fravær av forsterkning. Kondisjonal diskriminasjon er nødvendig for å respondere hensiktsmessig i ulike situasjoner.

Den siste varianten av stimuluskontroll som beskrives av Eikeseth og Smith (2013) som er relevant for etablering av et funksjonelt intraverbalt repertoar er sammensatt stimuluskontroll. Dette oppstår når en respons er under kontroll av to eller flere stimuli, eller flere elementer ved en stimulus. Eksempler på dette i forhold til intraverbal atferd kan være

når den foranledigende verbale stimulus er «si en rød frukt». Her vil responsen kontrolleres av både frukt og farge, individet kan ikke respondere bare i henhold til et av dem uten å risikere å svare galt. For eksempel vil banan, kiwi og appelsin være korrekte svar på «frukt» og brannbil og paprika være korrekte svar på noe som er rødt, men ingen av svarene vil være korrekte i henhold til den sammensatte stimulusen som utgjør spørsmålet.

Å kunne diskriminere ulike komplekse verbale stimuli er ifølge Eikeseth og Smith (2013) en prerequisitt for kompleks intraverbal atferd. Men i tillegg til dette må målpersonen også lære å produsere komplekse verbale responser, noe som kan være svært krevende. Det kan derfor være hensiktsmessig å etablere disse ferdighetene hver for seg for å gjøre prosessen lettere. Å trene på diskriminasjon av komplekse verbale stimuli er mulig selv for personer med begrenset eller lite vokalt repertoar gjennom *lyttertrenting* (Eikeseth & Smith, 2013). Ved lyttertrenting presenteres målpersonen for en verbal kondisjonal stimulus og responderer ofte ved å peke på et av flere bilder (Petursdottir, Carr, Lechago, & Almason, 2008). For eksempel kan målpersonen få se tre bilder, et av en banan, et av en bil og et av en lege. Målpersonen får oppgitt en kategori for eksempel «frukt» og bare om han da peker på bildet av bananen blir forsterkning formidlet, bildet av bananen vil få funksjon som en  $S^D$ . Om målpersonen peker på bildene av bil og lege vil forsterkning holdes tilbake, disse bildene vil være  $S\Delta$  for seleksjonsrespons. Gjennom lyttertrenting er det mulig å trene diskriminasjon av komplekse verbale stimuli uten å samtidig måtte trene på å produsere komplekse verbale responser.

Smith et al. (2016) undersøkte effekten av lyttertrenting på emergens av intraverbal respondering hos fem barn som var mellom fem og 15 år gamle, alle med en autismediagnose. Hvert av barna trente ved lyttertrenting på å svare på spørsmål som «hvilket dyr er grått?», «hvilken drikke er hvit?» og «hvilket dyr flyr?» gjennom å peke på bilder, i disse tilfellene bilder av elefant, melk og fugl. Ved mestring under trening ble det gjennomført en test der

samme spørsmål ble stilt, men barnet måtte respondere vokalt og ikke ved å peke på bilder. Resultatene viste at fire av de fem barna fremviste emergente intraverbale responser til spørsmålene etter lyttertrening. Det siste barnet hadde også økning av emergente intraverbale responser etter lyttertrening, men mestring først etter at det var gjennomført tacttrening.

Lignende resultater hadde også Vallinger-Brown og Rosales (2014). De undersøkte effekten av lyttertrening og stimulusparing for å emergens av intraverbale responser hos tre barn med autisme. Under lyttertrening svarte barna på intraverbale spørsmål ved å trykke på et av flere bilder på en iPad®, forsterkning ble bare formidlet om barnet trykket på bildet som hørte til spørsmålet. For eksempel ved spørsmålet «pek på den som viser tiden» ble forsterkning formidlet om barnet trykket på bildet av en klokke. Denne metoden viste seg å for disse barna å være mer effektiv enn stimulusparing, der treningen bestod av at spørsmålet ble stilt samtidig som barnet trykket på bildet på skjermen, for eksempel om barnet trykket på bildet av klokken fikk det høre «klokke brukes til å vise tiden». For alle tre barna var det økning i emergent intraverbal respondering etter lyttertrening. Ett av barna hadde 100% mestring ved de fire første målingene etter lyttertrening, og den intraverbale responderingen holdt seg stabilt høy. For to av barna ble det gjennomført direkte trening på intraverbale responser for å øke frekvensen ytterligere.

Petursdottir, Olafsdottir og Aradottir (2008), Petursdottir og Haflidadottir (2009) og Petursdottir, Lepper og Peterson (2014) undersøkte effekten av lyttertrening for emergens av intraverbaler på fremmedspråk for normalfungerende barn i fire- til femårsalder. I alle tre studiene ble det målt intraverbaler der barnet svarer på spørsmål stilt på fremmedspråk med responser på morsmål, og spørsmål stilt på morsmål med responser på fremmedspråk. For alle barna i de tre studiene var det økning i intraverbal respondering etter lyttertreningen. Petursdottir et al. (2014) gjennomført to ulike eksperimenter. I eksperiment 1 ble det utelukkende gjennomført en pre- og en posttest etter vanlig lyttertrening, det er derfor

problematisk å trekke konklusjoner ut fra denne studien da vanlige trusler mot validitet ikke kan utelukkes. Likevel vises det for alle de fire barna en økning i emergente intraverbale responser etter den gjennomførte lyttertreningen. For tre av barna var det størst økning i korrekte responser under betingelsen der de ble stilt spørsmål på fremmedspråk og svarte på morsmål. For det siste barnet var det ingen forskjell mellom å svare på morsmål eller å svare på fremmedspråk. I eksperiment 2 ble det foretatt gjentatte målinger for hvert barn. For deltakerne i eksperiment 2 konkluderes det med at lyttertrening førte til noe økning i emergente intraverbale responser, men at lyttertrening ikke førte til måloppnåelse (5 av 6 korrekte responser) under betingelsen der barnet ble stilt spørsmål på morsmål og skulle svare på fremmedspråk, da dette bare forekom ved to av probemålingene for en av deltakerne. Petursdottir, Olafsdottir, et al. (2008) og Petursdottir og Hafliðadottir (2009) har lignende konklusjoner. Generelt var det høyere frekvens av korrekte responser når barna svarte på fremmedspråklige spørsmål med sitt eget morsmål, enn motsatt. Ved alle målinger i disse studiene var frekvensen av korrekte responser generelt høyere etter lyttertrening enn ved baselinemålingene.

Det er også gjennomført studier der det ikke er påvist klare effekter på emergent intraverbal respondering etter at lyttertrening er gjennomført. Petursdottir, Carr, et al. (2008) undersøkte blant annet effekten av lyttertrening på intraverbale kategoriseringsresponser hos tre normalfungerende barn. Etter lyttertrening var det ingen forskjeller i respondering fra baselinemålinger (0 korrekte responser). For ett av barna forekom det en emergent intraverbal respons ved to ulike målinger, men dette var etterfulgt av tre målinger med 0 korrekte responser.

En annen studie på kategoriseringsferdigheter ble gjennomført av Miguel, Petursdottir, og Carr (2005) de undersøkte effekten av multiple-tacttrening, lyttertrening (receptive discrimination training) og direkte intraverbal trening. Lyttertreningen ble gjennomført

lignende det som tidligere er beskrevet, barnet skulle velge ut et av flere bilder avhengig av en verbal stimulus. Multiple-tacttrening bestod av at barnet ble trent til å i første fase tacte/navngi 10 objekter fra en kategori og i andre fase trent til å både navngi objektet og kategorien objektet tilhørte. Det ble brukt stimuli fra tre ulike kategorier: musikkinstrumenter, verktøy, og kjøkkenutstyr for eksempel trommer, piano, hammer, drill, gaffel og mikrobølgeovn. Seks normalfungerende barn mellom tre og fem år deltok i studien. Hvert av barna fikk trening på to av kategoriene, tre av barna startet med multiple-tacttrening og fikk deretter lyttetrening, de andre tre startet med lyttetrening og fikk deretter multiple-tacttrening. For fem av de seks deltakeren ble det avslutningsvis gjennomført direkte trening på intraverbaler. Miguel et al. konkluderer med at verken multiple-tacttrening eller lyttetrening førte til en betydelig økning i intraverbal respondering og direkte trening på intraverbaler var nødvendig for at en slik effekt skulle oppstå.

En annen studie som viser uklare effekter av lyttetrening ble gjennomført av Kodak og Paden (2015). De gjennomførte lyttetrening med to barn med autismespekterforstyrrelser. Et av barna, Rosemary fikk trening i å svare på «hvem-spørsmål», det andre barnet, Oscar trente på spørsmål om hva ulike objekter brukes til. For Rosemary så ikke lyttetrening ut til å ha noen effekt på den intraverbale responderingen, da det etter trening ikke var noen endring i frekvensen av intraverbale responser i forhold til baselinemålingene. For Oscar økte frekvensen av emergente intraverbale responser til mellom 80-100% etter lyttetrening, fra 0% korrekte responser under baselinemålingene.

Resultatene og konklusjonene fra tidligere gjennomførte studier er altså varierende. Smith et al. (2016) beskriver to mulige forklaringer på at deltakerne i deres studie etter lyttetrening har økning i frekvens av korrekte emergente intraverbale responser, mens dette ikke var tilfelle i flere andre studier. For det første var den foranledigende stimulusen som ble presentert under trening identisk med den som ble presentert under testing. Spørsmålene som

ble brukt var de samme, mens responsene (pekerespons og vokalrespons) var ulike. For eksempel blir det under testing for kategorier i Miguel et al. (2005) spurt om barnet kan «Si noen [kategori], fortell meg så mange du kan» mens det under lyttertrening blir sagt «Pek på [kategori]». En annen mulig forklaring til de observerte forskjellene i effekt var at det ble sikret at barnet kunne ta alle stimuli som ble brukt, og respondere til dem med lytteratferd, i forkant av lyttertreningen.

Å svare på når-spørsmål er en type intraverbal respondering som sammen med lignende spørsmål (hvem, hva, hvor, hvordan) vurderes som viktig fordi dem inngår i vanlige former for kommunikasjon. Å besvare denne typen spørsmål er ofte forsinket hos barn med autismespekterforstyrrelser og det har vist seg å være en sammenheng mellom barnas språkutvikling og kompleksiteten på spørsmålene som stilles til barna (Goodwin, Fein, & Naigles, 2015). Når foreldrene tilpasser kompleksiteten på spørsmålene som stilles, for eksempel presenterer mange spørsmål som kan besvares med ja eller nei reduseres også antallet muligheter barna får til å trene på denne typen spørsmål. Manglende ferdigheter på dette området kan føre til at barnet får problemer med å delta i sosiale interaksjoner (Jahr, 2001) som for eksempel lek og samtaler med andre barn, noe som for denne gruppen er svært problematisk fordi det vil føre til enda færre muligheter for utvikling av sosiale ferdigheter og språk (Bishop, 2014). Typisk for programmer basert på anvendt atferdsanalyse er at denne typen spørsmål etableres ved at når-spørsmålet presenteres og deretter umiddelbart etterfølges av en prompt, som barnet skal imitere. For å sikre imitasjon av prompten og ikke spørsmålet oppfordres det gjerne til at spørsmålet stilles lavt, mens prompten blir presentert med høyere stemme. Prompten fades gradvis bort, slik at stimuluskontrollen overføres fra prompten til når-spørsmålet (Luce et al., 1996; McEachin & Leaf, 1999; Sundberg & Partington, 1998).

Med denne studien ønskes det å undersøke hvilken effekt lyttertrening har for emergent intraverbal respondering når det gjelder når-spørsmål for en gutt på fem år med



autismediagnose. Det ble under lyttertrening trent på å svare på tre sett med ulike spørsmål knyttet til aktiviteter, som for eksempel «når bruker du badebukse?» eller «når sover du?». Lyttertreningen ble gjennomført ved at responser ble forsterket når gutten pekte ut det av flere bilder som passet som svar på spørsmålet som ble stilt. I forkant ble det kontrollert at gutten kunne tacte de bildene som skulle brukes under lyttertrening. Det ble også målt om det etter lyttertrening oppstod emergente bidireksjonale intraverbale responser ved at det ble gjort endringer i hvordan spørsmålene ble stilt, sammenlignet med under lyttertrening og probemålinger, eksempler på disse spørsmålene var blant annet «hva har du på deg når du bader?» og «hva gjør du på natten?».

## **Metode**

### **Deltaker**

Deltaker var en gutt på fem år diagnostisert med barneautisme. ADOS (Autism diagnostic observation schedule, Lord et al., 2001) og ADR-I (Autism diagnostic interview-revised, Rutter, Le Couteur, & Lord, 2003) var benyttet som diagnoseredskaper og diagnosen var basert på kriteriene fra ICD-10. Deltaker hadde en IQ skåre på 71 målt med Stanford-Binet Intelligence Scales (Thorndike, Hagen, & Sattler, 1986). Se Tabell 1 for en oversikt over deltakers språklige ferdigheter. Deltaker gikk i en ordinær barnehage, og hadde spesialpedagogisk hjelp 12 timer i uken av støttepedagog.

Studien var innmeldt til og godkjent av Norsk senter for forskningsdata (NSD). Det ble innhentet samtykke fra deltakers foreldre, der det også ble gitt tillatelse til at deler av trening og testingen kunne filmes.

### **Kartlegging**

I forkant av studien ble det kontrollert at deltaker kunne tacte de aktuelle stimuliene som skulle brukes. Dette ble gjennomført i to trinn. Først ble gjennom samtale med deltakers støttepedagog vurdert hvilke bildestimuli som kunne være aktuelle å bruke. Deretter ble

aktuelle stimuli skrevet ut som bilder i ca. 7 x 10cm format. Bildene ble et og et presentert for deltaker mens trener sa «hva er dette?». Riktig respons ble notert om deltaker innen fem sekunder svarte ved å oppgi hvilken hendelse eller aktivitet det var bilde av uten ytterligere prompting. Bare de bildestimuliene deltaker kunne tacte ble brukt i studien.

Det ble også gjennomført en samtale med deltakers støttepedagog og hennes veileder for å gjøre en vurdering av hva som kunne være aktuelt å bruke som forsterkere ved gjennomføring av tiltaket. Basert på støttepedagogens tidligere erfaringer med deltaker i treningssituasjon ble det vurdert å være tilstrekkelig med ros, high five og lignende under trening/testing etterfulgt av leking eller lesing i en bok sammen med trener etter endt økt. Under leking ble det benyttet såpebobler og andre leker. Bøkene som ble valgt til formålet var for det meste pekebøker med lite tekst og med luker som deltaker selv kunne åpne, eller kortere historier og eventyr.

Den delen av ABLLS-r (Partington, 2012) som omhandler intraverbaler ble gjennomført i forkant av studien. Resultatet fra denne viste at deltakeren totalt mestret 21,35% av de 178 mulige ferdighetene som kan skåres i denne kategorien. I slutfasen av tiden studien pågikk ble hele ABLLS-r gjennomført, resultatene fra denne kan ses i Tabell 1. Skåringen av de intraverbale ferdighetene ble altså igjen gjennomført i slutfasen av denne studien.

Det ble videre kartlagt at deltaker kunne svare korrekt på noen spørsmål som senere ble brukt under testing for å kunne formidle forsterkning. Eksempler på slike spørsmål var: «hva heter du?», «hvor mange år er du?», «hva heter mammaen din?» eller spørsmål om hvilke lyder ulike dyr lager. Spørsmålene ble valgt ut i samråd med deltakers spesialpedagog.

### **Setting og materiell**

Studien ble gjennomført i deltakers barnehage, på et eget rom der han fra tidligere var vant til å gjennomføre en-til-en trening. I rommet var det et bord med benker i barnestørrelse som eksperimentator og deltaker kunne sitte ved under trening og testing. Under lyttertrening

ble det brukt tre sett med totalt ni ulike bilder som ble presentert for deltaker i en A5 perm. Hver side i permen var laminert og bestod av tre bilder fra det aktuelle settet. Hvert av bildene hang sammen med et spørsmål om aktiviteter/steder. Se Tabell 2 for oversikt over hvilke spørsmål, svar og bilder som ble anvendt. Bildene ble hentet fra Google Images™. Totalt i studien ble det brukt 27 ulike laminerte sider i A5permen. Bildet som under lyttertrening var det korrekte svaralternativet kunne forkomme lengst til høyre, i midten eller lengst til venstre. Plasseringen av bildene i forhold til hverandre på hver side i permen var semirandomisert ved at fordelingen i første omgang ble avgjort ved bruk av et online randomiseringsverktøy (Urbaniak & Plous, 2013) og deretter justert slik at hvert bilde i settet minst tre ganger forekom i hver av de ulike posisjonene. Fordeling til settene var også semirandomisert og ble gjennomført ved bruk samme online randomiseringsverktøy (Urbaniak & Plous, 2013). Bildene ble tilfeldig fordelt til settene, og deretter ble det kontrollert at det ikke var noen bilder i samme sett som var så like at de ble vurdert å kunne gjøre lyttertreningen vanskeligere for det settet. Dette innebar at bildene som skulle representere «morgen» med bilde av en soloppgang og bildet av «sommer» med en sol over en strand ble fordelt i hvert sitt sett på tross av at den første fordelingen førte til at de var fordelt til samme sett.

Ved tacttrening ble bildene brukt et bilde om gangen, uten bruk av permen. Bildene ble i forkant av tacttrening skrevet ut og laminert hver for seg, de hadde en størrelse på ca. 7x10 cm.

Ved filming ble kamerafunksjonen på en Sony Xperia Z3 compact brukt.

For å registrere data ble de brukt registreringsskjema under både trening og testing. Det ble også brukt fire ulike skjema for å registrere prosedyreintegritet under lyttertrening, tacttrening og de to ulike testene (Appendiks A-C).

### **Avhengig variabel og reliabilitet**

Antall korrekte- og ukorrekte responser og antall prompts ble registrert under treningsbetingelsene. Under testing ble det målt antall korrekte og ukorrekte responser.

Avhengig variabel var antall korrekte responser. Under lyttertrening ble dette definert som at deltaker pekte på korrekt bilde innen fem sekunder etter at den kondisjonale stimulusen ble presentert. Under testing av emergent intraverbal atferd samt emergente bidireksjonale intraverbaler ble korrekt respons definert som at deltaker oppga korrekt svar innen fem sekunder etter at den verbale stimulusen ble presentert (se Tabell 2).

Reliabilitet ble målt ved at to observatører uavhengig av hverandre observerte samme økt og skåret ut fra de samme definisjonene med bruk samme registreringsskjema.

Reliabilitetsmålinger ble gjennomført ved lytter- og tacttrening og ved baseline og tester for emergente intraverbaler og emergente bidireksjonale intraverbaler. Observatørene fikk i forkant en innføring i hvordan registrering ved bruk av registreringsskjemaene skulle gjennomføres. Enkelte økter ble filmet for sikre muligheten for flere reliabilitetsmålinger. Filmene ble skåret i etterkant.

Mellomobservatørenighet (IOA) ble beregnet pr. økt ved at antallet observasjoner der observatørene var enige ble dividert på det totale antallet observasjoner i økten (enige+uenige) og deretter multiplisert med 100%.

IOA ble målt for totalt 50% av øktene under lyttertrening: 16.6% av øktene for sett 1, 50% av øktene for sett 2 og 100% av øktene for sett 3. Reliabilitetsmålingene var under lyttertrening gjennomsnittlig 100%, da observatørene var 100% enige ved alle målinger som ble gjennomført for alle sett.

Ved tacttrening for sett 1, ble det ved 71.43% av øktene målt IOA, gjennomsnittlig enighet 97.78 % med en range fra 88.8% til 100%.

Under baseline og test for emergente intraverbale responser ble IOA målt for totalt 60% av øktene (55.5% ved testing for sett 1; 61.5% for sett 2 og 53.8% for sett 3). Enigheten mellom observatørene under test for emergente intraverbaler var 100% ved alle gjennomførte målinger for alle sett.

Ved pre- og posttester for emergente bidireksjonale intraverbaler ble det målt IOA for totalt 75% av øktene. Reliabilitetsmålingene viste 100% enighet mellom observatørene ved de målte testene.

### **Prosedyre**

**Lyttertrening.** Det ble gjennomført to treningsøkter på ca. 2 minutter fire-fem ganger i uken over ca. 2 uker. Maksimalt to treningsøkter pr. dag. Under trening lå en side i permen med tre bilder på bordet fremfor deltaker. Eksperimentator presenterte et av «når» spørsmålene og deltakers målrespons var å peke på det av bildestimuliene som hørte til som korrekt svar på spørsmålet (se Tabell 2). Hvert av spørsmålene i det aktuelle settet ble stilt tre ganger pr. treningsøkt, slik at hver av treningsøktene bestod av totalt ni spørsmål. Ved korrekt respondering bladde trener til neste side i permen slik at tre nye bilder lå fremfor deltaker. Ved ukorrekt eller manglende respondering (mer enn fem sekunder fra presentasjon av kondisjonal stimulus) ble deltaker promptet ved at trener pekte på korrekt bildestimuli og deretter presenterte den kondisjonale stimulusen på nytt slik at deltaker fikk mulighet til å respondere korrekt. Ved korrekt respons ble ros formidlet, ved ukorrekt respons ble kondisjonal stimulus igjen presentert og promptet formidlet. Promptede responser ble registrert som ukorrekte. Deltakers responser ble forsterket etter et FR1 skjema, med ikke-atferdsspesifiserende ros, high five og lignende. Etter endt økt fikk deltaker tilgang til leker eller lesing med eksperimentator. Mestringskriteriet ble satt til at deltaker hadde åtte av ni responser korrekt to påfølgende treninger. Etter mestring ved FR1, ble forsterking under

trening formidlet etter ett VR3 skjema. Kriteriet for mestring ved VR3 var det samme som ved FR1.

Ved manglende mestring ved testing for emergente intraverbaler etter lyttertreningen, ble lyttertrening igjen innført i tråd med denne prosedyrebeskrivelsen, med unntak av at det ikke ble benyttet et FR1 forsterknings skjema. Forsterkning ble formidlet etter VR3 skjema, til det var oppnådd mestring (minst åtte av ni korrekte responser i to påfølgende økter).

**Testing.** Testing ble gjennomført på samme måte ved baseline og probemålinger. Det ble gjennomført to ulike tester. Begge testene ble gjennomført i samme rom som det ble gjennomført trening, ved bordet der trener og deltaker satt overfor hverandre. Testene ble gjennomført under ekstinksjon for de responsene det ble testet for. Forsterkning under testbetingelsen ble formidlet etter et VR3 skjema der deltaker mellom test-spørsmålene fikk mulighet til å svare på spørsmål det under kartlegging var klarert at han kunne svare på. Svar på disse spørsmålene ble forsterket med ikke-atferdsspesifiserende ros, high five og lignende. Når testen var gjennomført fikk deltaker leke eller lese med eksperimentator.

**Test av emergente intraverbaler.** Test for emergente intraverbaler ble gjennomført ved baseline, etter mestring av lytter- eller tacttrening på et VR3 forsterknings skjema og i opprettholdelsesfasen. Ved test for emergente intraverbaler ble spørsmålene stilt på samme måte som under lyttertrening, men deltaker hadde ikke permen med bildestimuli fremfor seg, og måtte svare vokalt fremfor å peke på bilder. Ved korrekt respons ble neste spørsmål presentert. Svarte deltaker feil eller ikke svarte i løpet av fem sekunder ble samme spørsmål presentert en gang til. Uavhengig av deltakers svar på det andre spørsmålet gikk trener etter dette videre til neste spørsmål, samme spørsmål ble dermed maksimalt stilt to ganger under samme test.

Ved noen riktige responser ble testing gjennomført til stabilitet. Stabilitetskriteriet var to påfølgende tester med samme resultat. Ved ingen korrekte responser under test ble testen

gjennomført en gang til. Om det etter to tester ikke forekom noen korrekte responser ble lyttertrening igjen gjennomført på samme måte som beskrevet tidligere. Ved mestring under lyttertrening ble det foretatt en ny test for intraverbaler. Om det nå forekom korrekte responser fortsatte testing til stabilitet. Ved ingen korrekte responser ble det gjennomført tacttrening.

Mestringskriteriet under test for intraverbaler var to påfølgende målinger med 90% eller høyere korrekt respondering. Ved mestring under test ble lyttertrening igangsatt for neste sett. Ved mestring for sett 3 ble det ved opprettholdelsesmålinger lagt inn distraktorer ved at det ble stilt flere andre spørsmål som «hvem sier kykkeliky?», «hvor mange år er du?» og lignende. Rekkefølgen på spørsmålene ble endret ved hver gjennomført test og mikset på tvers av de tre settene. Forsterkning ble også i denne opprettholdelsesfasen formidlet for korrekt respondering på andre spørsmål etter et VR3 forsterkningskjema.

**Test av emergente bidireksjonale intraverbaler.** Ved første måling med 100% korrekt respondering på test for emergente intraverbaler ble det gjennomført en test for emergente bidireksjonale intraverbaler. De stimuli som tidligere hadde vært bildestimuli/ responser under test var nå en del av spørsmålet. Eksempler på slike spørsmål var «hvem kommer når det er en tyv som stjeler?», «hva har du på deg når du bader?» eller «hva gjør du når det er natt?». Se Tabell 3 for en oversikt over spørsmålene som ble brukt ved disse testene. Prosedyren for testing av emergente bidireksjonale intraverbaler var den sammen som ved testing av emergente intraverbaler. Test for emergente bidireksjonale intraverbaler ble gjennomført en gang før lyttertrening startet og deretter en gang pr sett i etter måloppnåelse ved test i tiltaksfasen.

**Tacttrening.** Ved ingen korrekte responser ved to påfølgende tester for emergente intraverbaler, ble det gjennomført flere økter med lyttertrening, til måloppnåelse, deretter ble

det gjennomført ny test for emergente intraverbaler. Om det heller ikke nå forekom noen korrekte responser ble tacttrening gjennomført.

Under tacttrening ble deltaker stilt de samme spørsmålene som ble brukt under testene for intraverbaler. Samtidig som spørsmålet ble stilt viste eksperimentator deltaker bildet som hørte til spørsmålet (se Tabell 2). For eksempel bildet av en som pusset tennene til spørsmålet «når trenger du tannkrem?». Forsterkning ble i utgangspunktet formidlet etter hver korrekte respons (FR1). Mestringskriteriet var åtte av ni korrekte responser to påfølgende treningsøkter. Etter mestring ble gjennomsnittlig hver tredje respons forsterket (VR3) frem til test. Som under lyttertrening ble det heller ikke under tacttrening formidlet atferdsspesifiserende ros, men generell ros, high five og lignende, etterfulgt av leking eller lesing i slutten av økten.

### **Prosedyreintegritet**

Prosedyreintegritetsmålinger ble gjennomført for å undersøke hvorvidt trening og testing ble gjennomført i tråd med prosedyrebeskrivelsene. Trening og testing ble bare gjennomført av eksperimentator eller deltakers støttepedagog etter at hun hadde observert gjennomføring av trening og testing. Prosedyreintegritet ble målt ved at en observatør direkte observerte utvalgte trening- (lytter- og tacttrening) og testøkter (emergente intraverbaler og bidireksjonale emergente intraverbaler), når dette var praktisk mulig.

Observatøren hadde et registreringsskjema med beskrivelser av eksperimentators atferd og i hvilken rekkefølge de ulike momentene i treningen og testingen skulle gjennomføres. For eksempel under lyttertrening stod det på skjemaet: «eksperimentator legger tre bilder (en side i permen) på bordet», «stiller et når-spørsmål» og «ved korrekt respons, gir generell ros». Se Appendiks A-C for registreringsskjemaer brukt ved prosedyreintegritetsmålinger under trening og testing. Ved endring av forsterkningskjema under trening (fra FR1 til VR3) ble den delen av skjemaene som omhandler forsterkning



endret og det ble registrert som korrekt når det ikke ble formidlet forsterkning for hver trial. Observatørene fikk i forkant en innføring i hvordan de skulle registrere ved bruk av skjemaene. For å sikre flere målinger av prosedyreintegritet ble noen test- og treningsøkter filmet og skåret i etterkant.

Prosedyreintegritet ble beregnet ved å skåre antallet korrekte implementerte responser hos den som gjennomførte trening eller testing, delt på antallet mulige korrekte responser pr økt, multiplisert med 100%.

Totalt ble prosedyreintegritet målt ved 50% av øktene med lyttertrening, 16.6% av øktene med lyttertrening i sett 1, 50% for sett 2 og 100% for sett 3.

Prosedyreintegritetsmålingene for lyttertrening var gjennomsnittlig 100% for både sett 1, sett 2 og sett 3. Det forekom ingen målinger der prosedyreintegriteten ikke ble skåret til 100%.

Ved tacttrening for sett 1 ble prosedyreintegritet målt ved 71.43% av øktene, resultatet var 100% ved alle gjennomførte målinger.

Under baseline og testing for emergente intraverbaler ble prosedyreintegritet målt for totalt 35% av øktene. Prosedyreintegritetsmålingene under test for emergente intraverbaler ble målt til 100% ved alle gjennomførte målinger.

Ved pre- og posttest for emergente bidireksjonale intraverbaler ble prosedyreintegritet målt ved totalt 50% av øktene. Prosedyreintegritetsmålingene var 100% ved alle de gjennomførte målingene.

## **Design**

Det ble brukt en multiple probe design på tvers av tre sett med spørsmål.

## **Resultat**

### **Trening**

For sett 1 ble mestring oppnådd etter 36 trials med lyttertrening (18 trials ved FR1 og 18 trials ved VR3), men på grunnlag av manglende måloppnåelser ved gjentatte tester for

emergente intraverbaler, også etter ytterligere 18 trials med lyttertrening (VR3) ble det gjennomført tacttrening. Under tacttrening ble det totalt gjennomført 63 trials før mestring, 45 trials ved FR1 og 18 trials ved VR3.

For sett 2 ble det gjennomført 36 trials før mestring ved lyttertrening, 18 ved FR1 og 18 ved VR3. For sett 3 ble det også gjennomført 36 trials, 18 på hvert forsterknings skjema. Ved lyttertrening for sett 2 og 3 forekom det ingen ukorrekte responser og mestring kunne ifølge prosedyren ikke ha blitt oppnådd med færre trials. For sett 2 og 3 var det ikke nødvendig med tacttrening da 100% korrekt og stabil respondering under test for intraverbaler forekom umiddelbart etter lyttertrening.

### **Emergente intraverbaler**

Figur 1 viser effekten av lytter- og tacttrening på emergente intraverbale responser hos deltaker før og etter lytter- og tacttrening for sett 1, 2 og 3. Baselinemålingene viste stabil fraværende forekomst av korrekte responser, med 0% korrekte responser i alle baselinemålingene for sett 2 og 3. For baselinemålingene for sett 1 var det en forekomst av en korrekt respons på spørsmålet «når bruker du fargeblyanter?» da deltaker svarte «tegner», dette var etterfulgt av tre målinger med ingen korrekte responser. Etter lyttertrening økte umiddelbart andelen korrekte responser til 100% for sett 2 og 3. For sett 1 hadde ikke lyttertrening denne umiddelbare effekten og etter ytterligere lyttertrening med fraværende effekt under test ble tacttrening igangsatt. Etter tacttrening hadde deltaker en umiddelbar økning til 100% korrekt respondering under test for dette settet.

I opprettholdelesesmålingene som ble gjennomført etter mestring på alle tre sett der spørsmålene var mikset og det var lagt inn distraktorer i form av andre spørsmål var det to målinger med 66.6% korrekt respondering for sett 2, og en måling med 66.6% korrekt respondering for sett 3, ved de resterende målingene hadde deltakeren 100% korrekte

responser. Ved de to siste testene som ble gjennomført med fire ukers mellomrom hadde deltaker 100% korrekt respondering på alle spørsmål i alle sett.

### **Emergente bidireksjonale intraverbaler**

Tabell 8 viser effekten av lytter- og tacttrening på emergente bidireksjonale intraverbaler for sett 1, 2 og 3. Ved pretest forekom det en korrekt respons for sett 3.

Ved post-test for sett 1 forekom det en korrekt respons, og dermed en økning fra 0% korrekte responser til 33.3% for dette settet. For sett 2 forekom det ved posttest tre korrekte responser, og deltaker hadde dermed 100% korrekt respondering for dette settet etter lyttertrening. Posttest for sett 3 viste en korrekt respons, som for øvrig var den samme responsen («hva har du på deg i bilen?», «bilbelte») som forekom ved pretesten.

### **ABLIS-r**

Ved kartleggingen som ble gjennomført før studien startet ble det skåret at deltaker hadde mestring på 21.35% av de intraverbale ferdighetene på ABLIS-r (Partington, 2012). Ved studiens avslutningsfase ble det gjennomført en ny vurdering og registrert en skåre på 44.51% for de intraverbale ferdighetene.

## **Diskusjon**

Målet med denne studien var å undersøke effekten av lyttertrening for etablering av emergente intraverbale responser hos en fem år gammel gutt med autisme. Resultatene fra studien viste at lyttertrening kan ha hatt effekt på etableringen av emergente intraverbale responser for to av de tre settene det ble trent på. Det var også en økning i antallet korrekte responser ved testing for emergente bidireksjonale intraverbaler. For det første settet det ble trent på førte ikke lyttertrening alene til økning i intraverbale responser, men etter tacttrening økte andelen korrekte responser umiddelbart til 100%. For sett 2 og 3 ble det etter lyttertrening en umiddelbar økning til 100% korrekt respondering ved testene for emergente intraverbaler. Resultatene fra denne studien støtter dermed tidligere studier som konkluderer

med effekt av lyttertrening på emergente intraverbale responser (Petursdottir & Haflidadottir, 2009; Petursdottir et al., 2014; Petursdottir, Olafsdottir, et al., 2008; Smith et al., 2016; Vallinger-Brown & Rosales, 2014). Deltaker hadde også en økning av andelen intraverbale ferdigheter som ble skåret som mestret ved ABLLS-R (Partington, 2012), når målinger før og etter studien ble gjennomført sammenlignes.

At tacttrening hadde effekt på intraverbal respondering for sett 1 står i kontrast til funnene til Miguel et al. (2005), som konkluderer med at tacttrening ikke hadde effekt på intraverbal respondering, verken når den ble gjennomført før eller gjennomført etter lyttertrening. Likevel kan det ikke utelukkes at de observerte forskjellene i effekt skyldes hvilken type intraverbaler det ble trent på eller andre forhold ved deltaker og prosedyre. Det er heller ikke mulig, ut fra denne studien å konkludere med at det var tacttreningen som førte til den observerte effekten for sett 1. Tacttreningen blir i denne studien bare innført for et av de tre settene, og dermed studert ut fra en AB-design der det ikke er muligheter for mer enn en demonstrasjon av effekt. For å kunne trekke konklusjoner i forhold til effekten av tacttrening vil det være nødvendig å studere dette med en design som gjør det mulig å utelukke vanlige trusler mot validitet, for å i større grad kunne utelukke at de observerte endringene skyldes noe annet enn den gjennomførte intervensjonen.

Under lyttertrening for sett 2 og 3 startet deltaker også å tacte bildestimuliene, samtidig som han pekte, selv om dette ikke var nødvendig for å oppnå forsterkning. Dette kan muligens forklares av tacttreningen som ble gjennomført for sett 1. Det kan ikke utelukkes at effekten av lyttertreningen som ses for sett 2 og 3 er påvirket av tacttreningen som ble gjennomført for sett 1, noe som igjen vil stå i kontrast til funnene til Miguel et al. (2005). Likevel var det ingen endring i antallet korrekte responser under baselinemålingene for sett 2 og 3, før tiltaket ble igangsatt, noe som kan tyde på at lyttertreningen var nødvendig for å frembringe den observerte endringen i respondering.

For sett 1 var det etter tacttrening økning i korrekte responser ved målingen av de emergente bidireksjonale intraverbalene. For sett 2 var det en økning til 100% korrekte emergente bidireksjonale intraverbaler etter lyttertrening. For sett 3 var det ingen forskjell på målingene under pre- og posttestbetingelsene, og lyttertrening så ikke ut til å ha effekt på emergente bidireksjonale intraverbaler for dette settet. Det kunne vært interessant å gjennomføre flere tester av de emergente bidireksjonale intraverbalene for alle sett for å i noe større grad vurdere effekten av tact- og lyttertrening, og å kunne utelukke en del trusler mot validitet, for eksempel at manglende resultater ved test for sett 1 og 3 var tilfeldige.

Resultatene i denne studien støtter opp om konklusjonene til Petursdottir, Olafsdottir, et al. (2008) når de beskriver at lytter- og tacttrening kan gi økning i emergente intraverbale responser, men at disse responsene ikke nødvendigvis alltid blir bidireksjonale. For sett 2 og muligens også sett 1, kan det tyde på at det har oppstått en emergent bidireksjonalitet når deltaker etter lytter- og tacttrening svarer korrekt på spørsmål han tidligere ikke mestret og som det under trening ikke har blitt trent på. For sett 3, ser det derimot ut til at dette ikke er tilfelle. Men fordi deltaker allerede i baseline svarte rett på et av de tre spørsmålene er det for dette settet bare to muligheter for å demonstrere tiltakets effekt, og når det bare ble gjennomført en post-test måling er det problematisk å trekke konklusjoner.

Funnene i denne studien står i kontrast til funnene til Petursdottir, Carr, et al. (2008), Miguel et al. (2005) og delvis Kodak og Paden (2015) som konkluderte med ingen, eller minimal effekt av lyttertrening på emergente intraverbale responser. Som beskrevet tidligere presenterer Smith et al. (2016) mulige forklaringer på de store forskjellene i konklusjoner for de ulike studiene. I denne studien, ble det i likhet med Smith et al. bare anvendt stimuli som deltaker i forkant av studien kunne tacte ble. Videre var deltaker vant til å jobbe med bilder under en-til-en trening ved å fremvise lytterresponser og gjorde dette nesten umiddelbart når lyttertrening ble igangsatt. Det ble også i denne studien brukt identiske spørsmål under lytter-

og tacttrening som det ble brukt ved testing, med unntak av testing for emergente bidireksjonale intraverbaler. Dette kan være mulige forklaringer til at lyttertrening så ut til å ha effekt for deltaker i denne studien.

Deltaker lærte seg tidlig spørsmålene som ble brukt under testing for emergente intraverbaler. Under baselinemålinger startet deltaker å gjenta/ imitere eksperimentators spørsmål fremfor å besvare dem, for eksempel kunne deltaker si «når står du opp» når eksperimentator sa «når står du opp?». Deltaker kunne også gjenta spørsmålene utenfor eksperimentsituasjonen, for eksempel når han møtte eksperimentator i barnehagen. Det kan tenkes at den manglende effekten av lyttertrening som oppstod for sett 1 var et resultat av at deltaker i baselinemålingene hadde lært å gjenta eksperimentators spørsmål i testsituasjonen, fremfor å prøve å besvare dem. Først da tacttrening ble igangsatt for sett 1 opphørte imitasjonen under test for sett 1. Imitasjonen av eksperimentators spørsmål fortsatte å forekomme i sett 2 og 3, frem til også disse settene var mestret under lyttertrening. Testsituasjonen ble dermed slik at deltaker svarte korrekt på de spørsmålene han hadde trent på ved tact- eller lyttertrening og svarte ved å imitere på utrente spørsmål, noe som ble registrert som ukorrekt respondering.

Likevel var det bare et tilfelle av imitasjon som respons på spørsmålene som ble brukt for å teste emergente bidireksjonale intraverbaler. Ved testene for sett 2 og 3 forekom det ingen imitasjoner av eksperimentators spørsmål. Dette kan tyde på at deltaker lærte seg de spørsmålene som ble brukt under testen for emergente intraverbaler, da denne testen ble gjennomført relativt ofte, men at han ikke lærte seg spørsmålene som ble brukt ved testing for emergente bidireksjonale intraverbaler.

Det kan ikke utelukkes at effektene som ses i denne studien kan forklares av andre årsaker enn den gjennomførte lyttertreningen. Dette er i denne studien problematisk å konkludere med effekten av lyttertrening når det var nødvendig med tacttrening for sett 1 og

det er dermed bare en demonstrasjon (sett 2) og en replikasjon (sett 3) av effekt, og ikke to replikasjoner som er ønskelig ved et multiple probe design. Likevel kan det argumenteres for at det var tre muligheter for demonstrasjon av effekt, der effekt bare ble fremvist i to av de tre tilfellene.

Hvorvidt resultatene som ses her også vil fremkomme ved gjennomføring av lignende trening med andre barn er et empirisk spørsmål. Det er nødvendig å gjennomføre flere studier som denne, med andre deltakere for å kunne vurdere lyttertrening som metode og i hvilken grad det bør være den anbefalte metoden for etablering av intraverbale responser. Likevel er det mulig ut fra denne studien å konkludere om mulige effekter av metoden for deltaker, noe som har verdi sett fra et anvendt atferdsanalytisk ståsted (Baer et al., 1968).

Videre kan det heller ikke utelukkes at effekten som ses etter lyttertrening i denne studien ikke ville fremkommet om det hadde blitt trent på en annen type intraverbal respondering. Petursdottir, Carr, et al. (2008) fant ingen effekt av lyttertrening på emergente intraverbale responser i forhold til kategoriseringsferdigheter. Det kan ikke konkluderes med at lyttertrening vil fungere for enhver type av intraverbale responser, selv om det ser ut til å muligens ha hatt effekt på når-spørsmål for deltaker i denne studien.

Det kan heller ikke trekkes konklusjoner i forhold til om de intraverbale responsene vil forekomme og opprettholdes i naturlige settinger basert på denne studien, da det ikke ble foretatt noen målinger av dette. Det hadde vært interessant å undersøke om deltaker begynte å bruke de lærte responsene i andre settinger enn på treningsrommet. Generalisering til andre typer spørsmål eller oppgaver virker usannsynlig basert på data fra denne studien.

Baselinemålingene viste at det ikke var en økning i antall korrekte responser, til tross for mestring etter trening på andre sett, noe som kan indikere at deltakers responser ikke kom under kontroll av en generalisert kondisjonal diskriminasjon der han ble i stand til å svare på

ulike typer «når-spørsmål». Likevel kan målingene av de emergente bidireksjonale intraverbalene muligens tyde på en viss form for generalisering, da spesielt for sett 2.

En mulig styrke ved denne studien var at det ble gjennomført hyppige målinger av prosedyreintegritet som viste at prosedyrebeskrivelsen ble etterfulgt. Dette vil være en styrke om det skal gjøres vurderinger og review av lyttertrening som metode fordi det viser at resultatene fra denne studien ble oppnådd etter lyttertrening som ble gjennomført i tråd med den oppgitte prosedyrebeskrivelsen. Dermed er det mulig å konkludere i forhold til effekten av denne fremgangsmåten, i større grad enn det vil være mulig i studier der det ikke er gjennomført slike målinger (Wolery, 2011). Likevel kan målinger av prosedyreintegritet ikke førte til at man kan utelukke andre trusler mot validitet, da den observerte effekten kan skyldes noe annet enn selve lyttertreningen til tross for at treningen ble gjennomført i tråd med beskrivelsene av prosedyren.

Det ble også gjennomført prosedyreintegritetsmålinger av testene som ble gjennomført. Dette er en styrke da høye skårer viser at testene som ble gjennomført var like og at de observerte forskjellene i resultater på testene ikke skyldes at eksperimentator gjennomførte testene forskjellig. Det hadde vært ønskelig at det ble gjennomført flere prosedyreintegritetsmålinger for testene for sett 1 og ved pre-testen for måling av de emergente bidireksjonale intraverbalene, dette ble ikke gjort på bakgrunn av manglende tilgang til observatører og mulighet for filming.

Under tacttrening for sett 1 oppgav deltaker samme svar som han også hadde oppgitt under kartleggingen av hvilke stimuli som skulle brukes (se Tabell 2), likevel var det noe variasjon i svarene som ble avgitt under testene når bildestimuliene ikke lenger var tilstede. For eksempel svarte deltaker både «skrive, tegne og fargelegge» på spørsmålet «når bruker du fargeblyanter». Deltaker la også til flere ord i sine responser for noen av spørsmålene, for eksempel svarte han ofte «til å tegne» eller «til å pusse tennene». Dette er noe som kan



diskuteres å være problematisk da svarene ikke nødvendigvis er svar på «når-spørsmål», men responser som kan være svar på spørsmål om objekters funksjon, som «hva bruker du fargeblyanter til» eller «hva bruker du tannkrem til». Dette var bare tilfelle for disse to spørsmålene og det forekom ikke ved noen andre spørsmål at deltaker gav svar som var knyttet til objekters funksjon. Det var også en viss variasjon i svaret på spørsmålet «når kan du gå med shorts?» der deltaker vekslet mellom å svare «sommer» og «sommerferie». Begge disse responsene ble registrert som korrekte, da begge ble vurdert å være adekvate og funksjonelle i en anvendt setting og inneholdt ordet som var definert som målresponsen (se Tabell 2).

Det er videre en mulighet for at det var forskjeller på vanskelighetsgraden på de ulike settene. Det kan ikke utelukkes at det var egenskaper ved spørsmålene som gjorde sett 1 vanskeligere enn de to andre settene, og at dette kan forklare hvorfor det ikke ble måloppnåelse under testing for dette settet. Likevel hadde deltaker under lyttertrening allerede ved første økt mestring for alle sett, og prompting var bare nødvendig ved den aller første trial på det første settet, etter dette pekte deltaker 100% av gangene på korrekt bildestimuli etter presentasjon av kondisjonal stimulus. Dette kan indikere at det ikke var noen større forskjeller på vanskelighetsgraden av spørsmålene.

Spørsmålene deltaker besvarte ukorrekt under testing i opprettholdelsesfasen var spørsmålene «når kan du gå med shorts» og «når bruker du badebukse» der deltaker i totalt tre tilfeller besvarte disse spørsmålene med svaret som hørte til det andre spørsmålet («bader» på «når kan du gå med shorts» og «sommer» på «når bruker du badebukse»). Det kan diskuteres at disse spørsmålene er problematiske da svarene deltaker avgir er adekvate, selv om de i denne studien er definert som ukorrekte. Å snakke om å bruke badebukse på sommeren og shorts til bading vil ikke nødvendigvis være ukorrekt i andre settinger enn på treningsrommet i denne studien. Disse to spørsmålene står derfor i kontrast til andre spørsmål i denne studien,

for eksempel vil det trolig aldri være korrekt å besvare «når ringer du brannvesenet» med «pusser tenner» eller «når trenger du tannkrem» med «brann». Dermed kan det muligens diskuteres at disse to spørsmålene er noe mer krevende enn de andre syv spørsmålene.

Vanskelighetsgraden av spørsmålene kan også diskuteres i forhold til testing for emergente bidireksjonale intraverbaler. Spørsmålene som ble brukt ved denne testen var generelt sett mye lengre og mer komplekse enn spørsmålene som ble brukt ved testen for emergente intraverbaler. Ved test for emergente intraverbaler inneholdt spørsmålene mellom tre og seks ord, gjennomsnittlig fire ord. Ved testene for emergente bidireksjonale intraverbaler inneholdt spørsmålene mellom fem og ni ord, med et gjennomsnitt på 6.5 ord. Det er mulig at dette har påvirket i hvilken grad deltaker har oppfattet spørsmålet og fått med seg alle detaljer som var nødvendig for å avgi et adekvat svar. For eksempel svarte deltaker aldri korrekt på spørsmålet «hvem kommer når det er en tyv som stjeler?» selv om han ved 100% av testene for emergente intraverbaler svarte korrekt på «når kommer politiet?». Det er mulig at spørsmålet i dette tilfellet ble for komplekst og langt for at deltaker skulle være i stand til å besvare spørsmålet. Det kan også være tilfelle at svaret som krevdes var for komplisert for deltaker å avgi, for eksempel besvarte han alle spørsmål om «når kommer brannvesenet?» med «brann» eller «det brenner» etter at lyttertrening var gjennomført for sett 3, men han besvarte aldri spørsmålet «hva gjør man når det brenner?» med «ringer brannvesenet» eller noe lignende. Det er fullt mulig at svarene som kreves på dette spørsmålet var for kompliserte for deltaker og dette kan muligens være en årsak til de varierende resultatene ved testene for de emergente bidireksjonale intraverbalene. For sett 2, som deltaker hadde 100% mestring ved denne testen er det mulig at det var egenskaper ved spørsmålene som gjorde det enklere for deltaker å gi adekvate svar for dette settet, for eksempel var det i dette settet de to spørsmålene med færrest ord (fem) av alle spørsmålene brukt ved testing for emergente bidireksjonale intraverbaler.

Det må også diskuteres at det er mulig at lyttertreningen som ble gjennomført i denne studien ikke var den mest effektive. Det kan ikke utelukkes at det for eksempel kunne være hensiktsmessig å gjennomføre flere trials med trening før testing for sett 1. I denne studien så det ut til å være tilstrekkelig med 36 trials lyttertrening for sett 2 og 3, men til tross for at det ble gjennomført dobbelt så mange trials for sett 1 førte ikke dette til mestring under test. Det er mulig at lyttertrening gjennomført på en annen måte ville gitt andre resultater, for eksempel kan det tenkes at det for deltaker hadde vært hensiktsmessig med flere trials for sett 1 og færre trials for sett 2 og 3. Det kan også tenkes at andre metoder for gjennomføring av treningen ville gitt andre resultater for eksempel i forhold til de emergente bidireksjonale intraverbalene. Det kan heller ikke, som beskrevet tidligere utelukkes at tacttreningen for sett 1 var nødvendig for at lyttertreningen skulle ha effekt for sett 2 og 3. Det kan derfor videre være interessant å undersøke videre hvordan lytter- og tacttrening bør legges opp for å få oppnå mest mulig effektive tiltak, da mest mulig effektive tiltak er en grunnleggende del av anvendt atferdsanalytisk praksis (Baer et al., 1968; Van Houten et al., 1988). Det hadde også vært nyttig med en systematisk vurdering og sammenligning av ulike metoder for etablering av intraverbaler, for å kunne konkludere når det gjelder til hvilke metoder som bør foretrekkes i anvendte settinger.

For deltaker i denne studien er det vanskelig å vurdere sosial validitet da det ikke ble foretatt målinger av i hvilken grad de intraverbale responsene ble generalisert til naturlige settinger. Deltakers støttepedagog og øvrig barnehagepersonell fikk ved starten av opprettholdelsesfasen tilgang til spørsmålene som var anvendt under trening og testing. Personalet ble oppmuntret til å benytte disse spørsmålene i situasjoner der det naturlig var passende å anvende dem. Det ble likevel ikke foretatt noen målinger av i hvilken grad spørsmålene ble brukt og adekvate svar forekom.

Deltaker var svært entusiastisk og ba om få gå på treningsrommet så snart han så eksperimentator ankomme barnehagen. Han kunne også be gjentatte ganger om å få bli stilt spørsmål under forberedelser eller avslutting av trening- eller testøkter. Det så med andre ord ut som at det for deltaker var gøy å delta i studien, noe som bør anses å være av verdi. Hvilken nytte deltaker hadde av treningen på de ni når-spørsmålene kan derimot diskuteres. Ved gjennomføring av lignende studier hadde det vært nyttig å også gjennomføre målinger av i hvilken grad adekvate svar på når-spørsmål fortsetter å forekomme utenfor treningsrommet. Kartlegging av intraverbale ferdigheter ved ABLLS-R (Partington, 2012) viste økning i mestring av intraverbale ferdigheter fra 21.35% før studien startet til 44.51% ved studiens slutt. I hvilken grad dette skyldes gjennomføringen av denne studien er det ikke mulig trekke konklusjoner om, men dette kan indikere at deltakelse i studien i alle fall ikke hadde negative konsekvenser for deltaker.

For fagfeltet kan det derimot vurderes svært nyttig at det blir gjennomført studier som denne. Som beskrevet tidligere er det knyttet en rekke utfordringer til manglende språklige ferdigheter (Beitchman et al., 2014; Bishop, 2014; Botting & Conti-Ramsden, 2000; Flensburg-Madsen & Mortensen, 2017; Silva et al., 1987) og det er av stor anvendt verdi å undersøke hvilke tiltak det foreligger evidens for at er effektive, da dette er en del av fundamentet for anvendt atferdsanalyse (Baer et al., 1968; Van Houten et al., 1988). Som beskrevet tidligere er det i forhold til etablering av intraverbale usikkerhet knyttet til metodevalg (Daar et al., 2015) og studier som denne kan trolig forsvares selv om verdien for den enkelte deltaker kan være noe begrenset.

For videre forskning vil det være interessant å undersøke metoden for flere deltakere. Tidligere gjennomførte studier har hatt noe varierende prosedyrer, for eksempel om spørsmålene har vært de samme ved trening og testing, og om deltaker kunne ta aktuelle stimuli på forhånd, noe som har vært foreslått som en mulig forklaring på de varierende

resultatene (Smith et al., 2016). Det hadde vært interessant å undersøke i hvilken grad disse faktorene påvirker resultatet. Videre vil det være av stor anvendt verdi å finne frem til de mest effektive tiltakene for å sikre best mulig opplæring for personer som er i behov av dette.

### **Konklusjon**

Resultatene fra denne studien viste at det etter lyttertrening oppstod emergente intraverbale responser for to av de tre settene det ble trent på. Lyttertrening så ut til å være en effektiv måte å etablere intraverbale responser. For det første settet det ble trent på oppstod det ikke emergente intraverbaler etter lyttertrening aleine, men etter tacttrening hadde deltakeren 100% mestring for dette settet. Det kan konkluderes med at lyttertrening var mulig effektivt for to tredjedeler av spørsmålene det ble trent på. Resultatene viste videre at lyttertrening førte til emergente bidireksjonale intraverbaler for halvparten av spørsmålene det ble testet for, sammenlignet med en pretestmåling. Videre forskning er nødvendig for å undersøke som lyttertrening kan defineres som en effektiv metode for etablering av intraverbale responser.

### Referanser

- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97.  
doi:10.1901/jaba.1968.1-91
- Bader. [Bilde] (2017). Hentet fra [http://www.unicornostyle.com/web/wp-content/uploads/2017/09/nuotobambini\\_banner.jpg](http://www.unicornostyle.com/web/wp-content/uploads/2017/09/nuotobambini_banner.jpg)
- Barbera, M. L. (2007). *The verbal behavior approach: how to teach children with autism and related disorders*. London: Jessica Kingsley Publishers.
- Beitchman, J. H., Brownlie, E. B., & Bao, L. (2014). Age 31 mental health outcomes of childhood language and speech disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(10), 1102-1110.e1108. doi:10.1016/j.jaac.2014.07.006
- Bil. [Bilde]. Hentet fra <https://carsjpg.com/wewant-any-car/>
- Bishop, D. V. M. (2014). *Uncommon understanding: development and disorders of language comprehension in children* (Classic ed. ed.). London: Psychology Press.
- Botting, N., & Conti-Ramsden, G. (2000). Social and behavioural difficulties in children with language impairment. *Child Language Teaching and Therapy, 16*(2), 105-120.  
doi:10.1177/02656590000160020
- Brann. [Bilde]. Hentet fra [https://d2g8igdw686xgo.cloudfront.net/23325608\\_1503962079.7317.jpg](https://d2g8igdw686xgo.cloudfront.net/23325608_1503962079.7317.jpg)
- Daar, J. H., Negrelli, S., & Dixon, M. R. (2015). Derived emergence of WH question–answers in children with autism. *Research in Autism Spectrum Disorders, 19*, 59-71.  
doi:https://doi.org/10.1016/j.rasd.2015.06.004
- Eikeseth, S., & Smith, D. P. (2013). An analysis of verbal stimulus control in intraverbal behavior: implications for practice and applied research. *The Analysis of Verbal Behavior, 29*(1), 125. doi:10.1007/BF03393130

- Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 439-450.  
doi:10.1080/15374410902851739
- Flensburg-Madsen, T., & Mortensen, E. L. (2017). Associations of early developmental milestones with adult intelligence. *Child development*. doi:10.1111/cdev.12760
- Goodwin, A., Fein, D., & Naigles, L. (2015). The role of maternal input in the development of wh-question comprehension in autism and typical development. *42*(1), 32-63.  
doi:10.1017/S0305000913000524
- Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*(1), 185-241.  
doi:10.1901/jeab.1996.65-185
- Jahr, E. (2001). Teaching children with autism to answer novel wh-questions by utilizing a multiple exemplar strategy. *Research in Developmental Disabilities, 22*(5), 407-423.  
doi:10.1016/S0891-4222(01)00081-6
- Kodak, T., & Paden, A. R. (2015). A comparison of intraverbal and listener training for children with autism spectrum disorder. *Analysis of Verbal Behavior, 31*(1), 137-144.  
doi:10.1007/s40616-015-0033-3
- Laraway, S., Snyckerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*(3), 407-414. doi:10.1901/jaba.2003.36-407
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2001). *Autism diagnostic observation schedule (ADOS): Manual*: WPS.
- Luce, S. C., Maurice, C., & Green, G. (1996). *Behavioral intervention for young children with autism: a manual for parents and professionals*. Austin, Tex: Pro-Ed.

- Løvaas, O. I. (2003). *Opplæring av mennesker med forsinket utvikling: grunnleggende prinsipper*. Oslo: Gyldendal akademisk.
- McEachin, J., & Leaf, R. (1999). *A work in progress: behavior management strategies and a curriculum for intensive behavioral treatment of autism*. New York: DRL Books.
- Miguel, C. F., Petursdottir, A. I., & Carr, J. E. (2005). The effects of multiple-tact and receptive-discrimination training on the acquisition of intraverbal behavior. *The Analysis of Verbal Behavior*, 21(1), 27-41.
- Morgen. [Bilde] (2017). Hentet fra <https://www.fiveclipart.com/wp-content/uploads/2017/03/horizon-clip-art-horizon-clip-art-horizon-clip-art-mountain-horizon-clipart-free-horizon.jpg>
- Natt. [Bilde] (2010). Hentet fra [http://1.bp.blogspot.com/\\_PSSZ5svJ6ro/TPMJAUdiBFI/AAAAAAAAAiU/mfIovGFz1Os/s1600/Natt%2Bover%2BBeverdalen...JPEG](http://1.bp.blogspot.com/_PSSZ5svJ6ro/TPMJAUdiBFI/AAAAAAAAAiU/mfIovGFz1Os/s1600/Natt%2Bover%2BBeverdalen...JPEG)
- Partington, J. W. (2012). *Kartlegging av grunnleggende språk- og læreferdigheter - revidert: instruksjoner for skåring og utvikling av opplæringsplan: ABLLS-R Guide* (K. Larsen, Trans. Rev. utg. ed.). Pleasant Hill, CA: Behavior Analysts, Inc.
- Petursdottir, A. I., Carr, J. E., Lechago, S. A., & Almason, S. M. (2008). An evaluation of intraverbal training and listener training for teaching categorization skills. *Journal of Applied Behavior Analysis*, 41(1), 53-68.
- Petursdottir, A. I., & Haflidadottir, R. D. (2009). A comparison of four strategies for teaching a small foreign-language vocabulary. *Journal of Applied Behavior Analysis*, 42(3), 685-690. doi:10.1901/jaba.2009.42-685
- Petursdottir, A. I., Lepper, T., & Peterson, S. (2014). Effects of collateral response requirements and exemplar training on listener training outcomes in children. *The Psychological Record*, 64(4), 703-717. doi:10.1007/s40732-014-0051-x



Petursdottir, A. I., Olafsdottir, A. R., & Aradottir, B. (2008). The effects of tact and listener training on the emergence of bidirectional intraverbal relations. *Journal of Applied Behavior Analysis, 41*(3), 411-415.

Pusser tenner. [Bilde] (2013). Hentet fra <http://misoldent.cl/wp-content/uploads/2016/12/odontopediatria-centro-odontologico-de-canarias-730x400.jpg>

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the ucla young autism project model. *Journal of autism and developmental disorders, 39*(1), 23-41.  
doi:10.1007/s10803-008-0596-0

Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism diagnostic interview-revised. *Los Angeles, CA: Western Psychological Services, 29, 30.*

Silva, P. A., Williams, S., & McGee, R. (1987). A longitudinal study of children with developmental language delay at age three: Later intelligence, reading and behaviour problems. *Developmental Medicine & Child Neurology, 29*(5), 630-640.  
doi:10.1111/j.1469-8749.1987.tb08505.x

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, N.J: Prentice-Hall.

Smith, D. P., Eikeseth, S., Fletcher, S. E., Montebelli, L., Smith, H. R., & Taylor, J. C. (2016). Emergent intraverbal forms may occur as a result of listener training for children with autism. *Analysis of Verbal Behavior, 32*(1), 27-37. doi:10.1007/s40616-016-0057-3

Sommer. [Bilde] (2016). Hentet fra <http://freedesignfile.com/upload/2016/03/Hot-summer-holiday-background-with-funny-sun-vector-05.jpg>

Sundberg, M. L. (2014). Verbal behavior. In H. H. Cooper (Ed.), *Applied behavior analysis* (2 ed., pp. 536-557). Upper Saddle River, NJ: Pearson.

- Sundberg, M. L., & Michael, J. (2001). The benefits of skinner's analysis of verbal behavior for children with autism. *Behavior Modification*, 25(5), 698-724.  
doi:10.1177/0145445501255003
- Sundberg, M. L., & Partington, J. W. (1998). Teaching language to children with autism and other developmental disabilities. *Pleasant Hill, CA: Behavior Analysts*.
- Sundberg, M. L., & Sundberg, C. A. (2011). Intraverbal Behavior and Verbal Conditional Discriminations in Typically Developing Children and Children with Autism. *Analysis of Verbal Behavior*, 27, 23-43.
- Tegner. [Bilde] (1998). Hentet fra  
<https://static1.squarespace.com/static/592dadb6d482e95f540dcd20/5930a0436a49631d9a0ff148/594dbe47099c01a39ed59dcd/1511749031537/Child+Drawing.jpg?format=2500w>
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet intelligence scale*:  
Riverside Publishing Company.
- Tyv. [Bilde] (2009). Hentet fra [https://accountingweekly.com/wp-content/uploads/2015/06/35646432\\_thumbnail-640x480.jpg](https://accountingweekly.com/wp-content/uploads/2015/06/35646432_thumbnail-640x480.jpg)
- Urbaniak, G. C., & Plous, S. (2013). Research randomizer (Versjon 4.0) [Computer software].  
Hentet 9. februar, 2018, fra <http://www.randomizer.org/>
- Vallinger-Brown, M., & Rosales, R. (2014). An investigation of stimulus pairing and listener training to establish emergent intraverbals in children with autism. *Analysis of Verbal Behavior*, 30(2), 148-159. doi:10.1007/s40616-014-0014-y
- Van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *Journal of Applied Behavior Analysis*, 21(4), 381-384. doi:10.1901/jaba.1988.21-381

- Wolery, M. (2011). Intervention research: The importance of fidelity measurement. *Topics in Early Childhood Special Education, 31*(3), 155-157. doi:10.1177/0271121411408621
- World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines* (Vol. 1): World Health Organization.

Tabell 1

*Språklige ferdigheter*

Reseptivt språk	Vokal imitasjon	Be om	Benevning	Intraverbaler	Spontan vokalisering
96,15 %	100 %	54,54 %	61,26 %	44,51 %	100 %

*Note.* Resultater fra kartlegging ved ABLLS-r. Resultatene er oppgitt i antall prosent av totale muligheter for skåring i den aktuelle ferdighetskategorien. ABLLS-r ble gjennomført etter at tiltaket var igangsatt.

Tabell 2

*Når-spørsmålene, målresponser og bildestimuli som ble anvendt.*

Når-spørsmål	Målrespons	Bildestimuli
<b>Sett 1</b>		
Når kommer politiet?	Tyv	
Når bruker du fargeblyanter?	Tegner	
Når bruker du tannkrem?	Pusser tenner	
<b>Sett 2</b>		
Når sover du?	Natt	
Når står du opp?	Morgen	
Når bruker du badebukse?	Bader	
<b>Sett 3</b>		
Når kan du gå med shorts?	Sommer	
Når ringer du brannvesenet?	Brann	
Når bruker du bilbelte?	Bilen	

*Note.* Tabellen gir en oversikt over når-spørsmålene som ble brukt ved lytter- og tacttrening og ved testing for emergente intraverbale responser, responsene som var definert som korrekte og bildene som ble brukt ved lytter- og tacttrening.

Tabell 3

*Spørsmål brukt ved testing for emergente bidireksjonale intraverbaler*

Spørsmål	Korrekt respons
Sett 1	
Hvem kommer når det er en tyv som stjeler?	Politiet
Hva trenger du når du tegner?	Fargeblyanter
Hva trenger du når du pusser tenner?	Tannkrem
Sett 2	
Hva gjør du om natten?	Sover
Hva gjør du på morgenen?	Står opp
Hva har du på deg når du bader?	Badebukse
Sett 3	
Hvilke klær bruker du om sommeren?	Shorts
Hva har du på deg i bilen?	Bilbelte
Hva gjør man når det brenner?	Ringer brannvesenet

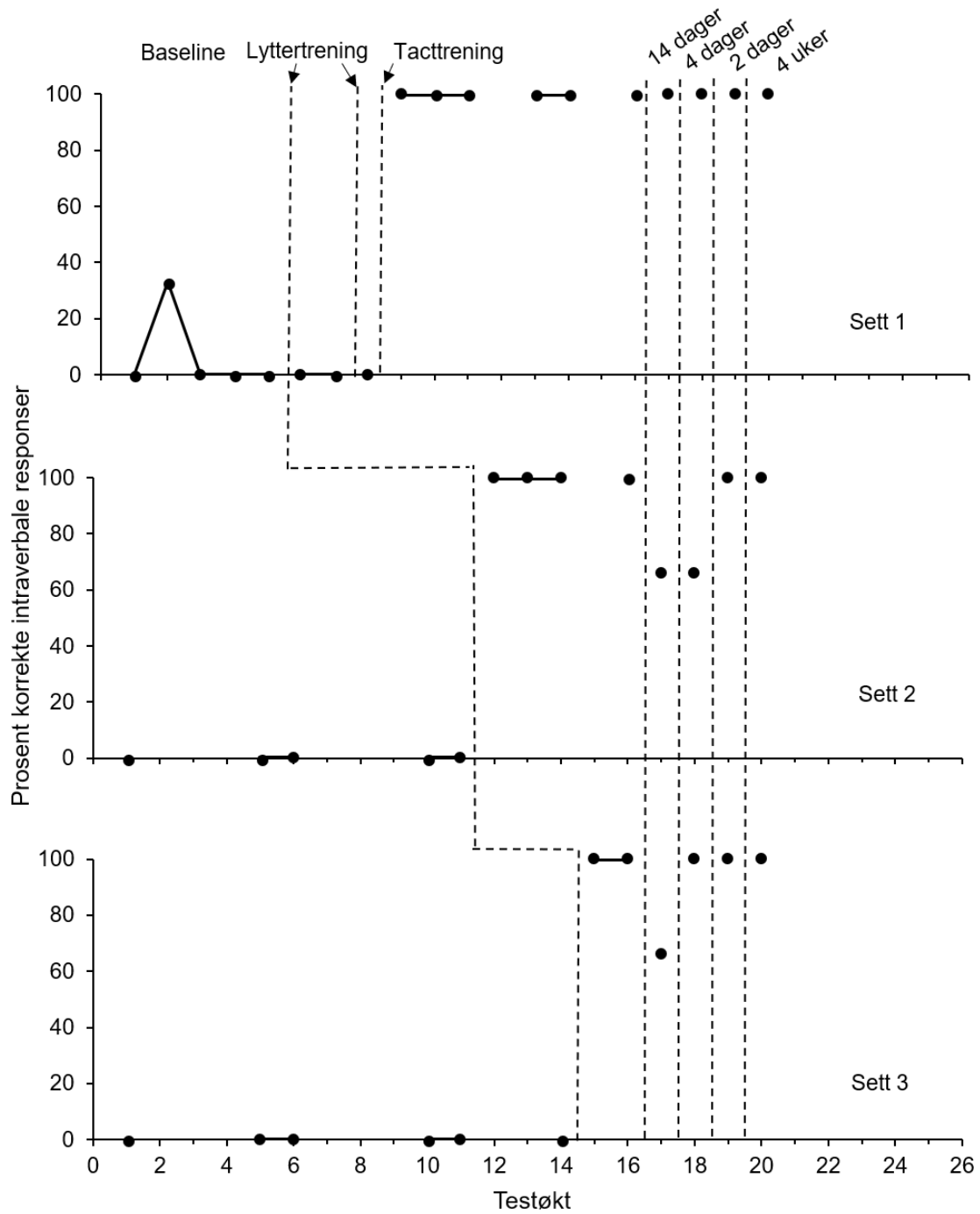
*Note.* Tabellen gir en oversikt over spørsmål brukt ved pre- og posttest for emergente bidireksjonale intraverbaler og responser som ble definert som korrekte.

Tabell 4

*Resultater ved test for emergente bidireksjonale intraverbaler*

Sett	Pretest	Posttest
Sett 1	0	33 %
Sett 2	0	100%
Sett 3	33%	33 %

*Note.* Antall korrekte responser ved pre-og posttest for emergente bidireksjonale intraverbaler oppgitt i prosent.



Figur 1. Emergente intraverbale responser under baseline, etter lytter- og tacttrenting og i opprettholdelsesfasen for sett 1, 2 og 3.



Appendiks A

Prosedyreintegritet lyttertrening

Dato:
Treningsøkt:

Målatferd		Riktig =R eller ✓ Feil= F eller -
Legger 3 bilder på bordet (en side i permen)		
Sier spørsmålet: «når..?»		
	<b>Rett respons:</b> generell ros	
	<b>Feil respons:</b> Peker på rett bilde. Gjentar «når..?» Ros ved rett.	
Legger ny side (med tre bilder) på bordet		
«når..?»		
	<b>Rett respons:</b> generell ros	
	<b>Feil respons:</b> Peker på rett bilde. Gjentar «når..?» Ros ved rett.	
Ny side		
«når..?»		
	<b>Rett respons:</b> generell ros	
	<b>Feil respons:</b> Peker på rett bilde. Gjentar «når..?» Ros ved rett.	
Ny side		
«når..?»		
	<b>Rett respons:</b> generell ros	
	<b>Feil respons:</b> Peker på rett bilde. Gjentar «når..?» Ros ved rett.	

## Appendiks B

### Prosedyreintegritet Tacttrening

Dato:	Øktnr:	Ja	Nei
Stiller spørsmålet som står i registreringsarket og viser bildet			
Ved feil svar: sier navnet på bildet			
Ved korrekt svar- ros			

Dato:	Øktnr:	Ja	Nei
Stiller spørsmålet som står i registreringsarket og viser bildet			
Ved feil svar: sier navnet på bildet			
Ved korrekt svar- ros			

Dato:	Øktnr:	Ja	Nei
Stiller spørsmålet som står i registreringsarket og viser bildet			
Ved feil svar: sier navnet på bildet			
Ved korrekt svar- ros			

Dato:	Øktnr:	Ja	Nei
Stiller spørsmålet som står i registreringsarket og viser bildet			
Ved feil svar: sier navnet på bildet			
Ved korrekt svar- ros			

Dato:	Øktnr:	Ja	Nei
Stiller spørsmålet som står i registreringsarket og viser bildet			
Ved feil svar: sier navnet på bildet			
Ved korrekt svar- ros			

Dato:	Øktnr:	Ja	Nei
Stiller spørsmålet som står i registreringsarket og viser bildet			
Ved feil svar: sier navnet på bildet			
Ved korrekt svar- ros			

### Appendiks C

#### Prosedyreintegritet

#### Tester: emergente intraverbaler og bidireksjonalitet

Dato:	Øktnr:	Ja	Nei
Ingenting liggende på bordet			
Stiller et og et spørsmål.			
Ingen tilbakemelding på spørsmålene i testen.			
Ca hvert tredje spørsmål er et annet spørsmål- ros for svar på disse			

Dato:	Øktnr:	Ja	Nei
Ingenting liggende på bordet			
Stiller et og et spørsmål.			
Ingen tilbakemelding på spørsmålene i testen.			
Ca hvert tredje spørsmål er et annet spørsmål- ros for svar på disse			

Dato:	Øktnr:	Ja	Nei
Ingenting liggende på bordet			
Stiller et og et spørsmål.			
Ingen tilbakemelding på spørsmålene i testen.			
Ca hvert tredje spørsmål er et annet spørsmål- ros for svar på disse			

Dato:	Øktnr:	Ja	Nei
Ingenting liggende på bordet			
Stiller et og et spørsmål.			
Ingen tilbakemelding på spørsmålene i testen.			
Ca hvert tredje spørsmål er et annet spørsmål- ros for svar på disse			

Dato:	Øktnr:	Ja	Nei
Ingenting liggende på bordet			
Stiller et og et spørsmål.			
Ingen tilbakemelding på spørsmålene i testen.			
Ca hvert tredje spørsmål er et annet spørsmål- ros for svar på disse			

Dato:	Øktnr:	Ja	Nei
Ingenting liggende på bordet			
Stiller et og et spørsmål.			
Ingen tilbakemelding på spørsmålene i testen.			
Ca hvert tredje spørsmål er et annet spørsmål- ros for svar på disse			