



METOD- OCH PROCESSUTVECKLING

PUBLICERAD: 27 MARS 2024

NEDLADDAD: 5 APRIL 2026

Single Case Experimental Design

Innehåll

1. Bakgrund	3
2. Syfte	4
2.1 Avgränsning	4
3. Resultat	5
3.1 Single Case studier	5
3.2 SCED	6
3.3 Litteratur om guidelines, rekommendationer och bedömningsmallar	20
4. Referenser	23

Observera att det är möjligt att ladda ner hela eller delar av en publikation.
Denna pdf/utskrift behöver därför inte vara komplett. Hela publikationen och
den senaste versionen hittar ni på www.sbu.se/374

1. Bakgrund

Med anledning av en fråga från Socialstyrelsen kring vilka studiedesigntyper som kan vara rimliga att inkludera när det gäller vetenskapligt kunskapsunderlag i nationella riktlinjer startade SBU ett internt arbetsprojekt (Dnr. 2021/563) kring studiedesigner som inkluderar ett eller ett fåtal personer eller fall. I mars 2022 bestämdes att SBU skulle fokusera på att öka kunskapen inom myndigheten när det gäller studiedesignen single case experimental design (SCED) samt analyser baserade på data från sådana studier.

2. Syfte

Beskrivning av

- studiedesignen Single Case Experimental Design,
- hur analys baserad på resultat från primärstudier med Single Case Experimental Design kan se ut,
- de verktyg och principer som finns för att bedöma tillförlitligheten av resultat från primärstudier med Single Case Experimental Design.

2.1 Avgränsning

Beskrivning av studiedesignerna fallstudier och fallserier¹, samt beskrivning av metaanalytiska metoder som utgår från resultat från primärstudier med Single Case Experimental Design exkluderas i denna delrapport. Detsamma gäller hur GRADE [1] anser att man bör kommunicera och använda sig av resultat med mycket låg tillförlitlighet.

¹ En fallstudie innebär en presentation av tillstånd, symptom, kliniska och praktiska fynd, behandling och uppföljning av en personuell patient eller insats till enskild person och uppföljning av denna. Fallstudien baserar resultat på observation, inte nödvändigtvis upprepad sådan, av *naturligt* förekommande händelser och är också många gånger deskriptiva till sin karaktär (medan resultat som baseras på SCED ofta beskrivs numeriskt och grafiskt [5]). Fallstudier har ofta använts för att uppmärksamma nya eller sällsynta sjukdomar och tillstånd eller för att studera sociala fenomen. De är även viktiga för information om olika typer av biverkningar eller oönskade utfall. Fallstudier är inte kontrollerade men kan användas till att generera hypoteser som sedan undersöks i kontrollerade studier. En fallserie är en rapport av flera fallstudier.

3. Resultat

3.1 Single Case studier

3.1.1 Analytisk enhet

Single Case-studier är primärstudier där det metodologiska tillvägagångssättet är att prospektivt studera ett enskilt fall eller liten grupp av fall över tid. Den grundläggande analysenheten är således det enskilda fallet, vilket kan vara en enskild person, en organisatorisk eller samhällelig enhet, liksom en enskild händelse. Härefter används huvudsakligen benämningen fall. En Single Case-studie kan även baseras på fler fall än ett, även om det sällan är fler än tre [2].

3.1.1.1 Olika typer och benämningar

Det finns både experimentella och icke-experimentella Single Case-studier [3] [4].

Icke-experimentella, eller kvasiexperimentella, Single Case-studier [3] [4] består av en bakgrundsmätning av utfallet i fråga (baslinje, *eng.* baseline) (A), följt av insats och mätning av utfall under tiden insatsen pågår (B). Detta är den enklaste studiedesignen (AB) och kallas även “phase change without reversal” och skiljer sig från en före-och eftermätningstudie genom att mätning av utfallet sker under tiden insatsen pågår [4]. Denna design medger inga slutsatser av kausal natur, däremot kan den vara användbar för att pröva om en insats är genomförbar [5].

Experimentella Single Case-studier baseras på upprepade faser av mätningar och i vissa designer införande respektive borttagande av en insats (exv. ABA, ABAB). Det finns olika benämningar på experimentella Single Case-studier [2]:

- Single Case Experimental design
- Single Subject Experimental Design
- Single Subject Research Design
- N-of-1-studies
- N-of-1 trials [4]
- ‘personalized (N-of-1) trials’ [3]
- Small-N-design
- Multiple Case Design

- Single Case Design
- Single Systems designs

Projektgruppen har valt att utgå ifrån benämningen Single Case Experimental Design (SCED), eftersom den tydliggör att det är ett eller ett fåtal (single) fall (case) som studeras och därmed att det kan vara både person, enhet eller händelse, samt att det är en studie där en planerad insats aktivt införs (experimental) under flera faser (minst ABA).

3.2 SCED

3.2.1 Historik

Den metodologiska ansatsen SCED utvecklades under 1920-60 talen inom psykologi och beteendeforskning. Idag används den inom flera områden såsom medicin, folkhälsa, och utbildning [3]. Inom det medicinska området, uppstod den så kallade N-of-1 trial i mitten av 1980-talet, vilket motsvarar SCED [6]. Ansatsen N-of-1 trial utvecklades på grund av de begränsningar som RCT upplevdes ha när det gällde att applicera resultat till en enskild person och fatta beslut om behandling för denna [7] [8]. N-of-1 trial bedömdes vara användbar för att identifiera en optimal insats för en person, såsom exempelvis behandling av sällsynta sjukdomar.

What Works Clearinghouse² inkluderar sedan 2020 Single case-studier som en av tre standardmetoder i deras standardhandbok, de övriga två är RCT och RDD (regression discontinuity design) [9].

Vad gäller användning i nutid av SCED visade en undersökning från 2008 avseende 21 tidskrifter inom utbildning och psykologi att 44 % av studier om insatser i dessa tidskrifter hade använt SCED [10] [11].

² What Works Clearinghouse (WWC) finansieras av Institute of Education Sciences (IES) inom det amerikanska utbildningsdepartementet som etablerades 2002. WWC:s arbete leds av ett team av personal på IES och genomförs i samverkan med företag med expertis inom utbildning, forskningsmetodik och spridning av utbildningsforskning, (informationen hämtad 231003 från <https://ies.ed.gov/ncee/wwc/WhoWeAre>.)

3.2.2 Syfte och frågeställningar i primärstudier med SCED-design

Syftet med primärstudier med SCED är att undersöka om en insats har effekt för ett enskilt fall, såsom en person eller specifikt sammanhang.

SCED-studier kan användas för olika typer av forskningsfrågor, till exempel:

- Är insats X effektiv för tillståndet/beteendet Y hos personen P i sammanhanget S1?
- Är insats X effektiv för tillståndet/beteendet Y hos personen P i sammanhanget S2?
- Är insats X mer effektiv än insats Z för tillståndet/beteendet Y hos personen P i sammanhanget S1?
- Är effekten för tillståndet/ beteendet Y större om man lägger till insats Z till insats X, än vid endast insats X, för personen P i sammanhang S1?

3.2.3 Metodologiskt tillvägagångssätt

I SCED-studier karaktäriseras det metodologiska tillvägagångssättet av att

- ett enskilt fall eller liten grupp fall,
- resultatet mäts ofta och upprepat i flera faser, och att
- studiens genomförande ofta är sekventiellt genom att insatsen som utvärderas först tillämpas för att därefter dras tillbaka [2] [3] eller ett resultat av förväxlingsfaktorer.

I en SCED-studie sker således systematiska, upprepade mätningar av utfallet, det vill säga innan, under och efter aktiv introduktion av utvärderad insats. SCED-studien har på så sätt olika faser med (B) och utan insats (A). Fokus i SCED-studier är reproducering, där tanken är att en eventuell effekt av en insats på ett utfall synliggörs i fler än en mätning (A1A1A1B1B1B1) eller fler än en fas (A1B1A2B2A3). Syftet med dessa upprepade faser och mätningar är att minska risken för att eventuell effekt på ett utfall är slumpmässig [2] [3].

3.2.4 Utfall

Det primära utfallet i en SCED-studie benämns på engelska ”target behaviour”, ”repeated measure” eller ”target variable” [2], och är knutet till det fall som deltar i utvärderingen. Det saknas därför oftast underlag i form av psykometriska eller diagnostiska utvärderingar [2]. Rekommendationen är i stället att undersöka samstämmigheten i mätning av utfall mellan olika observatörer (*eng.* inter-observer agreement on the target behavior) [2].

Det finns också andra utfall som mäts i SCED-studier såsom jämförelsevariabler, generaliserade variabler, implementeringsdata och utfallsdata från etablerade standardiserade test [2]. Jämförelsevariabler är variabler som inte förväntas att förändras till följd av genomförandet av insatsen, medan generaliserande variabler är variabler som inte bedöms vara det primära utfallet men kan tänkas bli påverkat som en generell effekt av insatsen. Implementeringsdata är data som berör metodtrohet, det vill säga i hur stor grad den undersökta insatsen genomförs i enlighet med hur den är

tänkt att genomföras. Utfallsdata från etablerade standardiserade test kan användas för att visa effekten av insatsen bortom det primära utfallsmåttet i förhållande till en mer allmän eller mer välkänt bedömningsverktyg [2].

3.2.5 Termer och begrepp

I Tabell 3.1 görs en sammanställning av vanliga termer och begrepp i SCED-studier.

Tabell 3.1 Svenska och engelska termer och begrepp som används i SCED-studier.

Fas (<i>eng. phase</i>)	En definierad observationsperiod där effekten av undersökt insats på studerat utfall mäts eller bedöms upprepade gånger
A	Benämning på baslinjefas/jämförelsefas, dvs. den fas där aktuellt utfall mäts eller observeras för att sedan jämföras med utfall aktiv fas där undersökt insats aktivt förs in (insatsfas)
B, C, D etcetera	Benämning på aktiva faser där en eller olika insatser (B, C, D) införs
A1-B1-A2-B2-A3-C1	En studie som börjar med baslinjefas (A) följt av en aktiv fas där insats B introduceras (B), därefter dras insats B bort och en andra baslinjefas inträder (A), efter detta följer ytterligare en aktiv fas där insats B introduceras en andra gång (B) liksom dras tillbaka en andra gång (baslinjefas) (A), och studien avslutas med en aktiv fas där insats C introduceras
Target behaviour (<i>sv. primärt utfall</i>)	Det utfall i form av exempelvis beteende som primärt mäts och utvärderas i en studie. Kallas också Dependent variable eller Target variable
Target variable (<i>sv. primärt utfall</i>)	Samma som Target behaviour
Dependent variable (<i>sv. primärt utfall</i>)	Samma som Target behaviour
Independent variable	Insats som undersöks
Intervention	Insats som undersöks. (<i>eng. Independent variable</i>)
Experimental control (<i>sv. experimentell kontroll</i>)	Aktiv, avsiktlig och systematisk introduktion av en insats, behandling, förfarande eller program med efterföljande observation/mätning av utfallsresultat
Intern validitet	Den omfattning i vilken den insats som införs är orsaken till förändringar i studerat utfall
Extern validitet	Den omfattning i vilken undersökt insats och dess effekt på studerat utfall kan generaliseras till andra fall
Level (<i>sv. nivå</i>)	Datavärdet (oftast medelvärde eller median) inom en specifik fas
Trend	Lutningen på en straight line of best-fit för data inom en fas
Variability (<i>sv. variabilitet</i>)	Dataintervallet inom en fas eller spridningen av data (t.ex. SD) för trenden

3.2.6 SCED-designer och standarder för dessa

What Works Clearinghouse publicerade 2020 vilka design standarder för SCED-studier som ska vara uppfyllda för att uppnå tillräcklig intern validitet [9]. En effekt, eller avsaknad av sådan, ska enligt dessa standarder kunna påvisas vid minst tre olika faser (ABA), liksom vid minst tre (fem är att föredra) upprepade mätningar av utfall i varje fas (AAABBBAAA). Detta för att nå lägre sannolikhet för att en eventuell effekt beror av något annat än den undersökta insatsen. Förutsättningar för en studiedesign som omfattar tillbakadragande av en insats i en eller flera faser (ABA, ABAB) är att insatseffekten på utfallet är reversibel [5] liksom att den kan upptäckas i närtid [2]. Om dessa förutsättningar föreligger behöver de också kompletteras med en bedömning om det är etiskt rimligt att dra tillbaka eller avsluta en insats som bedöms ha god effekt för det fall, exempelvis en person, som studeras. Vidare skall varje utfall enligt What Works Clearinghouse standarder mätas systematiskt av mer än en person och samstämmigheten i mätning av utfall mellan olika personer ska beräknas för minst 20 procent av mätpunkterna i varje fas. Samstämmigheten skall uppnå en nivå som satts på förhand, till exempel Cohens κ på minst 0,6. Det finns också kriterier från What Works Clearinghouse som definierar vilken nivå av evidens (strong, moderate eller no evidence) en SCED-studie ger.

Det finns flera sätt att designa en SCED-studie och olika typer av design svarar på olika forskningsfrågor. Gemensamt är att fall studeras under olika faser där baslinjen (A) utgör den utgångspunkt mot vilken uppföljande mätning under och efter introduktion av en insats (B) jämförs. Studiedesigner som anses experimentella och medge en bättre intern validitet samt kunna ge stöd för ett orsak-verkan-samband är Withdrawal/Reversal Design eller ABAB-trial, multiple baseline design, alternating-treatments design, och changing-criterion designs [4]. I var och en av dem finns möjlighet att introducera randomisering i designen, även om detta inte är vanligt förekommande. [4]. Detta i likhet med RCT-studier men där RCT syftar till att fastställa orsakssamband baserat på genomsnittliga resultat på gruppnivå avser primärstudier baserade på dessa SCED-designer att fastställa orsakssamband för en enskilt fall (individnivå) [3].

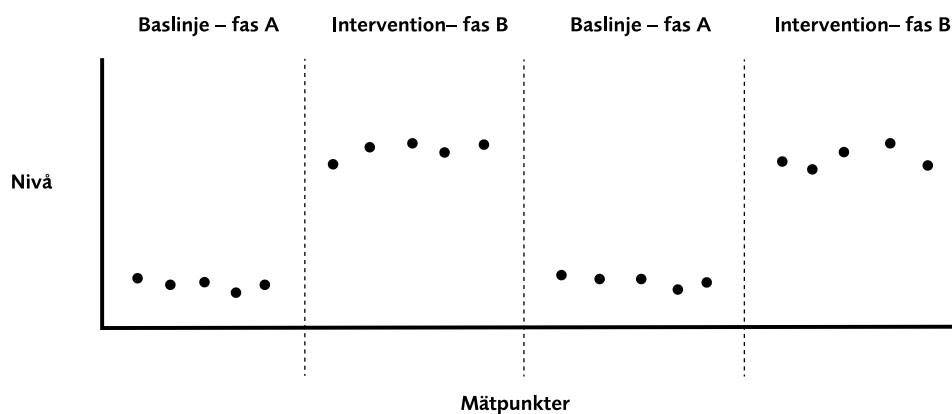
3.2.7 SCED med Withdrawal/Reversal Design (WRD)

I det fall insatsens effekt på utfallet bedöms som reversibel, upptäckbar i närtid och etiskt rimlig att dra tillbaka är en design av typen ABA eller ABAB möjlig att använda [5]. ABA innebär en baslinjemätning av det primära utfallet (A1) följt av insats och mätning av utfallet (B1), och därefter tillbakadragande av insats och därpå följande mätning av utfallet en andra gång (A2). I ABAB-designen införs insatsen en andra gång (B2), med syfte att nå reproducering, med därpå följande mätning av utfallet.

Reproducering anses ge starkare stöd för insatsens effekt [5]. Dessa upprepade faser innebär att det fall, exempelvis en person, som ingår i studien fungerar som sin egen jämförelse [3]. Det är viktigt att värdena i baslinjefasen stabiliserats innan man går vidare med en insatsfas. Det krävs som ovan nämnt minst tre datapunkter för att kunna säga att stabilitet har uppnåtts och det bör inte finnas en trend som pekar på att deltagaren förbättras, eftersom det då blir svårt att avgöra om en eventuell förändring beror på insatsen. I Figur 3.1 ses en schematisk bild av en WRD-studie med ABAB-design.

N-of-1-Trials är den medicinska benämningen för SCED-studier [4], och är ett exempel på WRD trial design. Det finns en särskild guide, CONSORT Extension for N-of-1 Trials [12] (CENT), framtagen för hur sådana studier skall rapporteras.

Figur 3.1 Schematisk bild av SCED med WRD/ABAB-design. Hypotesen är att insatsen ökar nivån av utfallet. I första fasen (baslinje – A) mäts utfallet vid fem tillfällen. Därefter introduceras insatsen (vid första streckade linjen) och den andra fasen (insatsen – fas B) börjar. Även här mäts utfallet vid fem tillfällen. Därefter tas insatsen bort (vid andra streckade linjen) och den tredje fasen påbörjas (baslinje – A). Man mäter återigen utfallet vid fem tillfällen. Vid tredje streckade linjen introduceras insatsen igen och fjärde fasen börjar. Som i tidigare faser mäts utfallet vid fem tillfällen.



SCED med Alternating-treatments design

Alternating-treatments design (sv. alternerande och samtidiga insatser) innebär att flera insatser utvärderas samtidigt genom att snabbt alternera insatserna [4]. Vanligtvis görs en snabb växling mellan olika insatser vid ett tillfälle, exempelvis tre typer av fotleds-ortoser och dess respektive effekt på gångkapacitet hos en person som prövas vid ett och samma besök hos en arbetsterapeut [2]. Även denna studiedesign innebär att personen utgör sin egen kontroll samt förutsätter att insatseffekten på utfallet är reversibel [5], men också omedelbar [2].

3.2.8 SCED med Multiple baseline design

När en insatseffekt inte bedöms vara reversibel och där insatsens effekt på utfallet är synlig först på sikt [2], eller om insatsen inte bedöms som etiskt rimlig att ta bort efter ett första införande kan multiple-baseline design (MBD) (sv. multipel baslinjedesign) användas [5]. MBD innebär flera samtidiga baslinjedesigner (AB, AB, AB) vilket eliminerar behovet av att återgå till baslinjen (ABA) och MBD är därför lämpad för utvärdering av insatser som inte är reversibla eller har effekter på lång sikt, som till exempel vissa rehabiliteringseffekter. I MBD design introduceras en insats (AB1) vid en viss tidpunkt (t1) (AB1t1) som kompletteras med en förskjutet (t2, t3) introduktion av samma insats (AB1t2, AB1t3) med avseende på en viss parameter, såsom person (p), sammanhang (s) eller beteende (b) [4]. Insatsen introduceras förskjutet och sekventiellt för exempelvis flera olika personer (AB1t1p1, AB1t2p2, AB1t3p3), eller för en person i flera olika sammanhang (AB1t1s1, AB1t2s2, AB1t3s3), eller olika målbeteenden hos en person (AB1t1b1, AB1t2b2, AB1t3b3) [2]. MBD kan därmed ses som flera AB-designer (eller ABC, ABCD), med lika många AB-designer som det finns personer, sammanhang eller beteenden. Idén är att visa att förändring endast sker när insatsen är riktad till personen, genomförs i ett visst sammanhang eller gäller ett visst eller flera specifika beteenden. MBD som avser en insats effekt på olika sorters beteenden hos en person såsom exempelvis effekt av kognitiv minnesträning på beteendet att komma ihåg att ta med nycklar, låsa dörren respektive äta lunch [2]. Personen, sammanhanget eller målbeteendet utgör i detta fall inte sin egen kontroll [2].

Den mest använda formen av MBD är den för person som parameter [2]. Minst 3 personer skall ingå i en sådan studie. Studien startar med en A-fas (A1a, A1b, A1c) utan någon insats under vilken utfallet mäts upprepade gånger för varje person (a,b,c), tills en av personerna påbörjar insatsen (A1aB1a), medan de andra fortsätter utan insats (A1b, A1c). Insatsen appliceras sedan efter en fördröjning på en andra person (B1b) och slutligen på den tredje personen (B1c) och alla personer mäts upprepat under hela

undersökningstiden med avseende på utfallet [2]. Helst väljs den vilken ordning som respektive person ska påbörja insatsen liksom efter hur många baslinjemått via en randomiseringsprocedur.

3.2.9 SCED med Changing-criterion design

Changing-criterion design (CCD) utgår ifrån ett antal hierarkiskt baserade kriterienivåer som implementeras på ett sekventiellt sätt hos exempelvis en person [4]. Kriterienivåer kan avse längd på den fas där en insats införs liksom olika styrka, grader eller doser av den insats som undersöks och genom sådana kriterienivåer kan ökning eller minskning av ett beteende mätas [13]. Logiken bygger på idén att effekten på utfallet kommer att variera i takt med faslängd eller insatsstyrka och på så sätt påvisa effekten av undersökt insats [13]. CCD särskiljs på så sätt från andra tillvägagångssätt genom att designen inte kräver att en insats dras tillbaka, eller en utökad baslinje, snabb synlighet av effekten på utfallet, liksom inte eller insatser över flera personer, beteenden och sammanhang [13].

3.2.10 Analys av SCED-studier

Det finns två olika sätt att analysera data från SCED-studier, visuellt och statistiskt [10].

3.2.11 Visuell analys [14]

Det vanligaste sättet att analysera SCED-studier är visuellt. Här läggs vikten på klinisk och praktisk snarare än statistisk signifikans. Analysen innefattar fyra steg och ett antal variabler bedöms i varje steg. I det första steget utvärderas baslinjen, och huruvida baslinjen är stabil. I det andra steget bedöms om det finns mönster i data inom faserna. De tredje steget består av att jämföra varje fas, dels med de närliggande faserna, dels med faserna som innehåller samma insats, för att undersöka om en förändring mellan insatsfas och baslinjefas sammanfaller med en effekt på utfallet. Därefter, i det fjärde steget, utvärderas om det sammantaget är så att en effekt (eller återgång till baslinje) visas vid minst tre olika tillfällen. Bland de variabler som används för att bedöma datamönster inom och mellan faser finns till exempel nivå (*eng. level*) trend och variabilitet (*eng. variability*). Visuell analys innebär en risk för typ 1 fel.

3.2.12 Statistisk analys

Det finns flera olika sätt att statistiskt analysera SCED-studier [10] (se Tabell 3.2).

Tabell 3.2 Olika statistiska analysmodeller som kan användas i en SCED-studie.

Kvasi-statistiska	Split-middle trend line, Binomial distribution test, Standard deviation band
icke-parametriska metoder	Wilcoxon matched-pairs signed-ranks test, Friedman two-way analysis of variance
Parametriska metoder	t-test, F test
Tidsserieanalyser	C statistic, Auto-regressive integrated moving average

De har dock alla brister. En sådan är att i en SCED-studie är data ofta inte oberoende eftersom det görs många mätningar med exempelvis samma person, sammanhang eller beteende. Det innebär att en förväxlingsfaktor införs. Det medför, liksom en visuell analys, en risk för typ 1 fel och kan också ge mer extrema p-värden. Statistisk analys av SCED-data är ett fält under utveckling.

3.2.13 Verktyg och principer för att bedöma tillförlitligheten av resultat från SCED-studier

Det finns ett flertal olika verktyg för att granska tillförlitlighet till resultat i SCED-studier. Fokus i denna rapport är de två verktygen:

1. SCD-RoB-verktyget [15]
2. RoBiNT [6] [16]

3.2.14 SCD RoB-verktyget

SCD RoB-verktyget [15] (se Tabell 3.3) är direkt baserat på Cochrane:s ramverk för risk för bias (RoB) granskning och designat för användning inom området intellektuell funktionsnedsättning. Även om det kan tillämpas på flera typer av SCED, är det inte en fullständig bedömning av alla källor till risk för bias för alla typer av SCED. Det finns speciella design- och giltighetsproblem för var och en av undertyperna av SCED som också måste åtgärdas. De som använder SCD RoB-verktyget använder i allmänhet också kompletterande metoder för att titta på andra former av bias. Till exempel har SCD RoB kompletterats med RoBiNT [17] eller What Works Clearinghouse SCED-designstandarder.

Tabell 3.3 SCD-RoB domäner, direkt från [15] / Table 3.3 Risk of Bias domains and Item Descriptions.

Selection Bias	The systematic differences between baseline characteristics of the participants compared.	Sequence generation Participant selection	The procedures used to allocate participants to intervention conditions or the order of the conditions to which participants are exposed. The criteria and the process used to include and select participants appropriate for the research.
Performance Bias	The systematic differences between participants in the care provided or in the exposure to factors other than the intervention under investigation.	Blinding of participants and personnel Procedural fidelity	The methods used to ensure members of the research team remain unaware of when the intervention is implemented to whom. The quality of the description for each experimental condition and the reporting of evidence indication sufficient adherence to the intervention under investigation. This includes the procedures used to train implementers.
Detection Bias	The systematic differences between participants in how outcomes are determined.	Blinding of outcome assessors Selective outcome reporting Dependent variable reliability Data sampling	The procedures used to ensure the individuals collecting outcome data are unaware of the study conditions and research purpose. The completeness of the data reported for all participants who began the study including those who withdrew and for each of the dependent variables. The procedures and reporting of agreement or reliability indices for the outcome variables. The extent to which the amount data collected for the research was sufficient to determine the level and trend of the data patterns in each condition to support the determination of a functional relation.
Other Sources of Bias	Researchers note additional sources of bias relevant to research area or not covered in the preceding domains.		

3.2.15 RoBiNT

RoBiNT [16] tidigare kallad SCED SCALE [6] är baserad på PEDro skalan och är inte relaterad till Cochrane:s RoB-verktyg, trots namnets likhet. Även om namnet antyder att det är specifikt för N-av-1-trials är detta verktyg utformat för att användas för de flesta SCED-designer. RoBiNT mäter 15 aspekter av intern och extern validitet (se Tabell 3.4), vilket överensstämmer väl med de standarder som utvecklats av What Works Clearinghouse [9].

Tabell 3.4 RoBiNT Standarder och poängsättning. Direkt taget från [16].

1. Design	Does the design of the study meet requirements to demonstrate experimental control? At least three demonstrations of the treatment effect (e.g., A-B-A-B; 6-phase multiple-baseline; Horner et al., 2005; Kratochwill et al., 2010; 2013); 1 point awarded for two demonstrations of the treatment effect
2. Randomisation	Was the phase sequence and/or phase commencement randomised?
3. Sampling behaviour (all phases)	Were there a sufficient number of data points (as defined) in each of baseline and intervention phases? At least 5 data points in every phase (Horner et al., 2005; Kratochwill et al., 2010; 2013); 1 point for less than 5 data points in any phase but at least 3 data points in each phase
4. Blinding patient/therapist	Were the participants and therapists blinded to the treatment condition (phase of study)?
5. Blinding assessors	Were assessors blinded to treatment condition (phase of study)? Single-case reports often use an assessor who is independent of the therapist (score 1 point), however, to score 2 points the assessor needs to be blind to the phase of intervention
6. Inter-rater reliability	Was IRR adequately conducted for the required proportion of data, and did it reach a sufficiently high level (as defined)? Although Kazdin (2011) notes that there are no precise rules for the frequency with which agreement should be checked, the standards described by Kratochwill et al. (2013) specify 20% of data sampled, analysed and reported per condition, and percent agreement (or equivalent) 80% or higher; percent agreement of 70% to 79%, even if no other information on sampling is provided, score 1 point. Use of machine-generated data, free from human judgement, or "reasonably objective measures" (as defined in the manual) are awarded 2 and 1 points, respectively
7. Treatment adherence	Was the intervention delivered in the way it was planned?
External validity and interpretation subscale	
8. Baseline characteristics	Were the participant's relevant demographic and clinical characteristics, as well as characteristics maintaining the condition adequately described? Recommended practice is that these conditions and variables should be considered and evaluated prior to commencing the experiment, and the way that they inform the intervention should be articulated in the report (required for 2 points). A functional analysis is one way of achieving this end, but other less formal procedures may be used as well. The evaluation goes beyond provision of demographic, medical and functional status variables, or a clinical profile of test scores (score 1 point).
9. Therapeutic setting	Were both the specific environment and general location of the investigation adequately described?
10. Dependent variable (target behaviour)	Was the target behaviour defined, operationalised, and the method of its measurement adequately described? Reports should provide an operational definition of the target behaviour (score 1 point), use a precise and repeatable measure of the target behaviour, and specify what constitutes a positive/negative response. All three elements are required to score 2 points
11. Independent variable	Was the intervention described in sufficient detail, including the number, duration and periodicity of sessions?

(intervention)

12. Raw data record	Were the data from the target behaviour provided for each session? Reports which provide a complete record of the raw data at a session by-session level (2 points). Reports which provide incomplete data, e.g., multiple probe designs; aggregated data (1 point). Only the former are awarded 2 points; whereas the latter score 1 point. Ad-hoc selection of data, whether provided in graphed or tabular format (0 points)
13. Data analysis	Was a method of data analysis applied and rationale provided for its use? Controversy remains about whether the appropriate method of analysis in single-case reports is visual or statistical. Nonetheless, 2 points are awarded if systematic visual analysis is used according to steps specified by Kratochwill et al. (2010; 2013), or visual analysis is aided by quasi-statistical techniques, or statistical methods are used where a rationale is provided for their suitability. One point is awarded if systematic/aided visual analysis is incomplete/ not conducted for every phase change or no rationale is provided for statistical analysis
14. Replication	Was systematic and/or inter-subject replication incorporated into the design? Replication refers to the repetition of the entire experiment, which may be direct intersubject and/or systematic replication. This item does not refer to replication of the experimental effect, which is covered by item 1 (design). For the award of 2 points at least three replications are required (i.e., original + 3 replications; Barlow et al., 2009); one or two replications (i.e., original + 1 or 2 replications) are awarded 1 point
15. Generalisation	Were generalisation measures taken prior to, during, and at the conclusion of treatment? Generalisation measures, whether response generalisation to other behaviours or stimulus response to other settings, need to be intentionally programmed into the design of the study, rather than a passive expectation of a "train and hope" approach (Stokes & Baer, 1977). Generalisation measures need to be evaluated throughout all phases of the experiment (Schlosser & Braun, 1994) in order to score 2 points. An accepted level of generalisation that would score 1 point is to evaluate generalisation measures prior to and at the conclusion of treatment

3.2.16 Andra RoB-verktyg för SCED

Under 2018 publicerade tidskriften *Research in Developmental Disabilities* ett specialnummer om tillförlitligheten till resultat från SCED-studier. I detta specialnummer [18] diskuteras verktygen

- CEC, som är specifikt för specialundervisning
- What Works Clearinghouse
- SCARF, Single-Case Analysis and Design Framework

när det gäller utvärdering av metodologisk kvaliteten hos SCED-studier.

I samma specialnummer introducerar Schlosser och kollegor [19] CSEDARS, ett verktyg för att skatta intern validitet hos SCED-studier där två eller flera insatser utvärderas med avseende på icke-reversibla

målbeteenden, och metodologiska standarder diskuteras av Ganz och Ayres [20].

I Wendt och Miller [21] nämns ytterligare verktyg såsom

- Certainty Framework (evidenssäkerhet), kan användas för forskning i grupp eller enstaka ämne
- EVIDAAC (Evidence in Augmentative and Alternative Communication) Scale (10 till 19 poäng), utveckling med stöd av U.S. Department of Education - National Institute on Disability and Rehabilitation Research
- Single-Subject Scale
- Romeiser Logan

3.2.17 Styrkor och svagheter

Förespråkare för SCED argumenterar till exempel för att primärstudier med SCED-design ger bättre underlag än en RCT-studie att bedöma klinisk och praktisk relevans i förhållande till hur värdefull en behandling eller insats är en enskild patient eller person. Detta till skillnad från en RCT-studie där den kliniska och praktiska relevansen avser populationsnivå. SCED-studier har i jämförelse med RCT-studier ofta inte heller krav på homogenitet. De kan därför vara extra värdefulla för enskilda kliniker eller i enskilda praktiska sammanhang inom socialtjänsten där en behandling eller insats prövas på en enskild person eller en liten grupp personer.

För att SCED-studier ska ha intern validitet ska de uppfylla den designstandard som WWC publicerat. Det finns flera faktorer som kan påverka den interna validiteten negativt, till exempel risken att de återkommande mätningarna i sig påverkar utfallet, eller att det sker en naturlig förändring över tid som felaktigt antas bero på insatsen. Det vanligaste är att man försöker stärka den interna validiteten med replikering, men även randomisering av till exempel när en insats ska introduceras kan användas.

Extern validitet är en svaghet med SCED-studier. Även här rekommenderas replikering. Ett sätt benämns på engelska 'direct inter-subject replication' där utvärderingen av aktuell insats upprepas så exakt som möjligt med andra deltagare som har liknande karaktäristika som de i originalutvärderingen. Ett annat sätt kallas 'systematic replication' där vissa aspekter som sammanhang eller typ av deltagare systematiskt enligt en fördefinierad plan varierar.

Varken deltagaren eller den som ger insatsen kan blindas. Däremot finns möjligheten att blinda den som mäter/bedömer utfall, så att den personen inte vet i vilken fas deltagaren befinner sig. Ett alternativ, där till exempel

ett beteende ska observeras, är att videofilma alla sessioner och sedan visa dem för bedömaren i slumpmässig ordning.

3.3 Litteratur om guidelines, rekommendationer och bedömningsmallar

Guidelines för hur studier baserade på single case design skall rapporteras finns i

- The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement [4]
- The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016: Explanation and elaboration [10]
- CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement [12]

Rekommendationer om val av studiedesign samt genomförande liksom analys av data från SCED-studier kan hittas i böcker såsom

- J. Barker, P. McCarthy, M. Jones, A. Moran. Single case research methods in sport and exercise psychology. Routledge, London, United Kingdom (2011).
- D.H. Barlow, M.K. Nock, M. Hersen. Single case experimental designs: Strategies for studying behavior change. (3rd ed.), Pearson, Boston, MA (2009).
- D.L. Gast, J.R. Ledford. Single case research methodology: Applications in special education and behavioral sciences. (2nd ed.), Routledge, New York, NY (2014).
- Guyatt, G., Jaeschke, R., & McGinn, T. (2002). N of 1 randomized controlled trials. In G. Guyatt, D. Rennie, M. O. Meade, & D. J. Cook (Eds.), *User's guides to the medical literature* (pp. 179–192). New York, NY: McGrawHill Medical and American Medical Association.
- R.H. Horner, E.G. Carr, J. Halle, G. McGee, S. Odom, M. Wolery. The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71 (2005), pp. 165-179.
- A.E. Kazdin. *Single-case research designs: Methods for clinical and applied settings*. Oxford University Press, New York, NY (2011).
- C.H. Kennedy. *Single-case designs for educational research*. Pearson, Boston, MA (2005).
- R. Kratochwill, J.H. Hitchcock, R.H. Horner, J.R. Levin, S.L. Odom, D.M. Rindskopf, W.R. Shadish. Single-case intervention research design standards. *Remedial and Special Education*, 34 (2013), pp. 26-38.

- T.R. Kratochwill, J.R. Levin. Single-case intervention research: Methodological and statistical advances. American Psychological Association, Washington, DC (2014).
- D.L. Morgan, R.K. Morgan. Single-case research methods for the behavioral and health sciences. Sage, Los Angeles, CA (2009).
- T.C. Riley-Tillman, M.K. Burns. Evaluating educational interventions: Single-case design for measuring response to intervention. Guilford Press, New York, NY (2009).
- K.J. Vannest, J.L. Davis, R.I. Parker. Single case research in schools: Practical guidelines for school-based professionals. Routledge, New York, NY (2013) [4].

Rekommendationer om val av studiedesign samt genomförande liksom analys av data från SCED-studier går att hitta Journal of Behavioral Education 2012, Remedial and Special Education 2013, the Journal of School Psychology and Neuropsychological Rehabilitation 2014 samt Aphasiology in 2015 [4].

Rekommendationer om metodologisk kvalitet finns exempelvis i:

- R.H. Horner, E.G. Carr, J. Halle, G. McGee, S. Odom, M. Wolery. The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71 (2005), pp. 165-179.
- T.R. Kratochwill, J.H. Hitchcock, R.H. Horner, J.R. Levin, S.L. Odom, D.M. Rindskopf, W.R. Shadish. Single-case intervention research design standards. *Remedial and Special Education*, 34 (2013), pp. 26-38.
- D.M. Maggin, A.M. Briesch, S.M. Chafouleas, T.D. Ferguson, C. Clark. A comparison of rubrics for identifying empirically supported practices with single-case research. *Journal of Behavioral Education*, 23 (2014), pp. 287-311.
- J.D. Smith. Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods*, 17 (2012), pp. 510-550. [4].

Litteratur som tar upp bedömning av risk för bias i Single Case-studier finns i:

- Tate RL, McDonald S, Perdices M, Togher L, Schultz R, Savage S. Rating the methodological quality of single-subject designs and n-of-1 trials: introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychol Rehabil.* 2008 Aug;18(4):385-401. doi: 10.1080/09602010802009201. PMID: 18576270,
- R.L. Tate, M. Perdices, U. Rosenkoetter, D. Wakim, K. Godbee, L. Togher, S. McDonald. Manual for the critical appraisal of single-case

reports using the Risk of Bias in N-of-1 Trials (RoBiNT) Scale.
Unpublished manuscript. University of Sydney, Australia (2013)

- R.L. Tate, M. Perdices, U. Rosenkoetter, D. Wakim, K. Godbee, L. Togher, S. McDonald. Revision of a method quality rating scale for single-case experimental designs and N-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation*, 23 (2013), pp. ,619-638.
- R.L. Tate, U. Rosenkoetter, D. Wakim, L. Sigmundsdottir, J. Doubleday, L. Togher, ..., M. Perdices. *The Risk of Bias in N-of-1 Trials (RoBiNT) Scale: An expanded manual for the critical appraisal of single case reports*. Author, Sydney, Australia (2015) [4].

4. Referenser

1. Zeng L, Brignardello-Petersen R, Hultcrantz M, Siemieniuk RAC, Santesso N, Traversy G, et al. GRADE guidelines 32: GRADE offers guidance on choosing targets of GRADE certainty of evidence ratings. *J Clin Epidemiol*. 2021;137:163-75. Available from: <https://doi.org/10.1016/j.jclinepi.2021.03.026>.
2. Krasny-Pacini A, Evans J. Single-case experimental designs (SCEDs) to assess intervention effectiveness in rehabilitation: a practical guide. *Annals of Physical and Rehabilitation Medicine*. 2018;61(3):pp. 164-79. Available from: <https://doi.org/10.1016/j.rehab.2017.12.002>.
3. Epstein L. H, Dallery J. The Family of Single-Case Experimental Designs. *Harvard Data Science Review*. 2022(Special Issue 3). Available from: <https://doi.org/10.1162/99608f92.ff9300a8>.
4. Tate R. L, Perdices M, Rosenkoetter U, Shadish W, Vohra S, Barlow D. H, et al. The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *Journal of Clinical Epidemiology*. 2016.
5. Manolov R, Gast D. L, Perdices M, Evans J. J. Single-case experimental designs: Reflections on conduct and analysis. *Neuropsychological Rehabilitation*. 2014.
6. Tate RL, McDonald S, Perdices M, Togher L, Schultz R, Savage S. Rating the methodological quality of single-subject designs and n-of-1 trials: introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychol Rehabil*. 2008;18(4):385-401. Available from: <https://doi.org/10.1080/09602010802009201>. PMID: 18576270.
7. Guyatt G, Sackett D, Adachi J, Roberts R, Chong J, Rosenbloom D, et al. A clinician's guide for conducting randomized trials in individual patients. *Canadian Medical Association Journal*. 1988(139):497–503.
8. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy-randomized trials in individual patients. *N Engl J Med*. 1986;Apr 3; 314(14):889-92. Available from: <https://doi.org/10.1056/NEJM198604033141406>. PMID: 2936958.
9. What Works Clearinghouse 2020:What Works Clearinghouse Standards Handbook, Version 4.1. <https://ies.ed.gov/ncee/wwc/handbooks>: Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Washington, DC: U.S. . [
10. Tate R. L, Perdices M, Rosenkoetter U, McDonald S, Togher L, Shadish W, et al. The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology*. 2016:10–31. Available from: <http://dx.doi.org/10.1037/arc0000027>.
11. Shadish WR, Sullivan KJ. Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods*. 2011(43):971–80. Available from: <https://doi.org/http://dx.doi.org/10.3758/s13428-011-0111>.

12. Vohra S, Shamseer L, Sampson M, Bukutu C, Schmid C. H, Tate R, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *British Medical Journal*. 2015;350:h1738. Available from: <https://doi.org/http://dx.doi.org/10.1136/bmj/h1738>.
13. Klein L A, Houlihan D, Vincent J L, Panahon C J. Best Practices in Utilizing the Changing Criterion Design. *Behav Anal Pract*. 2015;10(1):52-61. Available from: <https://doi.org/10.1007/s40617-014-0036-x>. PMID: 28352507; PMCID: PMC5352619.
14. Kratochwill T. R, Hitchcock J, Horner R. H, Levin J. R, Odom S. L, Rindskopf D. M, et al. Single-case designs technical documentation. http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf; 2010.
15. Reichow B, Barton E. E, Maggin D. M. Development and applications of the single-case design risk of bias tool for evaluating single-case design research study reports. *Res Dev Disabil*. 2018(79):53-64. Available from: <https://doi.org/10.1016/j.ridd.2018.05.008>. Epub 2018 Jun 27. PMID: 29958733.
16. Tate R.L , Perdices M, Rosenkoetter U, Wakim D, Godbee K, Togher L, et al. Revision of a method quality rating scale for single-case experimental designs and N-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation*. 2013;23:619-38. Available from: <https://doi.org/http://dx.doi.org/10.1080/09602011.2013.824383>.
17. Müller AR, Brands MMMG, van de Ven PM, Roes KCB, Cornel MC, van Karnebeek CDM, et al. Systematic Review of N-of-1 Studies in Rare Genetic Neurodevelopmental Disorders. *Neurology* 2021;16(11):529-40. Available from: <https://doi.org/10.1212/WNL.0000000000011597>. Epub 2021 Jan 27. PMID: 33504638; PMCID: PMC8032375.
18. Zimmerman K. N, Ledford J. R, Severini K. E, Pustejovsky J. E, Barton E. E, Lloyd BP. Single-case synthesis tools I: Comparing tools to evaluate SCD quality and rigor. *Research in Developmental Disabilities* 2018(79):19-32. Available from: <https://doi.org/10.1016/j.ridd.2018.02.003>.
19. Schlosser R. W, Belfiore P. J, Sigafos J, Briesch A. M, Wendt O. Appraisal of comparative single-case experimental designs for instructional interventions with non-reversible target behaviors: Introducing the CSCEDARS (“Cedars”). *Research in Developmental Disabilities* 2018(79):33-52.
20. Ganz JB, Ayres KM. Methodological standards in single-case experimental design: Raising the bar. *Res Dev Disabil*. 2018(79):3-9. Available from: <https://doi.org/10.1016/j.ridd.2018.03.003>. Epub 2018 Apr 12. PMID: 29655508.
21. Wendt O, Miller B. Quality Appraisal of Single-Subject Experimental Designs: An Overview and Comparison of Different Appraisal Tools. *Education and Treatment of Children*. 2012;35(2):235-68. Available from: <https://doi.org/10.1353/etc.2012.0010>.