

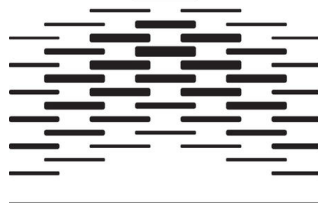
**MASTEROPPGAVE**  
**Læring i komplekse systemer med fordypning i**  
**atferdsanalyse**  
**2012**

Artikkel 1: Empirisk validert behandling

Artikkel 2: Er rate en kontrollerende variabel for bevaring av  
ferdigheter?

Frode Hammerø

Fakultet for helsefag  
Avdeling for atferdsvitenskap



**HØGSKOLEN I OSLO  
OG AKERSHUS**

Empirisk validert behandling

Frode Hammerø

Høgskolen i Oslo og Akershus

Avdeling for atferdsvitenskap

Authors note

Vil med dette takke min veileder Torunn Lian for tydelig og konkret veiledning gjennom hele prosessen. Ønsker også å rette en takk til min tidligere arbeidsgiver, Helse Stavanger, og min nåværende arbeidsgiver, Helse Fonna, som har dekket utgifter og lagt til rette for arbeid med studiene. Sist, men ikke minst, vil jeg takke min samboer for hennes tålmodighet og forståelse. Mange kvelder og helger har gått med til skolearbeid de siste årene, og uten støtte på hjemmebane, ville dette ikke vært mulig å gjennomføre.

## Abstract

På 90-tallet satt Task Force rapportene for alvor empirisk validert behandling på dagsorden. Divisjon 12 i den amerikanske psykologforeningen ble etablert med formål om å promotere og spre effektiv behandling. Tross kritikk av rapportene, er det i dagens samfunn økende grad av forventninger om effektiv behandling. Kunnskap om hvilke behandlingsformer som er effektive er viktig for enkeltindivider og samfunnet i sin helhet. En del klinisk praksis synes imidlertid å være uberørt av den stadig voksende forskningsbasen av empirisk validert behandling. Manglende implementering av forskningsstøttede metoder vil kunne medføre at klienter mottar behandling som er mer omfattende og mindre effektiv enn annen behandling. Dette synes etisk betenkelig. Artikkelen belyser gapet mellom klinisk praksis og forskning. Det rettes også søkelys mot andre utfordringer som manglende konsensus på kriterier for validering av behandling. Avslutningsvis gjøres det rede for empirisk validerte metoder innen atferdsanalysen, med et særlig fokus på forfatterens interessefelt; tidlig og intensiv opplæring av barn med autisme.

*Nøkkelord:* empirisk validert behandling, atferdsanalytisk behandling, autisme

### **Empirisk Validert Behandling**

I takt med at velferdsstatens utgifter øker, stilles det mer krav til dokumentasjon og effektiv behandling i alle deler av samfunnet. I økende grad har også pasienter og brukere forventninger om effektiv behandling. Parallelt med denne trenden utvikles det stadig nye behandlingsmetoder. Noen av disse kan vise til god effekt, andre ikke.

Krav til effektiv psykologisk behandling økte i styrke allerede tidlig på 90-tallet. Initiativet kom i utgangspunktet fra forsikringsselskaper, men også fagpersoner innenfor psykologien hadde begynt å studere effekter av terapi (e.g., Eysenck, 1952; Gendlin & Rychlak, 1970; Rachman & Wilson, 1980; Smith & Glass, 1977). I 1993 opprettet daværende president i den amerikanske psykologforeningen (APA), David Barlow, en egen avdeling for promotering og spredning av empirisk validert behandling. Avdelingen fikk navnet 'Division 12 APA Task Force', og ble satt sammen av kliniske psykologer fra ulike teoretiske retninger som psykodynamisk teori, interpersonell psykoterapi og kognitiv atferdsterapi. Mandatet var å utarbeide en liste over empirisk validerte behandlinger. Jeg vil senere i artikkelen komme tilbake til divisjon 12, og listene over validert behandling som ble publisert i kjølvannet av opprettelsen av avdelingen.

Kunnskap om hvilke behandlingsmetoder som er effektive er ikke bare viktig for klienter og deres pårørende, men også for fagpersoner, beslutningstakere og samfunnet generelt. Baer, Wolf, og Risley (1968) har beskrevet syv dimensjoner som kjennetegner god praksis innenfor anvendt atferdsanalyse, deriblant anvendt og effektiv. Begrepet anvendt henspiller på at den atferd det intervenseres i forhold til må være viktig for klienten og/eller samfunnet. Med dimensjonen effektiv retter Baer og kollegaer oppmerksomheten mot at den intervensjon som iverksettes må være effektivt nok til at det bidrar til atferdsforandringer som er sosialt betydningsfulle.

Min interesse for effektiv behandling har bakgrunn i mange års arbeid med opplæring av barn med autisme. For denne gruppen er det utviklet effektive og godt dokumenterte opplæringsmetoder som bidrar til at barnas funksjonsnivå kan forbedres betraktelig. Egen erfaring er imidlertid at mange av disse barna ikke får den opplæringen de kunne profitert på. Noen ganger skyldes dette manglende kompetanse i hjelpeapparatet, andre ganger bevilges det ikke tilstrekkelig med ressurser. Ikke sjelden velges også effektive metoder vekk til fordel for fagpersoners personlige faglige preferanser. Det siste henger ofte sammen med manglende metodiske føringer fra arbeidsgiver.

Denne artikkelen vil omhandle empirisk validert behandling. Jeg vil starte med en begrepsavklaring hvor jeg redegjør for de ulike begrep som benyttes innenfor dette feltet og avklarer hvilke jeg vil bruke ut over i artikkelen. Deretter vil jeg gi en kort historisk gjennomgang av det arbeidet som ligger til grunn for de kriteriene som i dag i stor grad benyttes for å avklare hvorvidt en behandling er validert. I forlengelsen av dette vil jeg gjengi noe av kritikken og innvendingene som oppstod i arbeidet med utviklingen av disse kriteriene og de påfølgende publiseringene av lister over empirisk validert behandling. I den avsluttende delen av artikkelen presenterer jeg empirisk validerte behandlingsmetoder innenfor atferdsanalysen, og trekker i dette inn mitt erfaringsfelt, behandling av barn med autisme.

### **Empirisk Validert Behandling – en Begrepsavklaring**

Evidensbasert behandling, empiriske støttet behandling og empirisk validert behandling er ulike betegnelser på behandling eller opplæring som er vist effektiv gjennom forskning. I denne artikkelen vil empirisk validert behandling benyttes gjennomgående. Mens evidensbasert behandling, empirisk støttet behandling og empirisk validert behandling referer til det samme, er det viktigere å trekke et skille mellom empirisk validerte behandling og evidensbasert praksis. Evidensbasert praksis er et mer omfattende begrep enn empirisk validert behandling. En slik praksis skal integrere best tilgjengelig forskningsevidens med

klinisk ekspertise, og det skal også tas hensyn til karakteristikkene med pasienten, kultur og preferanser (APA, 2006).

Evidensbasert praksis innebærer altså at kliniske vurderinger i større grad vektlegges ved valg av behandling. En viktig forskjell mellom evidensbasert praksis og empirisk validert behandling er at førstnevnte åpner opp for flere kilder til evidens enn randomiserte kontrollerte studier (RCT). RCT regnes som gullstandarden innenfor behandlingsstudier. I en slik studie sammenliknes gruppen som blir gitt behandlingen man ønsker å evaluere med en kontrollgruppe. Kontrollgruppen kan være en gruppe som mottar en annen type behandling, ingen behandling (for eksempel venteliste) eller placebo. Gruppene må bestå av personer med mest mulig lik karakteristika. Fordelingen av forsøkspersoner til de to gruppene skjer tilfeldig ( gjerne ved loddtrekning), for å unngå at egenskaper ved forsøkspersoner påvirker fordelingen. Evidensbasert praksis baserer seg også på RCT, men vektlegger også annen evidens som klinisk observasjon, kvalitativ forskning og systematiske casestudier (APA, 2006). Kazdin (2008) hevder det finnes lite vitenskapelig evidens for evidensbasert praksis i kliniske kontekster.

Empirisk validert behandling referer på den andre siden til spesifikke behandlinger som er vist effektive gjennom kontrollerte forsøk (Kazdin, 2008). Kazdin (2011) påpeker at det ikke finnes en enkelt definisjon eller et sett av kriterier som kan avklare hvorvidt en behandling er validert. Likevel er det et vanlig kriterium at det må foreligge minimum to gruppestudier med randomiserte kontrollerte forsøk eller et høyt antall N=1 studier for at en spesifikk behandling skal kunne få merkelappen empirisk validert (Chambless et al., 1998; Kazdin, 2011). N=1 design er vitenskapelige design hvor det foretas flere målinger innen samme deltaker. Før iverksetting av tiltak gjennomføres en baselineperiode, hvor det registreres atferdsnivå før intervensjon. Variablene manipuleres systematisk, slik at man viser

at endringene skyldes tiltaket og ikke andre variabler. I slike eksperimenter er deltakeren sin egen kontroll gjennom at deltakers atferd endres i takt med introduksjon av tiltaket.

Begrepsavklaringen viser at det benyttes ulike begreper til å beskrive samme fenomen. Som gjort rede for bør det imidlertid trekkes et tydelig skille mellom empirisk validert behandling og evidensbasert praksis. Evidensbasert praksis åpner opp for flere kilder til evidens, og vil i mange tilfeller ikke være validert i henhold til Chambless og kollegaers (1998) kriterier for empirisk validert behandling. Videre vil jeg kort beskrive det som kan betegnes som starten på validert behandling innenfor psykologien; opprettelse av divisjon 12 i den amerikanske psykologforeningen og påfølgende publikasjoner.

### **Task Force Rapport**

I forlengelse av arbeidet med å promotere og spre empirisk validert behandling, publiserte i 1995 divisjon 12 den første av flere rapporter om validert behandling. Rapporten ble kalt 'Task Force', og inneholdt i tråd med mandatet en liste over empirisk validerte behandlinger. I denne første rapporten (Task Force, 1995) ble det vist til 25 eksempler på validerte behandlinger. En stor andel av disse var basert på atferdsanalyse eller kognitiv atferdsterapi. I årene som fulgte ble det utarbeidet to nye oppdateringer av listen (Chambless et al., 1996; Chambless et al., 1998), og med den siste rapporten hadde antall behandlinger økt til 71. Avdelingen har også i senere tid utarbeidet en nettside som stadig oppdaterer listen med metoder som oppfyller kriteriene for empirisk validert behandling. I årene etter 'Task Force' har det blitt publisert en rekke oversikter over validerte behandlingsformer, noe som har gjort det enklere for fagpersoner og klienter å få oversikt over hva som finnes av effektiv behandling.

Chambless og kollegaer (1998) deler inn empirisk validert behandling i kategoriene: veletablerte behandlinger og trolig effektive behandlinger. For å møte standarden for veletablerte behandlinger, må det foreligge minimum to gode gruppestudier med randomiserte

kontrollerte forsøk, eller ni eller flere N=1 studier med gode eksperimentelle design. RCT må være statistisk signifikant bedre enn medikamenter, psykologisk placebo eller en annen behandling. Alternativt må RCT være på nivå med en annen empirisk validert behandling. N=1 design må benytte gode eksperimentelle design og vise bedre effekt enn medikamenter, psykologisk placebo eller en annen behandling. Felles for både RCT og N=1 design er at det må være benyttet behandlingsmanualer, karakteristikker hos klientutvalget må være spesifisert og minst to ulike forskere eller forskningsteam må ha vist effekt av behandlingen.

For å imøtekomme kriteriet for trolig effektiv behandling, må det foreligge to eksperimenter som viser at behandlingen er mer effektiv sammenliknet med en kontrollgruppe på venteliste. Til forskjell fra kravet for veletablerte behandlinger, er disse studiene enten ikke RCT, eller har de ikke blitt gjennomført av to uavhengige forskergrupper. I tilfeller hvor en eller flere RCT møter kriteriene for veletablert behandling, men der effekten ikke har blitt demonstrert av to ulike forskere eller forskningsteam, vil behandlingen bli definert som trolig effektiv. Dersom det er benyttet gode eksperimentelle design og behandlingen er sammenliknet med enten medikamenter, psykologisk placebo eller med annen behandling, vil også et lavt antall N=1 studier kunne oppfylle standarden for denne gruppen.

Som vist satt 'Task Force' rapporten empirisk validert behandling på dagsorden innen psykologien. Det ble understreket viktigheten av kvalitetsmessig god forskning, og kriterier for validering av behandling ble utarbeidet. Rapporten ble imidlertid ikke like godt mottatt i alle fagmiljøer, og jeg vil videre presentere reaksjoner i kjølvannet av publisering av den første rapporten.

**Kritikk av 'Task Force' rapporten.** 'Task Force' rapporten var utvilsomt kontroversiell blant mange klinikere innenfor psykologien. Derfor ble en del av innvendingene imøtegått allerede ved publisering av den første rapporten. En av kommentarene på publiseringen var at anbefalingene på behandlinger skulle bli forkastet



ettersom flere behandlinger burde vært inkludert. Gruppen bak 'Task Force' (heretter 'gruppen') påpeker at listen ikke er utfyllende. Dette ble også utbedret gjennom publikasjoner med oppdateringer av listen (Chambless et al., 1996; Chambless et al., 1998).

Det ble videre hevdet at listen var et middel for å fremme kognitiv atferdsterapi som tilnærming på bekostning av andre tilnærminger. Task Force (1995) understreker imidlertid at gruppen er sammensatt av profesjonelle klinikere og forskere fra ulike teoretiske ståsted. Rapporten ble også kritisert for å ha benyttet behandlingsmanualer som et kriterium for empirisk validert behandling. Det ble hevdet at manualer bare kan utvikles for kognitiv atferdsterapi. 'Task Force' hevder denne kritikken er basert på en misoppfattelse av hva som defineres som en behandlingsmanual. Gruppen peker på at en behandlingsmanual er en tydelig beskrivelse av behandlingen, og er av den oppfatning at dette bør kunne forelegges innenfor alle tilnærminger.

Andre fremmet også påstander om at all psykoterapi er like effektiv. Av denne grunn vil det derfor ikke være behov for å fokusere på behandlinger som er empirisk validerte. Dette er fortsatt en oppfatning hos noen klinikere. Med en hypotese om at effektstørrelsen mellom ulike psykoterapier er null, gjennomførte Wampold og kollegaer (1997) en metaanalyse som bekreftet hypotesen. En vanlig tolkning av slike resultater er at uspesifiserte faktorer som for eksempel forholdet mellom terapeut og klient bidrar til eventuelle variasjoner i behandlingsutbyttet (e.g., Blatt, Zuroff, Hawley, & Auerbach, 2010; Chambless & Ollendick, 2001). Task Force (1995) påpeker at slike påstander er basert på metaanalyser som har blitt overgått av senere evidens.

Noen kritiserte rapporten for å anbefale behandling som virker. Argumentet var at det alltid må gjøres mer forskning. Svaret i 'Task Force' var at dette ville være å frarøve klienters muligheter til å få en behandling som kan hjelpe dem. Ifølge gruppen vil tilbakeholdelse av det man vet om effektiv behandling, kunne føre til at psykologi kommer dårligere ut en

psykiatri. Det ble også hevdet at det er farlig å publisere slike lister fordi de kan bli skrevet i sten og dermed gjøre det vanskelig for andre behandlere å få økonomisk støtte fra tredjeparter når behandlingen de tilbyr ikke er validert. Gruppen erkjenner at dette kan skje, men mener at et mer presserende spørsmål er hvorvidt psykoterapi i det hele tatt vil bli inkludert i helseplaner. Uten slike publiseringer vil man risikere at psykoterapi ekskluderes fra disse planene, og at mennesker dermed ikke får den hjelpen de trenger.

Jeg har nå kort beskrevet noen av reaksjonene på rapporten. Noen argumenterte for at all psykoterapi er like effektiv, andre hevdet listen var et middel for å fremme kognitiv atferdsterapi. Listen ble også kritisert for å anbefale behandling som virker, og for å ha utelatt enkelte empirisk validerte behandlinger. Temaet blir fortsatt debattert, og jeg vil nå ta for meg pågående diskusjoner mellom klinikere og empirikere.

### **Debatt mellom Empirikere og Klinikere**

Empirisk validert behandling debatteres stadig. Kazdin (2008) peker på at en av de vanligste innvendingene mot validert behandling er at betingelsene i kontrollerte studier er så forskjellige fra den kliniske praksis at funnene vanskelig kan generaliseres (e.g., Westen, Novotny, & Thompson-Brenner, 2004). Det hevdes at viktige komponenter som behandler, klient, behandling og kontekst ofte er vesentlig ulik klinisk praksis. Hunsley og Lee (2007) gjennomførte en sammenlikning av såkalte 'effectiveness' og 'efficacy' studier. 'Efficacy' studier er laboratoriestudier med strenge eksperimentelle betingelser, mens 'effectiveness' studier er studier gjennomført i og av det ordinære hjelpeapparat. Forfatterne konkluderer med at det er empirisk støtte for generalisering av resultater fra kontrollerte studier til klinisk praksis.

I en gjennomgang av studier som underbygger empirisk validert behandling, konkluderte Westen og kollegaer (2004) med at de fleste studier ekskluderer deltakere med komorbide lidelser, og at utvalget dermed ikke er klinisk representativt. Weisz, Weersing, og

Henggeler (2005) imøtegår denne kritikken og viser til metodologiske svakheter med studien. Blant annet hevdes det at studier som ikke er på linje med deres kritikk av empirisk validert behandling har blitt utelatt. Det vises også til en inkonsistent bruk av evidensstandard. Ifølge Weisz og kollegaer har forfatterne mindre strenge kriterier på funn som støtter deres syn enn på funn de kritiserer.

Flere har også kritisert bruken av diagnoser som en klassifisering av utvalget som inkluderes i studiene (Chambless & Ollendick, 2001). Noen hevder en slik klassifisering er umulig, fordi hver enkelt case er unik, mens andre har foreslått en annen klassifisering av utvalg, for eksempel basert på personlighetstrekk (Beutler, Kim, Davison, Karno, & Fisher, 1996) eller funksjonelle sammenhenger (Scotti, Morris, McNeil, & Hawkins, 1996). Det synes imidlertid nødvendig med en eller annen form for klassifisering. Uten dette vil klinikere vanskelig kunne vurdere hvorvidt funn fra forskning kan overføres til egne klienter (Chambless & Ollendick, 2001). I mangel på bedre alternativer benyttes fortsatt DSM som klassifiseringsgrunnlag.

Gullstandarden med bruk av randomiserte kontrollerte design har også måttet tåle kritikk. Noen argumenterer for bruk av kvalitativ forskning (e.g., Bohart, O'Hara, & Leitner, 1998). Innen humanvitenskapen skal behandler kunne fremvise kreativitet, og ofte sammen med klient, gjøre oppdagelser som kan føre til stadige modifiseringer og endringer. Ifølge Bohart og kolleger (1998) kan slike komplekse fenomen best studeres ved kvalitative metoder som dialog med klienten.

Andre peker på at generaliteten til empirisk validert behandling basert på RCT er svak, ettersom klienter ikke randomiseres til behandling i den kliniske praksis (Seligman, 1995). Det har også blitt hevdet at forskere i realiteten sammenlikner behandlingen med ingen behandling. Westen og Bradley (2005) påpeker at kontrollgruppen ofte er en ventelistegruppe, eller en gruppe som får behandling av ukvalifisert personell med lite ressurser. Ifølge

forfatterne opereres det også med spesielt designede behandlinger, uten noen teoretisk forankring, gitt av studenter som er klar over at behandlingen ikke skal virke.

Manualbasert behandling har også blitt gjenstand for kritikk. Blant annet hevdes det at terapeuter som benytter standardisert behandling ikke vil evne og fremvise fleksibilitet i behandling om nødvendig. Det har også blitt stilt spørsmål til hvorvidt standardisert behandling er mindre effektiv enn skreddersydde behandlingsprogrammer. Empirien tyder imidlertid ikke på at dette er tilfellet (Chambless & Ollendick, 2001).

Bohart og kolleger (1998) hevder at humanistisk behandling ikke kan bli presist spesifisert på en slik måte at den operasjonaliseres i en manual. Dette begrunnes med at nøkkelen i humanistisk behandling er kreativ improvisasjon. Terapeuten skal fremvise fleksibilitet og spontanitet i møte med klienten. I denne tradisjonen anses *terapeuter* for å være 'treatment of choice', ikke behandlingen eller spesifikke behandlingspakker, slik det vil være innenfor mer manualbasert behandling. Bakgrunnen for dette er at den terapeutiske alliansen ofte ses på som en hovedkomponent i behandlingen (Blatt et al., 2010). Ifølge Bohart og kollegaer (1998) er terapeutens rolle å møte klienten med empati i klientens verden, være deltaker i en konstruktiv dialog, samt fremvise sensitivitet til hva som hender i øyeblikket og til muligheter for produktiv interaksjon. Forfatterne påpeker at det verken er ønskelig eller forventet at ulike terapeuter skal fremvise samme atferd i møte med klientene.

Innenfor enkelte teoretiske tilnærminger hevdes det at empirisk validert behandling fokuserer på symptomer. Fra et psykodynamisk ståsted vil eksempelvis depresjon bli sett på som et symptom på mer sentrale underliggende problemer (Chambless & Ollendick, 2001). I en del klinisk praksis handler behandlingen gjerne om prosessen, som for eksempel mestring av livets utfordringer (Kazdin, 2008). Empirisk forskning benytter sjelden vage mål som for eksempel livskvalitet eller mestring som parameter på effekt, selv om nettopp slike målområder vil være fokus for en del samtaleterapi. Bohart og kollegaer (1998) påpeker at

målet innenfor en humanistisk modell ikke nødvendigvis er å lindre en lidelse, forbedre en tilstand eller løse et problem. Målet kan like gjerne være å etablere en trygg terapeutisk allianse, der klienten aktivt kan reflektere rundt sine handlingsmønstre. Gjennom interaksjon med terapeuten gis klienten mulighet til og utforske, evaluere og reflektere rundt egne valg. Denne interaktive prosessen kan, om nødvendig, føre til at klienten endrer sine livsvalg.

Som vist rettes det fortsatt kritikk mot empirisk validert behandling. Ulike betingelser i kontrollerte studier og klinisk praksis trekkes frem som en barriere for generalisering av forskningsresultater til klinisk praksis. Kontrollerte studier kritiseres for å utelukke komorbide lidelser. Manualbasert behandling hevdes å redusere behandlers fleksibilitet i behandling av klienter. Fra enkelte fagmiljøer blir også empirisk validert behandling beskyldt for å bedrive symptombehandling. Tilsynelatende finnes det imidlertid lite empiri som støtter disse påstandene. Empirikere retter imidlertid også hard skyts mot klinisk praksis. Videre vil jeg gjøre rede for denne kritikken.

**Kritikk av klinisk praksis.** Klinisk praksis kritiseres for fokuset på individuelle behandlingspakker. I klinisk terapi er det vanlig å skreddersy behandlingen til hver enkelt klient (Kazdin, 2008). Ofte velges en eklektisk tilnærming. En slik tilnærming vil bestå av komponenter fra mange ulike tilnærminger som settes sammen til en behandlingspakke tilpasset den enkelte klient. Kazdin (2008) viser til at det ikke finnes noen formell og replikerbar prosedyre på fra hvilke behandlinger komponentene skal hentes og i hvilken rekkefølge disse skal leveres. Således er det lite sannsynlig at en annen behandler ville skreddersydd den samme behandlingspakken til samme klient. Dette gir lav reliabilitet. På samme måte som kontrollerte studier kritiseres for manglende generalitet til klinisk praksis, kan skreddersøm kritiseres for det samme. Klienters ulikhet vil kunne gjøre det vanskelig å generalisere fra en klient til andre (Kazdin, 2008).

En annen bekymring hos empirikere er at klinisk progresjon måles også ofte ut fra terapeutens inntrykk (synsing), heller enn data (Kazdin, 2008). Dette medfører at man ikke har objektive evalueringer av kliniske tilfeller. Uten data kan man vanskelig si noe om effekt av behandling. Stuart & Lilienfeld (2007) påpeker også at evaluering underveis er kritisk for å levere etisk ansvarlig behandling. Systematiske målinger vil ikke bare gi en pekepinn på effekt av behandlingen, men også være et nyttig verktøy for å måle fremgang underveis i behandlingen. Dette vil gi behandler et grunnlag for å gjøre eventuelle kursendringer underveis. Uten slike målinger vil man risikere at behandling som ikke virker, eller er direkte skadende, ikke avsluttes som en konsekvens av beholders feiltolkninger (Stuart & Lilienfeld, 2007).

Empirikere påpeker at kun et fåtall av de prosedyrer som benyttes i klinisk praksis er empirisk validerte (Kazdin, 2011). Dette tross store fremskritt innen forskningen, og stadig nye publikasjoner med oversikter over behandlinger med forskningsstøtte. Det har med andre ord aldri vært enklere å tilegne seg oppdatert kunnskap om effektive behandlingsmetoder. Kazdin (2011) hevder også at de prosedyrene som faktisk benyttes innenfor klinisk praksis, sjelden evalueres for effekt.

Parallelt med en manglende implementering av empirisk validerte behandlingsmetoder, utvikles det stadig nye behandlingsmetoder uten noen form for vitenskapelig belegg (Kazdin, 2011). Ifølge Kazdin (2008) fantes det allerede i 2000 over 550 ulike terapier for barn og unge, og tallet er stadig økende. Flere har påpekt at det er viktig med empiri før implementering og at det er uheldig med prematur spredning av behandling som ikke er empirisk validert (Heinicke, Carr, LeBlanc, & Severtson, 2010).

I den siste tiden har også skadelige effekter av behandling fått økt oppmerksomhet. Lilienfeld (2007) har utarbeidet en liste over potensielt skadelige behandlinger. Et eksempel er programmet 'Scared Straight'. Gjennom eksponering for fengselslivets realiteter, skal

ungdom i faresonen skremmes fra en fremtidig kriminell karriere. RCT har vist signifikant økning i arrestasjoner hos gruppen som har tatt del i programmet sammenliknet med en kontrollgruppe. Såkalt 'rebirthing' er et annet program som er påvist skadelig. Programmet er basert på en teori om at aggressivitet hos barnet skyldes en fødselstraume, og at dette kan bedres ved å gjenskape traume. Terapeuten vil gjerne pakke barnet inn i et teppe, sitte på barnet og presse gjentakende på barnet i et forsøk på å simulere fødselen. Flere barn har blitt kvelt under slike behandlinger (Lilienfeld, 2007).

Oppsummert kritiseres klinisk praksis for fokuset på individuelle behandlingspakker, måling av fremgang ut fra 'synsing' og for å sjelden benytte empirisk validert metoder. Det uttrykkes også bekymring for utvikling av et stadig økende antall nye behandlingsmetoder uten forskningsstøtte. Gapet mellom klinikere og empirikere kan tilsynelatende virke stort. Det må imidlertid presiseres at skillet mellom klinikere og empirikere ikke alltid er like tydelig som beskrevet her. Fokuset vil videre dreies mot debatten om empirisk validert behandling innad i forskningsmiljøet.

**Debatt blant empirikere.** Som vist kritiseres empirisk validert behandling blant en del klinikere. Det pågår imidlertid også en diskusjon om validert behandling blant empirikere. Denne dreier seg ikke om hvorvidt man bør benytte empirisk validert behandling, men om hvilke kriterier som må ligge til grunn for å validere behandling. Flere har stilt til spørsmål til hvorvidt noen av kriteriene som benyttes for å definere empirisk validert behandling er tilstrekkelig stringente (e.g., Biglan & Ogden, 2008).

Kazdin (2011) adresserer problemet med manglende konsensus på kriterier for empirisk validert behandling. Et økende antall offentlige og private instanser bruker ressurser på å identifisere empirisk validert behandling og spre dem til relevante mottakere. Når ikke alle instansene opererer med de samme kriteriene, vil det være variasjon i hvilke behandlinger som betegnes som validerte. Samme behandling vil dermed kunne fremstå som empirisk

validert med et sett av kriterier, og uten dokumentasjon ved bruk av et annet sett med kriterier (Westen & Bradley, 2005). Denne inkonsistensen generer forvirring både blant pasienter og behandlere på søken etter empirisk validert behandling (David & Montgomery, 2011).

Biglan og Ogden (2008) viser eksempelvis til en avdeling i det amerikanske helsedirektoratet ('The Substance Abuse and Mental Health Services Administration') som benytter enighet blant eksperter som et kriterium for empirisk validert behandling. Dette betyr at en behandling uten noen form for eksperimentell evidens likevel vil kunne definert som empirisk validert. Et annet eksempel er APA-rapporten fra 2006 som åpner opp for kliniske hensyn i valg av behandling. Biglan og Ogden (2008) advarer mot at dette gir klinikere store muligheter til å velge vekk empirisk validert behandling med begrunnelse i deres kliniske ekspertise. Lohr (2011) påpeker at evidensbegrepet i Task Force (2006) åpner opp for så godt som alle mulige kliniske syn og meninger, noe som gjør behovet spesifisering og definerings av evidensbegrepet viktig.

David og Montgomery (2011) foreslår standardiserte kriterier for empirisk validert behandling. Et sett av felles kriterier vil sette standarden for hvilken forskning som gjennomføres, samt bidra til at feltet fremstår med større enighet om hvilke behandlinger som kan defineres som empirisk validerte. Dette vil kunne være til stor hjelp for både profesjonelle og klienter på søken etter forskningsbasert behandling.

Andre viser til at det ensidige fokuset på spesifikke empirisk validerte behandlingspakker er en svakhet ved nåværende evalueringssystemer (e.g., David & Montgomery, 2011; Lohr, 2011). David og Montgomery (2011) foreslår et nytt evalueringssystem som ikke bare evaluerer evidens for gitte behandlinger, men også for den teoretiske forankringen for disse behandlingene. Dette vil styre forskning i riktig retning og forhindre at kvasivitenskapelige behandlingspakker vil få innpass på lister over empirisk validert behandling.



Kazdin (2008) foreslår at det bør forskes mer på hvilke variabler innen empirisk validert behandling som gir effekt. En randomisert kontrollert studie som undersøker effekt av to behandlinger kan benyttes til å vise en kausal relasjon mellom en intervensjon og terapeutisk endring. Dette betyr imidlertid ikke nødvendigvis at man vet hvorfor man fikk etablert en endring. Mer fokus på hvilke komponenter av behandlingen som gir effekt, vil gi en bedre klinisk praksis på lang sikt. Dette er i tråd med Baer, Wolf, og Risleys (1968) kriterier for god atferdsanalytisk praksis. Forfatterne påpeker at god atferdsanalytisk praksis skal være analytisk. Med dette menes at man skal kunne demonstrere at tiltaket påvirker atferden man ønsker å endre. Selv om det er enklere å vise kontroll over atferden i laboratoriumsetting, skal dette likevel etterstrebnes i anvendt atferdsanalyse. Kazdin (2008) peker også på identifisering av moderatorer som en måte å bedre feltet ytterligere på. Moderatorer kan for eksempel være karakteristikker ved pasient, terapeut og kontekst, som gjør behandlingen mer effektiv for noen enn for andre.

Tross debatten mellom klinikere og forskere, og uenighet om kriterier, er antallet empirisk validerte behandlinger for en rekke ulike dysfunksjoner økende (Kazdin, 2011). Stadig flere oversikter over validert behandling publiseres og gjøres tilgjengelig for allmennheten. Dette gjør det enklere for både klienter/pårørende og profesjonelle å identifisere behandling som har støtte i forskningen.

Jeg har ovenfor satt lys på diskusjonen rundt kriterier for empirisk validert behandling. Det er også presentert forslag til hvordan forskning kan forbedres. Flere påpeker problemet med at det ikke er en unison konsensus på hvilke kriterier som skal benyttes. Dette kan genere forvirring blant klinikere, pårørende og beslutningstakere. Det synes derfor nødvendig å utarbeide et felles sett av kriterier. Ulike empirikere peker også på at behandling kan gjøres mer effektiv gjennom å forske på hvilke variabler som har effekt, evaluere evidens for

teoretisk forankring og identifisere moderatorer. Videre vil jeg kort gjøre rede for hvorfor det er viktig å ta i bruk empirisk validerte metoder.

### **Behov for empirisk validert behandling**

Det er flere tungtveiende grunner til å benytte seg av validerte metoder. Som tidligere nevnt har ulike psykologiske retninger forskjellige syn på mål for behandling. Det er likevel plausibelt å tro at klienten oppsøker hjelp for å bli bedre. For klienten og dens pårørende, vil det derfor være avgjørende at behandlingen de blir tilbudt faktisk virker.

Det er også etisk riktig å benytte dokumentert behandling. Empirisk validert behandling er per definisjon behandling som har vist at den virker, og rangeringer av forskningsbasert behandling bør således være styrende for hvilken behandling klienter tilbys. I motsatt tilfelle kan terapeutens faglige preferanser, og kanskje enda mer bekymringsfullt; personlig interesse eller overbevisning, være avgjørende for hvilken behandling klienten mottar. Fagpersoners valg trenger på ingen måte å samsvare med hvilken behandling som til enhver tid har mest empirisk belegg. Det er en fare for at terapeuters valg av behandling kan bli preget av tilfeldigheter, som for eksempel hvilke behandlingsmetoder terapeuten er trent i, og/eller føler seg kompetent på. Dette kan medføre at klienter tilbys behandling som både er mer omfattende og mindre effektiv enn annen behandling.

Bruk av effektive behandlingsformer er ikke bare viktig for det enkelte individ og dets pårørende, men også i et samfunnsperspektiv. Ved tidlig bruk av effektive behandlingsformer har samfunnet muligheter til å oppnå en kraftig reduksjon i prevalens av de mest vanlige atferdsproblemene og psykiske lidelser hos barn og unge (Biglan & Ogden, 2008). Samfunnsøkonomisk vil det være svært gunstig om flere unge mennesker hjelpes med å gis de nødvendige ferdigheter til å bli produktive og ansvarsfulle voksne (Biglan & Ogden, 2008).

For beslutningstakere vil det være formålstjenlig å støtte veldokumenterte metoder, med begrunnelse i at tiltakene vil være kostnadsbesparende på sikt. Kortere og mer effektiv

behandling vil frigjøre ressurser som kan benyttes til behandling av andre klienter. Dette vil kunne bidra til mer kostnadseffektive behandlingsinstitusjoner, og reduksjon av pasienter på venteliste. I økende grad vil det også være juridiske grunner til å benytte empirisk validerte metoder. Stadig flere offentlige instanser er gjennom ulike lover og regelverk forpliktet til å tilby validert behandling.

Jeg har nå kort argumentert for bruk av empirisk validerte behandlingsmetoder. Dette synes viktig både i et individ og et samfunnsperspektiv. Fokuset vil nå rettes mot spesifikke empirisk validerte behandlingsmetoder innenfor atferdsanalysen.

### **Empirisk Validert Behandling innen Atferdsanalysen**

Som nevnt, viste 'Task Force' rapportene en stor overvekt av behandlinger basert på atferdsanalyse eller kognitiv atferdsterapi. Senere publiseringer har vist lignende resultater (e.g., Chambless & Ollendick, 2001; Nathan & Gorman, 2007). Atferdsanalytiske teknikker er regnet som en effektiv behandling mot en rekke psykiske lidelser.

Atferdsanalyse er en godt dokumentert behandling mot ulike angstlidelser. For agorafobi er in vivo eksponering ansett å være empirisk validert behandling (Chambless & Ollendick, 2001). Behandlingen begynner vanligvis med at klienten lager et angsthierarki av fryktede situasjoner eller aktiviteter. Klientene blir deretter oppfordret til å oppsøke disse plassene eller situasjonene og bli værende i dem inntil angsten forsvinner (Barlow, Allen, & Basden, 2007). Dette gjøres med en gradvis tilnærming. Også mot en rekke spesifikke fobier er eksponering in vivo rangert som en empirisk validert behandling (e.g., Chambless & Ollendick, 2001; Nathan & Gorman, 2007).

En liknende prosedyre, eksponering med responsprevensjon, er den best dokumenterte metode ved tvangslidelser (e.g., Chambless et al, 1998; Chambless & Ollendick, 2001). I tillegg til å eksponeres for situasjonen hvor tvangslidelsene forekommer, hindres klienten i å

utføre tvangshandlingene (responsprevensjon). Eksponering regnes også som en empirisk validert behandling mot posttraumatisk stress syndrom (Nathan & Gorman, 2007).

Atferdsanalyse er en validert behandling for depresjon (Chambless & Ollendick, 2001; Nathan & Gorman, 2007). Felles for atferdsanalytisk behandling av depresjon, er at depresjon ses på som en reduksjon i atferdsformer som leder til positive forsterkere. Med dette som utgangspunkt, fokuseres det på å øke positive dagligdagse aktiviteter, forbedre sosiale og kommunikative ferdigheter, øke adaptive ferdigheter og redusere negative livserfaringer (Craighead, Sheets, Brosse, & Ilardi, 2007). Noen rangerer også atferdsanalytisk behandling av schizofreni som en empirisk validert behandling (e.g., Nathan & Gorman, 2007).

Atferdsanalytisk behandling av schizofreni innebærer vanligvis bruk av tegnøkonomisystemer og sosial ferdighetstrening (Kopelowicz, Liberman, & Zarate, 2007).

De senere år er det også blitt utviklet flere atferdsanalytiske psykoterapier med godt empirisk belegg. Disse kalles ofte for tredjegenasjons atferdsanalyse. Aksepterings – og forpliktningsterapi (ACT) er en empirisk validert behandling for kronisk smerte. I stedet for å prøve å unngå eller kontrollere negative følelser, tanker, minner eller andre private hendelser, oppfordrer ACT til å akseptere disse hendelsene uten å forsøke å undertrykke dem. Gjennom øvelser vises klient hvordan språk ikke trenger å kontrollere atferd. Ut fra pasientens verdier settes konkrete mål og klienten forplikter seg til å handle på en slik måte at målene kan oppnås (Hayes & Bissett, 2000).

Dialektisk atferdsterapi (DBT) er en annen empirisk validert psykoterapi rettet mot borderline personlighetsforstyrrelser (Nathan & Gorman, 2007). DBT kan deles inn i tre steg. Første steg er at klienten hjelpes til å tilegne seg grunnleggende ferdigheter. Neste steg er å redusere posttraumatisk stress gjennom eksponering, før det til slutt fokuseres på å øke selvrespekt og definere individuelle mål (Hayes & Bissett, 2000).

En rekke behandlinger basert på anvendt atferdsanalyse er empirisk validert. Metodene presentert ovenfor er kun eksempler. En empirisk validert behandling som ikke er nevnt, er tidlig og intensiv opplæring basert på anvendt atferdsanalyse. Jeg vil avslutningsvis i denne artikkelen presentere studier som støtter metoden, argumentere for bruk av metoden, samt skissere utfordringer knyttet til empirisk validering av delkomponenter av behandlingen.

**Tidlig og intensiv opplæring av barn med autisme.** Som nevnt innledningsvis har min interesse for feltet bakgrunn i arbeid med barn med autisme. Tidlig og intensiv opplæring basert på anvendt atferdsanalyse (TIOBA) er en empirisk validert metode som systematisk etablerer ferdigheter innen språk, lek, sosial interaksjon og selvstendighet hos førskolebarn med autisme.

Ole Ivar Lovaas publiserte i 1987 en studie som vekket stor oppsikt innen autismebehandling. Resultatene viste at 9 av 19 barn (47%) i studien oppnådde normal intellektuell og skolemessig fungering etter 2 eller flere år med intensiv atferdsorientert behandling. Kun 2% av kontrollgruppen (N=40) oppnådde samme resultat. Studien har imidlertid blitt utsatt for mye kritikk. Kritikken har blant annet adressert det faktum at forsøkspersonene ikke ble randomisert til de to gruppene (Schopler, Short, & Mesibov, 1989).

Det er siden forsøkt gjort flere replikasjoner Lovaas-studien ved bruk av gullstandarden RCT (e.g., Sallows & Graupner, 2005; Smith, Groen, & Wynn, 2000). Smith, Groen og Wynn (2000) rapporterte liknende resultater på IQ-score, men ikke like mange barn oppnådde det som Lovaas definerte som normalfungering. Selv om få studier har kunnet matche tallene til Lovaas, er det en stadig økende mengde av studier som støtter atferdsanalytisk behandling av barn med autisme.

I en gjennomgang av 34 effektstudier på TIOBA (9 av disse med kontrollgruppe) viser Eldevik og kollegaer (2009) til gode resultater både på IQ og adaptiv atferd for barn som mottar intensiv atferdsorientert behandling. I en nyere metaanalyse av 11 effektstudier med

kontrollgruppe, konkluderer Peters-Scheffer, Didden, Korzilius, og Sturmey (2011) med at gruppene som fikk TIOBA fikk bedre resultater enn kontrollgruppene på IQ, non-verbal IQ, adaptive ferdigheter og på reseptivt og ekspressivt språk. Rogers og Vismara (2008) oppgir TIOBA som eneste intervensjon som imøtegår Chambless og kollegaers (1998) kriterier for å definere en behandling som veletablert.

Barns fremgang ved bruk av TIOBA er en tungtveiende grunn i seg selv for å tilby denne behandlingen. Studier tyder også på at intensiv atferdsorientert behandling vil kunne være kostnadsbesparende for samfunnet på lang sikt. Det kan enkelt tenkes at økt selvstendighet vil redusere utgifter knyttet til blant annet bemannet bolig og ulike støtte- og trygdeordninger. Både Jacobson, Mulick, og Green (1998) og Chasson, Harris, og Neelly (2007) estimerer store kostnadsbesparelser for henholdvis Pennsylvania og Texas i USA ved bruk av TIOBA. Peters-Scheffer, Didden, Korzilius, og Matson (2012) gjør en liknende beregning for Nederland, og konkluderer med at intensiv atferdsorientert behandling sparer samfunnet for betydelige kostnader. Det er imidlertid viktig å påpeke at to førstnevnte studiene baserer sine beregninger på de studiene som har vist best effekt av TIOBA.

Tidlig og intensiv atferdsorientert behandling tilbys i økende grad både i Europa og USA. Tross noen grunnleggende kjennetegn og kriterier for behandlingen, vil mange viktige elementer variere fra fagmiljø til fagmiljø. Det opereres med ulike anbefalinger på timeantall og ressurser, det finnes en rekke forskjellige behandlingsmanualer og ulike kvalitetssikringssystem. Mye av dette er forsket på, men med et økende antall tjenester og institusjoner som tilbyr atferdsorientert behandling, kan det stilles spørsmål til hvorvidt fagpersoner, pårørende og beslutningstakere er kjent med hva som bør anses som en gullstandard.

Innholdet i behandlingen varierer også mye. Atferdsanalyse består av mange ulike teknikker og strategier, som alle kan benyttes innenfor denne type behandling. Selv om

TIOBA er en empirisk validert metode, dukker det også stadig opp nye strategier og teknikker som tas i bruk, uten at disse nødvendigvis har empirisk belegg. Det kan være flere grunner til dette. Private konsulentfirma vil gjerne profilere seg på noe annet enn hva det offentlige hjelpeapparat kan tilby. Faren med et økende antall private firma som konkurrerer om de samme midlene, er at det vil komme flere og flere ulike tilnærminger og varianter på markedet. Dette trenger på ingen måte bare gjelde private firma. Det kan også tenkes at konkurranse mellom ulike offentlige fagmiljøer vil kunne føre til implementering og spredning av teknikker og strategier uten vitenskapelig belegg.

### **Avsluttende Merknader**

Empirisk validert behandling er stadig mer i vinden. Grunnet uenighet blant empirikere om hva som skal betegnes som evidens, kan imidlertid feltet fremstå som noe uryddig og uoversiktlig. For at feltet skal bevare integritet og troverdighet, er det viktig at det oppnås konsensus om hvilke kriterier som skal ligge til grunn for å definere behandling som empirisk validert.

Det debatteres hvordan feltet kan forbedres. Økt fokus også på evidens for den teoretiske forandringen til spesifikke metoder, kan være en måte å redusere spredning av kvasivitenskapelige behandlingspakker. Mer forskning rettet mot hvilke variabler innenfor en behandlingspakke som gir effekt (komponentanalyser), vil bidra til mer effektiv behandling. Identifisering av moderatorer vil kunne gi kunnskap om hvorfor behandlingen fungerer bedre på noen enn på andre. Dette vil være avgjørende for å tilby riktig behandling til riktig klient.

Det er viktig at det bedrives forskning og at forskningen sammenfattes og gjøres tilgjengelig for både fagpersoner, pårørende og beslutningstakere. Av individhensyn, etiske og samfunnsøkonomiske grunner bør det utelukkende benyttes empirisk validerte metoder. Prematur spredning av metoder, teknikker og strategier uten vitenskapelig belegg er uheldig, og bør unngås.

## Referanser

- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*, 271-285.
- 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
- Barlow, D. H., Allen, L. B., & Basden, S. L. (2007). Psychological treatments for panic disorders, phobias and generalized anxiety disorder. In P. E. Nathan & J. M. Gorman (Eds.), *A guide to treatments that work* (3 ed.). New York: Oxford University Press
- Beutler, L. E., Kim, J. E., Davison, E., Karno, M., & Fisher, D. (1996). Research contributions to improving managed health care outcomes. *Psychotherapy*, *33*(2), 197-206. doi: 10.1037/0033-3204.33.2.197
- Biglan, A., & Ogden, T. (2008). The evolution of evidence-based practices. *European journal of behavior analysis* *9*(1), 1-15.
- Blatt, S. J., Zuroff, D. C., Hawley, L. L., & Auerbach, J. S. (2010). Predictors of sustained therapeutic change. *Psychotherapy Research*, *20*(1), 37-54. doi: 10.1080/10503300903121080
- Bohart, A. C., O'Hara, M., & Leitner, L. M. (1998). Empirically violated treatments: Disenfranchisement of humanistic and other psychotherapies. *Psychotherapy Research*, *8*(2), 141-157. doi: 10.1080/10503309812331332277
- Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., . . . Woody, S. R. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, *51*(1), 3-16.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual review of psychology*, *68*, 685-716. doi: 10.1146/annurev.psych.52.1.685
- Chasson, G. S., Harris, G. E., & Neely, W. J. (2007). Cost comparison of early intensive behavioral intervention and special education for children with autism *Journal of child and family studies* *16*(3), 401-413. doi: 10.1007/s10826-006-9094-1
- Craighead, W. E., Sheets, E. S., Brosse, E. S., & Ilardi, S. S. (2007). Psychosocial treatments for major depressive disorder. In P. E. Nathan & J. M. Gorman (Eds.), *A guide to treatments that work* (3 ed.). New York: Oxford University Press.
- David, D., & Montgomery, G. H. (2011). The scientific status of psychotherapies: A new evaluative framework for evidence-based psychosocial interventions. *Clinical Psychology: Science and Practice*, *18*(2), 89-99. doi: 10.1111/j.1468-2850.2011.01239.x
- 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
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, *16*, 319-324.
- Gendlin, E. T., & Rychlak, J. F. (1970). Psychotherapeutic processes. *Annual Review of Psychology*, *21*(1), 155-190. doi: 10.1146/annurev.ps.21.020170.001103
- Hayes, S. C., & Bissett, R. T. (2000). Behavior psychotherapy and the rise of clinical behavior analysis. In J. Austin & J. E. Carr (Eds.), *Handbook of applied behavior analysis*. Reno: Context Press



- Heinicke, M. R., Carr, J. E., LeBlanc, L. A., & Severtson, J. M. (2010). On the use of fluency training in the behavioral treatment of autism: A commentary. *The Behavior Analyst*, 33(2), 223-229.
- Hunsley, J., & Lee, C. M. (2007). Research-informed benchmarks for psychological treatments: Efficacy studies, effectiveness studies, and beyond. *Professional Psychology: Research and Practice*, 38(1), 21-23. doi: 10.1037/0735-7028.38.1.21
- Jacobson, J. W., Mulick, J. A., & Green, G. (1998). Cost benefit estimates for early intensive behavioral intervention for young children with autism - general model and single state case *Behavioral Interventions*, 13, 201-226.
- Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, 63(3), 146-159. doi: 10.1037/0003-066X.63.3.146
- Kazdin, A. E. (2011). Evidence-based treatment research: Advances, limitations, and next steps. *American Psychologist*, 66(8), 685-698. doi: 10.1037/a0024975
- Kopelowicz, A., Liberman, R. P., & Zarate, R. (2007). Psychosocial treatments for schizophrenia. In P. E. Nathan & J. M. Gorman (Eds.), *A guide to treatments that work* (3 ed.). New York: Oxford University Press
- Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science*, 2(1), 53-70. doi: 10.1111/j.1745-6916.2007.00029.x
- Lohr, J. M. (2011). What is (and what is not) the meaning of evidence-based psychosocial intervention? *Clinical Psychology: Science and Practice*, 18(2), 100-104. doi: 10.1111/j.1468-2850.2011.01240.x
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, 55(1), 3-9 doi: 10.1037/0022-006X.55.1.3
- Nathan, P. E., & Gorman, J. M. (2007). *A guide to treatments that work*. New York: Oxford University Press.
- Peters-Scheffer, N., Didden, R., Korzilius, H., & Matson, J. (2012). Cost comparison of early intensive behavioral intervention and treatment as usual for children with autism spectrum disorder in the Netherlands. *Research in Developmental Disabilities*, 33(6), 1763-1772. doi: 10.1016/j.ridd.2012.04.006
- Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2011). A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 5(1), 60-69. doi: 10.1016/j.rasd.2010.03.011
- Practice, A. P. T. F. o. E.-B. (2006). Evidence-based practice in psychology. *American Psychologist*, 61(4), 271-285.
- Procedures, T. F. o. P. a. D. o. P. (1995). Training in and dissemination of empirically-validated psychological treatments. *The Clinical Psychologist*, 48(1), 3-23.
- Rachman, S., & Wilson, G. T. (1980). *The effects of psychological therapy* (2nd ed.). New York: Pergamon Press
- Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child & Adolescent Psychology*, 37(1), 8-38. doi: 10.1080/15374410701817808
- Schopler, E., Short, A., & Mesibov, G. (1989). Relation of behavioral treatment to "normal functioning": Comment on Lovaas. *Journal of Consulting and Clinical Psychology*, 57(1), 162-164. doi: 10.1037/0022-006X.57.1.162
- Scotti, J. R., Morris, T. L., McNeil, C. B., & Hawkins, R. P. (1996). DSM-IV and disorders of childhood and adolescence: Can structural criteria be functional? *Journal of Consulting and Clinical Psychology*, 64(6), 1117-1119

- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The consumer reports study. *American Psychologist*, *50*(12), 965-974. doi: 10.1037/0003-066X.50.12.965
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32* (9), 752-760. doi: 10.1037/0003-066X.32.9.752
- Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation*, *105*(4), 269-285.
- Stuart, R. B., & Lilienfeld, S. O. (2007). The evidence missing from evidence-based practice. *American Psychologist*, *62*(6), 615-616. doi: 10.1037/0003-066X62.6.615
- Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatments. *The Clinical Psychologist*, *48* (1), 3-23.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "All must have prizes". *Psychological Bulletin*, *122*(3), 203-215. doi: 10.1037/0033-2909.122.3.203
- Weisz, J. R., V.R., W., & Henggeler, S. W. (2005). Jousting with straw men: Comment on Westen, Novotny, and Thompson-Brenner (2004). *Psychological Bulletin*, *131*(3), 418-426.
- Westen, D., & Bradley, R. (2005). Empirically supported complexity: Rethinking evidence-based practice in psychotherapy. *Current Directions in Psychological Science*, *14*(5), 266-271. doi: 10.1111/j.0963-7214.2005.00378.x
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials *Psychological Bulletin*, *130*(4), 631-663. doi: 10.1037/0033-2909.130.4.631

Er rate en kontrollerende variabel for bevaring av ferdigheter?

Frode Hammerø

Høgskolen i Oslo og Akershus

Avdeling for atferdsvitenskap

Abstract

Fra presisjonsopplæringstradisjonen hevdes det at flytbasert trening produserer ferdigheter som bevares bedre enn ferdigheter som trenes til prosentvis mestring. Det synes imidlertid uklart om det finnes empirisk støtte for denne påstanden. Tre eksperimenter sammenliknet effekt av to ulike betingelser; tempo og presisjon. Det ble kontrollert for antall trials og feedback. Sammenfattede resultater for de tre eksperimentene viser liten forskjell på bevaring av ferdigheter for de to betingelsene. Funnene indikerer at tempo ikke er en kontrollerende variabel for bevaring av ferdigheter.

*Nøkkelord:* empirisk validert behandling, presisjonsopplæring, bevaring

### **Er Rate en Kontrollerende Variabel for Bevaring av Ferdigheter?**

Trenden i samfunnet er at det i økende grad stilles krav til at fagpersoner må gjøre rede for effektiviteten av det som promotes. Evidensbasert medisin har gått foran og rangert ulike behandlinger etter styrken på evidens. Innenfor psykologien ser man en økende interesse for, og spredning av, empirisk validerte behandlingsprogram. Utdanningssystemet har lenge vært preget av såkalt metodefrihet, men også innenfor opplæring synes nå interessen for empirisk validerte metoder å være tiltakende.

Det finnes mange og tungtveiende grunner for at også skoler og barnehager bør implementere opplæringsprogrammer som er vist effektive gjennom forskning. For det enkelte individ vil effektiv opplæring være avgjørende for å stå best mulig rustet til å møte samfunnets mange utfordringer. Kompetente enkelindivider vil gi store samfunnsøkonomiske gevinster, i form av produktive mennesker som bidrar til fellesskapet. Effektive opplæringsprogrammer vil også kunne hjelpe på flere av skolens utfordringer. Frafall som skyldes faglig utilstrekkelighet, vil kunne forebygges ved implementering av forskningsstøttede opplæringsprogram. Behovet for spesialpedagogisk hjelp vil kunne reduseres, og nødvendig spesialpedagogisk hjelp vil kunne gjøres mer effektiv. Lærerstanden har måttet tåle mye kritikk for manglende kompetanse. Det er nærliggende å tro at utdanningsinstitusjoners og læreres status vil øke som en følge av en mer effektiv opplæring av elevene.

En opplæringsmetode som har fått stadig mer oppmerksomhet de siste årene er Presisjonsopplæring. Metoden vektlegger tempo og har som mål å fremme atferd med flyt (Lindsley, 1992). Læringsutbytte ved bruk av presisjonsopplæringsstrategier oppsummeres vanligvis med akronymet RESA (retention, endurance, stability, application) som kan oversettes med bevaring, utholdenhet, stabilitet og anvendelse. Bevaring viser til at ferdigheter som fremvises med en viss frekvens opprettholdes på samme nivå etter trening.

Utholdenhet handler om å kunne arbeide med ferdigheten over lengre tid uten at det oppleves som anstrengende. Stabilitet vil si å kunne prestere på samme nivå i mer distraherende omgivelser enn under innlæring, mens anvendelse henspeiler til bruk av det som er lært i nye omgivelser, eller som en del av mer komplekse ferdigheter (Binder, 1996).

Innen presisjonsopplæring vil man typisk ha et høyt antall repetisjoner på hver ferdighet som trenes, og treningen kjennetegnes gjerne ved korte, tidsavgrensede økter. Den avhengige variabelen som benyttes er frekvens per tidsenhet. Dette har sin opprinnelse i Skinners fri operant studier. I en fri operant betingelse kan deltaker fritt avgi responser uten medvirkning fra eksperimentator/trener. Dette i motsetning til en 'discrete trial' betingelse, hvor eksperimentator/trener presenterer diskriminativ stimulus og formidler planlagte konsekvenser kontingent på responser (Cooper, Heron, & Heward, 2007).

Tempotrening eller flytbasert trening er en hovedkomponent innefor presisjonsopplæring. Flyt er innen denne tradisjonen definert som høy frekvens av riktige responser (Binder, 1996). Bruk av antall riktige responser per tidsenhet som avhengig mål skiller seg fra eksempelvis 'Discrete Trial Teaching' (DTT) og andre opplæringsmetoder som bruker prosentvis mestring.

Den grunnleggende antakelsen innen presisjonsopplæring er at det er en nær sammenheng mellom høyfrekvent og korrekt respondering, og at dette læringsutbyttet kan beskrives ved RESA. Et særlig viktig argument for bruk av presisjonsopplæringsstrategier er at ferdighetene skal bevares bedre ved bruk tempoprosedyrer og frekvensmål enn ved bruk av tradisjonelle mestringskriterer som for eksempel prosentvis korrekt (e.g., Binder, 1996; Bucklin, Dickinson, & Brethower, 2000; Holding, Bray, & Kehle, 2010; Kubina & Wolfe, 2005).

Flere presisjonslærere kritiserer den utstrakte bruken av nøyaktighetskriterier i opplæring. Blant annet hevdes det at bruk av prosentvis korrekt som mål på mestring fører til

at pedagoger ikke kan skille mellom kompetent og ikke-kompetent atferd (Binder, 1996; Binder, Haugton, & Bateman, 2002). Binder (1996) betegner dette som en av de største prosedyremessige hindringene for å etablere effektive undervisningsprosedyrer, og peker på nøyaktighetskriterier som en årsak til at elever mislykkes i skolen (Binder, Haugton, & Bateman, 2002). Holding et al. (2010) hevder at utelukkende bruk av nøyaktighetskriterier kan føre til at lærere avanserer i pensum prematurt, uten av ferdighetene har blitt en del av et repertoar som er anvendelig i naturlige situasjoner.

Det synes imidlertid ikke helt klart at det finnes empirisk støtte for påstandene om bedre bevaring av ferdigheter ved bruk av presisjonsopplæringstrategier (e.g., Doughty, Chase, & O'Shields, 2004; Heinicke, Carr, LeBlanc, & Severtson, 2010). Flere studier har sammenliknet effekt på bevaring ved bruk av frekvensmål eller nøyaktighetskriter. Singer-Dudek og Greer (2005) peker på at en generell svakhet med flere av studiene er at deltakerne i tempobetingelsen vanligvis har fått flere trials (læringsmuligheter) enn i nøyaktighetsbetingelsen (e.g., Olander, Collins, McArthur, Watts, & McDade, 1986). Av denne grunn kan det være vanskelig å isolere rate som en kontrollerende variabel for læringsutbyttet, ettersom effekten kan tilskrives antall trials.

Doughty og kollegaer (2004) belyser det samme problemet i en gjennomgang av empiriske artikler som sammenlikner nøyaktighetskriterier og frekvensmål. Studien viser at 79% av artiklene som ble evaluert, ikke kontrollerte for trening (trials). I 22% av disse artiklene, ble dette beskrevet som en mulig årsak til resultatet. Halvparten av de artiklene som faktisk kontrollerte for trening, rapporterte kun om tid brukt på trening, ikke antall trials. For eksempel viste Bucklin, Dickinson, og Brethower (2000) bedre bevaring og anvendelse av en delferdighet ved flyttrening enn ved nøyaktighetstrening. Gruppen som benyttet ratebyggingsprosedyrer fikk imidlertid flere trials enn gruppen som fikk nøyaktighetstrening.

Noen få empiriske studier har kontrollert for antall trials. Resultatene er imidlertid motstridende. Shirley og Pennypacker (1994) fikk bedre resultater ved bruk av frekvenskriterier enn nøyaktighetskriterier på test på bevaring av rettskriving hos to 8.klassinger med lærevansker. Kelly (1995) viste høyere bevaring av ordbilder hos førskolebarn med spesielle behov i flytbetingelsen enn i mestringsbetingelsen. Hanratty og Greer (2000) fikk derimot ingen forskjeller i en replikasjon av denne studien. Holding et al. (2011) sammenliknet effekt av flyttrening og DTT hos fire barn med autisme. Barna ble trent på ekspressiv benevning av substantiv. Samtlige av barna viste bedre bevaring av ferdighetene trent i flytbetingelsen.

Peladeau, Forget, og Gagne (2003) sammenliknet standard overlæringsprosedyrer med flytbygging hos 168 studenter. Studien viste at gruppen som ble eksponert for tempoprosedyrer hadde noe høyere responsrate enn studentene som benyttet overlæringsprosedyrer, men at dette verken påvirket akademiske resultater eller bevaring av ferdighetene 6 måneder etter. Fox og Ghezzi (2003) fikk små forskjeller mellom en overlæringsbetingelse og en flytbetingelse ved trening av konseptuell forståelse hos 41 studenter. Forfatterne understreker at lav eksperimentell styrke i denne delen av eksperimentet hindrer dem i å trekke noen konklusjoner. Mathews (2010) sammenliknet benevning av klokkeslett ved bruk av henholdsvis ratebygging og sakte trening (deltakerne ble promptet til å avgi responsene først etter 10 sekunder etter presentasjon av stimulus) hos tre voksne mennesker med utviklingshemming. Resultatene viste mer eller mindre lik bevaring av ferdighetene på posttest etter 2, 4 og 8 uker.

### **Eksperiment 1**

Formålet med dette eksperimentet var å undersøke hvorvidt DTT og høy rate av korrekt respondering genererte differensiert utkomme på bevaring av ferdighetene som ble trent. Dette er en problemstilling som bør adresseres for å sikre mest mulig effektiv



opplæring. Bruk av frekvenskriterium som mål på mestring vil i mange tilfeller innebære at man må trene lenger på en gitt ferdighet enn hva tilfellet er om man benytter prosentvis korrekt som kriterium. Dette vil være både tid – og ressurskrevende. Dersom flytbasert trening genererer bedre bevaring av ferdighetene vil det kunne være hensiktsmessig å operere med frekvenskriterier. Ved ingen eller små forskjeller vil det kunne stilles spørsmål til legitimitet av denne ekstra treningen. I så tilfelle vil tid og ressurser benyttet til å oppnå frekvensmålet i stedet kunne brukes til å etablere andre viktige ferdigheter.

Ekspérimentet vil sammenlikne effekten av to ulike betingelser; tempo (flytbygging) og presisjon (DTT). Bevaring av ferdighetene vil benytte som parameter på effekt. Det vil kontrolleres for både antall trials og ‘feed back’.

### **Metode**

**Deltaker.** Deltakeren i denne studen var en typisk utviklet gutt på 9 år. Under eksperimentperioden gikk han i 5.klasse. Deltaker hadde ingen tidligere erfaring som deltaker i læringseksperiment, og hadde heller ikke erfaring med tempotrening. Gjennom engelskundervisning hadde deltaker blitt eksponert for enkelte av stimuliene benyttet i studien.

**Setting.** Forsøkene ble gjennomført rett etter skoleslutt hjemme hos eksperimentator. For at ikke deprivasjon på mat ikke skulle påvirke innsatsen, ble deltaker gitt mat og drikke før oppstart av de daglige øktene. Forsøkene ble gjennomført i et rom på 40 kvadratmeter. Deltaker var plassert på enden av langsiden av bordet, mens eksperimentator satt like ved på kortsiden. Det foregikk ingen annen aktivitet i rommet under gjennomføring av forsøkene.

**Registrering.** Antall riktige responser per tidsenhet ble benyttet som avhengig mål i tempobetingelsen. I presisjonsbetingelsen var avhengig mål prosentvis korrekte responser. All datainnsamling ble gjort ved bruk av skriftlig materiell. Standard endringsskjema ble benyttet til å registrere antall riktige i tempobetingelsen. For å ta tiden på treningsintervall i

tempobetingelsen ble det brukt en timer. Det ble benyttet en klokke til å måle samlet tid brukt på betingelsen. I presisjonsbetingelsen ble prosentvis riktige responser chartet i et søylediagram. Det ble også i denne betingelsen brukt en klokke til å måle samlet tid brukt på betingelsen. All registrering ble gjort manuelt.

Det ble gjennomført interscorerelabilitetstest på totalscore for hele pretest og posttest, og for 1/3 av treningsøktene (tre av ni økter i tempobetingelse og en av tre økter i presisjonsbetingelse). To ulike observatører regnet ut totalscoren for hver økt og laveste totalscore ble dividert med høyeste totalscore (Cooper et al., 2007). Resultatet ble multiplisert med 100 for å finne prosentvis enighet mellom observatørene.

**Målatferd.** Ferdighetene som ble trent var betingede diskriminasjoner med norske ord som utvalgsstimuli og engelske ord som sammenlikningsstimuli. Oppgavesettet i begge betingelser hadde et multiple choice format (se Appendix A og B) med 10 utvalgsstimuli per ark. På samme linje som det norske substantivet ble det gitt tre svaralternativer på engelsk. Målatferd i begge betingelser var å sette en prikk i ruten til det tilsvarende engelske substantivet fra utvalget på tre. Deltaker ble på forhånd informert om at det ikke kunne settes mer enn en prikk på hver svarlinje.

**Pretest.** For å måle deltakers ferdighetsnivå før igangsetting av intervensjon, ble det gjennomført en pretest dagen før treningsstart. Resultatene fra pretest dannet grunnlag for fordeling av stimuli til henholdsvis tempo og presisjonsbetingelse.

Oppgavesettet under pretest var i et multiple choice format. Hver stimulus ble presentert separat, ved at deltaker fikk presentert en papirremse med en norsk utvalgstimulus og engelske sammenlikningsstimuli (Appendix C). Deltaker merket den riktige oversettelsen. Når deltaker hadde avgitt en respons, tok trener hånd om papirremsen for å forhindre at deltaker sammenliknet svaret med tidligere avgitte responser. Det ble utarbeidet tre ulike versjoner av stimulusmateriell, med ulik plassering av sammenlikningsstimuli for hver av

versjonene. Det ble valgt ut utvalgsstimuli som ikke lignet på det ekvivalente norske ordet.

Under pretest ble det ikke gitt feedback.

Det ble innhentet fire målepunkter på hver av de 30 glosene, til sammen 120 målepunkter. Alle målepunktene ble innhentet samme dag. På bakgrunn av resultatene på pretest, ble det utarbeidet to ulike oppgavesett med 10 engelske gloser i hvert av settene. Til hvert av settene ble det fordelt to gloser med 0% mestring, to gloser med 25% mestring, en glose med 50% mestring, en glose med 75% mestring og fire gloser med 100% mestring. De to oppgavesettene ble deretter trent i hver sin betingelse (tempo eller presisjon).

**Stimuli.** Etter pretest ble det utarbeidet tre ulike versjoner av stimulusmateriell til hver av betingelsene med 10 utvalgsstimuli og 30 sammenlikningsstimuli på hvert ark (multiple choice format). Både rekkefølgen på utvalgstimuli og plassering av sammenlikningsstimuli ble randomisert for hver av disse versjonene. Appendix A viser en versjon av stimulusmateriell for presisjonsbetingelse. Appendix B viser en versjon av stimulusmateriell for tempobetingelse. Sammenlikningsstimuli benyttet i pretest ble også brukt under trening i de to betingelsene.

All tekst var i skriftstørrelse 12. Deltaker benyttet en gul tusj til å merke riktig svar, mens eksperimentator brukte en blå tusj til å merke inkorrekte svar. Stimuli i eksperiment 1 var en blanding av kjente og ukjente gloser. Dette for å unngå et høyt antall feil innledningsvis i treningen. For i størst mulig grad sikre at deltaker ikke fikk erfaring med oppgavene mellom intervensjon og posttest, ble det ikke inkludert gloser som var en del av deltakers pensum.

**Prosedyre.** To ulike treningsprosedyrer ble arrangert for henholdsvis tempobetingelsen og presisjonsbetingelsen.

**Tempobetingelse.** Før igangsetting av trening ble det definert et frekvensmål og et stabilitetskriterium. Frekvensmålet ble satt til 25 riktige responser innenfor et 30 sekunders

tidsintervall. Dette frekvensmålet måtte innfris over to påfølgende økter (stabilitetskriterium). Deltaker kunne fritt avgi responser (fri operant) innenfor tidsintervallet. Tempobetingelsen bestod av 3 daglige økter, der hver økt hadde en varighet på 30 sekunder. Eksperimentator chartet resultatene sammen med deltaker. Av motivasjonelle forhold ble kun rette responser chartet.

I forkant av betingelsene ble deltaker informert om hvilken betingelse som var den gjeldene. I tempobetingelsen oppfordret eksperimentator deltaker til å respondere så fort som mulig, og forsøke å slå sin egen rekord. Før hver treningsøkt ble følgende instruks formidlet: 'nå skal du jobbe så fort du klarer, og prøve å slå rekorden'. Det ble presisert for deltaker at responsen ikke kunne endres når den først var avgitt. En timer ble plassert foran deltaker og timeren ringte når økten var ferdig (etter 30 sekunder). Det ble også målt samlet tid benyttet på betingelsen. For å unngå at deltaker skulle 'skynde seg' i *presisjonsbetingelsen*, ble deltaker ikke informert om denne målingen.

Deltaker ble presentert for sammenstiftede ark med 10 norske substantiv og tre engelske svaralternativer i en tabell på hvert ark (Appendix B). Som nevnt, var alle sammenlikningsstimuli riktige responser på andre oppgaver deltaker ble eksponert for under pretest. Det vil si at dersom den norske utvalgsstimulus var 'flaske' og sammenlikningsstimuli var: bottle – cap – church, ville også 'kork' og 'kirke' ha vært riktige respons på noen av oppgavene fra pretest.

I etterkant av hver økt (timing) ble det gjennomført en feedbackprosedyre. Eksperimentator gikk gjennom alle de avgitte responsene med deltaker. Eksperimentator sa for eksempel "Bestikk er cutlery". På korrekte responser ble det formidlet potensielle sosiale forsterkere som "kjempebra, flott, glimrende, korrekt eller lignende". Ved inkorrekte eller uavgitte responser ble det formidlet korreksjoner, for eksempel ved at eksperimentator sa; "Bestikk er cutlery", "Hva er bestikk?", samtidig som eksperimentator pekte på korrekt rute.

Deltaker gjentok ordet fra ruten eksperimentator pekte på, og satt deretter en prikk i riktig rute.

**Presisjonsbetingelse.** Presisjonsbetingelsen bestod av en daglig økt og ble alltid gjennomført i etterkant av de tre øktene med tempotrening. Foruten at oppgavesettet bestod av andre utvalgs – og sammenlikningsstimuli, var arkenes oppsett identisk med tempobetingelse (Appendix A). Oppgavesettet bestod av samme antall responser som deltaker hadde avgitt i foregående tempobetingelse. Arkene var sammenstiftet, og eksperimentator markerte på forhånd hvor langt deltaker skulle jobbe.

Deltaker ble i forkant av betingelsen informert om at det ikke var noe tempokrav i denne betingelsen, og at mestring ble målt i nøyaktighet. Følgende instruks ble formidlet før hver økt: 'nå kan du bruke så lang tid du ønsker på å svare'. Som i tempobetingelsen ble det påpekt at avgitte responser ikke kunne endres. Øktene ble ikke timet, og det var ingen tidsbegrensing som i tempobetingelsen. Det ble imidlertid målt samlet tid benyttet på betingelsen, uten at deltaker ble informert om dette.

Til forskjell fra tempobetingelse, formidlet trener feedback kontingent på hver respons, umiddelbart etter responsen var avgitt. Feedback ble formidlet på samme måte som i tempobetingelsen. På korrekte responser ble det formidlet potensielle sosiale forsterkere som 'kjempebra, flott, glimrende, korrekt eller lignende'. Ved inkorrekte eller uavgitte responser ble det formidlet korreksjoner, for eksempel ved at eksperimentator sa; 'Torsk er cod', 'Hva er torsk?', samtidig som eksperimentator pekte på korrekt rute. Deltaker gjentok ordet fra ruten eksperimentator pekte på, og satt deretter en prikk i riktig rute.

Eneste forskjell på feedbackprosedyre var *tidspunkt* på formidling av feedback. Mens feedbackprosedyren i tempobetingelsen ble gjennomført etter endt timing (deltaker avga mellom 15-27 responser under timings), ble feedback formidlet kontingent på hver respons i presisjonsbetingelsen.

**Bevaringstest/follow-up.** Fire uker etter siste treningsdag ble det gjennomført en posttest på bevaring av ferdighetene. Mestringskriteriet ble satt til 100%. Dette innebar at kun responser som ble mestret på alle de fire forsøkene under bevaringstest ble regnet som mestret. Bevaringstest ble gjennomført på samme måte som pretest, altså ved at deltaker ble gitt en og en papirremse med en utvalgsstimulus og tre sammenlikningsstimuli. Det ble benyttet samme stimulusmateriell som under pretest (Appendix C). Den eneste forskjell fra pretest var at det kun ble testet for responser som var fordelt til en av de to betingelsene. Dette betyr at deltaker ikke ble testet på de ti glosene som ikke var inkludert i noen av betingelsene. Glosene fra de ulike betingelsene ble presentert i randomisert rekkefølge, slik at det var tilfeldig om glosene ble hentet fra tempo – eller presisjonsbetingelsen.

**Design.** Det ble daglig alternert mellom de to ulike betingelsene; tempo og presisjon. For å kunne kontrollere for antall trials (forsøk/læringsmuligheter), ble alltid tempobetingelsen gjennomført først. Det ble benyttet et ”yoked schedule” mellom de to betingelsene ved at antall responser i tempobetingelsen ble bestemmende for antall responser i presisjonsbetingelsen. I etterkant av tempobetingelsen ble antall trials summert opp, og samme antall trials ble presentert i påfølgende presisjonsbetingelse. Det vil si at dersom deltaker hadde hatt 60 trials fordelt på tre timings under tempobetingelsen, ville deltaker bli gitt 60 trials i påfølgende presisjonsbetingelse.

Balansering av antall trials sikret samtidig at planlagte konsekvenser formidlet av trener ble balansert over betingelsene, ettersom deltaker fikk feedback (enten kontingent eller etter endt timing) på alle avgitte responser, uavhengig om responsen var korrekt eller inkorrekt. Det vil imidlertid være en forskjell på samlet antall potensielt forsterkende tilbakemeldinger og antall korreksjoner, da dette var avhengig av deltakers prestasjoner. I den grad det lot seg gjøre, ble det også balansert for hvor mange ganger hver stimulus ble brukt som sammenlikningsstimulus. Vekslingen mellom intervensjonene fortsatte inntil deltaker

nådde frekvensmålet over to påfølgende økter i tempobetingelsen. For å balansere antall trials ble syklusen avsluttet med presisjonsbetingelsen.

I presisjonsopplæring chartes alltid elevens resultater. Vanligvis gjøres dette sammen med eleven. Charting av egne resultater hevdes å være motiverende for eleven, og for å unngå at motivasjonelle forhold skulle utgjøre en forskjell på betingelsene, ble resultatene chartet sammen med deltaker i begge betingelser. I tempobetingelsen ble resultatene chartet på ”Standard Celeration Chart” (SCC), mens resultatene i presisjonsbetingelsen ble chartet i et søylediagram.

### **Resultat og Diskusjon**

Resultat fra interscorerelibilitetstest viser høy enighet mellom observatørene. Observatørere endte på 100% enighet på treningsøktene i både tempo –og presisjonsbetingelse, samt på posttest. På pretest var det 99% enighet mellom observatørene.

Deltaker nådde frekvenskriteriet i tempobetingelsen på forsøk 2 og 3 på tredje treningsdag. På forsøk 2 hadde deltaker 26 riktige og ingen feil. Under forsøk 3 fikk deltaker igjen 26 riktige, men denne gang med en feil. I påfølgende presisjonsbetingelse oppnådde deltaker en mestringsprosent på 98%.

Figur 1 viser resultat på bevaringstest etter en måned. Figuren viser ingen forskjell på bevaring av ferdigheter som ble trent i henholdsvis tempo – eller presisjonsbetingelse.

Deltaker oppnådde 90% mestring på oppgavene fra begge betingelsene.

Det ble målt samlet tid benyttet på trening, feedback og charting under hver økt i begge betingelser. Øktene i tempobetingelsene hadde en samlet varighet på 29 minutter, mens øktene i presisjonsbetingelsen hadde en varighet på 16 minutter, noe som betyr at det ble benyttet 13 minutter mer på tempobetingelsen.

I dette eksperimentet kan antall utvalgsstimuli kan være en potensiell validitetstrussel mot resultatene. Testformatet (multiple choice med tre sammenlikningsstimuli) gjorde at

deltaker hadde en relativt høy prosentvis sjanse til gjette seg frem til riktig svar. Deltaker hadde også blitt eksponert for noen av stimuliene tidligere. Til sammen bidro dette til at noen gloser hadde delvis eller full mestring på pretest. Det kan dermed hevdes at det ble inkludert for få utvalgstimuli til å kunne trekke noen slutninger om effekt av de ulike prosedyrene.

Sammenlikningsstimuli brukt i pretest ble også inkludert i trening. Rekkefølgen på sammenlikningsstimuli ble randomisert, men det ble benyttet identiske sammenlikningsstimuli under pretest, trening og posttest. Grunnen for dette var at det ville gjøre en sammenlikning av resultatene enklere. Pretest bestod av 30 utvalgsstimuli, mens kun 20 stimuli ble fordelt til en av to betingelsene. Dette betyr at 10 stimuli som ikke ble trent på, likevel ble beholdt som sammenlikningsstimuli i treningssettene. Utrente sammenlikningsstimuli vil dermed kunne ha blitt raskere etablert som S-delta enn sammenlikningsstimuli som var riktige responser på utvalgsstimuli som var inkludert i et av treningssettene. 7 av 10 responser i presisjonsbetingelsen hadde en sammenlikningsstimulus som ikke var korrekt respons i noen av settene, mens 2 av 10 gloser i tempobetingelsen hadde en sammenlikningsstimulus som ikke var inkludert i noen av settene. Dette kan ha medført at ferdighetene i presisjonsbetingelsen var enklere å etablere.

Det ble derfor gjennomført et nytt eksperiment hvor alle sammenlikningsstimuli i multiple choice oppgavene var en riktig respons på en av de andre oppgavene i samme oppgavesett. For å unngå høy mestring under pretest, ble det valgt ut ferdigheter som var ukjente for deltaker.

## **Eksperiment 2**

Eksperiment 2 var en systematisk replikasjon av eksperiment 1. Formålet med dette eksperimentet var å undersøke effekter av tempotrening og presisjonstrening når alle sammenlikningsstimuli i multiple choice oppgaven var riktig i en av oppgavene i oppgavesettet. I tillegg til tester for bevaring av ferdigheter ville jeg i dette eksperimentet



undersøke effekter av de to betingelsene på utholdenhet og når det ble krevd en annen responsmodalitet.

### **Metode**

Setting, registrering og design var likt som i eksperiment 1. I dette eksperimentet ble det gjennomført interscorerelabilitetstest på totalscore for 50% av responsene på pre –og postester og for 50% av treningsøktene i begge betingelser.

**Deltaker.** Deltakeren i dette eksperimentet var en 29 år gammel høyskoleutdannet kvinne. Deltaker hadde ikke tidligere deltatt i læringseksperimenter, og hadde ingen erfaring med tempotrening. Stimuli som ble benyttet i eksperimentet var ukjente for deltaker.

**Målatferd.** Ferdighetene som ble trent var betingede diskriminasjoner med norske ord som utvalgsstimuli og spanske ord som sammenlikningsstimuli. På samme linje som utvalgstimulus ble det gitt tre spanske substantiv som sammenlikningsstimuli. Målatferd i begge betingelser var å sette en prikk i ruten til det tilsvarende spanske ordet fra utvalget på tre. Deltaker ble på forhånd informert om at det ikke kunne settes mer enn en prikk på hver svarlinje.

**Pretest.** Også i eksperiment 2 ble det innhentet fire målepunkter på 30 gloser på pretest. Testformat var imidlertid ulikt fra eksperiment 2. Det ble gjennomført to pretester. Pretest 1 var en generaliseringstest. Deltaker fikk presentert fire sammenstiftede A-4 ark, hvor norske ord var randomisert for hvert ark. Deltaker skulle *skrive* det tilsvarende spanske ordet ved siden av det matchende norske ordet (Appendix D). Denne testen ble designet med henblikk på å evaluere hvorvidt man fikk en overføring av ferdighetene til en annen responsmodalitet (les/skriv) enn den som ble trent (les/merk).

Pretest 2 (Appendix E) var en multiple choice oppgave. Deltaker fikk presentert fire sammenstiftede A-4 ark med to kolonner, med henholdsvis 30 norske og 30 spanske gloser i separate kolonner på hvert ark. Deltaker satt strek fra det norske ordet til det tilsvarende

spanske. For hvert av de fire arkene, var rekkefølgen på glosene randomisert i begge kolonner. Også i eksperiment 1 ble det gjennomført en multiple choice test som pretest, men denne endringen i testformat førte til at deltaker fikk langt flere sammenlikningsstimuli i eksperiment 2, noe som igjen reduserte sannsynligheten for at deltaker skulle kunne gjette seg frem til korrekt respons.

Som i eksperiment 1 ble glosene fordelt jevnt til de to oppgavesettene. Deltaker hadde 0% mestring på alle stimuli som ble fordelt til de to betingelsene. Frekvenskriteriet ble oppjustert til 27 riktige responser på 30 sekunder, og stabilitetskriteriet ble økt til tre påfølgende økter. Disse endringene ble gjort på bakgrunn av deltakers alder og utdanning.

**Stimuli.** Basert på resultater fra pretest, ble det utarbeidet tre ulike versjoner av stimulusmaterieell til hver av betingelsene. Hvert ark inkluderte 10 utvalgsstimuli og 30 sammenlikningsstimuli i et multiple choice format. Både rekkefølgen på utvalgsstimuli og plassering av sammenlikningsstimuli var randomisert for hver av disse versjonene. Appendix F viser en av de tre ulike versjonene av stimulusmaterieell fordelt til presisjonsbetingelse og appendix G viser en av de tre versjonene av stimuli i tempobetingelse.

**Prosedyre.** Prosedyren for de to betingelsene var den samme som i eksperiment 1. Ferdighetene som ble trent i dette eksperimentet var imidlertid spanske gloser. Ettersom deltaker hadde liten læringshistorie med spansk, ville dette sannsynliggjøre lavere mestring på pretest enn under eksperiment 1. For å sikre dette ytterligere, ble det valgt ut gloser som på skrivemåte verken lignet det tilsvarende norske eller det tilsvarende engelske ordet. Det ble videre forsøkt å velge ut sammenlikningsstimuli som hadde cirka samme antall bokstaver. Dette for å unngå at lengden på ordet skulle få stimuluskontroll over responderingen.

For å unngå at enkelte sammenlikningsstimuli skulle bli lettere etablert som S-delta, ble det sikret at alle sammenlikningsstimuli var korrekte responser på andre utvalgsstimuli innen samme betingelse. Det vil si at når 'skap' (armario) ble presentert som utvalgsstimulus,

var sammenligningstimuli 'pluma' og 'cama' korrekte responser på andre utvalgsstimuli på samme ark.

**Posttest.** Det ble gjennomført tre ulike posttester. Mestringskriterium for bevaringstest og utholdenhetstest ble satt til 100% riktige. På generaliseringstest var også mestringskriteriumet 100%, men deltaker kunne ha inntil en uriktig bokstav på hvert skrevne ord. Posttest 1 var en generaliseringstest (Appendix I) og ble gjennomført etter henholdsvis en og to måneder etter siste treningsdagsdag. Testen var lik pretest 1, med unntak av det kunne testet for gloser som hadde blitt fordelt til en av betingelsene (se Appendix D). De 20 norske glosene fra de to betingelsene ble randomisert og presentert sammen i en kolonne på et A4-ark. Deltaker skrev det riktige spanske ordet ved siden av det tilsvarende norske ordet. Det ble utarbeidet fire ark der glosene var randomisert for hvert av arkene.

Posttest 2 (bevaringstest) var identisk med pretest 2 (Appendix E), og ble gjennomført etter henholdsvis en og to måneder etter siste treningsdag. Bevaringstestene ble presentert i etterkant av generaliseringstestene. For å unngå færre sammenlikningsstimuli enn på pretest, ble også de glosene som ikke var fordelt til noen av betingelsene inkludert i bevaringstestene. Det ble ikke foretatt noen ny randomisering fra pretest.

Etter 40 dager ble det gjennomført en utholdenhetstest. Det ble benyttet samme stimulusmateriell som under trening (Appendix F og G), med de samme tre versjonene av stimulusmateriell. Gloser fra begge betingelser ble testet hver for seg på et 1,5 minutters intervall. Deltaker ble oppfordret til å jobbe så fort som mulig i begge betingelser.

## **Resultat og Diskusjon**

Interscorereliabilitetsstest viste høy enighet mellom observatørene. Pretest multiple choice, skriftlig pretest og test av treningsøker viste alle 100% enighet. Resultater fra generaliseringstest etter en og to måneder viste en enighet på henholdsvis 100% og 92%, mens både bevaringstest etter en og to måneder oppnådde en enighet på 98%.

Deltaker nådde frekvenskriteriet i tempobetingelsen etter fire dager og til sammen 12 timede økter på 30 sekunder. Under de tre siste øktene hadde deltaker henholdsvis 28, 27 og 29 riktige responser. Deltaker hadde ingen feil på noen av disse forsøkene. I påfølgende presisjonbetingelse oppnådde deltaker en mestringsprosent på 98%.

Figur 2 viser resultat for bevaringstest for begge betingelsene etter en og to måneder. Figuren viser noe bedre bevaring av ferdighetene trent i tempobetingelsen etter en måned. Deltaker hadde 80% mestring på oppgavesettet fra tempobetingelse og 70% mestring på oppgavesettet fra presisjonsbetingelse. Etter to måneder viser imidlertid figuren noe bedre resultat på ferdigheter trent i presisjonsbetingelse enn ferdigheter trent i tempobetingelse. I presisjonsbetingelse oppnådde deltaker 80% riktige, mens 70% på ferdigheter trent i tempobetingelse.

Figur 3 viser resultat for generaliseringstest for begge betingelser etter en og to måneder. Figuren viser noe bedre generalisering av ferdigheter trent i tempobetingelse enn av ferdigheter trent i presisjonsbetingelse. Etter en måned oppnådde deltaker 30% mestring på ferdigheter fra tempobetingelse og 10% mestring på ferdigheter fra presisjonsbetingelse. Etter to måneder scoret deltaker 60% riktig på ferdigheter fra tempobetingelse og 40% på ferdigheter fra presisjonsbetingelse.

Figur 4 viser resultat på utholdenstest for de to betingelsene etter 40 dager. Som det går frem av figuren viser deltaker noe bedre utholdenhet på stimuli trent i tempobetingelse enn på ferdighetene trent i presisjonsbetingelse. Deltaker avga 71 responser (70 riktige og 1 feil) på stimuli fra tempobetingelsen. På gloser fra presisjonsbetingelsen avga deltaker 57 responser (alle riktige).

Det ble målt samlet tid benyttet på hver av betingelsene. Samlet tid inkluderte tid benyttet på trening, feedback og charting. I tempobetingelsen brukte deltaker til sammen 55 minutter, mens samlet tid i presisjonsbetingelsen var 32 minutter.

Resultatene fra bevaringstester på multiple choice oppgave viser liten forskjell på bevaring av ferdighetene trent i de to betingelsene for denne deltakeren. Generaliseringstesten viser imidlertid at ferdigheter trent i tempobetingelse har gitt noe bedre overføring enn ferdigheter trent i presisjonsbetingelsen for deltakeren. Resultatene fra utholdenhetstest viser også at trening på tempo har bidratt til bedre utholdenhet enn trening på presisjon for denne deltakeren.

Utholdenhetstesten kan være en validitetstrussel mot resultater på posttest etter to måneder. Ettersom deltaker hadde en høyere frekvens på gloser fra tempobetingelse under utholdenhetstest, fikk deltaker flere forsøk på gloser fra tempobetingelse (14 flere upromptede trials) enn fra presisjonsbetingelse. For å kontrollere antall trials, ble utholdenhetstesten utelatt i eksperiment 3.

### **Eksperiment 3**

Eksperiment 3 var en direkte replikasjon av eksperiment 2. Eneste forskjell fra eksperiment 2, var at det ikke ble gjennomført en test på utholdenhet, og at frekvenskriteriet ble nedjustert til 26 korrekte responser på 30 sekunder (på grunn av deltakers alder). Stabilitetskriteriet var for øvrig uendret.

#### **Metode**

Registrering, målatferd, stimuli og prosedyre var lik som i eksperiment 2. Pretestene (se Appendix D og E) hadde samme format og bestod av de samme stimuli som i eksperiment 2. Eneste forskjellen fra eksperiment 2 var hvilke gloser som ble fordelt til de to betingelsene, og at det ikke ble gjennomført en utholdenhetstest. Appendix J viser en av de tre versjonene av stimulusmateriell fordelt til presisjonsbetingelse, mens appendix K viser en av de tre versjonene av stimulusmateriell i tempobetingelse. Som i eksperiment 2 hadde deltaker 0%

mestring på alle gloser som ble fordelt til de to betingelsene. Bevaringstest (Appendix E) og generaliseringstest (Appendix I) ble også gjennomført på samme måte som i eksperiment 2.

I eksperiment 3 ble det gjennomført interscorerelabilitetstest for 1/3 av responsene avgitt i begge treningsbetingelser, og på 50% av resultater fra pre – og posttester.

**Deltaker.** 53 år gammel høyskoleutdannet kvinne. Deltaker hadde ikke tidligere deltatt i læringseksperimenter, og hadde ingen erfaring med tempotrening. Glosene som ble benyttet i eksperimentet var ukjente for deltaker.

**Setting.** Forsøkene ble gjennomført i et rom på 50 kvadratmeter hjemme hos deltaker. Deltaker satt på enden av langsiden på bordet, mens eksperimentator var plassert like ved på kortsiden. Det foregikk ingen annen aktivitet i rommet under gjennomføring av forsøkene.

### **Resultat og Diskusjon**

Interscorerelabilitetstest viste høy enighet mellom observatørene. Resultatet fra treningsøktene viste 100% enighet i begge betingelser. Observatørene var også 100% enige på resultater fra skriftlig pretest. Pretest multiple choice viste en enighet på 94%. Generaliseringstest etter en og to måneder oppnådde en enighet på henholdsvis 94% og 100%, mens bevaringstest etter en og to måneder scoret henholdsvis 100% og 97%.

Deltaker nådde frekvenskriteriet i tempobetingelsen etter til sammen fire dager og 14 timede økter på 30 sekunder. Siste treningsdag ble det gjennomført fem økter istedenfor de vanlige tre øktene. Deltaker avsluttet med tre påfølgende økter med 26 korrekte responser. På disse tre øktene hadde deltaker henholdsvis 1, 0 og 2 feil. I påfølgende presisjonsbetingelse oppnådde deltaker en mestringsprosent på 99%.

Figur 5 viser resultat for på bevaringstest for begge betingelser etter en og to måneder. Figuren viser bedre bevaring av ferdigheter trent i presisjonsbetingelse enn ferdigheter trent i tempobetingelse, både etter en og to måneder. Etter en måned hadde deltaker 70% mestring på responser trent i presisjonsbetingelse, og 40% mestring på responser trent i

tempobetingelse. Etter to måneder hadde deltaker 90% mestring på ferdigheter trent i presisjonsbetingelse, og 50% mestring på ferdighetene trent i tempobetingelse.

Figur 6 viser resultat på generaliseringstest for de to betingelsene. Figuren viser bedre generalisering av ferdigheter trent i presisjonsbetingelse enn ferdigheter trent i tempobetingelse. Etter en måned mestret deltaker 40% prosent av ferdighetene trent i presisjonsbetingelse og 20% av ferdighetene trent i tempobetingelse. Etter to måneder scoret deltaker 40% riktig på ferdighetene trent i presisjonsbetingelse og 30% på ferdighetene trent i tempobetingelse.

Måling av samlet tid benyttet på de to betingelsene viser at det ble brukt lenger tid på tempobetingelsen enn på presisjonsbetingelsen. Tempobetingelsen hadde en varighet på til sammen 1 time og 7 minutter, mens presisjonsbetingelsen varte i 27 minutter.

### **Generell Diskusjon**

Hovedformålet med dette eksperimentet var å undersøke hvorvidt DTT og høy rate av korrekt respondering genererte differensiert utkomme på bevaring av ferdighetene som ble trent. Dette ble gjort ved å sammenlikne effekten av de to ulike betingelsene tempo og presisjon.

I eksperiment 2 og 3 ble det i tillegg undersøkt hvorvidt tempotrening førte til bedre overføring av ferdigheter til en annen responsmodalitet enn presisjonstrening. I eksperiment 2 ble det arrangert en test for utholdenhet med henblikk på å undersøke effekten av de to betingelsene på utholdenhet.

Sammenfattede resultater på bevaringstester viser liten forskjell på bevaring av ferdigheter trent i de to betingelsene. Resultatene støtter støtter ikke påstander om at tempotrening genererer bedre bevaring av ferdighetene, og er således i tråd med funn av blant andre Peladeau et al (2003), Hanratty & Greer (2000), Fox & Ghezzi (2003) og Mathews

(2010). Resultatene er imidlertid motstridende til funn av Shirley og Pennypacker (1994), Bucklin og kollegaer (2000), Kelly (1995) og Holding og kollegaer (2011).

Resultatene indikerer at tempo ikke er den kontrollerende variabelen for bevaring av ferdighetene. Etersom det ble kontrollert for antall trials og feedback, er det nærliggende å konkludere med at antall trials er en kontrollerende variabel for bevaring. Et lavt antall deltakere og relativt få stimuli gjør imidlertid funnene noe begrenset. Det oppfordres til mer omfattende undersøkelser med et høyere antall deltakere og flere stimuli.

Frekvensmålet ble i alle tre eksperimentene satt ut i fra nivået til en person som allerede hadde flyt på ferdighetene. Dette er en vanlig strategi blant presisjonslærere (Johnson & Layng, 1996). All den tid det ikke er utarbeidet standardiserte frekvensmål for ferdighetene, vil det kunne hevdes at frekvenskriterier har blitt satt for lavt, og at man dermed ikke har fått full effekt av tempoprosedyren. For å unngå en slik kritikk vil det være en fordel at nye studier velger ut ferdigheter hvor det finnes standardiserte frekvensmål.

Sammenfattende resultater på generaliseringstest gir motstridende resultater. I eksperiment 2 viste deltaker bedre generalisering av ferdigheter trent i tempobetingelse, mens deltaker i eksperiment 3 viste markant bedre generalisering av ferdigheter trent i presisjonsbetingelse. Resultatene gjør det vanskelig å konkludere. Også her vil et lavt antall stimuli, i kombinasjon med få deltakere, gjøre funnene noe begrenset. Det oppfordres til videre undersøkelser for å avdekke hvorvidt tempotrening gir bedre generalisering av ferdigheter enn presisjonstrening.

En enkelt test på utholdenhet i eksperiment 2 viste noe bedre utholdenhet på ferdigheter trent i tempobetingelse enn ferdigheter fra presisjonsbetingelse. Funnene er imidlertid kun gjeldene for en enkelt deltaker, og forskjellen mellom betingelsene er heller ikke stor. Det oppfordres derfor til videre undersøkelser av hvorvidt tempotrening genererer bedre utholdenhet i arbeid med ferdigheter enn presisjonstrening. Deltaker hadde en høyere



frekvens på gloser fra tempobetingelse, noe som innebærer at deltaker fikk flere repetisjoner på gloser fra tempobetingelse før den siste bevaringstesten (2 mnd). Dette endrer imidlertid ikke konkluderende kommentarer knyttet til bevaring av ferdigheter, ettersom denne ekstra treningen vil ha vært til fordel for tempobetingelse.

Samtlige deltakere brukte lengre tid på tempobetingelse enn presisjonsbetingelse. Den mest plausible årsaken til dette er trolig feedbackprosedyren som ble gjennomført i etterkant av tempobetingelsen. Det gjøres imidlertid oppmerksom på at forskjellene i tidsbruk ikke kan tas til inntekt for presisjonstrening eller 'Discrete Trial Teaching'. Grunnen til forskjellene har sammenheng med hvordan eksperimentene ble arrangert for å kunne kontrollere for antall trials og feedback. I en tradisjonell presisjonsopplæringssetting ville det trolig kun blitt øvd på uriktige responser i etterkant av timede økter. Dette ville ha redusert tiden benyttet på tempobetingelsen betraktelig.

Forskjellig tidsbruk på de to betingelsene kan imidlertid ha påvirket resultatet. Ettersom det ble brukt mer tid på tempobetingelsen, har deltaker blitt visuelt eksponert for stimulusmateriellet over lenger tid enn i presisjonsbetingelsen, noe som *kan* ha ført til raskere læring i tempobetingelsen.

### **Avsluttende merknader**

Det ble gjennomført bevaringstester på ferdigheter trent i en av de to betingelsene, tempo eller presisjon. Resultatene viser liten forskjell på bevaring etter en og to måneder for deltaker 1 og 2. Deltaker 3 viser bedre bevaring av ferdigheter trent i presisjonsbetingelse.

Funnene har empirisk, konseptuell og klinisk verdi. Det er ikke ønskelig med prematur spredning av opplæringsmetoder som ikke har tilstrekkelig empirisk støtte. Eksempelvis har det blir argumentert for at ratebygging bør bli inkludert i pensum for atferdsanalytisk opplæring av barn med autisme (Kubina & Wolfe, 2005). Resultatene tyder på at det må

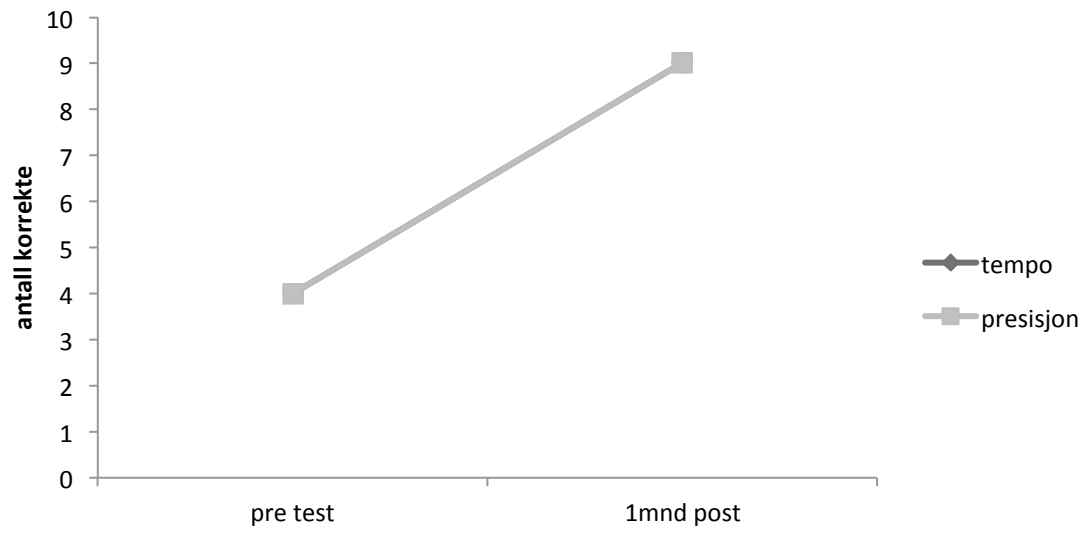
forskes mer på sammenheng mellom frekvens og bevaring før slike strategier implementeres i allerede godt dokumenterte opplæringsprogram.

Konseptuelt er det viktig at atferdsanalyse som vitenskap henger sammen. Ifølge Baer, Wolf, og Risley (1968) bør opplæringsstrategier forankres i grunnleggende atferdsprinsipper. På denne måte vil anvendt atferdsanalyse fremstå som en sammenhengende disiplin, ikke bare en samling av ulike teknikker. Hvorvidt frekvens er en dimensjon ved atferd som påvirker bevaring av ferdigheter, er således av konseptuell interesse.

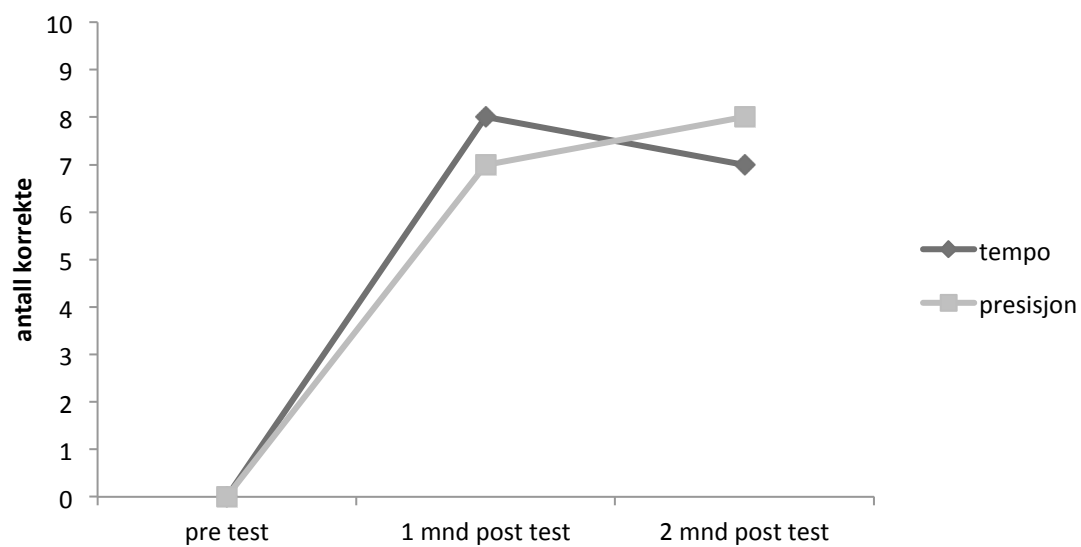
Klinisk er det viktig å evaluere hvorvidt det er kritisk å nå et frekvensmål, ettersom denne innsatsen vil være både tid – og ressurskrevende (Heinicke et al., 2010). Om frekvens ikke er en kontrollerende variabel for bevaring av ferdigheter, kan denne ekstra treningen i stedet rettes inn mot etablering av andre viktige ferdigheter. Dette vil bidra til mer effektiv opplæring.

## Referanser

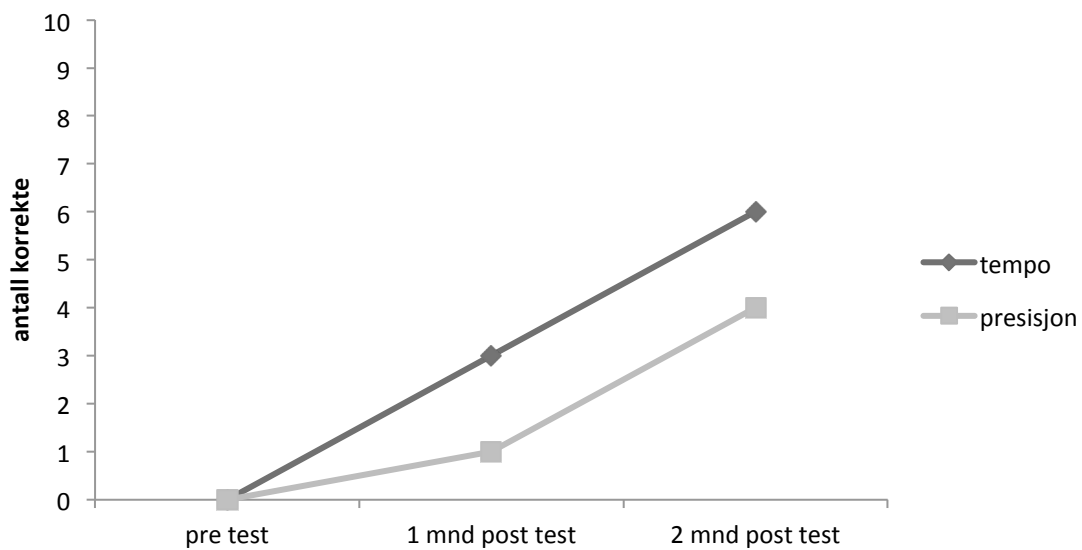
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97. doi: 10.1901/jaba.1968.1-91.
- Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19*(2), 163-197.
- Binder, C., Haughton, E., & Bateman, B. (2005). *Fluency: Achieving true mastery in the learning process*. University of Virginia: Professional Papers in Special Education.
- Bucklin, B. R., Dickinson, A. M., & Brethower, D. M. (2000). A comparison of the effects of fluency training and accuracy training on application and retention. *Performance Improvement Quarterly, 13*(3), 140-163. doi: 10.1111/j.1937-8327.2000.tb00180.x
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2 ed.). New Jersey: Pearson Education.
- Doughty, S. S., Chase, P. N., & O'Shields, E. M. (2004). Effects of rate building on fluent performance: A review and commentary. *The Behavior Analyst, 27*(1), 7-23.
- Fox, E. J., & Ghezzi, P. M. (2003). Effects of computer-based fluency training on concept formation. *Journal of Behavioral Education, 12*(1), 1-21. doi: 10.1023/A:1022373304577
- Hanratty, T., & Greer, R. D. (2000). *Is it rate or number of learn units that makes a difference? A comparison of two performance criteria*. Unpublished manuscript, Teachers College Columbia University, New York.
- Heinicke, M. R., Carr, J. E., LeBlanc, L. A., & Severtson, J. M. (2010). On the use of fluency training in the behavioral treatment of autism: A commentary. *The Behavior Analyst, 33*(2), 223-229.
- Holding, E., Bray, M. A., & Kehle, T. J. (2011). Does speed matter? A comparison of the effectiveness of fluency and discrete trial training for teaching noun labels to children with autism. *Psychology in the Schools, 48*(2), 166-183. doi: 10.1002/pits.20535
- Johnson, K. R., & Layng, J. (1996). On terms. *The Behavior Analyst, 19*(2), 281-288.
- Kelly, R. L. (1995). *A Functional analysis of the effects of mastery and fluency on maintenance*. (Unpublished doctoral dissertation), Columbia, New York.
- Kubina, R. M., & Wolfe, P. (2005). Potential applications of behavioral fluency for students with autism. *Exceptionality: A Special Education Journal, 13*(1), 35-44. doi: 10.1207/s15327035ex1301\_5
- Lindsley, O., R. (1992). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis, 25*(1), 51-57. doi: 10.1901/jaba.1992.25-51
- Mathews, T. (2010). *Teaching time telling and examining the relative effects of rate-building and rate-controlled practice on the retention and generalization of the time cues*. (Unpublished master's thesis), University of Waikato, Hamilton.
- Olander, C. P., Collins, D. L., McArthur, B. L., Watts, R. O., & McDade, C. E. (1986). Retention among college students: a comparison of traditional versus precision teaching. *Journal of Precision Teaching, VI*(4), 80-82.
- Peladeau, N., Forget, J., & Gagne, F. (2003). Effect of paced and unpaced practice on skill application and retention: How much is enough? *American Educational Research Journal 40*(3), 769-801. doi: 10.3102/00028312040003769
- Shirley, M. J., & Pennypacker, H. S. (1994). The effects of performance criteria on learning and retention of spelling words. *Journal of Precision Teaching, XII*(1), 73-86.
- Singer-Dudek, J., & Greer, R. D. (2005). A long-term analysis of the relationship between fluency and the training and maintenance of complex math skills. *The Psychological Record, 55*(3), 361-376.



Figur 1. Antall mestrede responser på bevaringstest for de to betingelsene i eksperiment 1.

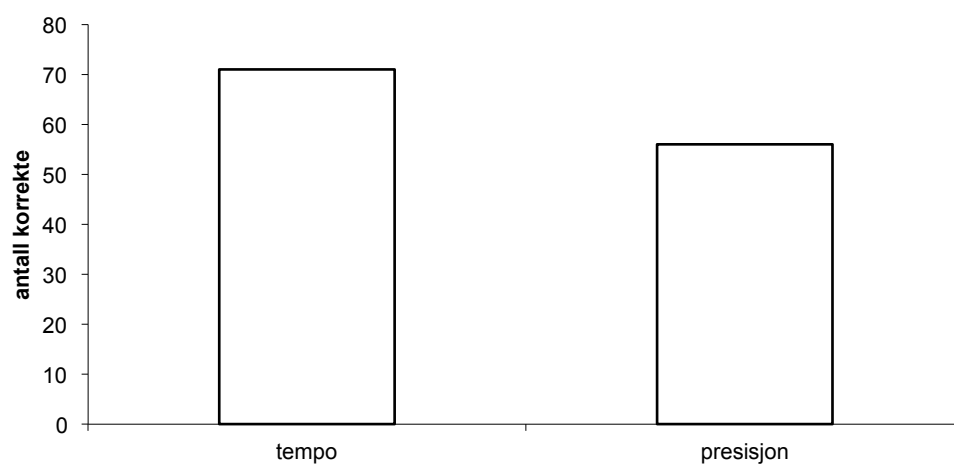


*Figur 2.* Antall mestrede responser på bevaringstest for begge betingelser i eksperiment 2.

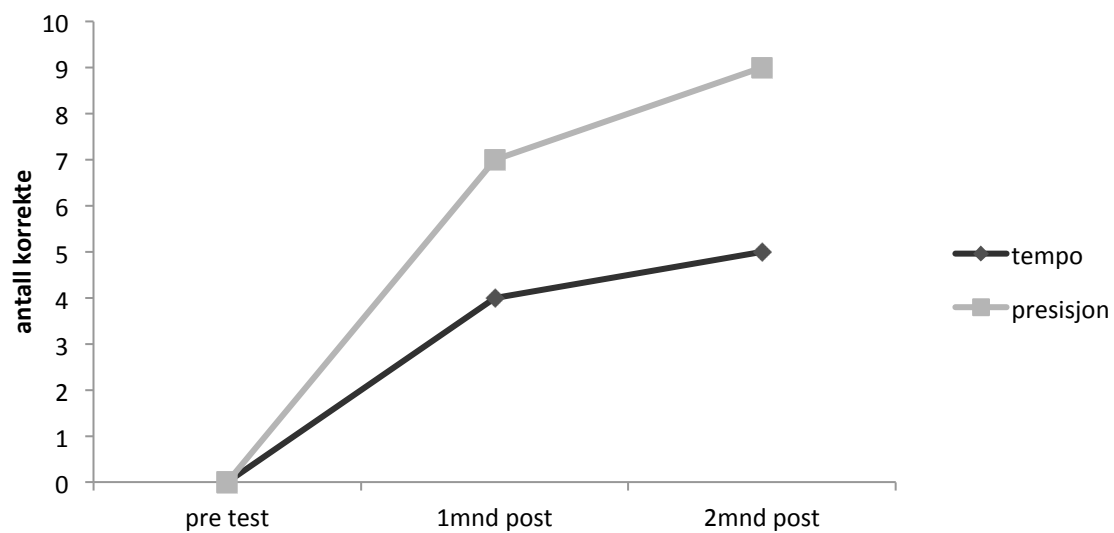


Figur 3. Antall mestrede responser på generaliseringstest for de to betingelsene i eksperiment

2.

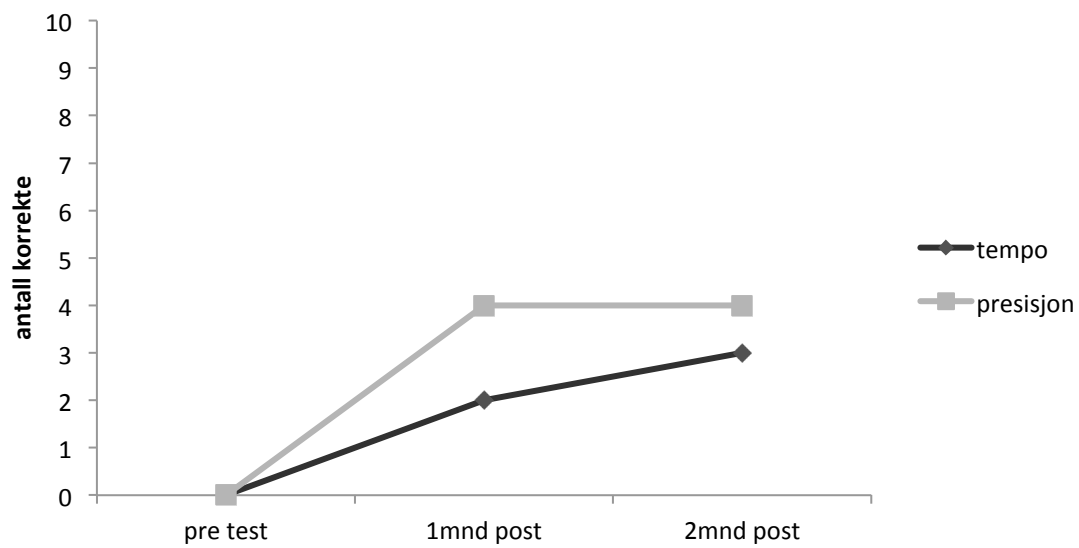


*Figur 4.* Antall mestrede responser på utholdenhetstest for de to betingelsene i eksperiment 2.



Figur 5. Antall korrekte responser på bevaringstest for begge betingelser i eksperiment 3.





Figur 6. Antall korrekte responser på generaliseringstest for begge betingelser i eksperiment

3.

## Appendix A. Stimulsmateriell presisjonsbetingelse eksperiment 1.

|                 |           |            |                |
|-----------------|-----------|------------|----------------|
| <b>Hanske</b>   | Glove     | Pipe       | Counter attack |
| <b>Kork</b>     | Cap       | Bin        | Can            |
| <b>Panser</b>   | Napkin    | Bonnet     | Office         |
| <b>Ramme</b>    | Frame     | Thumb      | Cap            |
| <b>Torsk</b>    | Bonnet    | Can        | Cod            |
| <b>Bosskurv</b> | Cutlery   | Bonnet     | Bin            |
| <b>Svamp</b>    | Frame     | Spunge     | Salmon         |
| <b>Knapp</b>    | Frame     | Button     | Post           |
| <b>Rør</b>      | Spunge    | Cabin      | Pipe           |
| <b>Tilskuer</b> | Spectator | Substitute | Button         |

## Appendix B. Stimulusmateriell tempobetingelse eksperiment 1.

|                    |         |         |         |
|--------------------|---------|---------|---------|
| <b>Hytte</b>       | Frame   | Cap     | Cabin   |
| <b>Laks</b>        | Lid     | Spunge  | Salmon  |
| <b>Bestikk</b>     | Cutlery | Napkin  | Thumb   |
| <b>Gardin</b>      | Cod     | Cabin   | Curtain |
| <b>Vifte</b>       | Bin     | Fan     | Lid     |
| <b>Tverrlegger</b> | Spunge  | Post    | Bar     |
| <b>Serviett</b>    | Bonnet  | Napkin  | Cutlery |
| <b>Kjeks</b>       | Bar     | Salmon  | Biscuit |
| <b>Lokk</b>        | Lid     | Frame   | Cod     |
| <b>Nål</b>         | Needle  | Cutlery | Cabin   |

## Appendix C. Pretest eksperiment 1.

|                    |                |            |           |
|--------------------|----------------|------------|-----------|
| <b>Kontring</b>    | Counter attack | Bar        | Napkin    |
| <b>Tilskuer</b>    | Substitute     | Fan        | Spectator |
| <b>Vifte</b>       | Bin            | Office     | Fan       |
| <b>Rør</b>         | Spunge         | Pipe       | Glove     |
| <b>Genser</b>      | Frame          | Sweater    | Can       |
| <b>Stolpe</b>      | Post           | Spectator  | Bonnet    |
| <b>Tverrlegger</b> | Post           | Spunge     | Bar       |
| <b>Innbytter</b>   | Cutlery        | Substitute | Cap       |
| <b>Hanske</b>      | Counter attack | Pipe       | Glove     |
| <b>Svamp</b>       | Spunge         | Frame      | Cabin     |
| <b>Bestikk</b>     | Thumb          | Napkin     | Cutlery   |
| <b>Boks</b>        | Sweater        | Can        | Spunge    |

|                 |            |                |         |
|-----------------|------------|----------------|---------|
|                 |            |                |         |
| <b>Bosskurv</b> | Can        | Bin            | Bonnet  |
|                 |            |                |         |
| <b>Ramme</b>    | Frame      | Fan            | Post    |
|                 |            |                |         |
| <b>Kork</b>     | Cap        | Bin            | Can     |
|                 |            |                |         |
| <b>Tommel</b>   | Substitute | Bar            | Thumb   |
|                 |            |                |         |
| <b>Panser</b>   | Bar        | Bonnet         | Napkin  |
|                 |            |                |         |
| <b>Hytte</b>    | Glove      | Cap            | Cabin   |
|                 |            |                |         |
| <b>Kontor</b>   | Office     | Cabin          | Cutlery |
|                 |            |                |         |
| <b>Serviett</b> | Napkin     | Counter attack | Bonnet  |
|                 |            |                |         |
| <b>Panne</b>    | Forhead    | Lid            | Salmon  |
|                 |            |                |         |
| <b>Støv</b>     | Cod        | Spunge         | Dust    |
|                 |            |                |         |
| <b>Torsk</b>    | Can        | Cod            | Bonnet  |
|                 |            |                |         |
| <b>Knapp</b>    | Frame      | Button         | Post    |
|                 |            |                |         |

|                   |         |         |         |
|-------------------|---------|---------|---------|
| <b>Laks</b>       | Lid     | Spunge  | Salmon  |
| <b>Kjeks</b>      | Salmon  | Bar     | Biscuit |
| <b>Nål</b>        | Cutlery | Needle  | Cabin   |
| <b>Gardin</b>     | Curtain | Cod     | Cabin   |
| <b>Sykepleier</b> | Nurse   | Forhead | Cutlery |
| <b>Lokk</b>       | Cod     | Frame   | Lid     |

Appendix D. Pretest 1 (generaliseringstest) eksperiment 2.

hus =

mais =

kylling =

tannpirker =

saus =

penn =

sko =

saks =

tallerken =

kniv =

gaffel =

skap =

seng =

katt =

hund =

serviett =

parasoll =

t-skjorte =

pute =

genser =

håndkle =

gris =

skje =

bok =

stol =

bord =

kjøleskap =

vindu =

toalett =

vann =

## Appendix E. Pretest 2 og posttest 2 (bevaringstest) multiple choice eksperiment 2.

hus  
mais  
kylling  
tannpirker  
saus  
penn  
sko  
saks  
tallerken  
kniv  
gaffel  
skap  
seng  
katt  
hund  
serviett  
parasoll  
t-skjorte  
pute  
genser  
håndkle  
gris  
skje  
bok  
stol  
bord  
kjøleskap  
vindu  
toalett  
vann

mojo  
plato  
chompa  
camiseta  
casa  
cerdo  
agua  
toalla  
callo  
tenedor  
cuchillo  
almohada  
mesa  
silla  
ventana  
frigorífico  
cuchara  
mondadientes  
armario  
gallina  
tijeras  
pluma  
sombrella  
aseo  
gato  
cama  
servilleta  
perro  
libro  
zapatillo



## Appendix F. Stimulusmateriell presisjonsbetingelse eksperiment 2.

|                   |              |              |              |
|-------------------|--------------|--------------|--------------|
| <b>Bord</b>       | Callo        | Libro        | Mesa         |
| <b>Mais</b>       | Cerdo        | Mesa         | Callo        |
| <b>Skap</b>       | Armario      | Cuchara      | Ventana      |
| <b>Bok</b>        | Callo        | Libro        | Cerdo        |
| <b>Gris</b>       | Libro        | Cerdo        | Toalla       |
| <b>Håndkle</b>    | Toalla       | Mesa         | Armario      |
| <b>Vindu</b>      | Armario      | Mondadientes | Ventana      |
| <b>Sko</b>        | Zapatico     | Ventana      | Mondadientes |
| <b>Skje</b>       | Toalla       | Cuchara      | Zapatico     |
| <b>Tannpirker</b> | Mondadientes | Zapatico     | Cuchara      |

## Appendix G. Stimulusmateriell tempobetingelse eksperiment 2.

|                  |          |          |          |
|------------------|----------|----------|----------|
| <b>Toalett</b>   | Aseo     | Perro    | Silla    |
| <b>Penn</b>      | Cama     | Aseo     | Pluma    |
| <b>Saks</b>      | Tenedor  | Gallina  | Tijeras  |
| <b>Hund</b>      | Cama     | Perro    | Silla    |
| <b>Stol</b>      | Pluma    | Chompa   | Silla    |
| <b>Genser</b>    | Chompa   | Tijeras  | Pluma    |
| <b>Gaffel</b>    | Tenedor  | Camiseta | Gallina  |
| <b>T-skjorte</b> | Tijeras  | Tenedor  | Camiseta |
| <b>Kylling</b>   | Camiseta | Gallina  | Chompa   |
| <b>Seng</b>      | Perro    | Cama     | Aseo     |

Appendix H. Skriftlig posttest (generaliseringstest) eksperiment 2.

mais =

saks =

skap =

saus =

kylling =

gaffel =

penn =

håndkle =

vindu =

hund =

seng =

gris =

sko =

pute =

katt =

skje =

genser =

stol =

toalett =

bord =

Appendix I. Posttest 1 (generaliseringstest) eksperiment 3.

Mais =

Penn =

Sko =

Saks =

Gaffel =

Skap =

Seng =

T-skjorte =

Genser =

Håndkle =

Gris =

Bok =

Stol =

Bord =

Vindu =

Toalett =

Kylling =

Skje =

Tannpirker =

Hund =

## Appendix J. Stimulusmateriell presisjonsbetingelse eksperiment 3.

|                   |              |              |              |
|-------------------|--------------|--------------|--------------|
| <b>Bord</b>       | Callo        | Libro        | Mesa         |
| <b>Mais</b>       | Cerdo        | Mesa         | Callo        |
| <b>Skap</b>       | Armario      | Cuchara      | Ventana      |
| <b>Bok</b>        | Callo        | Libro        | Cerdo        |
| <b>Gris</b>       | Libro        | Cerdo        | Toalla       |
| <b>Håndkle</b>    | Toalla       | Mesa         | Armario      |
| <b>Vindu</b>      | Armario      | Mondadientes | Ventana      |
| <b>Sko</b>        | Zapatico     | Ventana      | Mondadientes |
| <b>Skje</b>       | Toalla       | Cuchara      | Zapatico     |
| <b>Tannpirker</b> | Mondadientes | Zapatico     | Cuchara      |

## Appendix K. Stimulusmateriell tempobetingelse eksperiment 3.

|                  |          |          |          |
|------------------|----------|----------|----------|
| <b>Toalett</b>   | Aseo     | Perro    | Silla    |
| <b>Penn</b>      | Cama     | Aseo     | Pluma    |
| <b>Saks</b>      | Tenedor  | Gallina  | Tijeras  |
| <b>Hund</b>      | Cama     | Perro    | Silla    |
| <b>Stol</b>      | Pluma    | Chompa   | Silla    |
| <b>Genser</b>    | Chompa   | Tijeras  | Pluma    |
| <b>Gaffel</b>    | Tenedor  | Camiseta | Gallina  |
| <b>T-skjorte</b> | Tijeras  | Tenedor  | Camiseta |
| <b>Kylling</b>   | Camiseta | Gallina  | Chompa   |
| <b>Seng</b>      | Perro    | Cama     | Aseo     |

