

Experiment i offentlig sektor

Jens Dietrichson (jsd@vive.dk)

VIVE – The Danish Center for Social Science Research

KSP, Malmö 2025-09-05



VIVE

Outline

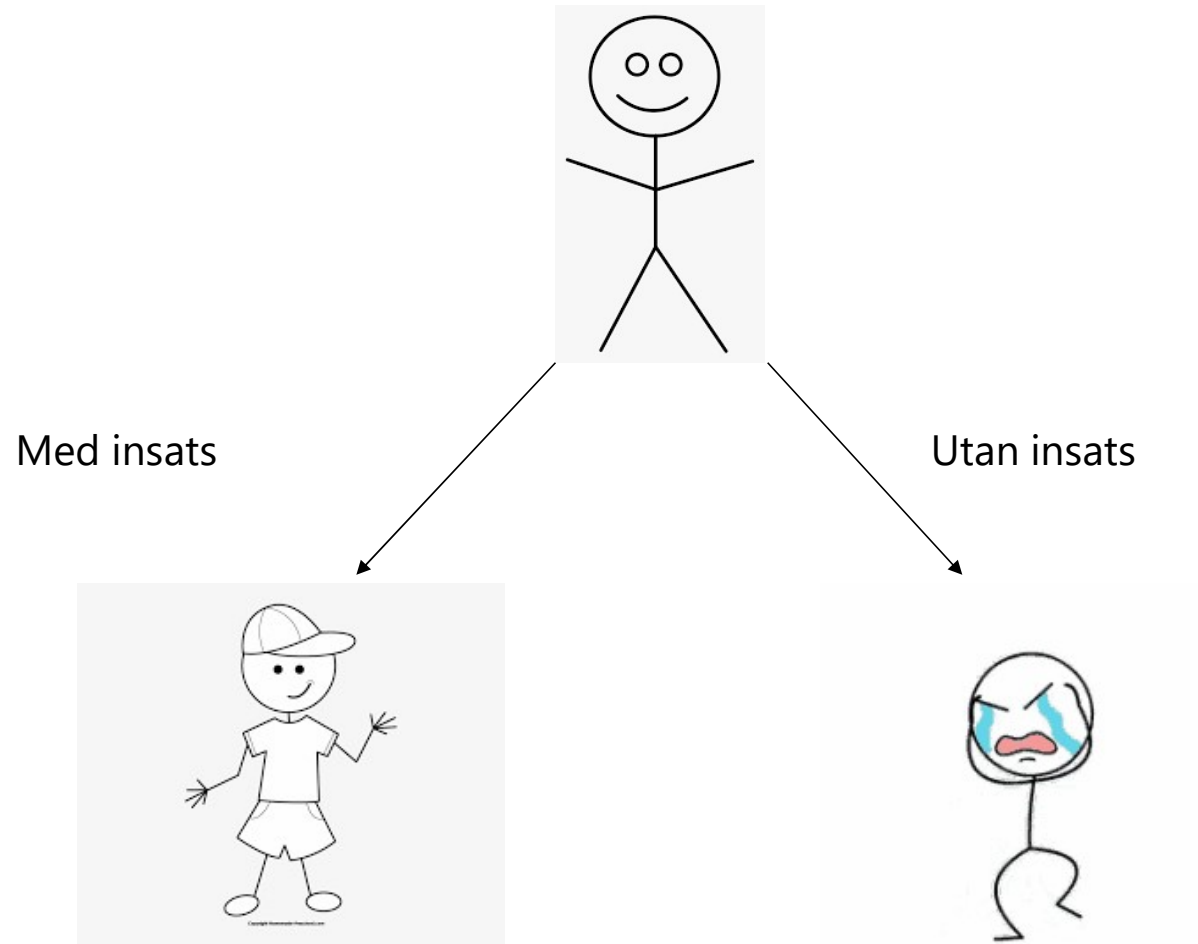
Presentation i 2 delar:

- > Del 1: argument för att experiment i allmänhet och randomiserade kontrollerade försök (RCTs) i synnerhet kan bidra till utvecklingen av offentlig sektor.
- > Del 2: exempel på barriärer och möjligheter för RCTs genom en jämförelse av Danmarks och Sveriges erfarenheter med fokus på grundskola och förskola.

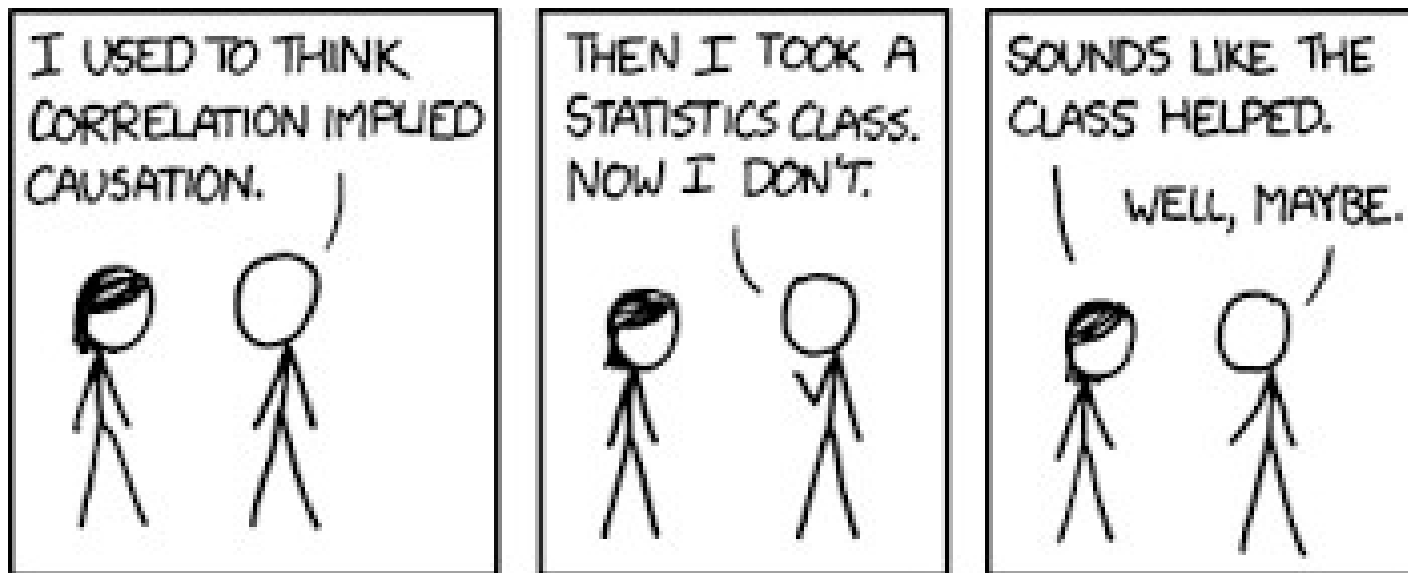
Vad är ett experiment?

- > SAOL: "försök i undersökande syfte". Inom samhällsvetenskapen förstås experiment ofta som en metod för att undersöka orsakssamband och effekter av insatser.
- > Många olika typer av experiment. Två kontraster som är viktiga för presentationen:
 - > Kontrollerade vs. naturliga experiment.
 - > Randomiserade vs. icke-randomiserade.
- > Naturliga experiment har den stora fördelen att de är just naturliga, de som utsätts för experimentet vet inte att de deltar i en forskningsstudie.
- > Nackdelen är att man kan få vänta länge på ett bra naturligt experiment. Att bedriva utveckling på basis av bara naturliga experiment är därför svårt och jag kommer därför att fokusera på kontrollerade experiment.

Definition av insatseffekter



Effektutvärderingens grundproblem



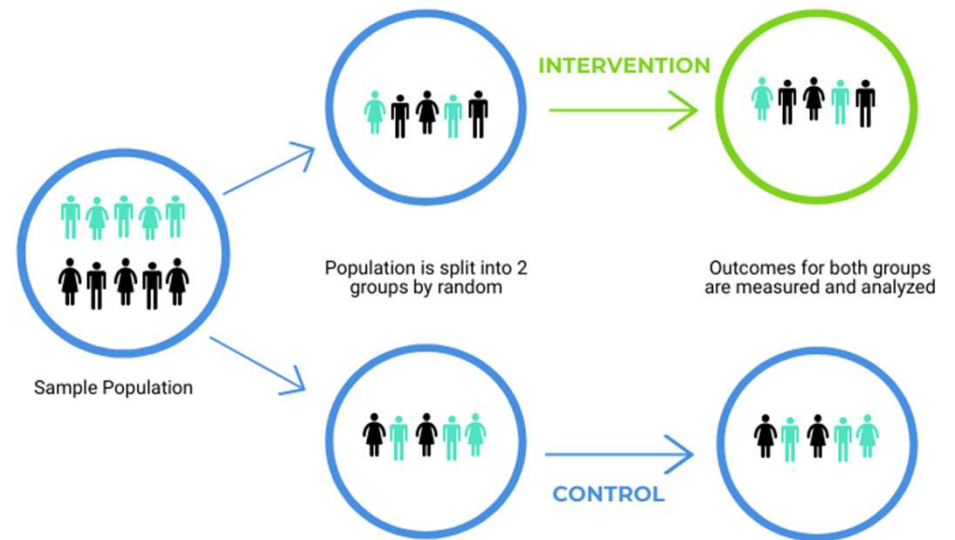
Källa: xkcd: <https://xkcd.com/552>

Experiment och effekter

- > När vi vet vad som skulle hända om någon inte får en insats, dvs. den kontrafaktiska situationen, är det okomplicerat att mäta effekten av insatsen.
- > Men när vi är osäkra på den kontrafaktiska situationen på individnivå, kan vi ofta inte mäta individuella effekter utan måste nöja oss med att jämföra med gruppgenomsnitt, dvs. skatta olika typer av genomsnittseffekter.
- > Genom att använda experiment av insatser försöker vi hitta en trovärdig kontrollgrupp som ger en god uppskattning av vad som skulle ha hänt med insatsgruppen, om de inte hade fått insatsen.
- > Även om vi är säkra på att insatsen kvalitativt förbättrar utfallen för insatsgruppen kan experiment vara nödvändiga för att skatta hur stor effekten är och hur kostnadseffektiv insatsen är.

Varför randomisera?

- > Eftersom orsakssamband inte kan observeras direkt (Hume, 1740), krävs det ofta starka antaganden om kontrollgruppen för att kunna skatta effekter.
- > RCTs "claim to fame" är att dessa antaganden inte är så starka: om insats- och kontrollgrupp väljs slumpmässigt är alla skillnader dem emellan också slumpmässiga.
- > De förväntade skillnaderna är därmed, i genomsnitt, noll, **för alla skillnader**.



Ett (halv)bra argument mot RCTs

- > *I motsats till de förväntade skillnaderna kan faktiska skillnader uppstå om man har otur.*
- > Risken för det reduceras med stickprovsstorleken, men ibland har man inte råd med tillräckligt stora stickprov.
- > "Minimization" eller matching kan vara bättre i riktigt små stickprov (men kan ofta kombineras med randomisering).
- > Om syftet med en studie inte bara är utvärdering av en enskild insats så kan randomisering vara värdefullt även om stickprovet är litet och analysen inte har tillräcklig statistisk styrka för den förväntade effekten. Många små studier kan kombineras med hjälp av meta-analys.

Bra argument mot kontrollerade försök (inkl. RCTs)

- > *Svarar inte på alla intressanta frågor.*
 - > T ex effekten av nationella reformer eller hur deltagare upplever en insats.
- > *Kontrollerade försök är till viss del artificiella:* deltagarna är oftast inte slumpvist utvalda, representerar inte den population vi egentligen är intresserade av, och vet typiskt att de deltar i en studie.
- > *Att insats- och kontrollgrupp är lika i genomsnitt gäller bara utan avhopp.*
 - > Eller om avhopparna inte har egenskaper som korrelerar med hur insatsen skulle påverkat deltagarna och det är en skillnad mellan insats- och kontrollgrupp i hur många som hoppar av.
- > *Kontrollerade försök är typiskt dyra och insatsperioden kan därför bli för kort.*

Mindre bra argument mot kontrollerade försök

- > *Kontrollerade försök är oetiska för att en grupp får och en grupp får inte en insats.*
- > När kontrollgruppen får det de skulle ha fått utanför försöket är argumentet i sig självt relativt svagt. Det är inte självklart dåligt att en grupp eventuellt får det bättre och en grupp är opåverkad.
 - > Det bygger på att kontrollgruppen väljer själv att vara med i studien (annars är det ett betydligt bättre argument!).
- > Det är ofta möjligt att designa försöket så att alla förväntas få det bättre än om studien inte gjorts:
 - > Kontrollgruppen kan få insatsen efter insatsgruppen (väntelistedesign).
 - > Kontrollgruppen kan kompenseras, t ex få betalt.

Halvtids-summering och en brasklapp

- > Kunskap om de kvantitativa effekterna av insatser kräver ofta en experimentell design.
- > RCTs har många attraktiva egenskaper. De har också några nackdelar, men dessa delas i stor utsträckning av icke-randomiserade kontrollerade försök. When in doubt, randomize!
- > Experiment kan dock verkligen inte svara på alla frågor om insatser som vi är intresserade av.
- > Kvalitativa studier är ett viktigt komplement till effektstudier (och tvärtom). **Evidens-informerad praktik kräver evidens från både kvantitativa och kvalitativa studier.**



Kvalitativ
ELLER
kvantitativ metod

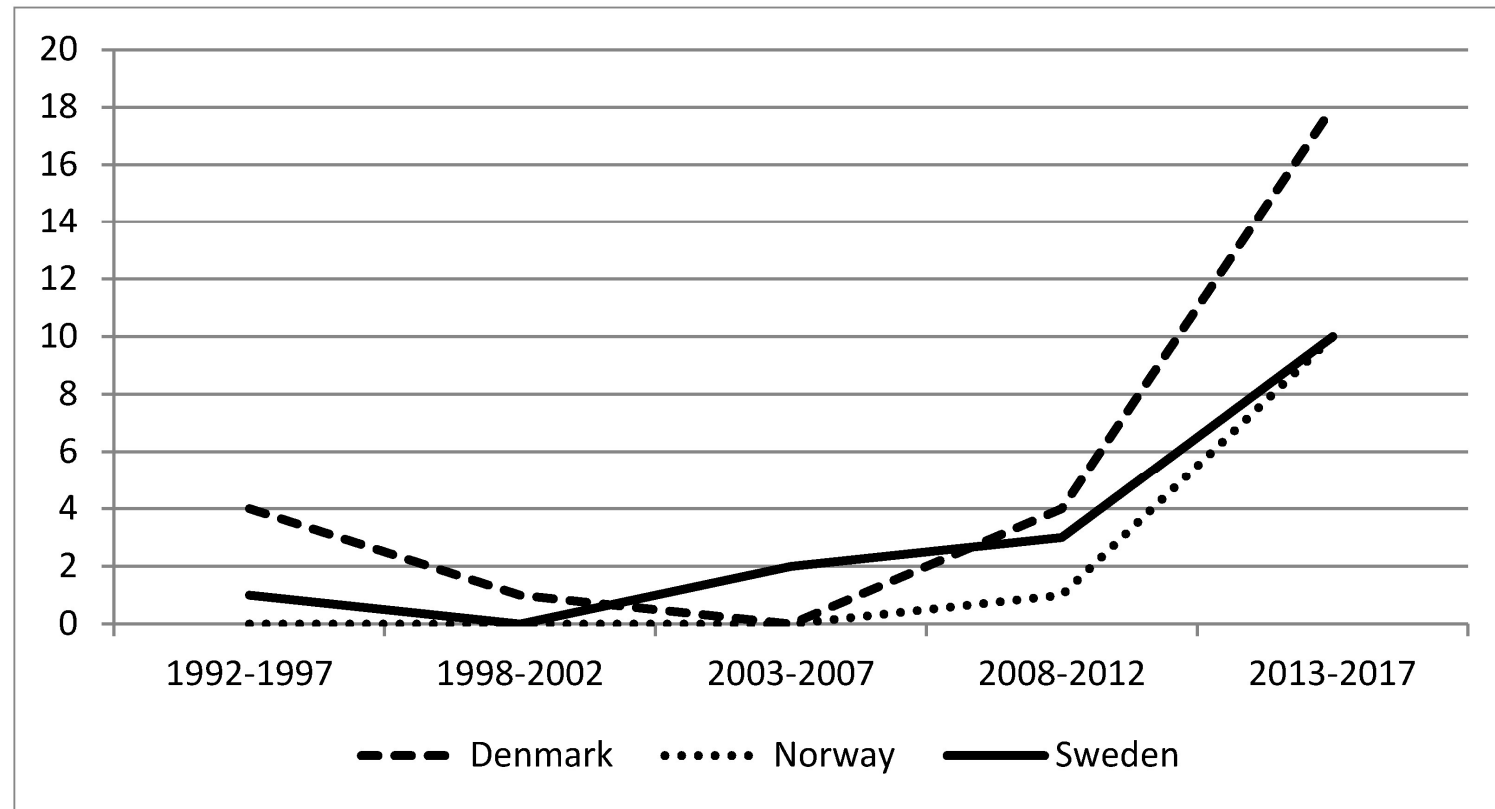
Kvalitativ
OCH
kvantitativ metod

VIVE

RCTs på kommunala ansvarsområden

- > Sundell och Olsson (2021) fann 528 svenska effektutvärderingar av beteendemässiga, psykologiska och sociala insatser i perioden 1990-2019, varav 456 (86%) var RCTs.
 - > Relativt många studier finansierade av landsting/regioner (123) och statliga myndigheter (82) men få av kommuner (12). Resten finansierat av fonder eller universitet.
 - > Trots mängden studier menade författarna att "Få av insatserna som socialtjänsten och psykiatrin erbjuder är dock utvärderade genom kontrollerade studier vilket gör att det är oklart om de är effektiva eller inte" (s. 5).
- > Pontoppidan et al. (2018) fann 17 RCTs genomförda i svenska grundskolor i perioden 1992-2017.
 - > 7 studier finansierade av statliga myndigheter och 2 studier finansierade av kommuner. Resten av fonder och universitet.
 - > Vad vet vi om svenska skolinsatsers effektivitet?

RCTs i skandinaviska grundskolor



Källa: Pontoppidan et al. (2018)

Varför har det gjorts relativt många RCTs i dansk skola (och förskola)?

- > Kräver en miljö som stöder användandet av metoden, finansiering, villiga deltagare och metodkunskaper (Sundell och Olsson, 2021).
- > Utvecklingen drevs inte av forskare inom utbildningsvetenskap/pedagogik, som dominerades av kvalitativa forskare. Utbildningspsykologer, ekonomer och statsvetare gjorde de första studierna.
- > Den politiska reaktionen på t ex PISA och andra internationella mätningar i Danmark (och Norge) handlade mycket om att man inte fick tillräckligt hög kvalitet för de resurser man lade ner. Diskussionen om kostnadseffektivitet ledde till ett fokus på effekter och därmed RCTs.
- > Ministerier och styrelser stöttade också med en del finansiering men kanske ännu mer den privata, icke-vinstdrivande Trygfonden har finansierat många danska randomiserade experiment i skola och förskola.
- > Många danska skolor och förskolor har varit och är villiga att delta i försök.

Varför relativt få RCTs i svensk skola?

- > Liknande dominans av kvalitativ forskning inom utbildningsvetenskap och pedagogik, men större andel svenska RCTs är gjorda av forskare från dessa fält.
- > Snäv evidens, men i mitt enda försök att rekrytera skolor till en RCT i Sverige var detta överraskande enkelt.
- > Problemen i svensk skola är, om man får tro internationella mätningar, större än i dansk skola. Reaktionen från politiskt håll har dock inte primärt handlat om kostnadseffektivitet.
- > I motsats till Danmark verkar det saknas stöd från ministerier och myndigheter med ansvar för utbildning.
- > Det har inte funnits någon privat motsvarighet till Trygfonden.

Bristande stöd från myndigheter

- > Skolverket skriver i en rapport om anti-mobbningsprogram (2011, s. 46): "Vedertagen praxis angående utvärdering av evidensbaserad programimplementering vilar mer eller mindre uteslutande på den randomiserade kontrollmodellen. Den är **inte möjlig** för en statlig myndighet i Sverige som varken har rätt att ingripa i skolornas verksamhet eller av etiska skäl skulle kunna **påbjuda skolor** att arbeta med ett antimobbningsprogram vars **effekter är osäkra**".
- > Bara 1 exempel på ett "policy-initierat" experiment i Sverige i Pontoppidan et al. (2018): Skolverkets *Insatser för att stärka utbildningens kvalitet för nyanlända elever*. Dvs. nästan alla var initierade av forskare.
- > Anmärkningsvärt stor andel av de RCTs som gjorts i skolan har fått stöd av t ex FORTE och Socialstyrelsen, som inte har direkt ansvar för utbildningsfrågor (Pontoppidan et al., 2018).

Bristande finansiering?

- > Skolforskningsinstitutet verkar mer positiva till RCTs men har stött relativt få. Av 64 finansierade projekt 2016-2024 är 16 kategoriserade som interventionsstudier, 4 av dessa nämner en kontroll/jämförelsegrupp, 1 är uttalat en RCT.
- > Vetenskapsrådet delar ut överlägset mest medel till utbildningsvetenskaplig forskning (163 mnkr 2024; Tholin, 2021). Få medel gick till interventionsstudier i 2024.
- > Projektansökningarna är forskarinitierade och det är forskare som väljer ut projekten, så svårt att veta om det är utbudet eller efterfrågan som är problemet.
- > Utbildningsvetenskap får väldigt lite forskningsmedel i relation till hur viktigt utbildning är för samhället och i relation till andra områden. Medicin och hälsa får t ex 5 gånger mer av Vetenskapsrådet (Tholin, 2021) och det finns ingen stor privat fond som fokuserar på studier om skola eller förskola.

Är bristen på RCTs ett problem?

- > Har ökningen av RCTs internationellt bidragit med ny och viktig kunskap?
 - > Läsinlärning: Från "Reading Wars" till "The Science of Reading".
 - > Formativa test som inlärningsmetod.
 - > Många insatsformer för barn med svårigheter fungerar **inte** väl, medan några gör det (t ex tutoring).
- > Vad vet vi med någon form av säkerhet från fält där RCTs är svåra att göra?
- > Det är mer tveksamt om dansk skola står starkare än svensk skola **på grund av** att Danmark har genomfört fler RCTs.
 - > Kumulativ inlärnings- och implementeringsprocess, som endast varit i gång i knappt ett decennium.
 - > På många områden kan (och bör) man också lära sig av andra länder, dvs. Sverige kan lära av Danmark och vice versa.
 - > På skolområdet saknar både Danmark och Sverige ett sammanhängande system för att fånga upp innovationer, testa om dessa fungerar och omsätta evidens till ny och bättre praktik.

Kommuner och RCTs

- > Borde kommuner delta i, finansiera och själva genomföra (finansiera, delta och administrera) RCTs?
- > Delta: ja.
- > Finansiera: finns tidigare exempel (2 i Pontoppidan et al., 2018). SKR-koordinerad finansiering?
- > Genomföra:
 - > Stora kommuner har tillräckligt många elever och skolor för att (ibland) göra RCTs med tillräcklig statistisk styrka. För små och mellanstora kommuner är detta svårt.
 - > Kräver specialiserad kompetens. Men det svåraste med RCTs rör implementeringen av insatserna, inte design och analys, och administrativ kompetens kan vara en kommunal fördel.
 - > Samarbete med forskare nog ändå att föredra och förmodligen viktigare att kunna använda sig av evidens än att själv generera den.

Sammanfattning

- > RCTs är ofta ett bra sätt att lära sig om effekten av insatser och ibland det bästa sättet.
- > När de passar är RCTs fullt möjliga att genomföra i offentlig sektor, inkl. alla områden kommunerna har huvudansvaret för.
- > Det är ett stort problem att det har gjorts och görs så få RCTs i svenska skolor och förskolor. De är få i relation till både närliggande områden (socialtjänst, psykiatri) och länder (Danmark, Norge).
- > Mitt argument är vi borde göra fler RCTs i skola och förskola, inte att vi bara ska göra RCTs. Nästan alla metoder har sitt existensberättigande och kan svara på vissa frågor och inte andra. RCTs är en pusselbit som saknas men knappast den enda.

Tack för uppmärksamheten!

jsd@vive.dk

VIVE

Referenser

Pontoppidan, M., Keilow, M., Dietrichson, J., Solheim, O. J., Opheim, V., Gustafson, S., & Andersen, S. C. (2018). Randomised controlled trials in Scandinavian educational research. *Educational Research*, 60(3), 311-335.

Tholin, J. (2021). Finansiering av utbildningsvetenskaplig forskning. Vetenskapsrådet, Stockholm.

Sundell, K., & Olsson, T. M. (2021). Svenska effektutvärderingar av beteendemässiga, psykologiska och sociala insatser 1990–2019. Forte.