

Bruk av Interteaching i undervisning: en gjennomgang av internasjonalt publiserte artikler (2014–2019)

Hanna Steinunn Steingrimsdottir og Erik Arntzen
OsloMet – Storbyuniversitetet

Interteaching er en undervisningsteknikk som legger vekt på studentaktiv læring. Teknikken baserer seg på veldokumenterte atferdsanalytiske prinsipper. I et tidligere litteratursøk av Sturmey et al. (2015) identifiserte forfatterne syv empirisk artikler som vurderte effekten av Interteaching sammenlignet med tradisjonell forelesning. Effekten var generelt positiv og teknikken ble beskrevet som *lovede*. Formålet med den presenterte artikkelen var å gjennomføre et oppdatert litteratursøk de siste fem årene, 2014–2019. Resultatet førte til totalt 12 artikler som oppfylte inkluderingskriteriene. Av de 12 artiklene som ble inkludert var en randomisert gruppestudie, seks studier der alternerende intervensjonsdesign ble brukt, fire kvasiekperimentelle studier og en artikkel med selvevaluering. Oppsummert viste tre artikler der Interteaching ble sammenlignet med andre undervisningsteknikker at studentene skåret bedre på testing ved bruk av Interteaching, derav to som viste en statistisk signifikant forskjell. Tre studier undersøkte effekten av å bruke Interteaching i nettbaserte kurs. Konklusjonen er at Interteaching kan nå beskrives som empirisk støttet teknikk.

Nøkkelord: Interteaching, studentaktiv læring, undervisning, eksperimentell design, sosial validitet, høyere utdanning, oversiktsartikkel.

Det finnes flere ulike måter å undervise på, fra tradisjonelle forelesninger til nettbaserte arrangementer der forelesninger og oppgaver er tilgjengelig på internett. Ved nettbaserte arrangementer styrer studenten når, hvor ofte og hvor lenge av gangen han eller hun hører på en forelesning eller jobber med oppgaver. Denne formen for undervisning har vært framtre­dende de siste månedene i forbindelse med COVID-19 pandemien. Universiteter og høyskoler i Norge har også annonsert at mye av undervisningen framover vil kunne foregå online (se for eksempel Norges Teknisk-Naturvitenskapelig Universitet, 2020).

Forfattermerknad: Vi takker to anonyme fagfeller for nyttige kommentarer. Det er ingen interessekonflikter forbundet med dette manuskriptet. All korrespondanse vedrørende manuskriptet rettes til Hanna Steinunn Steingrimsdottir <hannasteinunn@simnet.is>.

Selv om en tradisjonell forelesningsform, hvor mye informasjon deles på kort tid, sannsynligvis fortsatt er den mest brukte teknikken (Matheson, 2008) finnes det en del kritikk mot den (e.g., Barr & Tagg, 1995). Kritikken går blant annet ut på at studentene blir passive mottakere som hører på foreleseren over lang tid uten selv å delta aktivt i diskusjoner (Granmo et al., 2017). Faglæreren styrer undervisningen, hvor fort eller sakte han eller hun går igjennom pensumet, og det er faglæreren som styrer kommunikasjonen med studentene.

Undervisningsteknikker, der fokuset flyttes fra faglæreren til studentene og som skal motvirke noe av de ulempene som beskrives ved den tradisjonelle forelesningen, betegnes som studentaktiv læring

(Prince, 2004). Studier har vist hvordan slike undervisningsteknikker kan føre til et større læringsutbytte for studentene (e.g., Freeman et al., 2014), beskrevet som kunnskap, ferdigheter og generell kompetanse (se for eksempel Halvorsen et al., 2018). Sentralt innen atferdsanalyse er at læring er noe studentene gjør og ikke noe de passivt mottar, som gjør at en atferdsanalytisk tilnærming til læring blir i tråd med ønsker om at studenter skal være mer aktive i undervisning (Kunnskapsdepartement, 2017). Interteaching er én undervisningsteknikk som baseres seg på atferdsanalytiske prinsipper der fokus er nettopp studentaktiv læring (Boyce & Hineline, 2002). Boyce og Hineline påpekte hvordan forsterkningsbetingelsene ikke er optimale under den tradisjonelle forelesningen ved at forelesningen går for sakte for noen og for fort for andre. Ved bruk av Interteaching arrangeres kontingenser for at studentene skal møte forberedt og delta i diskusjoner eller samtaler. Samtalene varer i 30 til 40 minutter og skal ha utgangspunkt i hovedpoengene i en utvalgt tekst fra en lærebok eller artikkel som er annonsert på forhånd. Faglæreren lager sentrale spørsmål knyttet til teksten. Disse spørsmålene kan enten deles ut på forhånd eller i klasserommet. Faglæreren setter sammen to og to studenter (helst nye par i hver økt) og studentene diskuterer spørsmålene som er utlevert. Faglæreren følger med i de ulike diskusjonene og gir tilbakemeldinger, samt igangsetter diskusjoner hvis det er nødvendig.

I etterkant av sekvensen med diskusjoner leverer studentene et skjema til faglæreren der de har notert eget navn og hvem de har jobbet sammen med, de temaene de syntes var vanskelige i teksten, og det studentene ønsker at faglærer fokuserer på i undervisning. Notatet blir på den måten en viktig del av undervisningen. Faglæreren analyserer innholdet av notatet og tilpasser forelesningen sin i henhold til studentenes tilbakemeldinger. Boyce og Hineline påpeker at når forelesninger er basert på det studenter etterspør blir det større sannsynlighet for at

studentene deltar i forelesningene. Studentene kan også gjennomføre en quiz etter endt diskusjonssekvens og en poengsum over et visst kriterium kan bli en del av sluttkarakteren ved eksamen.

Halpern og Hakel (2003) har påpekt viktigheten av at undervisningsteknikker bør være empirisk støttet. Selv om det i den presenterte artikkelen er snakk om effektive undervisningsteknikker og vurderingen av disse, så benyttes ofte et system for å avgjøre hvilke behandlingsteknikker eller intervensjon som kvalifiserer for å være empirisk støttet. David Barlow ledet et arbeid tidlig på 90'tallet som skulle munne ut i en rapport om empirisk støttede behandlingsformer (Task Force on Promotion and Dissemination of Psychological Procedures, 1993). Senere har Chambless og Hollon (1998) beskrevet hva denne kvalifiseringen innebærer for gruppedesign (finnes tilsvarende kriterier for *single-case research* design). I vurderingen stilles det følgende spørsmål: Har behandlingen vist seg å være gunstig i kontrollerte studier eller eksperimenter? Har behandlingen vist seg å være nyttig i anvendte settinger? I så fall med hvilke deltakere og under hvilke betingelser. Har intervensjonen vist seg å være tilstrekkelig i betydningen at den er kostnadseffektiv i forhold til alternative intervensjoner? Chambless og Hollon argumenterer videre for at en intervensjon skal bli klassifisert som *virkningsfull* og da empirisk støttet, så må minst to kontrollerte eksperimenter eller studier gjort ved uavhengige forskningsgrupper ha vist effektene av intervensjonen. Dersom det bare én kontrollert studie eller et eksperiment, eller dersom all forskningen er gjort av en forskningsgruppe klassifiseres intervensjonen som *lovende*. Dette innebærer blant annet at det trengs replikasjoner som viser effekten av kontrollerte eksperimenter eller studier.

Boyce og Hineline (2002) etterlyste en empirisk vurdering av effekten av Interteaching på tross av at teknikken baserer seg på veldokumenterte atferdsanalytiske prinsipper som positiv og negativ forsterkning

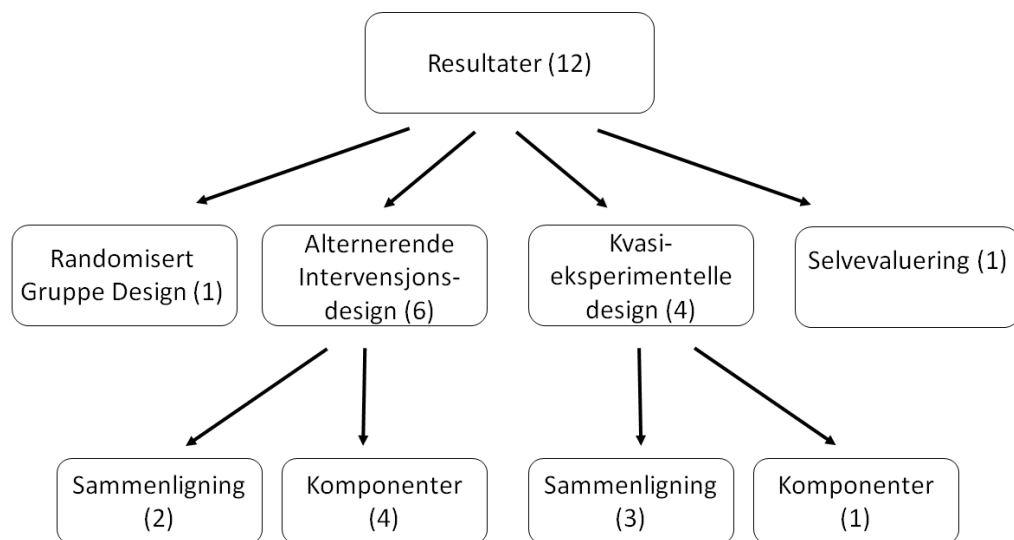
og motivasjonelle operasjoner. Siden 2002 har en del forskning blitt gjort for å evaluere effekten av denne teknikken. Sturmey et al. (2015) undersøkte artikler fram til 2014 og fant 17 artikler fra internasjonale fagfellevurderte tidsskrifter der Interteaching var uavhengig variabel og effekten var undersøkt i forhold til minst en avhengig variabel. Artikkene ble kategorisert i tre grupper (1) sammenligning av Interteaching med andre undervisningsteknikker, (2) artikler der ulike komponenter av Interteaching ble undersøkt og (3) kvasiekperimentelle studier. I syv av de 17 artiklene ble Interteaching sammenlignet med andre undervisningsformer, og seks artikler bekreftet at bruk av teknikken førte til bedre resultater. Uansett påpekte Sturmey et al. (2015) at selv om Interteaching er en *lovende* undervisningsteknikk kunne den ikke kategoriseres som empirisk støttet ettersom mesteparten av studiene var gjort av en forskningsgruppe. Formålet med den presenterte studien var å utvide den tidligere oversikten om bruken av Interteaching

ved å gjennomføre en femårs (2014–2019) oversikt av internasjonalt publiserte artikler. I tillegg vise til hvilke komponenter av Interteaching som er undersøkt og vurdering om denne teknikken er empirisk støttet.

Metode

Denne studien er en replikasjon og utvidelse av Sturmey et al. (2015) som gjennomførte et litteratursøk 14. mars 2014 i *PsychINFO* og *Google Scholar* databasene. Nøkkelordet forfatterne brukte den gang var «inter-teaching». Forfatterne hadde to inkluderingskriterier; (1) artiklene måtte beskrive en studie der Interteaching var uavhengig variabel og dokumentasjon av effekten på minst én avhengig variabel, og (2) artiklene var publisert i et fagfellevurdert internasjonalt tidsskrift.

I den presenterte studien ble de samme databasene (*PsychInfo* og *Google Scholar*) brukt i litteratursøket. Søket var begrenset til perioden 15. mars 2014–1. oktober 2019.



Figur 1. Kategorisering av de 12 identifiserte artiklene.

Notat. De 12 artiklene (øverste rad) ble kategorisert i fire kategorier (midterst rad) i henhold til grad av eksperimentell kontroll. De ble ytterligere kategorisert i henhold til innhold (nederste rad): sammenligning av Interteaching med annen undervisningsmetode (sammenligning) eller undersøkelse av effekten av ulike komponenter i Interteaching (komponenter).

I tillegg til de to overnevnte databasene ble *ERIC* også brukt. Litteratursøket ble gjennomført 1.10.2019. De samme inkluderingskriteriene og søkeord (inter-teaching) som Sturmey et al. ble benyttet. Et annet kriterium var at artiklene måtte være fagfellevurdert (dette betyr nødvendigvis ikke at tidsskriftene som artiklene er publisert i er oppført i det norske systemet; NSD publiseringskanaler). Artiklene som ble identifisert ble delt opp i følgende kategorier: (1) randomisert gruppedesign, (2) alternerende intervensjonsdesign, (3) kvasiekperimentelle design og (4) selvevaluering (se Figur 1).

Resultater

Litteratursøket førte til totalt 374 treff (*GoogleScholar* 355, *PsychINFO* 5, *ERIC* 14). Etter gjennomgang av artiklene gjensto 12 artikler fra de tre databasene som oppfylte inkluderingskriteriene beskrevet over (se Tabell 1). Artiklene var publisert i 10 ulike tidsskrifter for seks ulike utdanninger; psykologi (syv studier, derav tre atferdsanalyse) fysiologi, revisor, rehabilitering, eksperimentelle design og statistikk, og sykepleie (se Tabell 1). Tre av artiklene viste bruk av Interteaching i nettbaserte studier (se Tabell 1). I 11 av 12 artikler var studentene engelsktalende og i en artikkel var det studenter som hadde vietnamesisk som morsmål. Resultatene er kategorisert i henhold til designen benyttet i artiklene; randomisert gruppedesign, alternerende intervensjonsdesign, kvasiekperimentelle design og selvevaluering (se Figur 1).

Randomisert Gruppedesign

Kun én av de 12 artiklene som ble inkludert i litteratursøket brukte en randomisert fordeling av studenter i eksperimentell gruppe (Interteaching) og kontrollgruppe (undervisning som vanlig) som begge var nettbasert (Rieken et al., 2018). Studentene i begge gruppene fikk de samme ni oppgavene og fire quizer underveis i kurset, kategorisert i fire deler; prosjekter, kasusstudier, kritisk

analyse av artikler og quiz. En uavhengig hjelpelærer vurderte og ga karakterer. Resultatet på oppgavene og quizene ble sammenlignet mellom de to gruppene og viste en statistisk signifikant forskjell på tre av de fire delene (quiz, kritisk analyse av artikler og kasusstudier). For nærmere informasjon om skala og karakterresultater henvises leseren til artikkelen.

Alternerende Intervensjonsdesign

Av de 12 artiklene var alternerende intervensjonsdesign (*alternating treatment design*) benyttet i seks artikler (Felderman, 2016; Gayman et al., 2018; Rosales & Soldner, 2018; Rosales et al., 2014; Rosales et al., 2018; Soldner et al., 2017). Denne designen kan gi god informasjon om effekten av to eller flere undervisningsteknikker innen samme gruppe. Som vist i nederste rad i Figur 1 ble alternerende intervensjonsdesign brukt i to av studiene for å sammenligne effekten av bruk av Interteaching med en annen undervisningsteknikk (Felderman, 2016; Gayman et al., 2018). Figur 1, nederst rad, viser fire andre studier som undersøkte effekten av å endre ulike komponenter i Interteaching (Rosales & Soldner, 2018; Rosales et al., 2014; Rosales et al., 2018; Soldner et al., 2017).

Sammenligning av Ulike Teknikker

Felderman (2016) brukte en motbalansert (*counterbalanced*) alternerende intervensjonsdesign for å sammenligne effekten av Interteaching og tradisjonelle forelesninger på eksamenskarakter. Designen inkluderte to studentgrupper som vekslet mellom Interteaching (Gruppe 1) og tradisjonelle forelesninger (Gruppe 2) i løpet av semesteret. Ettersom det ikke var mulig å bruke en randomisert fordeling av studenter i gruppene ble en statistisk analyse av demografiske variabler gjennomført. Analysen viste at gruppene var like. Gruppe 1 fikk Interteaching mens Gruppe 2 fikk det samme materialet presentert via tradisjonell forelesning og motsatt i neste undervis-

Tabell 1. Resultat fra litteratursøket.

Forfatter(e)	År	Tittel	Tidsskrift	Studie
Felderman	2014	Preliminary analysis of Inter-teaching's frequent examinations component in the community college classroom	Journal of College Teaching & Learning*	Psykologi
Rosales, Soldner, & Crimando	2014	Enhancing the impact of quality points in Inter-teaching	Journal of the Scholarship and Teaching and Learning*	Psykologi (Atferdsanalyse)
Felderman	2016	A systematic replication comparing Inter-teaching and lecture in the community college classroom	Community College Journal of Research and Practice**	Psykologi
Byrne & Guy	2016	Inter-teaching within a human physiology course: A comparison of first- and second-year students' learning skills and perceptions	Advances in Physiology Education**	Fysiologi
Wheaton, O'Connell, & Yapa	2016	Inter-teaching: Improving the academic performance of auditing students in Vietnam	Australasian Accounting Business and Finance Journal and Authors**	Auditing (Vietnamese)
Soldner, Rosales, Crimando, & Schultz	2017	Inter-teaching: Application of an empirically supported behavioral teaching method in distance rehabilitation education	Rehabilitation Research, Policy, and Education*	Rehabilitering**
Slezak & Faas	2017	Effects of an Inter-teaching probe on learning and generalization of American Psychological Association (APA) style	Teaching of Psychology**	Eksperimentelle design og statistikk
Estridge & Owens	2018	Effectiveness of Inter-teaching with senior baccalaureate nursing students	Delta Kappa Gamma Bulletin*	Sykepleie
Rieken, Dotson, Carter, & Griffith	2018	An evaluation of Inter-teaching in an asynchronous online graduate-level behavior analysis course	Teaching of Psychology**	Psykologi (Atferdsanalyse)**
Rosales & Soldner	2018	An assessment of group size in Inter-teaching	Journal of the Scholarship of Teaching and Learning*	Psykologi
Gayman, Hammonds, & Rost	2018	Inter-teaching in an Asynchronous Online Class	Scholarship of Teaching and Learning in Psychology*	Psykologi**
Rosales, Soldner, & Zhang	2018	An Evaluation of the Pair Discussion Component of Inter-teaching	The Psychological Record **	Psykologi (Atferdsanalyse)

Notat. Artiklene er listet opp etter årstall. * Fagfelle-vurderte artikler, * Oppført i NSD databasen. ** Nettbasert kurs

ningsøkt. Begge gruppene fikk seks tester i løpet av semesteret, presentert etter hver økt, som kunne sammenlignes på tvers av gruppene. Eksamensspørsmålene var fra hele kurset og balansert slik det var like mange spørsmål fra hver undervisningsteknikk. Sensurering (tester og eksamen) ble gjennomført av en annen instruktør som ikke var kjent med hvilken undervisningsteknikk ble brukt. Resultatet viste at studentene fikk generelt bedre karakter på testene presentert etter *Inter-teaching* sammenlignet med tradisjonell forelesning, men det var kun statistisk forskjell på to av de seks testene. Videre viste resultatet liten forskjell mellom de to undervisningsteknikkene når det gjaldt eksamenskarakteren (ikke statistisk signifikant), men den forskjellen som fantes var i favør av *Inter-teaching*.

Gayman et al. (2018) brukte også en motbalansert alternerende intervensjonsdesign for å sammenligne *Inter-teaching* med vanlig undervisning i et nettbasert kurs. Undervisning foregikk i ni uker og som beskrevet av Felderman (2016) var det to ulike grupper som vekslet mellom tradisjonell forelesning og *Inter-teaching*. Resultatet viste en statistisk signifikant forskjell i karakter mellom de to undervisningsteknikkene der karakteren på tester var bedre for studentene etter sekvensene med *Inter-teaching* sammenlignet med forelesningssekvensene. Videre viste resultatet en statistisk signifikant forskjell i karakterer på eksamen med bedre karakterer på temaer som ble undervist ved bruk av *Inter-teaching* sammenlignet med tradisjonell forelesning. Forskjellene i karakterer påpekte forfatterne var relatert til at *Inter-teaching* fører til at studentene husker pensumet bedre sammenlignet med tradisjonell forelesning.

Analyse av Ulike Komponenter i *Inter-teaching*

Ekstrapoeng henviser til at hvis begge studentene i et par i en sekvens med *Inter-teaching* oppnår 80% korrekt på en posttest quiz presentert etter denne sekvensen, så

får studentene tildelt ekstra poeng som blir en del av eksamenskarakteren (jf. Boyce & Hineline, 2002). Rosales et al. (2014) undersøkte effekten av bruk av ekstrapoeng versus ingen ekstrapoeng på svarene på posttester. Forfatterne brukte en alternerende intervensjonsdesign i en klasse med 11 studenter. Presentasjon av de to betingelsene ble bestemt ved myntkast, men med den begrensningen at hver betingelse kun ble presentert to ganger på rad. Studentene fikk beskjed i forkant av undervisningsøkten om det var mulig å oppnå ekstrapoeng eller ikke. Resultatet viste at det å oppnå ekstrapoeng førte til en statistisk signifikant bedre karakter på posttestene sammenlignet med når det ikke var mulig å få slike ekstra poeng.

Soldner et al. (2017) undersøkte effekten av å diskutere spørsmålene i par versus at hver enkelt student jobbet med spørsmålene alene på quiz karakteren etter hver undervisningsøkt (totalt 13 quizer). Forfatterne brukte en alternerende intervensjonsdesign og myntkast for å bestemme hvilken betingelse som skulle presenteres. Resultatet for hele gruppa viste at det å jobbe sammen i par førte til statistisk signifikant bedre karakter sammenlignet med å jobbe alene. Videre analyse av hver enkelt students karakterer viste at 88% av studentene fikk bedre karakter etter å ha jobbet sammen i par sammenlignet med når de jobbet alene.

Rosales og Soldner (2018) undersøkte også effekten av diskusjonskomponenten i *Inter-teaching*. Det brukt en alternerende intervensjonsdesign der studentene ble eksponert for enten den ene betingelsen med par eller den andre med grupper (4–5 studenter). Hvilken betingelse som ble presentert til hver tid ble avgjort ved bruk av myntkast. Avhengig variabel var skår på quizer som ble presentert etter hver undervisningsøkt. Analysene av gjennomsnittsskår av alle quizene viste at diskusjon i par førte til litt høyere skår sammenlignet med å jobbe i grupper på 4–5 studenter; gjennomsnittskarakteren var B i par mens den var C i grupper. Analyse av individuelle skårer viste

at 28 av 35 studenter fikk bedre karakter etter å ha gjennomført diskusjon i par. Denne forskjellen var statistisk signifikant.

Diskusjonskomponenten ble undersøkt ytterligere av Rosales, et al. (2018). Forfatterne brukte en motbalansert alternerende intervensjonsdesign for å undersøke effekten på karakter på 10 quizer ved å sammenligne diskusjon i par versus i hele klassen. Resultatet viste at studentene fikk noe bedre karakter etter å ha jobbet sammen i par sammenlignet med en felles diskusjon i hele klassen, men forskjellen var ikke statistisk signifikant.

Kvasiekperimentelle Design

Fire eksperimenter brukte ikke randomisering i ulike grupper, såkalte kvasi-eksperimenter (se Figur 1). Tre av disse eksperimentene sammenlignet eksamenskarakter etter Interteaching med tradisjonell undervisning ved bruk av pre- og posttest, eller kun ved posttest (Estridge & Owens, 2018; Slezak & Faas, 2017; Wheaton et al., 2016). Ett av disse eksperimentene undersøkte effekten av ulikt antall tester i Interteaching (Felderman, 2014). I to av eksperimentene ble resultatet på eksamen sammenlignet fra to ulike studentgrupper i ulike kull (Estridge & Owens, 2018; Wheaton et al., 2016), mens i de to andre eksperimentene ble kullet delt opp i to parallelle klasser som fikk ulik undervisning (Felderman, 2014; Slezak & Faas, 2017).

Sammenligning av Ulike Teknikker

Estridge og Owens (2018) sammenlignet karakter på eksamen for to ulike kull, hvor tradisjonelle forelesninger ble brukt for det ene kullet og Interteaching for kullet som hadde oppstart et år senere. Resultatet viste en liten forbedring i karakterene, fra 84% til 88%; forskjell som var ikke statistisk signifikant.

Wheaton et al. (2016) sammenlignet eksamensresultater fra to ulike kull, et kull som hadde oppstart i 2011 og et annet som hadde oppstart i 2013. De uavhengige

variablene var tradisjonell forelesning og Interteaching. 2011 kullet fikk tradisjonelle forelesninger mens Interteaching ble brukt for 2013 kullet. Den avhengige variabelen var antall studenter som besto eksamen. Resultatet viste en signifikant forskjell mellom de to kullene der en større andel studenter besto ulike eksamener når Interteaching ble brukt sammenlignet med tradisjonelle forelesninger.

Slezak og Faas (2017) sammenliknet resultatet på en quiz etter at studenter hadde enten fått undervisning i bruk av APA style med Interteaching og tradisjonell forelesning. De to gruppene var sammenlignbare på ulike demografiske variabler. Resultatet viste at begge gruppene skåret statistisk signifikant bedre på posttest sammenlignet med pretest når analysen ble gjennomført for hver gruppe for seg. Det var ikke en statistisk signifikant forskjell mellom bruk av Interteaching og tradisjonell forelesning, men det nevnes at Interteaching gruppen fikk litt høyere karakter. To måneder senere fikk deltakerne en annen test for å undersøke langtidseffekten av undervisningen og da var det ingen forskjell mellom de to gruppene.

Analyse av Ulike Komponenter i Interteaching

I artikkelen til Felderman (2014) ble effekten av ulike antall tester på sluttkarakteren undersøkt. De 52 studenter som var påmeldt ble delt opp i to grupper. Demografiske variabler bekreftet at gruppene var like på mange forskjellige områder (for eksempel, alder, kjønn, antall år på skole). Begge gruppene fikk en pretest før undervisning startet. Gruppe 1 fikk seks tester i løpet av kurset, mens Gruppe 2 fikk 12 tester. Resultatet viste at Gruppe 2 fikk generelt litt bedre karakter på testene sammenlignet med Gruppe 1, men forskjellen var ikke statistisk signifikant. Gruppe 2 fikk også større endring fra pre- til posttest sammenlignet med pre- og posttest resultatet til Gruppe 1, men forskjellen var ikke statistisk signifikant.

Selvevaluering

Byrne og Guy (2016) studerte effekten av innføring av Interteaching i et kurs der studentgruppen var blanding av første og andre års studenter. Kurset ble undervist i fire moduler der Interteaching ble brukt i modul 2 og 4 mens tradisjonell forelesning ble brukt i modul 1 og 3. Forfatterne brukte selvevalueringsskjemaer der studentene ble bedt om å vurdere vanskelighetsgrad av kurset i tillegg til å sammenligne karakter på eksamen på tvers av studieår. Resultatet viste at første års studentene rapporterte at kurset var vanskeligere i starten av kurset sammenlignet med andre års studentene, som hadde den fordel å ha vært igjennom deler av pensumet tidligere i sitt studieløp. I tillegg rapporterte studentene lavere aktivitet sammenlignet med andre års studenter. Men i den andre sekvensen med Interteaching forsvant disse forskjellene og begge gruppene fikk sammenlignbare eksamensresultater på slutten av kurset.

Sosial Validitet

Sosial validitet ble undersøkt i seks av de 12 studiene. Felderman (2016) brukte et spørreskjema for å undersøke (a) hvilken undervisningsteknikk studentene likte best, (b) hvilken teknikk førte til størst læringsutbytte for studentene og (c) hvordan studentene likte foreleseren (fra 1 *veldig dårlig* til 7 *veldig god*). Resultatet viste liten forskjell av hvilken undervisningsteknikk studentene likte best; Interteaching (38%), ingen preferanse (32%), forelesning (30%). Ved evaluering av hvilken teknikk studentene mente de hadde hatt størst læringsutbytte (Interteaching, tradisjonell forelesning eller ingen forskjell) rapporterte 41% av studentene at de hadde hatt større læringsutbytte ved bruk av Interteaching dernerst ingen forskjell (36%) og til slutt tradisjonell forelesning. Faglæreren ble evaluert som *god* eller *veldig god* av 82% av studentene.

Studentene i Gayman et al. (2018) fikk et spørreskjema med 7-punkters Likert skala der de ble bedt om å evaluere (1) hvilken

undervisningsteknikk studentene likte best og (2) hvilken teknikk de synes de lærte mest av å bruke. Videre ble de bedt om å vurdere kvaliteten av (3) sekvensene med Interteaching, (4) sekvensene med standard forelesninger, (5) Interteaching spørsmålene, (6) diskusjonen, (7) oppfølgingsforelesningene etter sekvensene med Interteaching og (8) tradisjonelle forelesningene. Resultatet viste at studentene likte bedre Interteaching fremfor tradisjonell forelesning i det nettbaserte kurset og at de synes at de hadde større læringsutbytte av den teknikken sammenlignet med tradisjonell forelesning. Studentene rapporterte også bedre kvalitet i undervisning ved bruk av Interteaching.

Soldner et al. (2017) brukte også et spørreskjema for å undersøke hvilken fremgangsmåte studentene foretrakk når det gjelder diskusjoner, om de likte å jobbe i par, jobbe alene eller ingen preferanse. Resultatet viste at studenter likte bedre å ha en diskusjonspartnerne og mente at det også hadde ført til større læringsutbytte.

Rosales et al. (2018) undersøkte sosial validitet ved bruk av et spørreskjema. Studentene ble bedt om å vurdere om de likte bedre å diskutere spørsmålene i par, om de likte bedre å diskutere spørsmålene i hele klassen eller ingen preferanse. De ble også bedt om å vurdere hvilken av de to teknikkene de mente de hadde hatt størst læringsutbytte av eller om de ikke hadde noen preferanse. Resultatet viste at studentene likte bedre å ha diskusjoner i hele klassen og at de mente de hadde lært mer ved å diskutere spørsmålene i klassen.

På samme måte som den ovennevnte studien brukte Rosales og Soldner (2018) et spørreskjema for å undersøke om studenter likte bedre å diskutere spørsmålene i par eller å diskutere spørsmålene i en større gruppe med 4–5 studenter. Resultatet viste at til tross for at studentene fikk bedre karakter på quiz etter diskusjon av spørsmålene i par så rapporterte studentene at de foretrakk å jobbe i større grupper fremfor å jobbe to og to sammen. De rapporterte også at de mente

at de hadde fått større læringsutbytte når de jobbet i en større gruppe i motsetning til å jobbe i par.

Estridge og Owens (2018) brukte åpne spørsmål for å undersøke om studenter likte best Inter-teaching eller tradisjonell forelesning. De fant en forskjell der Inter-teaching gruppen rapporterte dypere forståelse av pensumet og mente den teknikken hadde ført til bedre studieferdigheter.

Diskusjon

Formålet med denne artikkelen var å gjennomføre en replikasjon av Sturme et al. (2015) for å identifisere nyere forskning om effektene av Inter-teaching, hvilke komponenter som er undersøkt og om teknikken er empirisk støttet. Det presenterte litteratursøket førte til identifisering av 12 internasjonalt publiserte artikler. Resultatet fra det presenterte litteratursøket viser at Inter-teaching er en godt egnet og effektiv studentaktiv undervisningsteknikk. Som beskrevet innledningsvis er det større fokus på studentaktive læringsformer i dag enn tidligere (Kunnskapsdepartementet, 2017; NKR, 2011; Prince, 2004). Et aspekt ved dette er at studentene skal tenke på det de leser som fører til såkalt dybdelæring av pensumet. Inter-teaching fører akkurat til slik atferd; studentene får tildelt spørsmål fra faglærere og skal diskutere viktige elementer av pensumet i par eller i større grupper og tenke dermed på det de har lest.

Vurdering av de Ulike Eksperimentene

Sturme et al. (2015) rapporterte to studier hvor det var brukt en randomisert fordeling av studenter i to grupper for å sammenligne Inter-teaching med andre undervisningsteknikker. Begge studiene viste en statistisk signifikant forskjell der Inter-teaching førte til høyere karakter sammenlignet med tradisjonell forelesning. I det presenterte litteratursøket fant vi en slik randomisert studie (Rieken et al., 2018). Resultatet viste at Inter-teaching førte til

statistisk signifikant bedre karakter på tre av fire oppgaver og tester som studentene fikk.

I en alternerende intervensjonsdesign er to eller flere intervensjoner sammenlignet med hverandre ved å veksle imellom teknikkene som skal brukes, for eksempel enten på ulike timer av dagen, ved bruk av ulike faglærere, eller ved introduksjon av ulike temaer. I denne artikkelen er designen benyttet for ulike grupper, men kan også benyttes når det er snakk om *single-case research* design (se for eksempel Kazdin, 2011 for nærmere beskrivelse av alternerende intervensjonsdesign). Sturme et al. (2015) rapporterte bruk av denne type design i tre studier mens seks av de 12 artiklene som ble identifisert i det presenterte litteratursøket benyttet en slik design (Felderman, 2016; Gayman et al., 2018; Rosales & Soldner, 2018; Rosales et al., 2014; Rosales et al., 2018; Soldner et al., 2017). To av disse sammenlignet Inter-teaching med andre undervisningsteknikker (Felderman, 2016; Gayman et al., 2018) og resultatene viste at Inter-teaching hadde den beste effekten.

Sturme et al. rapporterte to studier med bruk av kvasiekperimentelle design der resultatet viste en positiv effekt ved bruk av Inter-teaching. Resultatet i det presenterte litteratursøket viste at tre studier (Estridge & Owens, 2018; Wheaton et al., 2016; Slezak & Faas, 2017) brukte kvasiekperimentelle design for å sammenligne effekten av Inter-teaching med tradisjonell forelesning på pre- og posttest. Forfatterne i disse artiklene tok data på ulike demografiske variabler for å sikre at gruppene var sammenlignbare, samt redusere trusler mot indre validitet. Hovedresultatet fra alle artiklene viste at Inter-teaching førte til bedre karakter på eksamen, men forskjellen var kun statistisk signifikant i en av disse (Wheaton et al., 2016). I likhet med Sturme et al. sin konklusjon vil vi påpeke at resultatene fra disse studiene bør tas med forbehold ettersom den eksperimentelle designen er svak.

Mens ulike studier har sammenlignet effekten av Inter-teaching med andre under-

visningsteknikker (e.g., Felderman, 2016; Gayman et al., 2018; Rieken et al., 2018) var det flere som undersøkte ulike komponenter i Interteaching (Felderman, 2014; Rosales & Soldner, 2018; Rosales et al., 2014; Rosales et al., 2018; Soldner et al., 2017). Boyce og Hineline (2002) beskrev de ulike komponentene i Interteaching som for eksempel bruk av spørsmål, å jobbe i par, registrere og levere et notat, hyppige tester, gi ekstrapoeng for gode resultater på quiz, og at forelesning fokuserer på det studentene har identifisert som vanskelig eller ønsker ytterligere informasjon om. Spørsmålet er om alle disse komponentene er like viktige, samt hvor ofte man bør teste studentenes kunnskap. Er det for eksempel viktig å diskutere i par, å jobbe alene eller å diskutere i større grupper? Artiklene i det presenterte litteratursøket viste at det å diskutere spørsmålene i par førte til bedre karakter sammenlignet med det å jobbe alene eller gjennomføring av diskusjonene i større grupper. Videre viste søket at ekstrapoeng kontingent på gode quiz resultater for begge studentene i et par førte til høyere karakter på framtidige quiz (Rosales, et al., 2014).

Støttet av Empiri

En begrensning synes kanskje å være bruken av en såpass gammel publikasjon i vurdering av hva som kvalifiserer for empirisk støttet (Chambless & Hollon, 1998). En av grunnene til dette er at kvalifiseringssystemet ble benyttet av den artikkelen vi replikerte (Sturmey et al., 2015). For det andre brukes fremdeles dette systemet i dag for å vurdere hva som kvalifiserer for empirisk støttede behandlingsformer (Tolin et al., 2015). Tolin et al. har påpekt 11 omfattende forandringer som bør gjøres med dette kvalifiseringssystemet. Blant annet av vektlegging av oversiktsstudier framfor og enkeltstående studier, kvalitetsmessig gradering av studier, vurdering av klinisk signifikans i tillegg til statistisk signifikans, og vurdering av langtidseffekter i tillegg til kortidseffekter. Den tredje grunnen er at det ikke finnes noe tilsvarende kvalifise-

ringssystem for undervisningsteknikker, kun diverse oversikter over ulike strategier som er hevdet å være evidensbasert (Australian Society for Evidence Based Teaching, 2020).

En publikasjon som er interessant i vurdering av undervisningsteknikker er en studie gjort av Henderson et al. (2014). De spurte 73 faglærere i fysikk om hvorledes effekten av undervisning ble vurdert av dem selv og utdanningsinstitusjonene der de jobbet. Resultatene viste blant annet at institusjonene vurderte effektiviteten ut fra studentenes evaluering av undervisningen, mens faglærerne vurderte effektiviteten ut fra studentens utførelse på eksamen og usystematiske formative vurderinger (for eksempel resultater på flervalgstester). Få faglærere og utdanningsinstitusjoner brukte en praksis for å vurdere effektivitet av undervisning som var basert på forskningslitteratur. Forfatterne etterlyser et mer ensartet system for evaluering av effektiviteten ved undervisning og større vekt på en systematisk formativ vurdering.

Sosial Validitet

Kort beskrevet viser sosial validitet til at et gitt tiltak skal være gunstig og viktig for alle involverte parter (Foster & Mash, 1999; Schwartz & Baer, 1991; Wolf, 1979). Av den grunn må en se på sammenhengen mellom det å tilby undervisning som fører til god effekt, samt gi undervisning som evalueres som gunstig. Carter (2010) har argumentert for at selv om et behandlingsprogram fører til god effekt betyr det ikke at det nødvendigvis skal brukes. Det innebærer at dersom et program som er kjent for å være effektivt skal tas i bruk, så er det viktig at de som skal administrere det og brukerne evaluerer det som gunstig (brukermedvirkning). Sturmey et al. (2015) fant at det var en variasjon i hvor fornøyde studentene var med bruk av Interteaching, men at de som regel foretrakk Interteaching fremfor tradisjonell forelesning. Halvparten av studiene i det presenterte litteratursøket undersøkte sosial validitet og viser delvis det samme resultatet.

To av studiene (Estridge & Owens, 2018; Gayman et al., 2018) viste at studentene foretrakk Interteaching framfor tradisjonell forelesning og at de mente de fikk større læringsutbytte ved bruk av den teknikken, mens en studie (Felderman, 2016) viste liten forskjell i preferanse. Preferanse for Interteaching ble også funnet tidligere ved evaluering av Interteaching i Norge (e.g., Løkke et al., 2008). Vi mener et forsknings-spørsmål som er viktig å undersøke er hvilke variabler som påvirker at studenter foretrekker Interteaching framfor tradisjonell undervisning.

I tre av studiene (Rosales et al., 2018; Rosales & Soldner, 2018; Soldner et al., 2017) ble diskusjonskomponenten evaluert. Det som var interessant var at studentene rapporterte at de likte bedre å ha diskusjon i hele klassen fremfor i par. Videre at de likte bedre å jobbe i større grupper (4–5 studenter) i motsetning til å jobbe i par og at de likte bedre å jobbe i par enn å jobbe alene. Resultatene fra de ulike studiene viste derimot at det å jobbe i par førte til bedre karakter sammenlignet med den andre varianten i hver av de studiene. Rosales et al. (2018) fant at en av grunnene til at studentene likte større grupper fremfor å jobbe i par var at de ikke stolte faglig på medstudenten. De var bekymret for at den informasjonen de fikk fra medstudenten ikke var like god som fra en faglærer og i noen tilfeller var medstudent så dårlig forberedt at det påvirket diskusjonen i negativ retning.

Nettbasert Undervisning

Det er stadig flere som tilbyr nettbaserte kurs (Soldner et al. 2017) og som beskrevet i innledningen vil COVID-19 pandemien tvinge fram at mye undervisning blir basert på bruk av digitale plattformer. Men selv om nettbaserte kurs kan føre til større fleksibilitet for studentene er det viktig å følge med på læringseffektene. En nylig publisert studie viste at det å endre fra vanlig undervisning til nettbasert undervisning

førte til dårligere karakter hos studentene, større sannsynlighet for frafall og dårligere progresjon i videre studier (Bettinger et al., 2017).

Sturmey et al. (2015) fant ingen artikler der Interteaching ble brukt i nettbaserte studier, men de påpekte at slike studier var av stor interesse. Tre artikler i det presenterte litteratursøket undersøkte bruk av Interteaching i nettbaserte kurs (Gayman et al., 2018; Rieken et al., 2018; Soldner et al., 2017). Rieken et al. (2018) og Gayman et al. (2018) fant at det å bruke Interteaching i nettbasert undervisning førte til signifikant bedre karakter på fire av de avhengige variablene som ble undersøkt sammenlignet med tradisjonell undervisning. Soldner et al. (2017) viste også i sin studie at det å tilrettelegge for diskusjon i par i et nettbasert kurs førte til signifikant bedre karakter sammenlignet med å jobbe alene. Tilsammen viser disse tre studiene hvordan det er mulig å øke læringsutbyttet i nettbaserte studier.

Implikasjoner

I det presenterte litteratursøket var bruk av ekstrapoeng på karakter undersøkt i en studie (Rosales et al., 2014). Erfaringsmessig er slik bruk av ekstrapoeng, så vidt vi kjenner til, ikke inkludert i sekvenser med Interteaching i Norge (e.g., Arntzen & Hoiium, 2010; Løkke et al., 2008; Equivalence, 2020). Når læringsutbytter vurderes kan faglærere benytte enten formative eller summative vurderinger. Formative vurderinger brukes for å gi tilbakemelding på studentenes atferd underveis, mens summative vurderinger brukes for å evaluere hvor mye kunnskap studenten har tilegnet seg i slutten av kurset (Dixon & Worrell, 2016). Innføring av ekstrapoeng kontingent på prestasjoner på quizer kan dermed brukes som en type formativ vurdering. Studenten vil da få kontinuerlig tilbakemelding på egen presentasjon i løpet av kurset. Studentens atferd bringes under forsterkningsbetingelser der tiden mellom atferd utføres og konsekvens presenteres blir kortere, sammenlignet

med når summative vurderingene brukes der tiden mellom atferd utføres og konsekvens presenteres kan være veldig lang (Boyce & Hineline, 2002). For å øke verdien på den umiddelbare konsekvensen kan testresultater på quizer også bli en del av karakteren ved kurssets slutt. Dermed blir studentens atferd under større påvirkning av både umiddelbare og utsatte konsekvenser. I tillegg bør det undersøkes hvorvidt antall poeng har en betydning for utførelse på quizene. Hvor mange poeng bør gis for å få frem ønsket effekt?

Videre studier bør se på forholdstallet mellom faglærer og studenter i Interteaching. Vil det være slik at forholdstallet mellom faglærer og antall studenter i en sekvens med Interteaching påvirker læringsutbytte og karakterer på eksamen? Og i tilfelle hvordan? Hvor hyppig bør interaksjonen være mellom faglærer og studenter for få større læringsutbytte for studenten? Forholdstallet kan selvsagt også være en interaksjon med andre variabler som for eksempel lengden på sekvensen diskusjoner.

Diskusjonskomponenten i sekvensene med Interteaching ble undersøkt i tre av artiklene (Rosales & Soldner, 2018; Rosales et al., 2018; Soldner et al., 2017). Vi foreslår også at det gjennomføres en studie der de ulike sammensetningene (å jobbe alene, i par, grupper på 4-5 studenter, hele klassen) undersøkes ytterligere innen en studentgruppe, hvor studentene blir utsatt for ulike betingelser.

Oppsummering

Vi fant 12 artikler om Interteaching i en femårs periode etter den forrige oversiktsstudien av Sturmey et al. (2015). Til sammen er det 10 studier som anvender en eksperimentell design for å avdekke effekten av denne undervisningsteknikken (syv beskrevet av Sturmey et al. (2015) og tre i det presenterte litteratursøket). Av disse 10 studiene er Interteaching sammenlignet med tradisjonell forelesning og resultatene fra ni av 10 artikler viser at Interteaching fører til bedre

resultater, selv om forskjellene ikke alltid er statistisk signifikante. Ifølge det systemet som er brukt for å vurdere om en intervensjon eller behandling er empirisk støttet eller ikke (Chambless & Hollon, 1998), fant Sturmey et al. (2015) at Interteaching som undervisningsteknikk var innen kategorien *lovende*. Våre funn indikerer at denne klassifiseringen bør endres slik at Interteaching er en empirisk støttet teknikk i henhold til det nevnte systemet.

Referanser

- Australian Society for Evidence Based Teaching. (2020, Juni 12). *10 evidence-based teaching strategies – the core list*. <https://www.evidencebasedteaching.org.au/evidence-based-teaching-strategies>
- Arntzen, E., & Hoiu, K. (2010). On the effectiveness of interteaching. *The Behavior Analyst Today, 11*, 153–160. doi: 10.1037/h0100698
- Barr, R. B., & Tagg, J. (1995). From teaching to learning — A new paradigm for undergraduate education. *Change: The Magazine of Higher Learning, 27*, 12–26. doi:10.1080/00091383.1995.10544672
- Bettinger, E. P., Fox, L., Loeb, S., and Taylor, E. S. (2017). Virtual classrooms: How online college courses affect student success. *American Economic Review, 107*, 2855–2875. doi: 10.1257/aer.20151193
- Boyce, T. E., & Hineline, P. N. (2002). Interteaching: A strategy for enhancing the user-friendliness of behavioral arrangements in the college classroom. *The Behavior Analyst, 25*, 215–226. doi:10.1007/BF03392059
- Byrne, B., & Guy, R. (2016). Interteaching within a human physiology course: a comparison of first- and second-year students' learning skills and perceptions. *Advances in physiology education, 40*, 349–353. doi:10.1152/advan.00141.2015
- Carter, S. L. (2010). *The social validity manual: A guide to subjective evaluation*

- of behavioral interventions. Boston, MA: Academic Press.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting Clinical Psychology*, 66, 7–18. doi:10.1037/0022-006X.66.1.7
- Dixon, D. D., & Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory Into Practice*, 55(2), 153–159. doi:10.1080/00405841.2016.1148989
- Estridge, K. M., & Owens, J. (2018). Effectiveness of Interteaching with Senior Baccalaureate Nursing Students. *Delta Gamma Kappa Bulletin*, 84, 14–25.
- Equivalence. (2020, Juni 12). *Lectures*. <https://equivalence.net/wp/lectures/>
- Felderman, T. A. (2014). Preliminary analysis of Interteaching's frequent examinations component in the community college classroom. *Journal of College Teaching and Learning*, 11, 149–156. doi:10.19030/tlc.v11i4.8851
- Felderman, T. A. (2016). A systematic replication comparing Interteaching and lecture in the community college classroom. *Community College Journal of Research and Practice*, 40, 739–749. doi:10.1080/10668926.2015.1075445
- Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting Clinical Psychology*, 67, 308–319.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 8410–8415. doi:10.1073/pnas.1319030111
- Gayman, C. M., Hammonds, F., & Rost, K. A. (2018). Interteaching in an asynchronous online class. *Scholarship of Teaching and Learning in Psychology*, 4, 231–242. doi:10.1037/std0000126
- Granmo, S., Halvorsen, L. R. & Løkke, J. A. (2017). Intraverbaler som de viktigste byggeklossene i en atferdsanalytisk akademisk pedagogikk. *Norsk Tidsskrift for Atferdsanalyse*, 44, 87–92. Hentet fra: <https://nta.atferd.no/>
- Halpern, D. F., & Hakel, M. (2003). Applying the science of learning to the university and beyond: Teaching for long-term retention and transfer. *Change: The Magazine of Higher Learning*, 35, 36–41. doi:10.1080/00091380309604109
- Halvorsen, L. R., Løkke, J. A., & Granmo, S. (2018). Studentaktiv forskning som virkemiddel for å utdanne profesjonsutøvere som også kan atferdsanalyse. *Norsk Tidsskrift for Atferdsanalyse*, 45, 87–100. Hentet fra: <https://nta.atferd.no/>
- Henderson, C., Turpen, C., Dancy, M., & Chapman, T. (2014). Assessment of teaching effectiveness: Lack of alignment between instructors, institutions, and research recommendations. *Physical Review Special Topics - Physics Education Research*, 10, 010106. <https://doi.org/10.1103/PhysRevSTPER.10.010106>
- Kazdin, A. E. (2011). Single-case research designs: Methods for clinical and applied settings. New York: Oxford University Press.
- Kunnskapsdepartementet. (2017). Kultur for kvalitet i høyere utdanning 2016–2017 (Stortingsmelding 16). Oslo: Undervisningsdepartementet
- Løkke, J., Løkke, G., & Arntzen, E. (2008). Er samlæring (interteaching) i store klasser mer effektivt en tradisjonelle forelesninger? En effektstudie. *Norsk Tidsskrift for Atferdsanalyse*, 35, 131–139. Hentet fra: <https://nta.atferd.no/>
- Matheson, C. (2008). The educational value and effectiveness of lectures. *The Clinical Teacher*, 5, 218–221. doi:10.1111/j.1743-498X.2008.00238.x
- Norges Teknisk-Naturvitenskapelig Universitet. (2020, Juni 12). *Koronavirus (COVID-19): Retningslinjer og råd*.

- <https://www.ntnu.no/korona>
- NKR (2011). Nasjonalt kvalifikasjonsrammeverk for livslang læring (NKR). Oslo: Kunnskapsdepartementet
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*, 223–231. doi:10.1002/j.2168-9830.2004.tb00809.x
- Rieken, C. J., Dotson, W. H., Carter, S. L., & Griffith, A. K. (2018). An evaluation of Interteaching in an asynchronous online graduate-level behavior analysis course. *Teaching of Psychology*, *45*, 264–269. doi:10.1177/0098628318779275
- Rosales, R., & Soldner, J. L. (2018). An assessment of group size in Interteaching. *Journal of the Scholarship of Teaching and Learning*, *18*, 105–117. doi:10.14434/josotl.v18i2.22539
- Rosales, R., Soldner, J. L., & Crimando, W. (2014). Enhancing the impact of quality points in Interteaching. *Journal of the Scholarship of Teaching and Learning*, *14*, 1–11. doi:10.14434/josotlv14i5.12746
- Rosales, R., Soldner, J. L., & Zhang, L. (2018). An evaluation of the pair discussion component of Interteaching. *The Psychological Record*, *68*, 71–79. doi:10.1007/s40732-018-0269-0
- Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis*, *24*, 189–204. doi:10.1901/jaba.1991.24-189
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, New York: Houghton Mifflin Company.
- Slezak, J. M., & Faas, C. (2017). Effects of an Interteaching probe on learning and generalization of American Psychological Association (APA) Style. *Teaching of Psychology*, *44*, 150–154. doi:10.1177/0098628317692619
- Soldner, J. L., Rosales, R., Crimando, W., & Schultz, J. (2017). Interteaching: Application of an empirically supported behavioral teaching method in distance rehabilitation education. *Rehabilitation research, Policy, and Education*, *31*, 372–386. doi:10.1891/2168-6653.31.4.372
- Sturmey, P., Dalfen, S., & Fienup, D. M. (2015). Inter-teaching: A systematic review. *European Journal of Behavior Analysis*, *16*, 121–130. doi:10.1080/15021149.2015.1069655
- Task Force on Promotion and Dissemination of Psychological Procedures. (1993). Training in and dissemination of empirically-validated psychological treatments: Report and recommendation. *The Clinical Psychologist*, *48*, 3–23.
- Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D., & Thombs, B. D. (2015). Empirically Supported Treatment: Recommendations for a New Model. *Clinical Psychology: Science and Practice*, *22*(4), 317–338. Doi: 10.1111/cpsp.12122
- Wheaton, M., O'Connell, B., & Yapa, P. (2016). Inter-teaching: Improving the Academic Performance of Auditing Students in Vietnam. *Australasian Accounting, Business Finance Journal*, *10*, 3–16. doi:10.14453/aabfj.v10i4.2
- 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*, 203–214. doi:10.1901/jaba.1978.11-203

Use of Interteaching in education: A systematic review of internationally published articles (2014–2019)

Hanna Steinunn Steingrimsdottir and Erik Arntzen
OsloMet – Oslo Metropolitan University

Interteaching is a pedagogy technique where students work actively with the reading material during lectures. The technique is based on well-known behavior analytic principles. In an earlier literature review, Sturmey et al. (2015) identified seven articles employing experimental design where Interteaching was compared to traditional lectures. The effect of Interteaching was generally good and the authors concluded that this technique is a promising pedagogy technique. The purpose of the current study was to identify recent studies on Interteaching, published from 2014 to 2019. The result lead to in total 12 articles whereof randomized group design was used in one article, alternating treatment design in six articles, four quasi-experimental articles and one article using self-evaluation. Even though the results showed only statistically significant difference in two of the three studies, the results were generally in favour of Interteaching. Three studies evaluated the effect of using Interteaching in online studies. The conclusion is that Interteaching is empirically supported.

Keywords: Interteaching, student-active learning, teaching, experimental design, social validity, higher education, review paper.