

Fagfelleurderte artikler

Av
Henrik Herrebrøden

Tester og testtilbakemeldinger som direkte bidragsyttere til dypere læring blant studenter

Henrik Herrebrøden
Universitetet i Bergen,
Psykologisk fakultet,
Bergen, Norway.
E-post:
[h.herrebroden@
hotmail.com](mailto:h.herrebroden@hotmail.com)

Sammendrag

Denne artikkelen diskuterer forskning som tilsier at tester og testtilbakemeldinger kan bidra til dypere læring blant studenter. Testing har vist seg å forbedre langtidshukommelsen av innlært kunnskap sammenlignet med repetert lesing, noe som kalles for testeffekten. For at testeffekten skal optimaliseres, bør tester utformes med mål om å bidra til dyp prosessering av innlært kunnskap. Testtilbakemeldingers rolle i utdanning kan være å bidra til økt metakognitiv selvregulering blant studenter. For at selvregulering skal promoteres best mulig, bør testtilbakemeldinger inneholde utdypende informasjon. Samlet ser tester og testtilbakemeldinger ut til å være nyttige læringsverktøy og potensielle bidragsyttere til dypere læring blant studenter.

Nøkkelord: Test, testtilbakemelding, dyp læring, dyp prosessering, metakognisjon, testeffekten

Abstract

This article discusses research supporting that tests and test feedback are learning tools that may promote deeper learning among students. Testing improves long-term retention of knowledge compared to repeated study, a phenomenon known as the testing effect. In order to promote the testing effect, tests should facilitate deep processing of knowledge among students. Test feedback has the potential to increase students' metacognitive self-regulation. For this type of metacognition to be enhanced, feedback should contain elaborate information. Conclusively, tests and test feedback appear to be valuable learning tools and potential promoters of deeper learning.

Keywords: Test, test feedback, deep learning, deep processing, metacognition, testing effect

Innledning

Vurdering av studenter har vært et omdiskutert tema i forbindelse med kvalitetsreformen innen utdanning (f.eks. Dysthe, Raaheim, Lima & Bygstad, 2006). Sentrale spørsmål er hvordan og ikke minst

hvorfor studenter bør vurderes. I forbindelse med disse spørsmålene har det tradisjonelt blitt satt et hovedskille mellom summative og formative vurderingsfunksjoner (Dysthe et al., 2006; Johnson & Beal, 2005; Lauvås, Havnes & Raaheim, 2000). Summative vurderinger kan i hovedsak beskrives som vurdering *av* læring, der vurderingen skal fungere som et mål på hva studenter tidligere har lært. Formative vurderinger kan på sin side beskrives som vurdering *for* læring, der vurderingen skal bistå og forme fremtidig læring. I senere tid har formative aspekter ved vurderinger fått mer oppmerksomhet, og bruken av vurderinger har vist seg å kunne ha positive effekter på innsatsen og læringsprosessen blant studenter (Dysthe et al., 2006; Rød, 2012; Rød, Eiksund & Fjær, 2010). Jeg mener at fokuset på hvordan vurderinger kan bidra positivt i utdanningen er meget interessant. I den forbindelse savner jeg imidlertid et fokus på de mer direkte læringsfordelene ved bruk av vurderinger – uavhengig av om vurderingene er tiltenkt en summativ eller formativ rolle i utdanningen. Hvilken lærdom tar studentene med seg fra selve vurderingene? Kan vurderinger gi læring i seg selv, og ikke bare fungere som virkemidler for å evaluere fortidens læring eller forme fremtidens læring?

I denne review-artikkelen ønsker jeg å benytte et kognitivt perspektiv og argumentere for at vurderinger kan være direkte og virkningsfulle læringsverktøy. Med vurderinger mener jeg – i denne sammenheng – bruk av tester og testtilbakemeldinger. For å begrense artikkelens omfang og tydeliggjøre hvilke tester jeg snakker om, begrenser jeg bruken av ordet *test* til å inkludere prøver og andre kunnskapstester der studenten må demonstrere kunnskap uten å ha tilgang til kunnskapskilden (f.eks. flervalgsprøve uten tilgang til læreboken eller forelesningsnotatene). Med *testtilbakemelding* mener jeg enhver tilbakemelding (f.eks. skriftlig tilbakemelding) som adresserer en testrespons, og som gir en evaluering av utførelsen av oppgave(n) i testen. Tradisjonelt har slike vurderinger fått lite fokus som positive læringsverktøy, men snarere blitt fremstilt som evaluering av tidligere læring, som i verste fall kan ha negativ påvirkning på læringsprosessen (f.eks. Crooks, 1988; Hattie & Timperley, 2007; Roediger & Karpicke, 2006a). I senere år har imidlertid empirien bidratt til at tester og testtilbakemeldinger også anses som positive læringsverktøy innen utdanning (for sammendrag, se Hattie & Timperley, 2007; Roediger & Karpicke, 2006a). Denne artikkelen vil se nærmere på slik empiri og kognitiv teori som utforsker hvordan studenter kan oppnå *dyp læring* når de blir vurdert. Først vil jeg forklare nærmere hva dyp læring innebærer, før jeg ser nærmere på hvordan dette kan promoteres gjennom bruk av tester og testtilbakemeldinger.

Dyp kontra overfladisk læringstilnærming

Kontrasten mellom overfladiske og dype læringstilnærminger ble identifisert på 1970-tallet, da forskere undersøkte sammenhenger mellom ulike tilnærminger til læring og læringsutbytte (f.eks. Biggs, 1979; Entwistle, Hanley & Hounsell, 1979; Marton & Säljö, 1976). Biggs (1979) var en av de første som rapporterte om studenters studievaner. Han la merke til at enkelte studenter var mer opptatt av å memorere fakta enn å foreta en kritisk tilnærming til stoffet – en metode som gjenspeiler *overfladisk læringstilnærming*. Andre studenter gikk derimot dypere ved å analysere meningen bak ulike begreper og hvordan begreper var

relatert til hverandre – en fremgangsmåte forbundet med *dyp læringstilnærming*. I senere tid har slike læringstilnærminger blitt fremhevet som et kontinuum fremfor en dikotomi, der studenter kan vise ulik grad av strategier til ulike tider (se tabell 1). To kjennetegn har imidlertid blitt fremhevet som prevalente hos studenter som ofte befinner seg i den dypere enden av læringskalaen: dyp prosessering og metakognisjon (f.eks. Chin & Brown, 2000).

Tabell 1. Læringstilnærminger (reprodusert fra Kember, 1996, s. 347: «Table 1. The characteristics of discussed learning approaches as a function of intention and strategy components»)

Approach	Intention	Strategy
Deep	Understanding	Seeking comprehension
Intermediate 1	Primarily understanding	Strategic memorisation for examination or task after understanding reached
Understanding and memorising	Understanding and memorising	Seeking comprehension then committing to memory
Intermediate 2	Primarily memorising	Strategic attempt to reach limited understanding as an aid to memorisation
Surface	Memorising with-out understanding	Rote-learning

Note: Fra «The intention to both memorise and understand: Another approach to learning?» av D. Kember, 1996, *Higher Education*, 31(3), s. 347. Opphavsrett 2014 av Springer Science + Business Media. Reprodusert med tillatelse.

Dyp prosessering omhandler måten informasjon innhentes og lagres på, og er et element i en av de mest sentrale modellene innen hukommelsesteori: *the levels-of-processing approach* (Craik & Lockhart, 1972; Lockhart & Craik, 1990). Denne teorien ble formet på en tid da Atkinson og Shiffrins (1968) hukommelsesteori var dominerende – en teori som promoterte repetisjon som det viktigste elementet for læring. Craik og Lockhart (1972) vektla derimot at det ikke bare er antallet repetisjoner, men også *kvaliteten* i repetisjonene, som avgjør læringsutbyttet. Det spiller med andre ord mindre rolle om studenten leser pensum flere ganger, ifølge Craik og Lockhart, dersom han eller hun ikke prosesserer informasjonen dypt nok.

Overfladisk prosessering kjennetegnes ved at individet fokuserer på overfladiske, ofte sensoriske egenskaper ved stimuli (Craik & Lockhart, 1972). Dyp prosessering skjer derimot ved at individet fokuserer på de *semantiske aspektene* ved stimuli og at det er meningen bak begreper som er i fokus. Det er dette som skjer når studenten analyserer lesestoffet for mening, og prøver å forstå begreper og relatere dem til tidligere innlært kunnskap. Craik og Lockharts hypotese var at jo flere semantiske aspekter ved en stimulus som blir analysert, desto flere slike aspekter vil lagres sammen med stimulusen i langtidsminnet. I en akademisk kontekst betyr dette at minnet om et begrep blir styrket og utvidet, og mer tilgjengelig i det lange løp. I ettertid har Craik og Lockharts teori blitt både kritisert og modifisert (f.eks. Baddeley, 1978; Lockhart & Craik, 1990). Et resultat av dette er at prosesseringsdybde – i likhet med læringsdybde – kan anses som mer kompleks enn en

enkel dikotomi (Craik, 2002). Antakelsen om at dypere prosessering gir bedre langtidshukommelse enn overfladisk prosessering, har imidlertid blitt bekreftet i en rekke studier (f.eks. Craik & Tulving, 1975; Kapur et al., 1994; Schallert, 1976; Toichi & Kamio, 2002).

Den andre viktige komponenten av dyp læringstilnærming, *metakognisjon*, omhandler vår tenkning om egen tenkning og kunnskap, og hvordan vi forholder oss til dette (Paris & Winograd, 1990). Denne kognitive egenskapen består av to viktige komponenter som begge er gjensidig avhengige av hverandre, og fremtredende hos mange flittige studenter: metakognitiv overvåking og metakognitiv selvregulering. *Metakognitiv overvåking* (monitoring) er prosessen der studenten tilegner seg kunnskap om sin egen tenkning og kunnskap (Butler & Winne, 1995; Metcalfe & Shimamura, 1994). I tillegg til å overvåke egen tenkning og kunnskap direkte, kan studenten også overvåke ytre informasjon som sier noe om disse kognitive elementene, som f.eks. en tilbakemelding ved en eksamen (Bandura, 1989; Butler & Winne, 1995). Overvåkingen er det som tillater studenten å foreta en *metakognitiv selvregulering*, hvor studenten kan endre sin kognisjon og rette på kognitive feil (Butler & Winne, 1995). Studenten med høy metakognitiv aktivitet er med andre ord ofte en selvdreven student, ved at han eller hun evaluerer egen tenkning og kunnskap, samt er i stand til å korrigere seg selv og søke nødvendig informasjon og støtte.

Oppsummert ser dyp prosessering og metakognisjon ut til å være verdifulle egenskaper hos studenter, og viktige elementer i en dyp læringstilnærming. Interessant nok har vurderinger ofte blitt forbundet med overfladisk læring (f.eks. Entwistle, 1983). Det har blitt påpekt at tester kan oppmuntre til overfladisk prosessering av isolert faktakunnskap fremfor dyp forståelse (Crooks, 1988), mens testtilbakemeldinger kan brukes for å sammenligne seg selv med andre fremfor å drive selvregulering av egen kunnskap (Hattie & Timperley, 2007). Selv om dette kan være tilfellet for en del studenter, er det verdt å påpeke at tester og tilbakemeldinger i økende grad også har blitt fremhevet som bidragsytere til dypere læring (f.eks. Cox & Clark, 1998; Roediger & Karpicke, 2006a; Rushton, 2005). De neste seksjonene vil se nærmere på forskning som underbygger dette.

Testers bidrag til dypere læring: Testeffekten

Spitzer (1939) var en av de første som demonstrerte at tester kan være læringsverktøy. Hans studie viste at skoleelever som gjennomgikk en flervalgstest om innholdet i en lest artikkel, husket langt mer av innholdet ved en senere test enn elever som ikke ble testet. I dag blir et slikt positivt læringsutbytte kalt for *testeffekten* (f.eks. McDaniel, Anderson, Derbish & Morrisette, 2007). Mange potensielle forklaringer på testeffekten har blitt foreslått siden Spitzers studie (McDaniel & Masson, 1985; Roediger & Karpicke 2006a). En løsning er å bruke Atkinson-Shiffrin-modellen til å forklare at testing er positivt fordi det gir elevene en gjennomgang av innlært kunnskap (Thompson, Wenger & Bartling, 1978). Testing blir her bare sett på som en repetisjon som – i likhet med lesing – gir læring. Denne forklaringen virker imidlertid usannsynlig når man ser nærmere på studier som sammenligner effektene av testing og lesing (f.eks. Butler, 2010; Roediger & Karpicke, 2006b).

Roediger og Karpicke (2006b, Eksperiment 2) gjennomførte en studie der de sammenlignet effektene av lesing og testing i forbindelse med læring av innholdet i en tekst. Etter å

ha først lest teksten i en leseperiode, ble de 180 deltakerne i studien delt inn i tre grupper. Gruppe 1 fikk lese teksten i tre ytterligere leseperioder, gruppe 2 fikk lese teksten i to perioder og deretter ta en test, mens gruppe 3 ble testet tre ganger om innholdet i teksten. Testene ble gjennomført ved at testtakerne ble bedt om å skrive ned så mye de husket av teksten innen en ti minutters tidsfrist. Før læringsutbyttet skulle måles gjennom avsluttende tester, ble deltakerne spurt om hvor godt de trodde de ville huske teksten når de ble testet en uke senere. Gruppe 1, med deltakere som kun leste, var klart mer optimistisk enn de to andre gruppene.

Først ble halvparten av deltakerne i de ulike gruppene testet, fem minutter etter siste lese- eller testperiode. Her gjorde deltakere fra gruppe 1 det best (83 % av innholdet i teksten ble gjengitt på en «korrekt måte» (etter sensors bedømming)), fulgt av deltakere fra gruppe 2 (78 %) og 3 (71 %). Når de resterende deltakerne ble testet en uke senere, ble derimot rangeringen reversert. Her gjorde deltakerne fra gruppe 3 det best (61 %), fulgt av deltakere fra gruppe 2 (56 %) og gruppe 1 (40 %). Ved å sammenligne lesegruppen (gruppe 1) og den hyppigst testede gruppen (gruppe 3), ble det tydelig at lesing ga signifikant bedre hukommelse av kunnskapen fem minutter etter at teksten ble gjennomgått for siste gang, mens hyppig testing ga signifikant bedre hukommelse en uke senere. Det så altså ut til at testing ga bedre læringsutbytte på sikt – men mindre tillit til egen kunnskap – sammenlignet med lesing. En lignende testeffekt har blitt dokumentert i en rekke andre studier (f.eks. Butler 2010; Butler & Roediger, 2007; Karpicke & Roediger, 2008; Marsh, Roediger, Bjork & Bjork, 2007; McDaniel, Roediger & McDermott, 2007).

Vedrørende Roediger og Karpickes (2006b) studie, er det verdt å påpeke at læringsutbyttet ble målt gjennom avsluttende tester som var like de innledende testene som testgruppene tidligere hadde gjennomført. Testgruppene kan dermed ha dratt nytte av treningen de fikk med det aktuelle testformatet, og de kognitive kravene det stiller (Morris, Bransford & Franks, 1977). Likevel ser det ut til at læringsutbyttet fra tester ikke er begrenset til andre lignende tester, men at kunnskap fra én test kan overføres til flere ulike testformater (McDaniel, Anderson et al., 2007). Enda mer lovende, sett med akademiske øyne, er studier som indikerer at testing gir bedre overføring av lært kunnskap til nye kontekster, sammenlignet med repetert lesing (f.eks. Butler, 2010; Carpenter, 2012; Larsen, Butler, Lawson & Roediger, 2012). Det er imidlertid verdt å påpeke at mange studenter også kan oppnå godt læringsutbytte gjennom repetert lesing, og at utbytte av lesing varierer blant studenter (f.eks. Dempster, 1989; Hermida, 2009). Dette forandrer likevel ikke det faktum at tester later til å kunne være positive innslag i læringsprosessen. En forklaring på hvordan testing kan bidra positivt til langtidshukommelse og overføring av kunnskap, ser ut til å ligge i studenters prosessering av kunnskap under tester.

Repetert lesing kontra testing: Illusjon av kompetanse kontra dyp prosessering
Hvorfor kan repetert lesing gi overdreven optimisme over egen kunnskap i forhold til det reelle læringsutbyttet (f.eks. Roediger & Karpicke, 2006b)? En todelt forklaring kan bidra til en forståelse av det omvendte forholdet mellom forventet og reelt læringsutbytte, som ofte observeres i forbindelse med lesing. For det første kan en repetert lesing være ineffektiv om

den ikke promoterer en dyp forståelse av teksten (Callender & McDaniel, 2009). Repetert lesing kan da lede til at de samme assosiasjonene mellom teksten og dens innhold blir repetert, uten dypere prosessering (Callender & McDaniel, 2009; Fritz, Morris, Bjork, Gelman & Wickens, 2000). For det andre vil ordene i teksten vekke minner om kunnskap som har blitt dannet ved tidligere lesing av samme tekst, noe som gjør at kunnskapen ofte virker mer og mer kjent for hver gang teksten leses (Kornell, Castel, Eich & Bjork, 2010). Mange studenter tror at denne følelsen av å gjenkjenne kunnskap betyr at de forstår teksten på en dyp måte, og at de vil være i stand til å demonstrere god kunnskap ved en senere test når de ikke har den originale teksten foran seg (Koriat & Bjork, 2005; Nelson & Narens, 1990). En slik *illusjon av kompetanse* medfører at mange studenter tyr til overfladisk prosessering når de gjentar lesing av teksten, ettersom de tror at de allerede kan det de leser om (Greene, 1989; Jacoby, 1991; Koriat & Bjork, 2005; Son, 2004). Som vist i studien til Roediger og Karpicke (2006b), vil imidlertid mange av disse studentene oppdage at kunnskapen ikke er like lett å hente frem fra langtidshukommelsen når hjelpen fra ordene i originalteksten ikke er til stede.

Når ledetrådene fra ordene i teksten fjernes ved testing, må studenten derimot prosessere dypere for å hente inn relevant informasjon fra langtidshukommelsen (Carpenter & DeLosh, 2006; Jacoby, 1991). Studenten som mottar et testspørsmål om et begrep, vil ofte analysere dets mening og identifisere assosierte begreper i hukommelsen i et forsøk på å hente relevant kunnskap inn i bevisstheten (McDaniel & Masson, 1985). Denne semantiske analysen kan medføre to positive konsekvenser. For det første vil studenten ofte hente inn den relevante kunnskapen han trenger for å besvare spørsmålet på testen, noe som kan styrke det opprinnelige minnet om kunnskapen som ble dannet ved innlæring (McDaniel & Masson, 1985). For det andre vil en semantisk analyse kunne medføre at flere semantiske aspekter inkluderes i minnet om det aktuelle begrepet, og at det dermed skjer en semantisk utvidelse av det opprinnelige minnet (Craik & Lockhart, 1972; Runquist, 1982). Fordelen med denne utvidelsen er at de nye semantiske aspektene, som nå er assosiert med begrepet, også vil kunne fungere som ledetråder, og således trigge kunnskapen om begrepet i nye situasjoner der disse semantiske aspektene er til stede (Glenberg, 1979). Samlet vil en styrking og utvidelse av minner kunne medføre at studenten kan innhente relevant kunnskap fra langtidshukommelsen i flere situasjoner, og på tvers av kontekster, etter at han eller hun har blitt testet (Butler, 2010; Cooper & Monk, 1976; McDaniel & Masson, 1985).

Den beste måten å teste studenter på

Som beskrevet ovenfor kan dyp prosessering være en sentral årsak til at tester bidrar til dypere læring. Når man skal tilrettelegge for best mulig testing, vil det derfor være særlig ett spørsmål som gjør seg gjeldende: Hvordan kan tester i størst mulig grad bidra til dyp prosessering? En nøkkel ser ut til å ligge i formatet av testingen og kravene som stilles til hukommelsen.

Svært ofte blir flervalgsoppgaver brukt ved tester (f.eks. Karpicke & Roediger, 2008; Marsh et al., 2007; Spitzer, 1939). Selv om disse testene også kan bidra til bedre læring, ser det ut til at testeffekten ikke utnyttes optimalt under dette formatet (Butler & Roediger, 2007). Svaralternativene fungerer nemlig som ledetråder, og en godt forberedt student vil ofte ikke

ha store problemer med å gjenkjenne rett alternativ (McDaniel, Roediger & McDermott, 2007). Marsh et al. (2007) understreker derfor verdien av å gjøre flervalgstester krevende for at dette formatet skal kunne gi en god testeffekt. En annen måte å gjøre tester mer krevende på – i tillegg til å lage krevende oppgaver ved flervalgstester – er å fjerne svaralternativene og teste studentenes *frie hukommelse* gjennom åpne oppgaver (Kang, McDermott & Roediger, 2007).

Ved åpne oppgaver, som f.eks. essay- eller kortsvarsoppgaver, må studentene klare seg med ledetrådene som finnes i oppgaveteksten. Dermed er det den frie hukommelsen – og ikke gjenkjennelse av presentert kunnskap – som settes på prøve. Studenten må selv innhente kunnskapen fra hukommelsen, noe som gjør at han eller hun oftere må inngå i en semantisk analyse (McDaniel & Masson, 1985). Gay (1980) gjennomførte en studie der studenter enten tok seks kortsvarstester eller seks flervalgstester underveis i semesteret. Deretter ble studentene testet ved en avsluttende eksamen, bestående av kortsvars- og flervalgsspørsmål om de samme begrepene som ved de tidligere testene. I resultatene fra den avsluttende eksamenen viste de to studentgruppene ingen signifikante forskjeller ved flervalgsspørsmålene, mens studentene som hadde gjennomført innledende kortsvarstester gjorde det signifikant bedre ved kortsvarsspørsmålene. Andre studier støtter også læringsverdien av innledende kortsvarstester sammenlignet med flervalgstester, spesielt dersom læringsutbyttet måles ved avsluttende kortsvarstester (f.eks. Butler & Roediger, 2007; Kang et al., 2007), men også ved avsluttende flervalgstester (McDaniel, Anderson et al., 2007). Dette kan skyldes de økte kravene ved åpne testoppgaver, som antatt medfører dypere prosessering og dermed bedre langtidshukommelse av kunnskap (f.eks. McDaniel & Masson, 1985; Whitten & Leonard, 1980).

Konkluderende er det altså mye forskning som tilsier at testing kan bidra til dyp læring i utdanning. Mens repetert lesing ofte kan gi en illusjon av kompetanse, vil testing oftere utfordre studentene og medvirke til dypere prosessering. For å fasilitere dyp prosessering og testeffekten på en best mulig måte, bør testene inneholde åpne oppgaver slik at hukommelsesprosessen blir krevende og dyp prosessering kan promoteres.

Testtilbakemeldingers bidrag til dyp læring

Ovenfor ble det beskrevet hvordan testing ofte gir studentene en metakognitiv «oppvekker» ved at illusjonen av kompetanse fjernes, og at de da får et innblikk i hvor mye de *faktisk* kan og ikke hvor mye de *tror* de kan. Dette gir også en parallell til den metakognitive overvåkingen som ble beskrevet innledningsvis, og som utgjør en viktig del av dyp læring, ved at studenten evaluerer sin egen kunnskap (Butler & Winne, 1995). Det å bli testet kan i seg selv aktivere en slik overvåking (Koriat & Bjork, 2005). Dersom man i tillegg gir studentene en tilbakemelding etter en test, vil de få et enda bedre grunnlag for å vite hva de kan og ikke kan – og ikke minst hva de kan gjøre for å forbedre seg gjennom metakognitiv selvregulering (Butler & Roediger, 2008; Hattie & Timperley, 2007; Kang et al., 2007). Jeg vil nå se nærmere på empiri og relevant teori som belyser læringsutbyttet som kan oppnås gjennom ulike former for tilbakemeldinger. Først er det imidlertid verdt å understreke at det finnes en rekke faktorer som kan påvirke effekten av tilbakemeldinger. Type tilbakemelding, når tilbakemeldinger blir gitt samt individuelle forskjeller er bare noen av faktorene som spiller

inn (Hattie & Timperley, 2007). I denne artikkelen begrenser jeg mitt hovedfokus til å diskutere hvordan noen sentrale former for testtilbakemeldinger kan bidra til dyp læring.

Korrigerende tilbakemeldinger (corrective feedback), også kalt *tilbakemeldinger om utfall* (outcome feedback), forteller mottakeren noe om hvorvidt utfallet av en respons er korrekt, eller i hvilken grad det møter en gitt standard (Butler & Winne, 1995; Hattie & Timperley, 2007). Eksempler på slike tilbakemeldinger er en beskjed om hvorvidt et svar på et testspørsmål er rett eller galt, eller en presentasjon av korrekte eller passende svar (Butler & Roediger, 2008). Butler og Roediger (2008) gjennomførte en studie der deltakerne tok en flervalgstest under tre tilbakemeldingsbetingelser: en gruppe mottok korrigerende tilbakemeldinger umiddelbart etter hvert testsvar, en annen gruppe mottok korrigerende tilbakemeldinger når flervalgstesten var fullført, mens en tredje gruppe gjennomførte testen uten noen form for tilbakemelding. De korrigerende tilbakemeldingene inneholdt en indikasjon om hvorvidt svar var rette eller gale, en repetisjon av spørsmålene, deltakerens valgte svar samt en presentasjon av de korrekte svarene. En uke senere ble deltakerne testet med de samme testspørsmålene, men uten hjelp fra svaralternativer. Her gjorde deltakerne i de to tilbakemeldingsgruppene det signifikant bedre enn gruppen som ikke mottok tilbakemeldinger ved den første testen. Selv om ikke alle studier har vist like positiv effekt av korrigerende tilbakemeldinger, tilsier metaanalyser at majoriteten av studenter som mottar slike tilbakemeldinger vil oppnå en positiv regulering av kunnskap (f.eks. Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Butler & Winne, 1995; Lysakowski & Walberg, 1982).

En metaanalyse av Bangert-Drowns et al. (1991), der ulike korrigerende tilbakemeldinger ble sammenlignet, indikerer at det særlig er en presentasjon av korrekte eller passende svar som bidrar til læring blant studenter. Slike korrigerende tilbakemeldinger inngår ofte under betegnelsen *utdypende tilbakemeldinger* (elaborate feedback), siden de gir informasjon om mer enn bare mottakerens respons i seg selv, og disse bør benyttes fremfor å kun markere hvorvidt studenters svar er korrekte eller feilaktige (Bangert-Drowns et al., 1991; Butler & Winne, 1995; Kulhavy & Stock, 1989). Et annet element innen utdypende tilbakemeldinger, er *forklarende tilbakemeldinger* (Bangert-Drowns et al., 1991). Forklarende tilbakemeldinger oppgir mer enn bare korrekt svar; de sier noe om premissene bak svarene, og hvorfor en respons er korrekt eller mangelfull. Dette kan anses som et viktig element innen utdanning, der studenter ofte besvarer testspørsmål eller utfører responser hvor de benytter en mangelfull fremgangsmåte, eller mangler en dyp forståelse av hvorfor ulike løsninger er å foretrekke fremfor andre (Kulhavy & Stock, 1989; Mason & Bruning, 2001).

Chase og Houtmanfar (2009) sammenlignet effekten av å gi rett svar som tilbakemelding med det å gi rett svar med forklaring. Deltakerne var 210 studenter som underveis i semesteret gjennomførte ti flervalgstester i et psykologifag, der de ble testet for kunnskap fra ulike kapitler av læreboka. Studentene ble delt i to grupper. Den ene gruppen mottok tilbakemelding om hvilket svar som var rett ved hvert enkelt spørsmål, mens den andre gruppen mottok dette sammen med en mer utdypende forklaring, med informasjon fra og henvisning til relevante deler av læreboka som kunne forklare hvorfor det aktuelle svaret var rett. Ved halvveis fullført og endt semester, ble studentene gitt større eksamener som inkluderte mange av spørsmålene de tidligere hadde besvart ved

kapitteltestene. Ved disse eksamensspørsmålene viste studentene som hadde mottatt forklarende tilbakemeldinger signifikant større fremgang, sammenlignet med studentene som kun mottok tilbakemelding om rett svar. En interessant tilleggsanalyse viste også at forklarende tilbakemeldinger ga særlig fremgang ved de 25 prosent av spørsmålene som oftest ble besvart feil ved de innledende kapitteiprøvene, og som dermed kan anses som vanskeligst.

Flere studier fra ulike fagområder støtter konklusjonen om at forklarende tilbakemeldinger utgjør en nyttig funksjon for å fremme selvregulering og overføring av kunnskap til nye kontekster blant studenter (f.eks. Balzer et al., 1994; Birenbaum & Tatsuoka, 1987; Earley, Northcraft, Lee & Lituchy, 1990; Sassenrath & Garverick, 1965). Det å ikke bare få vite hvilket svar som er rett, eller hvilken løsning som kan være passende, men også hvorfor det er slik, ser ut til å gi studenter verdifull informasjon. Jeg vil nå se nærmere på teori som kan forklare hvorfor utdypende testtilbakemeldinger ofte kan bidra positivt i læringsprosessen.

Forklaring på effekten av testtilbakemeldinger: Fra operant betinging til metakognisjon

Tradisjonelt har forskning på testtilbakemeldinger, samt forklaringer på deres virkninger, vært preget av behaviorismens dominans (Kulhavy & Stock, 1989). Lenge ble tilbakemeldinger sett på som lite mer enn operant betinging, der det viktigste elementet var å belønne studenters gode responser ved tester (Kulhavy, 1977; Skinner, 1968). Bekreftelse ved gode responser anslås også i moderne tid som en av mekanismene som kan forklare noe av nytten ved tilbakemeldinger (f.eks. Butler, Karpicke & Roediger, 2008). Etter hvert som mer forskning – og også den kognitive revolusjon – har kommet til, har det imidlertid blitt vanskeligere å basere seg på denne forklaringen alene (for et sammendrag, se Kulhavy & Stock, 1989). Et av de fremste ankepunktene mot å forklare læringseffekten fra testtilbakemeldinger med belønning av korrekte responser, er at mange studenter lærer mest av å motta tilbakemeldinger etter feilaktige eller mangelfulle responser (f.eks. Anderson, Kulhavy & Andre, 1971; Bangert-Drowns et al., 1991). Metakognitiv selvregulering, der studenter retter på egne feil etter å ha overvåket egen kunnskap og tenkning i forhold til informasjon om en gitt standard, kan representere en mer sannsynlig årsak bak den positive læringseffekten forbundet med tilbakemeldinger (Butler & Winne, 1995).

Selvregulering starter ofte ved at studenten driver metakognitiv overvåking og blir oppmerksom på at det finnes en *diskrepans* mellom det reelle, kognitive prestasjonsnivået (som f.eks. demonstrert gjennom en testrespons) og den ønskede standarden for det kognitive prestasjonsnivået (som f.eks. representert ved en tilbakemelding som henviser til en mer korrekt testrespons) (Butler & Winne, 1995; Carver & Scheier, 1990; Kulhavy & Stock, 1989). Butler og Winne (1995) understreker at tilbakemeldinger er avgjørende for at studenter skal identifisere en slik diskrepans og iverksette selvregulering mot en standard: «For all self-regulated activities, feedback is an inherent catalyst» (s. 246). Butler og Winne påpeker også at de ikke setter noe skille mellom indre og ytre tilbakemeldinger. Med andre ord kan metakognitiv overvåking resultere i at tilbakemeldinger kommer fra studenten

selv, og man kan like gjerne rette på egen kunnskap etter å ha selv evaluert den og funnet den mangelfull, som etter å ha mottatt ytre kritikk fra en professor. En vanlig utfordring oppstår imidlertid ved at det ikke alltid er like lett å være objektiv og identifisere feil ved egen kunnskap og tenkning (f.eks. Baker, 1979). Mange studenter overvåker egen kunnskap og tenkning på en mangelfull måte, og de besitter heller ikke nok informasjon til å gi seg selv gode tilbakemeldinger basert på en god standard (Baker, 1979; Butler & Winne, 1995; Meyer, 1986). Ytre tilbakemeldinger kan her bistå den metakognitive overvåkingen med nyttig informasjon.

Ytre tilbakemeldinger som påpeker mangelfull kunnskap eller tenkning, medfører som regel at mottakeren blir mer bevisst sin egen overvåking av disse kognitive elementene (Butler & Winne, 1995). Etter å ha rettet oppmerksomheten mot egen tenkning og kunnskap, kan tilbakemeldinger bygge videre på denne bevisstgjøringen ved å bidra med flere former for nyttig informasjon som studenter kan bruke som standard for egen kognisjon, og som kan lede studenten mot metakognitiv selvregulering (Butler & Winne, 1995; Kulhavy & Stock, 1989). Dette krever imidlertid at tilbakemeldingene inneholder informasjon som er konkret nok til at studentene kan bruke den som standard (Earley et al., 1990; Kulhavy & Stock, 1989). Karakterer, samt korrigerende testtilbakemeldinger som ikke henviser til noe konkret informasjon utover hvorvidt responsen var god eller dårlig, ser ofte ut til å være for abstrakte til å gi veiledning om hvordan selvreguleringen kan gjennomføres (f.eks. Bangert-Drowns et al., 1991; Butler, 1988; Mueller & Dweck, 1998). Derimot kan utdypende tilbakemeldinger, der studenten henvises til mer spesifikk informasjon, gi retning til metakognitiv selvregulering.

Korrigerende tilbakemeldinger der studenten blir informert om en mangelfull respons ved en test og henvises til et mer korrekt svar, kan medføre at studenten metakognitivt *overskriver* egen kunnskap og erstatter feilaktige antakelser med ny og korrekt kunnskap i hukommelsen (Butler & Winne, 1995). Dersom testtilbakemeldingene utdypes enda mer til å forklare hvorfor kunnskapen er mangelfull, mottar studenten informasjon på et dypere nivå, som kan brukes som standard for selvregulering (Butler & Winne, 1995). Eksempelvis vil en matematikkstudent som får informasjon om hvilken formel som burde vært brukt ved en bestemt oppgave, få en dypere veiledning – og ha større mulighet for selvregulering – enn dersom han eller hun kun mottar informasjon om hvilket svar som er korrekt (Birenbaum & Tatsuoka, 1987). I prosessen av å bidra til metakognitiv selvregulering, kan slike forklarende tilbakemeldinger også medføre dypere læring og økt forståelse av mekanismene bak ulike responser, kontra korrigerende tilbakemeldinger som kun gir forslag til regulering på et overfladisk kunnskapsnivå (Butler & Winne, 1995; Tan, Biswas & Schwartz, 2006). Dette kan antakelig være noe av forklaringen på at studier har funnet god regulering og overføring av kunnskap etter forklarende tilbakemeldinger sammenlignet med korrigerende tilbakemeldinger alene (f.eks. Chase & Houmanfar, 2009; Sassenrath & Garverick, 1965). Det later til at dyp forståelse og metakognitiv selvregulering av kunnskap faciliteres når tilbakemeldinger er rike på spesifikk informasjon (f.eks. Earley et al., 1990; Sassenrath & Garverick, 1965; Tan et al., 2006).

Et eksempel på utdypende tilbakemeldinger rettet mot selvregulering: Metakognitive tilbakemeldinger

Gjennom utdypende tilbakemeldinger kan studenten motta både domenespesifikk informasjon (f.eks. fakta, formler) som kan brukes til å rette på egne feil, samt informasjon om de kognitive strategiene (f.eks. innhenting av informasjon, argumentasjon) som ble benyttet ved en testrespons, og hvordan disse kan forbedres (Narciss & Huth, 2004). Jeg vil nå beskrive et eksempel fra en studie (Roll, Aleven, McLaren & Koedinger, 2011) som demonstrerer tilbakemeldinger med en kombinasjon av domenespesifikk informasjon og råd om forbedring av kognitive strategier.

Roll et al. (2011) benyttet et dataprogram til å gi utdypende tilbakemeldinger til videregående matematikkelever mens de løste geometrioppgaver. Dataprogrammet ga tilbud om utdypende tilbakemeldinger ved at elevene kunne klikke på et spørsmålsteget og få tilgang til informasjon om relevante formler og eksempler på hvordan disse formlene kunne benyttes. I tillegg til slik domenespesifikk informasjon, mottok halvparten av elevene i studien *metakognitive tilbakemeldinger* (metacognitive feedback; Roll et al., 2011). Dette ble gjort ved at datamaskinen registrerte deres fremgangsmåter og måten de innhentet informasjon på ved ulike oppgaver, for deretter å gi tilbakemeldinger dersom disse var mangelfulle. Eksempelvis kunne datamaskinen registrere at flere elever ikke leste all instruksjon, og hoppet over mye viktig informasjon som de kunne trengt til å løse oppgavene på en korrekt måte. Disse elevene ble da oppmuntret til selvregulering gjennom følgende beskjed: «No need to hurry so much. Take your time and read the hint carefully. Consider trying to solve this step without another hint. You should be able to do so» (Roll et al., 2011, s. 271). Resultatene viste at elevene som mottok slike metakognitive tilbakemeldinger ble flinkere til å velge rett strategi og innhente informasjon, sammenlignet med elever som kun mottok utdypende tilbakemeldinger i form av domenespesifikk informasjon. Særlig interessant er det at datasystemet også registrerte at denne forbedringen vedvarte under testing med nye oppgaver, i måneden etter at studentene sluttet å motta slike tilbakemeldinger. Funn om at metakognitive tilbakemeldinger kan bidra til en forbedring av studenters selvregulering, som kan vedvare og overføres til nye oppgaver, er spennende dersom målet er å fremdyrke selvdrevne studenter. En rekke studier fra ulike felt støtter ideen om at tilbakemeldinger rettet mot kognitive prosesser og strategier kan fremme metakognitiv selvregulering (f.eks. Earley et al., 1990; Lee, Lim & Grabowski, 2010; Zeller-mayer, Salomon, Globerson & Givon, 1991).

Oppsummert ser testtilbakemeldinger ut til å kunne bidra til metakognitiv selvregulering blant studenter. For å promotere selvregulering, er det nyttig at tilbakemeldinger inneholder spesifikk informasjon som studenter kan bruke som standard for regulering av egen tenkning og kunnskap. For å fasilitere dyp forståelse av begreper og god selvregulering av egne responser, er utdypende tilbakemeldinger å foretrekke, gjerne med innslag av forklarende tilbakemeldinger. Forskning tyder på at utdypende tilbakemeldinger med fordel kan inneholde både domenespesifikk informasjon og råd om forbedring av kognitive strategier.

Oppsummering

Denne artikkelen har drøftet forskning og teori som belyser noen av de direkte, kognitive fordelene assosiert med tester og testtilbakemeldinger. Et argument for bruk av testing, er testeffekten. Tester bidrar positivt til læring og overføring av kunnskap, antakelig gjennom krav til hukommelsen som gir dyp prosessering. Videre har testtilbakemeldinger potensial til å gjøre studenter mer selvdrevne gjennom økt metakognitiv aktivitet. Tilbakemeldinger vil kunne bidra med verdifull informasjon som kan bidra til metakognitiv selvregulering blant studenter. Med utgangspunkt i dyp læring, har også råd for utformingen av tester og testtilbakemeldinger blitt diskutert. Hva gjelder tester, bør de være krevende og sette krav til studenters frie hukommelse for at dyp prosessering skal promoteres best mulig. Tilbakemeldinger bør være spesifikke og utdypende ved å gi informasjon og forklaringer som fremmer selvregulering.

Et overordnet mål med denne artikkelen har vært å belyse at tester med tilbakemeldinger kan gjøre mer enn å evaluere studenter og bidra til fremtidig læring. Jeg mener det finnes grunnlag for å påstå at tester og testtilbakemeldinger kan gi et mer direkte læringsbidrag – som del av en dyp tilnærming til læring.

Litteratur

- Anderson, R. C., Kulhavy, R. W. & Andre, T. (1971). Feedback procedures in programmed instruction. *Journal of Educational Psychology*, 62(2), 148–156.
- Atkinson, R. M. & Shiffrin, R. C. (1968). Human memory: A proposed system and its control processes. I K. W. Spence & J. T. Spence (Red.), *The psychology of learning and motivation: Advances in research and theory vol. 2* (s. 89–195). New York, NY: Academic Press.
- Baddeley, A. D. (1978). The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review*, 85(3), 139–152.
- Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Literacy Research*, 11(4), 366–374.
- Balzer, W. K., Hammer, L. B., Sumner, K. E., Birchenough, T. R., Martens, S. P. & Raymark, P. H. (1994). Effects of cognitive feedback components, display format, and elaboration on performance. *Organizational Behavior and Human Decision Processes*, 58(3), 369–385.
- Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology*, 25(5), 729–735.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A. & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213–238.
- Biggs, J. (1979). Individual differences in study processes and the quality of learning outcomes. *Higher Education*, 8(4), 381–394.
- Birenbaum, M. & Tatsuoka, K. K. (1987). Effects of “on-line” test feedback on the seriousness of subsequent errors. *Journal of Educational Measurement*, 24(2), 145–155.
- Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5), 1118–1133.
- Butler, A. C., Karpicke, J. D. & Roediger, H. L. III (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 918–928.

- Butler, A. C. & Roediger, H. L. III (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19(4–5), 514–527.
- Butler, A. C. & Roediger, H. L. III (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition*, 36(3), 604–616.
- Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58(1), 1–14.
- Callender, A. A. & McDaniel, M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology*, 34(1), 30–41.
- Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, 21(5), 279–283.
- Carpenter, S. K. & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, 34(2), 268–276.
- Carver, C. S. & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97(1), 19–35.
- Chase, J. A. & Houmanfar, R. (2009). The differential effects of elaborate feedback and basic feedback on student performance in a modified, personalized system of instruction course. *Journal of Behavioral Education*, 18(3), 245–265.
- Chin, C. & Brown, D. E. (2000). Learning in Science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching*, 37(2), 109–138.
- Cooper, A. J. R. & Monk, A. (1976). Learning for recall and learning for recognition. I J. Brown (Red.), *Recall and Recognition* (s. 131–156). New York: Wiley.
- Cox, K. & Clark, D. (1998). The use of formative quizzes for deep learning. *Computers and Education*, 30(3/4), 157–167.
- Craik, F. I. M. (2002). Levels of processing: Past, present ... and future? *Memory*, 10(5–6), 305–318.
- Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Behavior*, 11(6), 671–684.
- Craik, F. I. M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268–294.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438–481.
- Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review*, 1(4), 309–330.
- Dysthe, O., Raaheim, A., Lima, I. & Bygstad, A. (2006). *Undervisnings- og vurderingsformer. Pedagogiske konsekvenser av Kvalitetsreformen. Evaluering av kvalitetsreformen*. Delrapport 7.
- Earley, P. C., Northcraft, G. B., Lee, C. & Lituchy, T. R. (1990). Impact of process and outcome feedback on the relation of goal setting to task performance. *Academy of Management Journal*, 33(1), 87–105.
- Entwistle, N. (1983). *Styles of learning and teaching: An integrated outline of educational psychology for students, teachers, and lecturers*. Chichester: John Wiley and Sons.
- Entwistle, N., Hanley, M. & Hounsell, D. (1979). Identifying distinctive approaches to studying. *Higher Education*, 8(4), 365–380.

- Fritz, C. O., Morris, P. E., Bjork, R. A., Gelman, R. & Wickens, T. D. (2000). When further learning fails: stability and change following repeated presentation of text. *British Journal of Psychology*, 91(4), 493–511.
- Gay, L. R. (1980). The comparative effects of multiple-choice versus short-answer tests on retention. *Journal of Educational Measurement*, 17(1), 45–50.
- Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, 7(2), 95–112.
- Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(3), 371–377.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hermida, J. (2009). The importance of teaching academic reading skills in first-year university courses. *The International Journal of Research and Review*, 3, 20–30.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513–541.
- Johnson, W. L. & Beal, C. (2005). Iterative evaluation of a large-scale, intelligent game for language learning. I C.-K. Looi, G. McCalla, B. Bredeweg & J. Breuker (Eds.), *Artificial Intelligence in Education* (s. 290–297). Amsterdam: IOS Press.
- Kang, S. H. K., McDermott, K. B. & Roediger, H. L. III (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19(4–5), 528–558.
- Kapur, S., Craik, F. I. M., Tulving, E., Wilson, A. A., Houle, S. & Brown, G. M. (1994). Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect. *Proceedings of the National Academy of Sciences of the United States of America*, 91(6), 2008–2011.
- Karpicke, J. D. & Roediger, H. L. III (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968.
- Kember, D. (1996). The intention to both memorise and understand: Another approach to learning? *Higher Education*, 31(3), 341–354.
- Koriat, A. & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 31(2), 187–194.
- Kornell, N., Castel, A. D., Eich, T. S. & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, 25(2), 498–503.
- Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, 47(1), 211–232.
- Kulhavy, R. W. & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1(4), 279–308.
- Larsen, D. P., Butler, A. C., Lawson, A. L. & Roediger, H. L. III (2012). The importance of seeing the patient: Test-enhanced learning with standardized patients and written tests improves clinical application of knowledge. *Advances in Health Sciences Education*, 18(3), 409–425.
- Lauvås, P., Havnes, A. & Raaheim, A. (2000). Why this inertia in the development of better assessment methods? *Quality in Higher Education*, 6(1), 91–100.
- Lee, H. W., Lim, K. Y. & Grabowski, B. L. (2010). Improving self-regulation, learning strategy use, and achievement with metacognitive feedback. *Educational Technology Research and Development*, 58(6), 629–648.
- Lockhart, R. S. & Craik, F. I. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 44(1), 87–112.
- Lysakowski, R. S. & Walberg, H. J. (1982). Instructional effects of cues, participation, and corrective feedback: A quantitative synthesis. *American Educational Research Journal*, 19(4), 559–578.

- Marsh, E. J., Roediger, H. L. III, Bjork, R. A. & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin and Review*, 14(2), 194–199.
- Marton, F. & Säljö, R. (1976). On qualitative differences in learning: II. Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology*, 46(2), 115–127.
- Mason, B. J. & Bruning, R. (2001). *Providing feedback in computer-based instruction: What the research tells us*. Hentet 2. januar 2014 fra <http://dwb.unl.edu/Edit/MB/MasonBruning.html>.
- McDaniel, M. A., Anderson, J. L., Derbish, M. H. & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, 19(4–5), 494–513.
- McDaniel, M. A. & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(2), 371–385.
- McDaniel, M. A., Roediger, H. L. III & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, 14(2), 200–206.
- Metcalfe, J. & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. Cambridge: MIT Press.
- Meyer, L. A. (1986). Strategies for correcting students' wrong responses. *The Elementary School Journal*, 87(2), 227–241.
- Morris, C. D., Bransford, J. D. & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 519–533.
- Mueller, C. M. & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33–52.
- Narciss, S. & Huth, K. (2004). How to design informative tutoring feedback for multi-media learning. I H. Niegemann, D. Leutner & R. Brünken (Eds.), *Instructional design for multimedia learning* (s. 181–195). Münster: Waxmann.
- Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. I G. H. Bower (Red.), *The Psychology of Learning and Motivation*, (vol. 26, s. 125–173). San Diego, CA: Academic Press.
- Paris, S. G. & Winograd, P. (1990). How metacognition can promote academic learning and instruction. I B. F. Jones & L. Idol (Red.), *Dimensions of thinking and cognitive instruction* (s. 15–51). New Jersey, NJ: Lawrence Erlbaum Associates.
- Roediger, H. L. III, & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181–210. doi: 10.1111/j.1745-6916.2006.00012.x
- Roediger, H. L. III & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255.
- Roll, I., Aleven, V., McLaren, B. M. & Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*, 21(2), 267–280.
- Runquist, W. N. (1982). Accessibility of information to extralist cues following periods of disuse. *Journal of Verbal Learning and Verbal Behavior*, 21(5), 563–577.
- Rushton, R. (2005). Formative assessment: A key to deep learning? *Medical Teacher*, 27(6), 509–513.
- Rød, J. K. (2012). Hvordan øke læringsutbyttet fra forelesninger med begrensede ressurser? *Uniped*, 35(3), 43–56.
- Rød, J. K., Eiksund, S. & Fjær, O. (2010). Assessment based on exercise work and multiple-choice tests. *Journal of Geography in Higher Education*, 34(1), 141–153.
- Sassenrath, J. M. & Garverick, C. M. (1965). Effects of differential feedback from examinations on retention and transfer. *Journal of Educational Psychology*, 56(5), 259–263.

- Schallert, D. L. (1976). Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior*, 15(6), 621–632.
- Skinner, B. F. (1968). *The technology of teaching*. New York: Appleton-Century-Crofts.
- Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(3), 601–604.
- Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology*, 30(9), 641–656.
- Tan, J., Biswas, G. & Schwartz, D. L. (2006). Feedback for metacognitive support in learning by teaching environments. I R. Sun & N. Miyake (Red.), *Proceedings of the 28th Annual Meeting of the Cognitive Science Society* (s. 828–833). Mahwah, NJ: Lawrence Erlbaum.
- Thompson, C. P., Wenger, S. K. & Bartling, C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory*, 4(3), 210–221.
- Toichi, M. & Kamio, Y. (2002). Long-term memory and levels-of-processing in autism. *Neuropsychologia*, 40(7), 964–969.
- Whitten, W. B. & Leonard, J. M. (1980). Learning from tests: Facilitation of delayed recall by initial recognition alternatives. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 127–134.
- Zellermayer, M., Salomon, G., Globerson, T. & Givon, H. (1991). Enhancing writing-related metacognitions through a computerized writing partner. *American Educational Research Journal*, 28(2), 373–391.