

# Måleegenskaper ved et verktøy for måling av akseptabilitet og målingsrutiner i praksisfeltet. En oppdatering

Børge Strømgren<sup>1</sup>, Jon Løkke<sup>2</sup> og Stian Orm<sup>3</sup>

<sup>1</sup>OsloMet – storbyuniversitetet, <sup>2</sup>Høgskolen i Østfold, <sup>3</sup>Sykehuset Innlandet

**Innledning.** Sosial validitet og underdimensjonen akseptabilitet er viktig i anvendt atferdsanalyse og atferdsanalyse i praksis, og er knyttet til en av de sju dimensjonene i anvendt atferdsanalyse, den anvendte dimensjonen. Akseptabilitet er en viktig del av atferdsanalysens estetikk og omhandler normative vurderinger og atferdsanalysens anseelse. Vi har tidligere gjennomført en pilotstudie på måleegenskapene ved en skala for akseptabilitet som antydte god strukturvaliditet og test-retest reliabilitet. I denne studien undersøker vi strukturvaliditet til skalaen i et større og mer heterogent utvalg av ulike yrkesgrupper fra ulike tjenester og med ytterligere en klinisk vignett. **Metode.** Studenter ved høyskoler og deltakere på konferanser ble invitert til å delta gjennom en lenke til anonyme spørreskjema via Nettskjema. 175 besvarte hele eller deler av spørreskjemaene. **Resultater.** Konfirmerende faktoranalyse understøtter faktorstrukturen funnet i pilotstudien, alle faktorladninger var over .925. Likt som ved en tidligere undersøkelse skilte skalaen på vignetter med antatt høy og lav grad av akseptabilitet, det var ikke signifikante forskjeller mellom skalaskårene i denne studien og i den tidligere med samme skala og vignetter. Videre korrelerte utdanningsnivå og rapportert bakgrunnskunnskap om akseptabilitet, moderat for utforming av tiltak ( $r_s = .48$ ), og sterk for igangsetting ( $r_s = .50$ ) og evaluering av tiltak ( $r_s = .52$ ).

**Diskusjon.** Skalaen viser god strukturvaliditet og egner seg for måling av akseptabilitet. Fremtidige studier bør fokusere på *anvendelse og implementering av skalaen i klinisk praksis*.

**Nøkkelord:** sosial validitet, akseptabilitet, skala, normative premisser, faktoranalyse, strukturvaliditet

## Measurement properties of a tool for measuring acceptability and measurement routines in the field of practice. An update

**Introduction.** Social validity and its sub-dimension acceptability is important in applied behavior analysis and behavior analysis in practice. Acceptability is tied to behavior analysis in practice and to one of the seven dimensions in applied behavior analysis—the applied dimension. Acceptability is an important part of the esthetics of behavior analysis and deal with normative judgment and the standing of behavior analysis. We have previously carried out a pilot study on the measurement properties of a scale for acceptability which indicated good structure validity and test-retest reliability. In the current study we further investigate the structure validity of the scale with a larger and more heterogenous sample with different occupational groups from different human services. We also added a vignette in addition to the two from the pilot study.

**Method.** University students and participants at conferences were invited to participate thru a link to a anonymous questionnaire at Nettskjema. 175 replied to all on some parts of the questionnaire

Det er ingen konflikter med hensyn til dette manuskriptet, og forfatterne er enige om at det skal sendes til fagfelleurdert i Norsk Tidsskrift for Atferdsanalyse (NTA).

Setteredaktør har vært Tore Vignes, VID vitenskapelige høyskole.

Vi takker to ukjente fagfeller for nyttige bidrag til revidering av manuskriptet.

Henvendelser kan rettes til Børge Strømgren, OsloMet – storbyuniversitetet, [bstromgr@oslomet.no](mailto:bstromgr@oslomet.no)

**Results.** A confirmatory factor analysis supports the factor structure found in the pilot study. All factor loadings were over .925. As in the pilot study the scale separated vignettes with perceived high acceptability from the one with perceived low acceptability. There was no difference between the scale scores in the current study and the pilot study for the vignettes that were used in both studies. Level of education correlated moderately with reported background knowledge on acceptability ( $r_s = .48$ ) and strongly with implementation ( $r_s = .50$ ) and evaluation ( $r_s = .52$ ) of interventions.

**Discussion.** The scale shows good structure validity and is suitable for measuring acceptability. Future studies should focus on application and implementation of the scale in practice.

*Key words:* social validity, acceptability, scale, normative premises, factor analysis, structure validity

---

De siste årene har det vært økt oppmerksomhet rundt sosial validitet og akseptabilitet i atferdsanalytisk forskning og praksis (Eilifsen & Vandbakk, 2021; Løkke et al., 2020; Strømgren et al., 2020). Økt oppmerksomhet rundt sosial validitet og akseptabilitet kan antyde en normativ vending i det norske atferdsanalytiske miljøet (Salthe & Løkke, 2013). Normative vurderinger må hentes inn ved å spørre folk, og bruk av spørreskjemaer og statistiske analyser av spørreskjemaene, er den dominerende metoden. Det har tradisjonelt sett vært en del kritikk mot statistikk i det atferdsanalytiske miljøet (Arntzen & Løkke, 2015).

Baer (1977) argumenterte for at de endringene eller effektene som atferdsanalytikere, og egentlig alle andre klinikere, bør være opptatt av, kan sees ved visuell inspeksjon, og at statistiske analyser kan gjøre at små og klinisk ubetydelige effekter fremstår kraftfulle og verdt å skrive om. Det kan hende poengene til Baer er gode når det gjelder effektevaluering, men når det gjelder beskrivelser av hvordan folk opplever ting, så trengs det både indirekte målinger og statistiske analyser for å understøtte validitet og reliabilitet i målingene (Strømgren et al., 2020). Noe moderasjon av standpunktene om visuell inspeksjon og statistikk kan leses i Schwartz & Baer (1991): «Researchers should consider basic rules of test construction and statistics when developing social validity assessments» (s. 198).

Direkte målinger er observasjon av atferd, mens utgangspunktet for indirekte målinger er observasjon av verbal atferd om fenomenet som er av interesse. En ting er å observere folks angstanfall, mens en annen ting er å høre folk fortelle om det. I artikkelen er vi opptatt av indirekte målinger og statistiske analyser for å understøtte validiteten og reliabiliteten til en skala for måling av akseptabilitet (Strømgren et al.) (Appendix A).

Sosial validitet referer til hvorvidt et tiltak inkludert fremgangsmåtene, målene og effektene er akseptable slik det vurderes av personer i omgivelsene (Strømgren et al., 2020). Sammen med de to evalueringsområdene, evaluering av effekt og tilfredshet med mål for endring, er akseptabilitet ett av tre evalueringsområder for tiltak (Løkke & Salthe, 2012). I 1968 ble sosial validitet befestet som en sentral dimensjon ved anvendt atferdsanalyse (Baer et al., 1968). Sosial validitet er et viktig evalueringsområde for anvendt atferdsanalyse og atferdsanalyse i praksis, fordi terapeuten skal bidra til å løse problemer som er viktige for mottakeren og/eller omgivelsene (Baer et al.; Schwartz & Baer, 1991; Wolf, 1978).

Innenfor atferdsanalyse har det vært lite tradisjon for å måle subjektivt anliggende fenomener (Baer et al., 1968, 1987; Strømgren et al., 2020). Begrunnelse er at det er forskjell på selve fenomenet og snakking (verbal atferd) om fenomenet (Karlan & Rusch, 1982). Vi bruker selvskading og

snakking om selvskading som eksempel. Verbal atferd om selvskading, om det er positivt eller negativt og så videre, er under helt andre forsterkningsbetingelser enn selve atferden. Hvis en person forteller at hen tenker på å ta livet sitt, så kan vi som terapeuter innnta to ulike perspektiver på det som blir sagt. Det første perspektivet er substansielt og fokuserer på *semantikken, altså innholdet i det som blir sagt*. Det andre perspektivet er mer radikalbehavioristisk eller kontekstuel og fokuserer på *funksjonen til det som blir sagt*; eksempelvis hvorfor fortelles det nå, i denne konteksten, til meg.

Når vi er ute etter å måle subjektive fenomener, så har vi ikke lengre mulighet for å etablere eksperimentell kontroll over fenomenet, og målene knyttet til prediksjon og kontroll er vanskelig å oppnå. Samtidig er det viktig å presisere at det er fullt mulig å gjennomføre funksjonelle analyser av snakking og andre subjektive fenomener. Vi må ty til indirekte mål som selvrporterer for å få tak i folks opplevelser og meninger. Foreløpig er den mest systematiske målemetoden for måling av akseptabilitet å spørre folk ved bruk av standardiserte spørreskjemaer (Strømgren et al., 2020). Det er viktig at spørreskjemaene inneholder utsagn som virkelig dreier seg om målene med intervensjonen og teknikkene som brukes. Vi benytter komplekse statistiske analyser for å finne ut om de ulike utsagnene i et spørreskjema måler det samme fenomenet, for eksempel akseptabilitet ved teknikkene som brukes i en intervensjon. Det analysene ikke gir svar på er hva utsagnene måler. Hva sosial validitet er for noe og hvilke utsagn som kan måle fenomenet er noe som må diskuteres i fagfeltet, men det ligger utenfor formålet med artikkelen.

Flere ulike litteraturgjennomganger har funnet at sosial validitet er stemoderlig behandlet innenfor anvendt atferdsanalyse, kun 3 til 13% av artikler publisert i atferdsanalytiske tidsskrift inkluderer evalueringer av sosial validitet (Carr et al., 1999; Ferguson et al., 2019; Løkke et al., 2020). For eksempel

fant Ferguson et al. (2019) at kun 12% av 1209 artikler publisert i *Journal of Applied Behavior Analysis* (JABA) i perioden 1999 til 2016 inkluderte evalueringer av sosial validitet. Løkke et al. (2020) fant at kun 3% av artiklene publisert i *Norsk Tidsskrift for Atferdsanalyse* (NTA) i kategorien anvendt atferdsanalyse og atferdsanalyse i praksis mellom 2006 og 2020 inkluderte målinger av sosial validitet. Funnene antyder at den store majoriteten av atferdsanalytiske studier ikke inkluderer evaluering av sosial validitet, og dermed er det vanskelig å opprettholde at det er en sentral dimensjon i anvendt atferdsanalyse. Det kan antyde at det er for få lett tilgjengelige verktøy for måling av akseptabilitet innenfor anvendt atferdsanalyse og atferdsanalyse i praksis.

Når sosial validitet eller underdimensjonen akseptabilitet måles i atferdsanalytiske studier, måles det i hovedsak gjennom spørreskjemaer (Carter, 2010; Common & Lane, 2017; Miltenberger, 1990). For eksempel fant Spear et al. (2013) i en litteraturgjennomgang av måling av sosial validitet i single-subjekt studier av tiltak for elever med emosjonelle- eller atferdsvansker at 16 av 22 studier benyttet spørreskjema for måling av sosial validitet. Vi gjennomførte i 2020 en pilotstudie av et spørreskjema for måling av akseptabilitet (Strømgren et al., 2020). I Strømgren et al. (2020) sin pilotstudie gikk vi gjennom en del tidligere forskning på akseptabilitet med ulike skaler og utvikling av slike (Se for eksempel Carter, 2010; Common & Lane, 2017; Kazdin, 1980a; Kazdin, 1980b; Miltenberger, 1990). Vi henviser til Strømgren et al. 2020 for nærmere gjennomgang. Vi anbefaler også Eilifsen og Vandbakk (2021).

Spørreskjemaet vi benyttet i pilotstudien (Strømgren et al., 2020) ble oversatt og bearbeidet fra Eilers og Hayes (2015). Skjemaet er kort og konsist for å være mest mulig anvendbart i praksis. I pilotstudien fant vi at spørreskjemaet hadde god strukturvaliditet med én faktor, god indre konsistens og god test-retest reliabilitet, samt evne til å

diskriminere mellom tiltak med antatt god og dårlig akseptabilitet. I pilotstudien ble det pekt på behov for å gjennomføre en større studie med flere og mer heterogene utvalg av deltakere og flere kaser (Strømgren et al.).

Formålet med denne studien er å undersøke skalaene validitet med et større og mer heterogent utvalg og flere kaser enn det som ble gjort i Strømgren et al. (2020). Dette kan bidra til å etablere noe større sikkerhet om spørreskjemaet måler det det er tiltenkt å måle (akseptabilitet). I studien undersøker vi *strukturvaliditet* gjennom konfirmerende faktoranalyse, for å se at alle utsagnene i skjemaet måler det samme fenomenet, nemlig akseptabilitet av tiltak.

Videre ser vi på om skåringen i denne undersøkelsen av skalaen ligner skåringene fra forrige undersøkelse (Strømgren et al., 2020) med samme vignetter (Appendix C), og vi ser på om en ny vignett (Appendix B) som er antatt å ha god akseptabilitet skårer noenlunde likt som den tidligere brukte vignetten som var antatt å ha god akseptabilitet, en form for *innholdsvaliditet*. Vi undersøker også reliabilitet og sammenhengen mellom fagpersoners utdanningsnivå og kjennskap til akseptabilitet og sosial validitet.

## Metode

### Deltakere og prosedyre

Et utvalg på 171 personer ble invitert til å svare på en anonym nettbasert spørreundersøkelse via Nettskjema. Invitasjoner ble distribuert gjennom informasjonskanaler på universiteter/høgskole (OsloMet og Høgskolen i Østfold) som tilbyr bachelor- eller masterutdanninger vi anså som relevante. I tillegg ble invitasjoner annonsert på fagkonferanser vi anså som relevante (SOR-konferansen og NAFOs årsmøteseminar).

Nettskjemaet inneholdt demografiske spørsmål (alder, utdanningsnivå, profesjon og arbeidssted og antall år personen har arbeidet i yrket). Så spurte vi om personens egen oppfatning av sin bakgrunnskunnskap om sosial validitet og akseptabilitet. Videre

spurte vi om rutiner for vurdering av akseptabilitet ved utforming, igangsettelse og evaluering av tiltak. Siste del av undersøkelsen besto av tre vignetter/tiltaksbeskrivelser med en skala for vurdering av akseptabilitet.

Personene ble i invitasjonen informert om at det var helt frivillig å delta, at svarene var helt anonyme-ikke mulig å identifisere, og at de kunne avslutte *når som helst*. Denne informasjonen ble gjentatt i selve spørreskjemaet i tillegg til kontaktopplysninger for prosjektleder og hvordan opplysningen vil bli brukt. For å komme videre i skjemaet måtte personen trykke på samtykke-knapp.

### Design

Vi utviklet en ny klinisk vignett i tillegg til de to som vi brukte i Strømgren et al. (2020). Vignett «RRB» var ment å illustrere et reduksjonstiltak med antatt lav akseptabilitet, mens «ANG» var ment å illustrere et angstbehandlingstiltak antatt høy akseptabilitet. Den nye vignetten «STOTT» var ment å illustrere et tiltak med antatt høy akseptabilitet.

Invitasjonene inneholdt lenker til nettskjemaet og de inviterte fikk lenker til ulike skjemaer med hensyn på rekkefølgen av vignetter – en rekke er for eksempel ABC. Det var fire ulike konstellasjoner av vignetter; skjema ABC hadde vignetterrekkefølge «ANG», «RRB» og «STOTT», skjema BCA hadde «RRB», «STOTT» og «ANG», skjema CBA hadde «STOTT», «RRB» og «ANG», og endelig hadde skjema CAB «STOTT», «ANG» og «RRB». På samme måte som i Strømgren et al. (2020) søkte vi slik å utjevne eventuelle overføringseffekter fra den ene vignetten til den andre.

### Materiell

Rutiner for vurdering av akseptabilitet. Vi benyttet tre spørsmål til å finne ut hvorvidt deltakerne hadde rutiner for vurdering av akseptabilitet på arbeidsplassen sin: (1) i forbindelse med utarbeiding av tiltak, (2) i forbindelse med igangsetting av tiltak, og (3) i forbindelse med evaluering av tiltak.

Spørsmålene ble vurdert på en skala fra 1 – aldri til 10 – alltid.

**Skala for akseptabilitet.** Vi benyttet samme skala som i Strømgren et al. (2020). Kort oppsummert består den av 8 utsagn som vurderes på en skala fra 1 (Helt uenig) til 5 (Helt enig). Skalaen var i Strømgren et al. vurdert til å ha god faktorstruktur (faktorkorrelasjoner fra .80 til .92) og reliabilitet (test-retest:  $r_s = .79$ , og indre konsistens:  $\alpha = .97$ ).

### Statistiske analyser

Som i Strømgren et al. (2020) ble skårer fra alle vignetter ble analysert samlet, altså stablet opp på hverandre før analyse. All dataanalyse er gjort med SPSS 27 om ikke annet er angitt.

**Manglende data.** Analyser ble gjennomført med valgene «missing listwise» eller «missing pairwise». Dermed ble kun «cases» med ingen manglende data ble med i analysene, i SPSS kalt «valid cases», heretter kun kalt cases. Når det ikke mangler data, er antallet cases 171.

**Konfirmerende faktoranalyse (CFA).** I Strømgren et al. (2020) gjennomførte vi en eksplorative faktoranalyse (EFA). I denne studien vil vi se om denne kan styrkes ved å gjennomføre en konfirmerende faktoranalyse (CFA) på samme skala, men med nytt datasett fra «nye» personer som har skåret. Slik kan vi «bekrefte» en latent underliggende faktor. Analysen vil gjennomføres med statistikkprogrammet JASP 0.16.1 (JASP-Team, 2022) som bruker R-pakken lavaan (Rosseel, 2012), fordi den kan estimere ordinalskaler med *Diagonally Weighted Least Squares* (DWLS).

I tillegg til at skalaen ved en CFA bør gi faktorkorrelasjoner på samme nivå som ved EFA, vil vi også se på om data passer modellen med en faktor, såkalte «fit indices». Vi vil se på følgende indikatorer: a) Absolutte: *Normed Chi-Square* (NC =  $\chi^2/df$ : bør være  $\leq .3$  til  $\leq .5$ ), *Root Mean Square Error of Approximation* (RMSEA: bør være  $\leq .06$ , 90 % CI = [ $\approx .0$ ,  $\leq .08$ ]), og *Standardized Root Mean Square Residual* (SRMSR: bør være  $\leq .08$ ),

og b) Inkrementelle: *Comparative Fit Index* (CFI: bør være  $\geq .93$ ) og *Tucker-Lewis Index* (TLI: bør være  $\geq .93$ ).

**Korrelasjoner.** Vi vil se på sammenhengen mellom utdanningsnivå (videregående, universitet/høgskole opp til 4 år og universitet/høgskole over 4 år) og selvrappert grad av bakgrunnskunnskap om akseptabilitet og validitet. På grunn av ordinalskala for variabelen utdanningsnivå rapporterer vi Spearmans korrelasjon,  $r_s$ . Vi vil også se på sammenhengen mellom utdanningsnivå og rapportert forekomst av vurdering av akseptabilitet og sosial validitet i utarbeiding, igangsetting og evaluering av tiltak. Forekomsten ble rapportert på en skala fra 1 (Aldri) til 10 (Alltid).

### Rutiner for vurdering.

Vi spurte om hvilke rutiner deltakerne hadde for vurdering av sosial validitet og akseptabilitet (de som hadde vurdert).

## Resultater

### Deltakere

Vi fikk in 175 svar, fordelt på ulike utgaver av spørreskjemaene der vignettene var presentert i ulik rekkefølge. Deltakernes alder var fra 18 til 68 år,  $M = 39.10$  og  $SD = 11.54$ . De hadde vært i nåværende yrke i fra under ett til 38 år,  $M = 9.3$ ,  $SD = 8.03$ . For øvrige opplysninger, se Tabell 1.

### CFA

Resultatene er vist i Tabell 2. Det er gjennomgående høye faktorladninger og smale konfidensintervall for alle utsagnene, utsagn 5 har lavest faktorladning med oppunder .9.

«Fit indices». To av tre absolutte indikatorer viste «good fit»: *Normed Chi-Square* (NC) = 4.91, mens *Root Mean Square Error of Approximation* (RMSEA) = .09, 90% CI [.07, .11] (burde vært .06, 90 % CI = [ $\approx .0$ ,  $\leq .08$ ]), og *Standardized Root Mean Square Residual* (SRMSR) = .02. De inkrementelle indikerte også «good fit»: *Comparative Fit Index* (CFI) = 1.0, og *Tucker-Lewis Index*

Tabell 1. Oversikt over personenes utdanningsnivå, profesjon og type tjenester de yter.

<b>Utdanningsnivå</b>	<b>N</b>	<b>%</b>
Videregående	4	2.3
Universitet/h tom 4 år	91	53.2
Universitet/h. over 4 år	76	44.4
<b>Profesjon</b>	<b>N</b>	<b>%</b>
Vernepleier	119	69.6
Sosionom	5	2.9
Barnevernspedagog	6	3.5
Pedagog/spesialpedagog	18	10.5
Assistent/miljøarbeider	10	5.8
Psykolog/Psykologiviter	7	4.1
Atferdsanalytiker/læringspsyk	3	1.8
<b>Type tjeneste</b>	<b>N</b>	<b>%</b>
Bolig/hjemmetjenester funksj. neds.	70	40.9
Bolig/hjemmetjenester psyk. lid.	8	4.7
Avlastningsbolig	2	1.2
Skole	35	20.5
Barnehage	6	3.5
Psykisk helsevern for voksne	4	2.3
Psykisk helsevern for barn og unge	2	1.2
Habiliteringstjenesten for voksne	12	7.0
Habiliteringstjenesten for barn og unge	4	2.3
Annen spesialisthelsetjeneste	4	2.3
Barnevernsinstitusjon	3	1.8
Helse- og sosialforvaltning	2	1.2
Annet	19	11.1
<b>Annet</b>	<b>N</b>	<b>%</b>
Adm i privat helseforetak/privat firma div./privat omsorg div.	3	1.8
Brukerstyrt personlig assistent	1	.6
Dagtilbud for mennesker med utviklingshemming	1	.6
Demensomsorg/sykehjem	4	2.4
Fag og kvalitetskonsulent/Selvstendig konsulent	2	1.2
Har ikke noe jobb for øyeblikket	1	.6
Hydro / arbeidsledig	2	1.2
Bolig for personer med ervervet hjerneskade	1	.6
Pedagogisk fagsenter	1	.6
Pedagogisk psykologisk tjeneste/fagsenter	3	1.8
Universitet/høgskole	2	1.2

(TLI) = 1.0. Samlet sett synes data å passe modellen med en faktor. Dette indikerer at skalaen måler akseptabilitet på en konsistent måte.

### Validitet

Vi fant forskjeller i gjennomsnittsskåre og standardavvik mellom vignett med antatt lav akseptabilitet (RRB) og de to vignettene med antatt høy akseptabilitet (ANG og STOTT).

Tabell 2. Faktorladninger for hvert utsagn i skalaen, men nedre og øvre konfidensintervall.

Utsagn i skala	Estimat	95 % KI Nedre	95 % KI Øvre
1. Tiltaket er en akseptabel måte å endre deltakerens atferd	0.954	0.947	0.962
2. Jeg er villig til å bruke tiltaket for å endre deltakerens atferd	0.965	0.958	0.972
3. Jeg liker fremgangsmåten som brukes i tiltaket	0.968	0.961	0.975
4. Jeg mener at tiltaket vil være effektivt for å endre deltakerens atferd	0.925	0.916	0.935
5. Jeg tror tiltaket sannsynligvis fører til varige forbedringer	0.899	0.888	0.910
6. Jeg mener at tiltaket vil være fordelaktig for deltakeren	0.945	0.938	0.953
7. Jeg mener at fremgangsmåtene som blir brukt i tiltaket er velbegrunnede	0.935	0.926	0.943
8. Samlet sett synes jeg tiltaket er positivt	0.981	0.975	0.988

Note. Hvert utsagn vurderes på en Likertskala fra Svært enig = 5, Enig = 4, Nøytral = 3, Uenig = 1, Svært uenig = 1.

For RRB var  $M = 2.53$ ,  $SD = .99$ ,  $SE = .08$ , for ANG var  $M = 4.16$ ,  $SD = .84$ ,  $SE = .20$ , og for STOTT var  $M = 4.22$ ,  $SD = .72$ ,  $SE = .06$ . Det var signifikant forskjell mellom ANG og RRB,  $t(170) = 16.23$ ,  $p < .001$ , og mellom STOTT og RRB,  $t(170) = 19.28$ ,  $p < .001$ , men ikke mellom ANG og STOTT,  $t(170) = -0.83$ ,  $p = .41$ .

Som i Strømgren et al. (2020) antyder analysen at skalaen oppførte seg som forventet, de to vignettene med antatt høy akseptabilitet hadde en høyere gjennomsnittlig total skalaskåre enn vignetten med antatt lav akseptabilitet, med henholdsvis  $N = 171$ ,  $M = 33.23$ ,  $SD = 6.72$  for ANG,  $N = 169$ ,  $M = 33.57$ ,  $SD = 5.73$  for STOTT og  $N = 167$ ,  $M = 20.12$ ,  $SD = 8.89$  for RRB. Forskjellen mellom ANG i dette datasettet og ANG i Strømgren et al. sitt T2-datasett var ikke signifikant,  $t(226) = -1.46$ ,  $p = .15$ , likedan var heller ikke forskjellen mellom RRB i dette datasettet og RRB i Strømgren et al. sitt T2-datasett signifikant,  $t(222) = 0.16$ ,  $p = .99$ .

Dette antyder at skalaen kan måle grad av akseptabilitet og at høyere samlet skåre indikere høyere grad av akseptabilitet. Videre

ser det ut til at skalaen måler likt med både studenter og ferdig utdannede og slik sett er solid med ulike utdanningsnivåer eller erfaringsbakgrunn.

### Korrelasjoner

Det var en svak, men signifikant korrelasjon ( $r_s = .25$  Bca 95 % CI [.10, .40],  $p = .001$ ) mellom utdanningsnivå og rapportert bakgrunnskunnskap om akseptabilitet og sosial validitet. Når det gjelder utdanningsnivå og rapportert forekomst av å ha vurdert sosial validitet og akseptabilitet, var sammenhengen moderat for utforming av tiltak ( $r_s = .48$  Bca 95 % CI [.34, .61],  $p < .001$ ), og sterk for igangsetting av tiltak ( $r_s = .50$  Bca 95 % CI [.37, .62],  $p < .001$ ) og for evaluering av tiltak ( $r_s = .52$  Bca 95 % CI [.38, .65],  $p < .001$ ). Dette tyder på at kunnskap om sosial validitet og akseptabilitet øker med at økt utdanningsnivå.

### Rutiner for vurdering

Rapporterte rutiner for vurdering er oppgitt i Tabell 3. Det er tydeligvis diskusjoner i personalgruppen som er mest benyttede rutine i tillegg til samtaler med

Tabell 3. Rutiner for vurdering av sosial akseptabilitet og validitet i utforming, igangsetting og evaluering av tiltak.

<b>Rutine</b>	<b>N</b>	<b>%</b>
Ingen faste rutiner	63	36.8
Vi snakker med pårørende	63	38.0
Vi diskuterer i personalgruppen/tiltaksgruppen	117	68.4
Vi benytter et spørreskjema til bruker/pårørende	10	5.8
Vi intervjuer relevante personer i omgivelsene	14	8.2
Vi vurderer sosial akseptabilitet som et eget punkt i tiltaksbeskrivelsen	25	14.6
Vi benytter en standardisert skala for vurdering av sosial akseptabilitet som fylles ut av bruker/pårørende	7	4.1
Annet	2	1.2
<b>Rutine annet</b>	<b>N</b>	<b>%</b>
Sjekklisten til Løkke og Salthe (2012) som utgangspunkt	1	0.6
SME (Senter for Medisinsk Etikk) + Sjekklisten til Løkke og Salthe (2012) som utgangspunkt	1	0.6

pårørende. Mange har ingen rutiner. Det er noe forekomst av vurderinger innskrevet i tiltaksbeskrivelser, noe som vel er et krav dersom tiltaket har elementer av makt og tvang. Ellers er det noe forekomst av spørreskjema, standardisert skala og intervjuer. En har oppgitt å bruke sjekklisten til Salthe og Løkke og en har oppgitt å bruke en rutine fra Senter for Medisinsk Etikk i tillegg.

## Diskusjon

Formålet med studien var å understøtte validiteten til en kort og tilgjengelig skala for måling av akseptabilitet av atferdsanalytiske tiltak, og dermed stimulere til måling av akseptabilitet i atferdsanalytisk praksis. Vi gjennomførte en større studie med et mer heterogent utvalg og flere kasus, for å bidra til å etablere større sikkerhet rundt hvorvidt spørreskjemaet måler det det er tiltenkt å måle. Resultatene innebærer at spørsmålene i spørreskjemaet måler det samme fenomenet, som antas å være akseptabilitet. Videre fant vi at hele 37% av deltakerne ikke hadde faste rutiner for vurdering av sosial validitet og akseptabilitet ved tiltak. Ikke overraskende, var høyere utdanningsnivå korrelert med høyere selvrapportert kunnskap om sosial validitet og rutiner for å vurdere akseptabi-

litet på arbeidsplassen. Cirka 70% av deltakerne var vernepleiere og rundt 50% jobbet i tjenester for personer med funksjonsnedsettelse. Resultatene kan indikere at høyere andel høyskoleutdannede og vernepleiere i tjenester for personer med funksjonsnedsettelse er viktig for å sikre kvaliteten i tjenestene, men dette er usikkert fordi vi ikke har oversikt over responsrate. Samplingmetoden vi brukte, å vise en lenke til Nettskjema og invitere til å delta i undersøkelse, gir oss ikke anledning til å si noe sikkert om generaliteten av funnene. Vi kan ikke si at utvalget er representativt, og dermed ikke at de fleste med høyskoleutdanning og vernepleierutdanning har god, eller bedre kunnskap enn andre, om sosial validitet og akseptabilitet.

Begrepsvaliditet omhandler hvorvidt vi måler det konstruert som vi ønsker å måle (Løkke & Løkke, 2014). Det finnes uendelig mange operasjonaliseringer av psykologiske konstrukt slik som akseptabilitet, og å måle alle sider av fenomenet er ikke mulig, og heller ikke ønskelig. Å ha som intensjon å måle alle sider av et fenomen, er en form for essensialisme som vil stoppe videre forskning og kunnskapsutvikling (Stanovich, 2013). Vi må gjøre et utvalg og finne ut om utvalget vårt (av spørsmål eller atferder) operasjonaliserer det samme konstruert, og ikke ulike



konstrukt. Faktoranalyse bidrar til å underbygge at utvalget vårt operasjonaliserer det samme konstruktet. Med andre ord styrkes begrepsvaliditeten.

I forrige studie (Strømgren et al., 2020) undersøkte vi reliabiliteten til skalaen for måling av akseptabilitet og fant at den var god, mens vi i denne studien har fokusert på strukturvaliditet som bidrar til å styrke skalaens begrepsvaliditet. I denne studien hadde vi en ekstra vignett og et mer heterogent utvalg. At de statistiske analysene støttet den samme faktorstrukturen som vi fant i forrige studie antyder at utsagnene på skalaen representerer konstruktet akseptabilitet, og kan generaliseres til ulike populasjoner og situasjoner.

En styrke ved den foreliggende studien er et heterogent utvalg der mange ulike yrkesgrupper og tjenester er representert. Videre er det gjennomført etablerte statistiske analyser som støtter konklusjonene i artikkelen og vi la til en vignett med et ikke-atferdsanalytisk tiltak for å se om målingene kunne generaliseres til nye situasjoner. En begrensning ved studien er at veldig mange av deltakerne hadde høyere utdanning, så det mer uklart om målingene lar seg generalisere til yrkesgrupper uten høyere utdanning. Det er en svakhet siden en stor andel av de som er ansatt i tjenestene mangler høyere utdanning, og det er derfor viktig å få kunnskap om målingene kan generaliseres til den gruppen.

Samlet antyder funnene i denne og forrige studie (Strømgren et al., 2020) at skalaen måler akseptabilitet med god reliabilitet over tid. Skalaen er kort og enkel å ta i bruk, men vi kjenner ikke til betingelsene som skal til for at den skal brukes i praksis. Det trengs studier som undersøker implementering av skalaen i klinisk praksis, og som undersøker måleegenskapene når skalaen brukes i kliniske kasus og ikke med oppdiktede vignetter. Videre trengs det studier av brukertilfredsheten som undersøker hva tjenesteytere, klienter og pårørende synes om skalaen.

## Referanser

- Arntzen, E. & Løkke, J. A. (2015). Visuelle analyser av data—er det greit å ikke vite alt? *Norsk Tidsskrift for Atferdsanalyse*, 42(2), 97–105. <http://www.nta.atferd.no/journalissue.aspx?IdDocument=558>
- Baer, D. M. (1977). “Perhaps it would be better not to know everything”. *Journal of Applied Behavior Analysis*, 10(1), 167–172. <https://doi.org/10.1901/jaba.1977.10-167>
- Baer, D. M., Wolf, M. M. & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91–97. <https://doi.org/10.1901/jaba.1968.1-91>
- Baer, D. M., Wolf, M. M. & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 20(4), 313–327. <https://doi.org/10.1901/jaba.1987.20-313>
- Carr, J. E., Austin, J. L., Britton, L. N., Kellum, K. K. & Bailey, J. S. (1999). An assessment of social validity trends in applied behavior analysis. *Behavioral Interventions*, 14(4), 223–231. [https://doi.org/10.1002/\(sici\)1099-078x\(199910/12\)14:4<223::Aid-bin37>3.0.Co;2-y](https://doi.org/10.1002/(sici)1099-078x(199910/12)14:4<223::Aid-bin37>3.0.Co;2-y)
- Carter, S. L. (2010). Chapter 3 - Instruments for Evaluating Social Validity. I S. L. Carter (Red.), *The Social Validity Manual* (s. 37–67). Academic Press. <https://doi.org/10.1016/B978-0-12-374897-3.00003-9>
- Common, E. A. & Lane, K. L. (2017). 4 - Social Validity Assessment. I J. K. Luiselli (Red.), *Applied Behavior Analysis Advanced Guidebook* (s. 73–92). Academic Press. <https://doi.org/10.1016/B978-0-12-811122-2.00004-8>
- Eilers, H. J. & Hayes, S. C. (2015). Exposure and response prevention therapy with cognitive defusion exercises to reduce repetitive and restrictive behaviors displayed by children with autism

- spectrum disorder. *Research in Autism Spectrum Disorders*, 19, 18–31. <https://doi.org/10.1016/j.rasd.2014.12.014>
- Eilifsen, C. & Vandbakk, M. (2021). Sosial validitet og anvendt atferdsanalyse. *Norsk Tidsskrift for Atferdsanalyse*, 48(1), 47–62. <http://www.nta.atferd.no/journalissue.aspx?IdDocument=873>
- Ferguson, J. L., Cihon, J. H., Leaf, J. B., Van Meter, S. M., McEachin, J. & Leaf, R. (2019). Assessment of social validity trends in the journal of applied behavior analysis. *European Journal of Behavior Analysis*, 20(1), 146–157. <https://doi.org/10.1080/15021149.2018.1534771>
- JASP–Team. (2022). *JASP (Version 0.16.1) [Computer software]*. <https://jasp-stats.org/>
- Karlan, G. R. & Rusch, F. R. (1982). Correspondence between saying and doing: some thoughts on defining correspondence and future directions for application. *Journal of Applied Behavior Analysis*, 15(1), 151–162. <https://doi.org/https://doi.org/10.1901/jaba.1982.15-151>
- Kazdin, A. E. (1980a). Acceptability of alternative treatments for deviant child behavior. *Journal of Applied Behavior Analysis*, 13(2), 259–273. <https://doi.org/10.1901/jaba.1980.13-259>
- Kazdin, A. E. (1980b). Acceptability of time out from reinforcement procedures for disruptive child behavior. *Behavior Therapy*, 11(3), 329–344. [https://doi.org/10.1016/S0005-7894\(80\)80050-5](https://doi.org/10.1016/S0005-7894(80)80050-5)
- Løkke, J. A., Orm, S. & Strømgren, B. (2020). Kort rapport: Måling av sosial validitet i studier publisert i Norsk Tidsskrift for Atferdsanalyse. *Norsk Tidsskrift for Atferdsanalyse*, 47(1), 1–4. <https://nta.atferd.no/journalissue.aspx?IdDocument=819>
- Løkke, J. A. & Salthe, G. (2012). Sjekklister for målrettet tiltaksarbeid: fra normative og deskriptive premisser til tiltak og evaluering. *Norsk Tidsskrift for Atferdsanalyse*, 39(1), 17–32. <https://nta.atferd.no/journalissue.aspx?IdDocument=308>
- Miltenberger, R. G. (1990). Assessment of Treatment Acceptability: A Review of the Literature. *Topics in Early Childhood Special Education*, 10(3), 24–38. <https://doi.org/10.1177/027112149001000304>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Salthe, G. & Løkke, J. A. (2013). Om begreper. Normativt. *Norsk Tidsskrift for Atferdsanalyse*, 40(2), 209–215. <http://www.nta.atferd.no/journalissue.aspx?IdDocument=443>
- Schwartz, I. S. & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis*, 24(2), 189–204. <https://doi.org/10.1901/jaba.1991.24-189>
- Spear, C. F., Strickland-Cohen, M. K., Romer, N. & Albin, R. W. (2013). An examination of social validity within single-case research with students with emotional and behavioral disorders. *Remedial and Special Education*, 34(6), 357–370. <https://doi.org/10.1177/0741932513490809>
- Stanovich, K. E. (2013). *How to think straight about psychology* (10. utg.). Pearson.
- Strømgren, B., Løkke, J. A. & Orm, S. (2020). Måleegenskaper ved en skala for sosial akseptabilitet. En pilotstudie. *Norsk Tidsskrift for Atferdsanalyse*, 47(2), 63–79. <https://nta.atferd.no/journalissue.aspx?IdDocument=834>
- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11(2), 203–214. <https://doi.org/10.1901/jaba.1978.11-203>

## Appendiks A

### Bedømming av et tiltaks anseelse (akseptabilitet)

Oversatt og revidert fra Eilers & Hayes (*Research in Autism Spectrum Disorders*, 2015, s. 25)

av Jon Arne Løkke, Stian Orm & Børge Strømgren, 2020

	Svært enig	Enig	Nøytral	Uenig	Svært uenig
Tiltaket er en akseptabel måte å endre deltakerens atferd					
Jeg er villig til å bruke tiltaket for å endre deltakerens atferd					
Jeg liker fremgangsmåten som brukes i tiltaket					
Jeg mener at tiltaket vil være effektivt for å endre deltakerens atferd					
Jeg tror tiltaket sannsynligvis fører til varige forbedringer					
Jeg mener at tiltaket vil være fordelaktig for deltakeren					
Jeg mener at fremgangsmåtene som blir brukt i tiltaket er velbegrunnede					
Samlet sett synes jeg tiltaket er positivt					

Appendix B – Ny vignettVignett 3 (STOTT): Støttekontakt

Problem: Stine (11 år) er enebarn og bor fast hos mor. Mor har en psykisk lidelse som gjør at hun i lange perioder kan være svært passiv og isolerer seg hjemme. Dette går utover Stine som i stor grad blir hjemme med mor og i liten grad får deltatt på fritidsaktiviteter og annet utenfor hjemmet. Barnevernet er koblet på saken.

Tiltak: Barnevernet tenker at Stine trenger å komme seg ut av hjemmet og delta på fritidsaktiviteter sammen med jevnaldrende, spesielt i de periodene mor er syk. Det blir satt inn en støttekontakt 4 timer i uken for å ta med Stine på fritidsaktiviteter.

Effekt: Stine er nå ute 4 timer i uken sammen med støttekontakten. De drar på riding sammen og går på fritidsklubben sammen med andre jenter på Stine sin alder. Skolen opplever at Stine er gladere enn tidligere og hun er blitt bedre på å konsentrere seg på skolen. Når barnevernet snakker med Stine forteller hun at hun syntes det er veldig fint å være sammen med støttekontakten.

Begrunnelse: Gode relasjoner, kontakt med jevnaldrende og deltakelse i fritidsaktiviteter er viktig for barn og unges utvikling og velvære. Støttekontakt er et tiltak som kan settes inn for barn i familier med lite nettverk og som trenger flere opplevelser. Målet er at en støttekontakt skal bidra til en mer aktiv fritid, sosiale opplevelser og fungere som en positiv rollemodell.

**Appendiks C - Vigneter brukt i Strømgen et al. (2020).**Vignett 1 (RRB) : Redusere repeterende atferd

Problem: Kristian (13 år) har stor interesse for og bruker mye tid på å se på YouTube videoer av en amerikansk komiker, og blir sterkt påvirket av ting som sies av komikeren. Kristian tar det som sies bokstavelig, så når komikeren sier at juice er farlig, fører det til at Kristian ikke drikker juice på lang tid.

Tiltak: For å redusere Kristians tid på YouTube og interesse for den amerikanske komikeren, så skal personalet og foreldrene ikke snakke om eller kommentere komikeren. Og dersom Kristian tar det opp eller snakker om det, skal det ignoreres, og personalet skal ta opp et annet samtaletema. Tiltaket er detaljert beskrevet i en tiltaksbeskrivelse.

Effekt: Kristian bruker like mye tid på YouTube og komikeren, men han snakker ikke om komikeren med hverken personalet eller foreldrene. I starten maste Kristian mye for å få andre til å respondere på hans snakking om komikeren, men det har gradvis gitt seg.

Begrunnelse: Ekstinksjon (ignorering) benyttes ofte som tiltak når man ønsker å redusere forekomsten av en atferd hos barn, unge og voksne som har atferd som er utfordrende for omgivelsene.

Vignett 2 (ANG): Trening på å sitte alene med døren igjen

Problem: Kristian (13 år) er redd for å være alene i et rom med døren igjen. Det fører til at Kristian er avhengig av at en voksen er i nærheten av ham, hele tiden. Det går utover søsken fordi foreldrene må bruke mye tid med Kristian.

Tiltak: Kristian skal øve på å sitte alene med døren igjen. Treningen foregår ved gradvis eksponering styrt av tilbakemeldinger fra Kristian om hvor mye ubehag han føler ved å sitte alene med døren igjen. Når Kristian ikke føler ubehag trappes tiden oppover.

Effekt: Kristian tør stadig å sitte lengre tid alene med døren igjen. Det tok litt tid før Kristian ga tilbakemelding om at han kunne gå fra det første trinnet i eksponeringen til det andre, men deretter gikk det jevnt fremover.

Begrunnelse: Gradvis eksponering er en evidensbasert teknikk for å redusere angst og engstelse hos barn, unge og voksne som har angst, fobier eller er redd for ting.