

Kvalitetssystemer til stryk: En empirisk, konstruktiv studie

Vidar Gynnild

Professor, Institutt for pedagogikk og livslang læring, Norges teknisk-naturvitenskapelige universitet (NTNU)

vidar.gynnild@ntnu.no

SAMMENDRAG

NOKUT benytter ekspertpaneler for å vurdere om kvalitetssystemene i høyere utdanning holder mål. Denne undersøkelsen retter seg mot fem institusjonelle kvalitetssystemer som ikke ble godkjent, og analyserer sannsynlige årsaker med utgangspunkt i ekspertpanelenes evalueringsrapporter. Artikkelen består av to metodisk sett ulike deler. Utgangspunktet er en empirisk studie av fem evalueringsrapporter, mens andre del er konseptuell og konstruktiv i den forstand at jeg skisserer tankemodeller til mulig videreutvikling av kvalitetssystemene. Analysen viser at studiekvalitetsarbeidet skjer med utgangspunkt i intensjoner som ikke i tilstrekkelig grad følges opp. Rapporteringen framstår som beskrivende, uten dypere analyser, vurderinger og konklusjoner. Evalueringspanelene kommenterer dette, men bidrar ikke med konseptuelle løsninger som hjelper institusjonene. Vesentlige utfordringer ligger i overgangen fra beskrivelse til analyse og teoriarbeid basert på en systemforståelse av handlingsfeltet.

Nøkkelord

Kvalitetssystemer, Analyse, Rapportering, Evaluering, konseptuelle metoder

ABSTRACT

The Norwegian Quality Assurance Agency (NOKUT) employs expert panels to decide whether or not institutional quality systems should be approved. This study examines expert panels' reports of five failing institutions to investigate why their quality systems were disapproved. The article is made up of two parts – the first part is empirically based while the second suggests conceptual approaches that might support the amendment of existing quality systems. The study brings evidence of a significant gap between impeccable intentions at one end and harsh realities at the other end. Typically, self-reporting at the institutional level appears descriptive rather than analytical, providing few, if any helpful conclusions. Expert panels in their evaluation reports pointed to such shortcomings but

failed to provide conceptual contributions that would make any difference. Key issues remained conceptually unresolved, particularly the transition from *description* to *analysis* in view of a systems approach.

Keywords

quality systems, analysis, reporting, evaluation, conceptual models

INNLEDNING

I årene etter årtusenskiftet har «kvalitetssikring» av utdanningen blitt et sentralt begrep og en stor utfordring for norsk høyere utdanning. Nasjonale og lokale initiativ har vært forankret i et felleseuropeisk utdannings samarbeid med vekt på et enhetlig rammeverk for utdanningene, bruk av læringsmål, og generelt et perspektivskifte fra undervisning til læring (*Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG)*, 2015):

At the heart of all quality assurance activities are the twin purposes of accountability and enhancement. Taken together, these create trust in the higher education institution's performance. A successfully implemented quality assurance system will provide information to assure the higher education institution and the public of the quality of the higher education institution's activities (accountability) as well as provide advice and recommendations on how it might improve what it is doing (enhancement). Quality assurance and quality enhancement are thus inter-related. They can support the development of a quality culture that is embraced by all: from the students and academic staff to the institutional leadership and management (ESG, 2015, s. 7)

Universitets- og høyskoleloven (§ 1-6) fastsetter i tråd med dette at universiteter og høyskoler skal ha et tilfredsstillende system for intern kvalitetssikring. Innen rammen av lov og forskrift bestemmer institusjonene selv hvordan systemet skal utformes, og nasjonalt organ for kvalitet i utdanningen (NOKUT) har som oppgave å evaluere institusjonenes kvalitetssystemer etter kriterier de selv fastsetter (Universitetsloven § 2-1). NOKUT er et faglig uavhengig, statlig forvaltningsorgan som driver tilsyn, samtidig som det skal stimulere til utviklingsarbeid innen høyere utdanning. Målet er å bidra til at samfunnet kan ha tillit til at utdanningene holder et godt faglig nivå, samtidig som institusjonene blir stimulert til å utvikle gode kvalitetskulturer. For å oppnå dette, benytter NOKUT sakkyndige komiteer for å evaluere kvalitetssystemene lokalt. Dette ble gjennomført for første gang i 2003, mens praksis fra og med 2009 ble noe endret ved relativt større vekt på institusjonenes bruk og nytte av kvalitetssikringssystemene.

De sakkyndige komiteenes oppgave er å vurdere om kvalitetssystemene holder mål ut fra kriterier fastsatt i «Forskrift om tilsyn med utdanningskvaliteten i høyere utdanning» (Lovdata, 2011). Komiteene skal vurdere kvalitetssystemet og institusjonenes aktive bruk av det ut fra følgende kriterier: a) Stimulans til kvalitetsarbeid og kvalitetskultur: om institusjonen stimulerer til engasjement for kvalitetsarbeid blant ansatte og studenter og deres demokratiske organer, b) Mål, plan og ledelsesforankring: om mål, ansvar, prosesser og aktører som inngår i kvalitetssystemet er klart beskrevet (systemets strukturelle oppbyg-

ning), og hvordan systemet for kvalitetssikring utvikles i tråd med institusjonens behov, c) Innhenting av dokumentert informasjon om kvalitet i studiene: om sikring og vurdering av kvaliteten i hvert enkelt studium bygger på dokumentert informasjon som systematisk innhentes fra flere kilder, og om det er særskilte prosesser for å kvalitetssikre oppretting av nye studier, d) Analyse, vurdering og rapportering: om den informasjonen som systemet generer, vurderes og framstilles for ansvarlige fora og ledernivå, e) Bruk av kunnskap til kvalitetsforbedring: om tiltak for forbedringer iverksettes på grunnlag av de kvalitetsanalyser som gjøres (Lovdata, 2011, paragraf 6).

Utviklingsarbeid er et krevende felt blant annet fordi sentrale begreper med «kvalitet» i forstavelen ikke er entydig definert og blir brukt i flere ulike betydninger med utgangspunkt i taus kunnskap hos aktørene. Det samme gjelder begreper som «undervisning» og «læring». «Kvalitet» blir noen ganger tolket som noe spesielt, utsøkt eller originalt, en begrepsforståelse vi gjenfinner i det førindustrielle samfunnet. Som del av industrisamfunnets framvekst ble det satt nye, felles standarder for produksjon, og kvalitetsbegrepet fikk etter dette også en annen betydning. «Kvalitet» ble «det samme som at leve op til en norm. Variation bliver en uting, der skal elimineres» (Dahler-Larsen, 2013, s. 32). Til tross for at akademiske miljøer gjerne flagger begreper som «kreativitet» og «originalitet» høyt, er ikke dette nødvendigvis fenomener som gjenfinnes i kvalitetssystemene, som ofte er mer opp-tatt av felles standarder og «avvik» i forhold til regler og rutiner. Industrisamfunnets kvalitetsbegrep har inntatt akademia.

Med dette har det opprinnelige, faglig-kollegiale samarbeidet fått en institusjonalisert, parallell maktstruktur hjemlet i lov. Arbeid med kvalitetssikring er sentralt organisert og styrt, oftest fjernt fra de faglige aktørene i undervisningen. Økt byråkratisering og standardisering av prosedyrer for datainnsamling og rapportering har imidlertid møtt kritikk: «The agenda is managerial rather than academic, with accountability as a high priority; procedures are top-down and bureaucratic» (Biggs, 2001, s. 222). Forholdet mellom kontroll og utvikling er oftest også et stridstema: «While forces of accountability are strong, those devoted to improvement, including the promotion of innovation, are fragmented» (Newton, 2002, s. 3). John Biggs antyder at enkelte kvalitetsinitiativ faktisk kan ha motsatt effekt av den tilsiktete:

In short, some common QA [Quality Assurance] procedures have the opposite effect to that intended. This approach . . . is . . . not functionally concerned with the quality of teaching and learning, but with quantifying some of the presumed indicators of good teaching and good management (Biggs, 2001, s. 232).

Drift av kvalitetssystemene krever mye tid og ressurser, og behovet for dokumentasjon av de reelle effektene er nødvendig. Det kan i sin tur bidra til et mer velfungerende kvalitetsarbeid:

What is surprising is that, even though 'quality' has had such a profound impact on academics' lives and behaviours . . . this major transformation remains curiously underresearched and under-theorised. Indeed, if we were to stand back as an academic community and ask 'what has been learned? and by whom?', or 'what has improved?', though many would be quick to provide answers, in reality

we can point to very little research into how 'quality policy', or other areas of strategy designed to improve learning and teaching have been used, how this has impacted on academic practice (Newton, 2002, s. 1–2).

Denne artikkelen er skrevet på grunnlag av evalueringsrapportene fra fem institusjoner som fikk kvalitetssystemene sine underkjent av NOKUT. Artikkelen består av en empirisk og en konstruktiv del, som er tematisk nært koblet. Evalueringsrapportene utgjør datagrunnlaget for analysen, mens andre del eksemplifiserer konseptuelle modeller som grunnlag for utvikling.

DATA, METODE OG PROBLEMSTILLINGER

Datagrunnlaget for undersøkelsen er komplett i den forstand at alle rapportene fra de fem institusjonene, som fikk sine kvalitetssystemer underkjent i perioden 2012–2014, har vært analysert. De involverte institusjonene er ulike med hensyn til størrelse og faglig innretning, men var alle underlagt felles evalueringskriterier. Rapportene er skrevet med utgangspunkt i en felles struktur, og varierer i omfang fra 25 til 42 sider. Disse ble lest i sin helhet flere ganger for å få inntrykk av hvilke vesentlige ankepunkter som ble framsatt, og som må antas å ha bidratt til å underkjenne kvalitetssystemene. Typisk argumentasjon framsatt i rapportene ble deretter eksemplifisert med sitater. En slik praksis er metodisk sett ikke helt uproblematisk, fordi nyanser er lett å overse. På den annen side antas sitatene å styrke autentisiteten i teksten.

Som nevnt i innledningen, ble de sakkyndige komiteenes evaluering gjennomført med utgangspunkt i totalt fem kriterier. Av praktiske årsaker fokuserer undersøkelsen på de tre siste, omtalt som punkt c, d og e i innledningen. Følgende problemstillinger blir undersøkt:

- Hva kan forklare at kvalitetssystemene ved de fem institusjonene ble underkjent?
- Hvilke typer råd ble gitt institusjonene til forbedring av utdanningskvaliteten?
- Hvordan kan konseptuelle modeller og teori bidra til mer målrettet kvalitetsarbeid?

De tre evalueringskriteriene la grunnlaget for en komparativ studie av evalueringsrapportene med vekt på fellestrekk mellom institusjonene. Deretter ble komiteenes råd for å forbedre kvalitetssystemene undersøkt. I hvilken grad gis det likhetstrekk mellom de råd som blir gitt på tvers av institusjoner, og i hvilken grad kan det sannsynliggjøres at rådene som gis vil ha ønsket virkning? Undersøkelsen er gjennomført som en kasusstudie der konklusjonene har gyldighet for det undersøkte utvalget (Flyvbjerg, 2006; Yin, 1994). Men «that knowledge cannot be formally generalized does not mean that it cannot enter into the collective process of knowledge accumulation in a given field or in a society» (Flyvbjerg, 2006, s. 227).

Kvalitetssystemene ved alle institusjonene ble evaluert med utgangspunkt i et felles kriteriesett, og ingen ble evaluert av samme komite. Rekruttering til de sakkyndige komiteene ble foretatt etter følgende kriterier: a) alle skal ha erfaring fra kvalitetsarbeid eller evaluering; b) minst én skal ha erfaring fra ledelse ved en høyere utdanningsinstitusjon; c) minst én skal ha tilknytning til relevant utenlandsk institusjon; d) en student, som har erfaring fra

institusjonsstyre, andre sentrale tillitsverv eller fra arbeid i studentorganisasjon og e) minst én med professor- eller dosentkompetanse (Lovdata, 2011). Bredt sammensatte komiteer skal sikre upartiskhet: «Quality assurance must be independent from the political control in order to assure its legitimacy in the eyes of the universities» (Beerens, 2015, s. 43). Felles kriterier for evaluering styrker påliteligheten i rapportenes konklusjoner, samtidig som det er grunn til å anta at skjønnsutøvelsen også er farget av komiteenes erfaringer, preferanser og dømmekraft.

ANALYSE

Som venstre kolonne i Tabell 1 viser, framstår ulike former for studentevalueringer som den dominerende datainnsamlingsmetoden. Komiteene påpeker lav svarprosent som en viktig begrensning ved evalueringsordningene, men kommer ellers med få forslag til metoder, som kanskje også kunne fortelle noe om studentenes læring. Eksklusiv bruk av evaluering gjør institusjonene sårbare for metodekritikk, fordi dette er et format som setter store krav til ryddighet med hensyn til *kriterier* og *standarder*. Disse to dimensjonene blir utfordret i alle former for evaluering; kriterier definerer hva som skal evalueres, mens standarder refererer til nivå. Kriterier benyttes ofte i folkelig tale som en fellesbetegnelse, men analytisk handler det om to ulike dimensjoner. Dersom referansene er utilstrekkelige eller fraværende, skaper det i sin tur usikkerhet ved datatolkning med konsekvenser for tiltak. De metodiske utfordringene ved datainnsamlingen blir imidlertid ikke adressert av de sakkyndige komiteene, som i stedet synes å være mest opptatt av om det foregår «evaluering» eller ei.

Implisitt og eksplisitt er det institusjonenes «undervisningskvalitet» oppmerksomheten rettes mot, og komiteene utfordrer ikke bruken av «tilfredshetsundersøkelser» som metodisk tilnærming. Ett sted blir det imidlertid kommentert at datainnsamlingen er lite fokusert: «I første omgang fikk komiteen tilsendt mange tunge filer, men bare noen av disse dokumenterte det systematiske arbeidet for å sikre kvaliteten» (I-3, 25.04.13, s. 24).

Kategorien analyse, rapportering og vurdering (kolonne 2 i Tabell 1) framstår som et svakt punkt for alle institusjonene: «Mange av indikatorene har ingen analyser knyttet til seg – ingen vurderinger og ei heller noen konklusjoner med eventuelt foreslåtte tiltak» (I-2). Et annet gjennomgående trekk er at rapporteringen framstår som beskrivende og lite analytisk: «Rapportene fra fakultetene til rektor og fra rektor inneholder i liten grad analyser og vurderinger av undervisnings- og læringskvalitet, men angir snarere hvilke utfordringer fakultetene har med å innhente opplysninger om denne kvaliteten» (I-3). Rapporteringen blir dermed en øvelse fylt av gode intensjoner, men der positive effekter er vanskelig å få øye på: «De registrerer hva som rapporteres, men trekker få konklusjoner, foretar få vurderinger og peker på få tiltak som bør iverksettes for å sikre eller utvikle kvaliteten i utdanningen» (I-4).

Tabell 1. Fem institusjoner evaluert ut fra tre kriterier gitt av NOKUT

Institu- sjon	Innhenting og dokumentert informasjon om kvaliteten i utdanningen	Analyse, rapportering og vurdering	Bruk av kunnskap til kvalitetsforbedring
I	«... høgskolen innhenter dokumentert informasjon om kvalitet i studiene på en god måte. Komiteen vil likevel anbefale høgskolen å ... utvikle semesterevalueringene, stimulere til økt studentdeltakelse og å overveie ... et system for anonyme meldinger om avvik».	«Det analytiske nivået som er til stede i rapportene på programnivået, framtrer ... ikke med tydelighet i rapporten på institusjonsnivå. Komiteen mener at utslingsprosessen på institusjonsnivå svekker informasjonsstrømmen av akkumulerte og analytiske data til høgskolerådet og styret».	«Verken de siste årsrapportene om kvalitetsarbeidet eller en nærmere dokumentasjon om implementeringen og institusjonens bruk og nytte av systemet forelå». «På institusjonsnivået ... foretas det ingen samlet analyse som gir en trygg basis for videre tiltak».
II	«Hovedbildet er at underveisevalueringer gjennomføres i noe begrenset grad, og sluttevalueringer har kritisk lav svarprosent. Komiteen mener at dette i sum ikke gir tilstrekkelig innhenting av informasjon fra studenter».	«Mange av indikatorene har ingen analyser knyttet til seg – ingen vurderinger og ei heller noen konklusjoner med eventuelt foreslåtte tiltak». ... «Ut over kvalitetsrapportene har ikke komiteen kunnet registrere omfattende analyser og vurderinger som er presentert for ansvarlige fora».	«Når rapporteringssløyfer ikke har den nødvendige kvalitet, blir grunnlaget for å fatte beslutninger om tiltak for å forbedre kvaliteten vanskelig. I den grad en må prioritere ... , bør det kvalitetsarbeidet som foregår på siden av systemet tas inn i systemet. Høgskolen bør ... gjøre koplingen til plan- og budsjettarbeid tydeligere i kvalitetssystemet.»
III	«Referansegrupper og studenters individuelle skriftlige sluttevalueringer gir forskjellig informasjon. Dersom referansegruppene er den eneste kanalen en student har for å komme med sine synspunkter på kvaliteten i utdanningen, er dette ikke tilstrekkelig. Kvalitetssystemet må være klart gjennomskuelig for studentene».	«Rapportene fra fakultetene til rektor og fra rektor inneholder i liten grad analyser og vurderinger av undervisnings- og læringskvalitet, men angir snarere hvilke utfordringer fakultetene har med å innhente opplysninger om denne kvaliteten». ... «Det vanskeliggjør rett og slett interessante analyser».	«Det kom fram at systemet på grunnivået genererer en del informasjon som ikke bringes videre» «I rektors rapport fremkommer det ikke alltid hva som er gjort for å følge opp forhold som ble sett på som utfordringer i forrige års melding».
IV	«Det er ... studentevalueringer, dialogmøte med studenter, en godt utbygd utvalgsstruktur og et omfattende utredningsarbeid, særlig i forbindelse med praksisstudier» «For øvrig er det dokumentert innhenting av informasjon om kvaliteten i mastergradsstudier i begrenset omfang ...» « For forskerutdanningen er ... informasjon[en] sparsom og usystematisk»	«Rapportene er ofte refererende og lite målrettede. De registrerer hva som rapporteres ... men trekker få konklusjoner, foretar få vurderinger og peker på få tiltak som bør iverksettes for å sikre eller utvikle kvaliteten i utdanningen. Hverken intervjuene eller dokumentasjonene indikerer at rapportene har hatt stor effekt i organisasjonen. Dermed framtrer dagens rapportering, analyse og vurderingsarbeid, utover det som uføres lokalt, som sporadisk og usystematisk».	«Komiteens hovedinntrykk er at det i for liten grad fremkommer hvordan kunnskap generert av kvalitetsarbeid blir brukt til kvalitetsforbedring». ... «Høgskolen må både styrke og dokumentere denne delen av kvalitetsarbeidet bedre fremover».

Institu- sjon	Innhenting og dokumentert informasjon om kvaliteten i utdanningen	Analyse, rapportering og vurdering	Bruk av kunnskap til kvalitetsforbedring
V	«Institusjonen skaffer seg informasjon om undervisningskvalitet gjennom studentevalueringene, men har ... svært lite systematisert og formalisert ekstern informasjon om programkvalitet. ... ingen eksterne programevalueringer eller fagfellevurderinger av egne studier».	«Kvalitetssikringssystemet genererer svært mye rapportering uten at rapportene analyseres nærmere. Rapportene inneholder svake og mangelfulle analyser av utdanningskvalitet, og kvalitetsrapporten som avgis til styret, er i stor grad gjentakelser av de fakultetsvise kvalitetsrapportene».	«Høyskolen har ikke godtgjort hvorvidt tiltak iverksettes på grunnlag av kvalitetsanalyser, når analysene av utdanningskvaliteten er såpass mangelfulle. En uformell kvalitetskultur er i seg selv positivt ... Men i [NN sitt] tilfelle er ikke kvalitetssikringssystemet implementert på alle nivåer i organisasjonen, og for mye av ansvaret ligger på én person»

De sakkyndige komiteene rapporterer generelt om vesentlige mangler på flere punkt, slik at helhetsbildet blir negativt: «Høyskolen har ikke godtgjort hvorvidt tiltak iverksettes på grunnlag av kvalitetsanalyser, når analysene av utdanningskvaliteten er ... mangelfulle» (I-5). «Når rapporteringssløyer ikke har den nødvendige kvalitet, blir grunnlaget for å fatte beslutninger om tiltak for å forbedre kvaliteten vanskelig» (I-2). Dermed blir det ingen logisk sammenheng mellom datainnsamling, analyse og tiltak. Datagrunnlaget er en nødvendig, men ikke tilstrekkelig forutsetning for gode analyser. Alle faser av slike evalueringer er krevende, og det kan virke som institusjonene ikke har hatt tilstrekkelige kompetanse tilgjengelig.

De sakkyndige komiteene benytter på sin side kvalitetsbegrepet uten definisjon eller forklaring. Fra et kommunikasjonssynspunkt er dette utfordrende, idet ingen av partene settes i stand til å argumentere prinsipielt om «kvalitet». Dette motiverer for å undersøke hvilke råd de sakkyndige komiteene gir. Selv om «kvalitet» ikke er definert, vil komiteenes innspill, i hvert fall i noen grad, speile underliggende tanker og ideer om hvordan «utdanningskvalitet» best kan fremmes ved institusjonene. En oversikt over dette er gjengitt i Tabell 2.

Ved å sammenholde hvilke råd som blir gitt, finner vi flere fellestrekk. Blant disse er ønsket om å gjøre kvalitetssystemet mer transparent, vekt på bruk av studentevalueringer, opplæring av personell i ulike funksjoner og bedre rapportering. For I-3 sies det spesielt at analyser til «vurderinger og forslag til tiltak for forbedringer av kvaliteten i utdanningen må «gis større oppmerksomhet og dokumenteres bedre». Rapportene eksemplifiserer mangler, og gir bare generelle råd, som for eksempel «gjennomgå dagens kvalitetssystem», at det «må skapes tydeligere rutiner», rapportering om «kvaliteten i utdanningen må styrkes». Denne tilnærmingen er i tråd med retningslinjer gitt av NOKUT, som krever at det ved «evaluering av institusjonenes interne kvalitetsarbeid [skal] gjøres ... en helhetlig vurdering av kvalitetssikringssystemet og institusjonenes aktive bruk av det» (Lovdata, 2011). De råd som blir gitt, er av generell karakter og uttrykker gode intensjoner for systemforbedring. Det er imidlertid vanskelig å se hvordan institusjonene kan bli vesentlig «klokere» av de råd som blir gitt. Det er et udekket behov for støtte for institusjonenes utviklingsarbeid, slik det framgår i sitatet: «Analyser, vurderinger og forslag til tiltak for forbedringer av kvaliteten i utdanningen må gjennomgående gis større oppmerksomhet og dokumenteres bedre» (I-4, 17.06.14).

Tabell 2. Komiteenes råd med hensyn til videreutvikling av kvalitetsarbeidet

Inst.	Komiteenes råd om videreutvikling av kvalitetsarbeidet
I-1, 13.09.12	Får beskjed om å reetablere læringsmiljøutvalget, tydeliggjøre kvalitetssikringssystemet og utarbeide årsrapport for styret for å velge riktig tiltak. Studentmedvirkningen bør styrkes, og høyskolen bør «... overveie og etablere et funksjonelt og systematisert avvikssystem» (s. 20). Kvalitetssikringssystemet bør gjøres mer transparent, foreslår den sakkyndige komiteen.
I-2, 05.09.13	For å få i gang det systematiske kvalitetsarbeidet bør høyskolen gjennomgå dagens kvalitetssystem for å se om det kan tilpasses bedre til egne behov; ha stor oppmerksomhet på å få i gang studentevalueringer som gir reell informasjon fra studenter om utdanningskvalitet; gi opplæring til ledere, ansatte og studenter om systemet i seg selv og høyskolens forventninger til deltagelse i kvalitetsarbeidet; fortsette det påbegynte arbeidet med årlige kvalitetsrapporter; ha stor oppmerksomhet på det videre arbeidet med å utvikle LMU; vurdere å gjennomføre evaluering av hele studieprogram; vurdere å styrke arbeidet med industrielle referansegrupper, og det øvrige arbeid med å evaluere studienes relevans og evt. vurdere å utvikle indikatorer for relevans.
I-3, 25.04.13	Systembeskrivelsen er mangelfull. Det må skapes tydeligere rutiner for hvordan emner og program skal gjennomføres, spesielt med bruk av studentevalueringer. Komiteen anbefaler at NN innfører obligatoriske emneevalueringer og det må utarbeides selvstendige rapporter uavhengig av referansegruppene.
I-4, 17.06.14	Rapportering om kvaliteten i utdanningen må styrkes, spesielt på institusjonelt nivå. Kvalitetssystemet forventes å inneholde en mer presis ansvars- og rutinebeskrivelse, slik at systemet raskt kan operasjonaliseres. «Analyser, vurderinger og forslag til tiltak for forbedringer av kvaliteten i utdanningen må gjennomgående gis større oppmerksomhet og dokumenteres bedre. ... Høyskolen må avklare hva målsetningen med det omfattende rapport- og utredningsarbeidet som er utført rundt praksisstudier er. Høyskolen bør vurdere å knytte arbeidet tydeligere til et oppfølgende kvalitetsarbeid»
I-5, 25.08.14	«Høyskolen har ikke godtgjort hvorvidt tiltak iverksettes på grunnlag av kvalitetsanalyser, når analysene av utdanningskvaliteten er såpass mangelfulle» (s. 24). ... «Men ... kvalitetssikringssystemet [er ikke] implementert på alle nivåer i organisasjonen, og for mye av ansvaret ligger på én person» (s. 24).

Til komiteenes forsvar skal det sies at det ikke er enkelt å forbedre kvalitetssystemet på generelt grunnlag, uten å ta stilling til de prinsipielle tilnærminger ved denne typen arbeid. Skal fokus rettes mot «undervisning» eller «læring», eventuelt ut fra hvilke definisjoner av disse begrepene? Og er formålet «kvalitetssikring», «kvalitetsutvikling» eller begge deler? Disse utfordringene et forhold som tidligere er undersøkt på emnenivå (Gynnild, 2014b). Et viktig funn var i dette tilfellet at kun undervisningsperspektivet ble benyttet som grunnlag for å si noe om «kvaliteten», som ble oppfattet som et absolutt, snarere enn et relativt fenomen. Selve kvalitetsbegrepet forsvant som felt for undersøkelse, spesielt at all evaluering forutsetter kriterier, og at disse står sentralt i tolkning av resultatene. Én institusjon ble rådet til å «ha stor oppmerksomhet på å få i gang studentevalueringer som gir reell informasjon fra studenter om utdanningskvalitet» (I-2, 05.09.13), men uten metodisk veiledning. Fordi råd gis i generell form på systemnivå er faren stor for at institusjonene ikke klarer å omsette forslag i praksis.

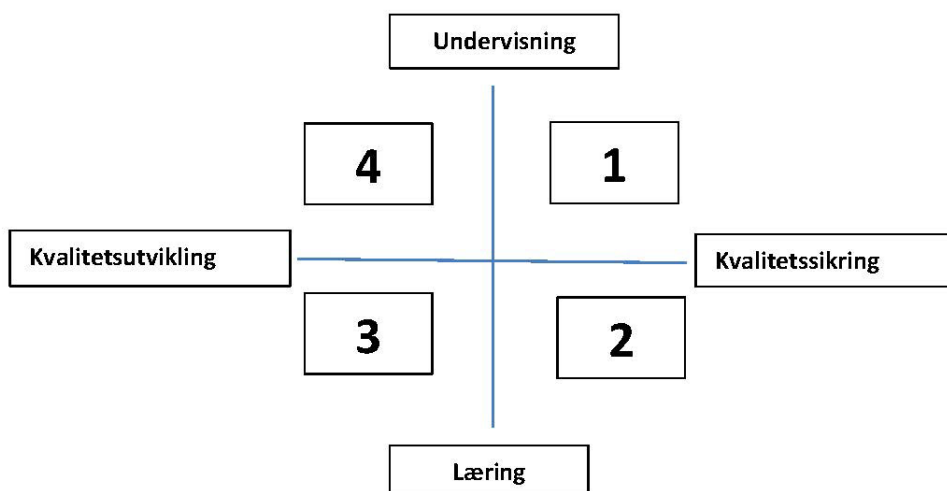
Den empiriske undersøkelsen dokumenterer noen av manglene ved kvalitetssystemene som ble underkjent, men de sakkyndige komiteene evner i mindre grad å imøtekomme rollen som utviklingsressurs for institusjonene. Komiteene bidrar med åpenbart velmente

råd, men det er tvilsomt om institusjonene føler seg tilstrekkelig godt hjulpet. Et generelt inntrykk er at verken institusjonene eller de «sakkyndige» dokumenterer tilstrekkelig innsikt og erfaring til å ivareta sine oppdrag. Fra institusjonenes side framstår kvalitetsrapporteringen som en pålagt og noe usikker øvelse med stort behov for et veiledende teoretisk rammeverk. Undersøkelsens konstruktive bidrag blir ut fra dette å introdusere begreper og modeller som støtte for praksis. Denne delen er av generisk art, og modellene antas å ha generell relevans i alt kvalitetsarbeid.

BEGREPER OG MODELLER

Et fellestrekk ved ekspertpanelenes råd er generelle forslag til hva de mener bør gjøres for å forbedre kvalitetssystemene, men de utdyper ikke *hvordan* dette kan gjennomføres. Et par eksempler illustrerer poenget: «Rapportering om kvaliteten i utdanningen må styrkes» og «Høgskolen bør vurdere å knytte arbeidet tydeligere til et oppfølgende kvalitetsarbeid» (se Tabell 2). Dette er neppe kontroversielle utsagn, men bidrar i seg selv ikke til at institusjonene står bedre rustet til å finne gode løsninger. Forslagene framstår som svakt faglig fundert, og bidrar ikke til dypere innsikt i utfordringenes egenart. Figur 1 eksemplifiserer en matrise med fire «mulighetsrom» avgrenset av de to aksene. Matrisen kan brukes som en veileder i arbeid med å identifisere målområder, men krever definisjoner av begrepene som avgrenser aktuell kvadrant. Kvadrant 1 avgrenses av «undervisning» og «kvalitetssikring», noe som betyr at disse begrepene må tydeliggjøres som grunnlag for ytterligere innsats. I hverdagsspråket er begrepsforståelsen ofte implisitt, men her velger vi å benytte følgende begrepsforståelse: «Teaching can be understood as any activity which is undertaken with the intention of helping somebody learn» (Goodyear, 2015, s. 30). Etter dette kan emnedesign og vurderingsopplegg inkluderes som del av undervisningen, ettersom begge setter vesentlige rammer for læring.

Figur 1. Matrise for analyse av kvalitetssystemer (Gynnild, 2014a)



Goodyears definisjon har undervisningens *formål* som sitt utgangspunkt, mens dens *former* beskrives i tre kategorier; «interactive forms of teaching that involve real-time exposition and other kinds of direct instruction, as well as facilitation of students' learning activities; pre-active forms of teaching: planning, design etc.; post-active forms of teaching: reflection, evaluation, assessment» (Goodyear, 2015, s. 32). Dette gir inntak til *hva* som kan adresseres i den første kvadranten, altså *kriteriene* for evalueringen. Å «kvalitetssikre» undervisningen forutsetter imidlertid både *kriterier* og *standarder*, som eksemplifisert i Tabell 3 nedenfor:

Tabell 3. Eksempel på vurderingsrubrikk til bruk i første kvadrant i Figur 1

Hva blir vurdert (kriterier)	Krav til godkjenning (standarder)	
Ulike former for undervisning etter Goodyear (2015)	Ikke tilfredsstillende 1, 2, 3, 4, 5	Tilfredsstillende 6, 7, 8, 9
Gjennomføring av interaktive undervisningsformer		
Emne- og oppgavedesign		
Tilbakemeldingsordninger til undervisere og veiledere		

Kvadrant 2 krever svar på spørsmål om krav til læring, eventuelt et minstemål for å bestå eller bli sertifisert. Denne situasjonen kjenner vi fra krav til førerprøven der det ikke gis en gradert karakter, men det settes visse standarder for godkjenning. I akademia blir spørsmålet *hva* og *hvor* godt studentene skal lære gitte tema, for eksempel forholdet mellom faglig dybde og bredde. Dette gir føringer for læringsmål og læringsformer, og til dokumentasjon av læringsresultat. Speiler vurderingsuttrykket reell måloppnåelse eller noe annet (Sadler, 2009)?

I Kvadrant 3 (Figur 1) rettes fokus mot læringsforbedringer, som innbefatter tema som læringsmål, undervisnings-læringsaktiviteter og vurdering. Her synes det nærliggende å tenke at forslag til tiltak begrunnes med utgangspunkt i en analyse av nå-situasjonen. Her trengs data med forklaringspotensial, i motsatt fall er det ikke gitt at tiltak gjør en forskjell. Til støtte gis det også en stor forskningslitteratur på sammenhenger mellom prosesser og læringsresultat (Biggs, 1987; Case & Marshall, 2009; Gynnild & Myrhaug, 2012; Marton, 1994).

Kvadrant 4 (Figur 1) retter oppmerksomhet mot kvalitetsutvikling av undervisningen. Begrepsforståelsen blir igjen sentral, spesielt *hva* det betyr å «forbedre» undervisningen. Det svaret som gis innen rammen av undervisningsparadigmet, er bruken av tilfredshetsstudier av undervisningen, der studentene selv setter standard for god og dårlig undervisning. Ut fra en annen begrepsforståelse kan tiltak være å lage et bedre *emnedesign* med utgangspunkt i læringsmål og prinsippet om «Constructive Alignment» (Biggs, 1996).

DISKUSJON

I Figur 1 rettes søkelyset mot to nærstående, men ulike fenomen, nemlig «undervisning» og «læring» der begrepsforståelsen ofte tas for gitt. Oppmerksomhetens innretning er imidlertid vesentlig i alt utviklingsarbeid, og dersom begrepsforståelsen er implisitt, fram-

står den ikke som tilgjengelig for refleksjon og dialog. Den største nytteverdien ved bruk av Figur 1 ligger trolig i dens bidrag til perspektivskifte og begrepstolkning. Overgangen fra «undervisning» til «læring» er prinsipielt sett banal, men erfaring viser at dette er krevende, slik litteraturen om terskelbegreper (threshold concepts) dokumenterer (Atherton, Hadfield, & Meyers, 2008). Et perspektivskifte bidrar til at «verden» ser annerledes ut, med behov for andre typer data, som så fører til andre analyser og tiltak. Når omverdenen oppfattes på nye og mer meningsfylte måter, er enkel retreat til tidligere forståelsesformer verken ønskelig eller sannsynlig.

I en kjent artikkel skiller forfatterne mellom det de omtaler som «The Instruction Paradigm» og «The Learning Paradigm» (Barr & Tagg, 1995). Forfatterne er opptatt av å vise at ulike perspektiv gir ulike føringer for læringssyn og rolleforståelse. Denne undersøkelsen tyder på at kvalitetsarbeidet har vært forankret i en implisitt didaktisk modell, som ikke blir nærmere problematisert av aktørene. Denne tilnærmingen motiverer for «fine tuning» av undervisningen i eksisterende form, men evalueringene bidrar ikke med vesentlige grep med potensial for å endre studieatferd, som grunnlag for bedre læring. *Studenttilfredshet med undervisningen* framstår som et gjennomgående tiltak. Litt spissformulert blir spørsmålet om studentene «liker» undervisningen. Praksis følger den underliggende begrepsforståelsen, som er individrettet snarere enn systemrettet. Dette var tema for sosialpsykologen Kurt Lewin allerede på 1930-tallet, og organisasjonslitteraturen har viet spørsmålet om sosial endring stor oppmerksomhet gjennom mange tiår. Psykologen Schein lanserte senere sin teori om «a force field of driving and restraining forces», og hevder som Lewin at «you cannot understand a system until you try to change it» (Schein, 1996, s. 34): «Unless we intervene, we will not learn what some of the essential dynamics of the system really are» (Schein, 1996, s. 35).

Oftest blir evaluering benyttet for læring/utvikling og/eller kontroll, men i vårt tilfelle kan den også oppfattes som *legitimering* av autoritetsforankrede krav. Dessverre gis det ingen direkte sammenheng mellom «god undervisning» og «god læring», og de mest effektive tiltak for endret studieatferd og bedre læring er knyttet til et undervisningsbegrep som ikke benyttes i det undersøkte datagrunnlaget, nemlig undervisning som *emne- og programdesign*. Motsatt av formidlingsbaserte teorier om læring gir designperspektivet tilgang til langt flere variabler, som er mer inkluderende når det gjelder forklaring av læringsresultat. Ensidig fokus på interaktiv undervisning reduserer kompleksiteten og dermed mulighetsrommet for bedre læring. En mulig forklaring er trolig forestillinger om lærerens historiske rolle, og et kunnskaps- og læringssyn betinget av formidling. Dersom «kundepersepektivet» settes i system, bidrar det imidlertid verken til selvkritikk eller selvinnrett, som begge er viktige forutsetninger for økt egeninnsats og bedre læring. En professor beskriver situasjonen som følger:

Den trenden som dei fleste av studentane syner, er at dersom dei ikkje kan løyse ei oppgåve i same stund som dei les oppgåva, konkluderer dei med at dette kan dei ikkje fordi det har dei ikkje blitt lært av foreleser. Eg tykkjer difor det er vanskeleg å vite korleis eg skal gripe dette for å få dei til å forstå at eigeninnsats er viktig og at mange problem må ein arbeide med over tid for å løyse. Spesielt er dette utfordrande sidan det er så mange omsyn frå universitetet si side som gjer at studentane trur at dei er kunder til ei teneste, og at kunden alltid har rett. Denne haldninga kjem moglegvis frå at

universitetet har blitt så fokusert på at der skal vera så mange høver for studentane å gjeve tilbakemelding ... at studentane har tolka dette som at det alltid er forelesarane som skal streve etter å tilpasse seg og gjere det mest mogleg behageleg for studentane. Dermed har studenten lært å alltid tenke «kva gjer foreleser feil med omsyn på undervisning?» i staden for «kva kan eg gjere for å lære?» [NN]

Institusjonenes omtale av egne kvalitetssystem er typisk beskrivende, uten dypere analyser, vurderinger og konklusjoner. De eksterne evalueringspanelene kommenterer dette, men bidrar ikke med forslag som hjelper institusjonene fram til funksjonelle løsninger. Hensikten med å introdusere begreper og modeller er å gjøre aktørene bedre i stand til å gjøre reflekterte valg.

Sierpinska (1994) suggests that people feel they have understood something when they achieve a sense of order and harmony, where there is a sense of a «unifying thought», of simplification, of seeing some underlying structure and that in some sense, feeling that the essence of an idea has been captured (Swan, 2014, s. 3).

En vesentlig konseptuell utfordring ligger i perspektivskiftet fra undervisning til læring. La oss si vi er opptatt av å undersøke og fremme et gitt faglig nivå i et studium, for eksempel med utgangspunkt i den andre kvadranten i Figur 1. Fokus er her rettet mot læring, spesielt hva som framtrer som sterke og svake sider, og vurderingsdata kan være en aktuell tilgang. Eller vi kan observere og gjøre intervju med studentene, jamfør følgende replikkveksling:

Student: Let's see, 30–40 % of my courses have been very memory based. It's like you get an answer that does not make sense to you at all, you are unable make a sound judgment as to whether or not the answer makes sense.

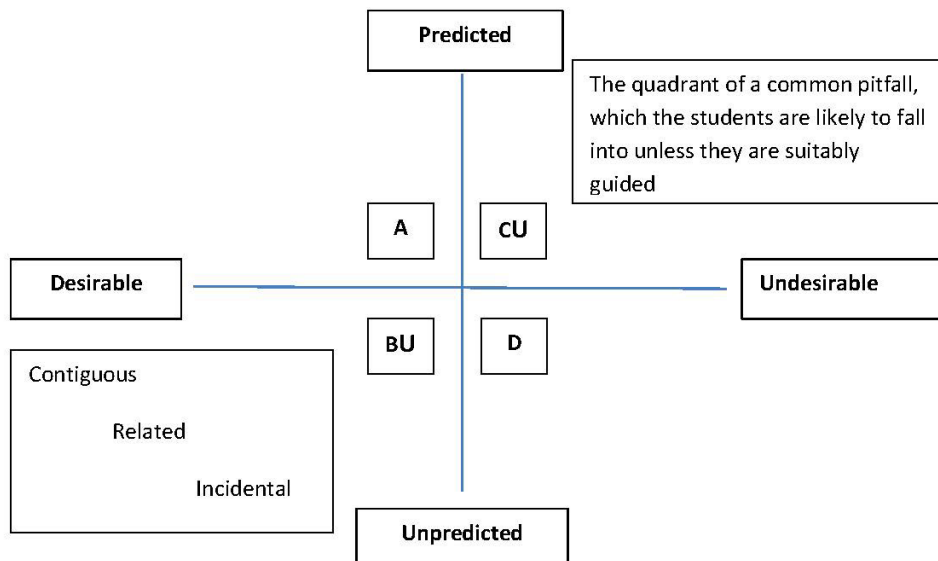
Interviewer: Do you also get exam questions of the same kind?

Student: Yes, very much so ... some of my professors assign the same tasks one year after the other. Most of my exams are like – yes, there are five kinds questions that keep reoccurring, and among these you can expect to get three or four. That's why you can specialize in advance on certain kinds of questions.

I dette tilfellet gir innspill fra kun én student grunn til uro i forhold til de krav som blir stilt, med det resultat at karakterene gir et feil bilde av reelt kompetansenivå. Ved den obligatoriske studentevalueringen var responsen svært positiv, og det ble konkludert med at undervisningen holdt «god kvalitet». Dette viser at spørsmål med utgangspunkt i implisitte didaktiske modeller dels avdekker og dels skjuler, avhengig av hvilke spørsmål som blir stilt. I eksemplet foran ble spørsmål rettet mot læring ikke adressert ved studentevalueringen, trolig med den begrunnelse at målet med evalueringen var å avdekke brist i «undervisningen». Eksemplet illustrerer betydningen av tilgang til ulike typer data. Å forbedre undervisningen vil, slik studenten uttrykker seg, bety arbeid med *emne- og oppgavedesign*. Dette kan sees som en nødvendig forutsetning for å få gyldige vurderingsdata for analyser i Kvadrant 2.

Et praktisk redskap med relevans for Kvadrant 2 og 3 er gjengitt i Figur 2 (Hussey & Smith, 2003). Oppmerksomheten rettes her mot mønstre eller tendenser for videre analyse.

Figur 2. Chart of Learning Outcomes (Hussey & Smith, 2003, s. 366)



Den konseptuelle tilnærmingen i Figur 2 kan benyttes ved all vurdering. Når perspektivene er forstått, kan sensor/faglærer ta notater med utgangspunkt i besvarelsene. Ved «Gap Analysis» trekkes læringsmålene inn som referanse for å sjekke måloppnåelsen. Det samme konseptet benyttes ved «learning analytics» som det knyttes stor interesse til, spesielt i forbindelse med bruken av digital teknologi i nye læringsplattformer. Vårt eksempel viser imidlertid at dette primært representerer en konseptuell tilnærming der bruken av teknologi er sekundær.

OPPSUMMERING

Ett av forskningsspørsmålene for denne undersøkelsen etterspurte konseptuelle modeller og teori som kan bidra til mer målrettet kvalitetsarbeid. I dette avsnittet skisserer forfatteren metodiske tilnærminger med antatt relevans for de fem institusjonene i undersøkelsen.

Fra metodemonomani til datamangfold

Studentevaluering av undervisning har vært et selvskrevet fenomen i kvalitetsarbeidet, jamfør følgende sitat: «Komiteen anbefaler at NN innfører obligatoriske emneevalueringer, og det må utarbeides selvstendige rapporter uavhengig av referansegruppene» (Tabell 2). Alt evalueringsarbeid er imidlertid krevende blant annet fordi behovet for *kriterier* og *standarder* alltid er en krevende utfordring, som sjelden gjøres eksplisitt. Denne undersøkelsen tyder dessuten på at mye av evalueringen dessverre har liten effekt, dersom en setter krav til bedre læring som mål. En åpenbar forklaring er i vårt tilfelle at oppmerksomheten i hovedsak er rettet mot virkemidler i stedet for mål. Dette perspektivet fanger ikke læringsmekanismene, hva som hemmer og hva som fremmer god læring. Undervisning oppfattes

som en individuell prestasjon som vurderes som mer eller mindre god eller dårlig, avhengig av forventninger og ønsker hos studentene. Oppmerksomhetens innretning styrer datainnsamlingen. Forfatteren argumenterer for reflekterte valg ut fra et teoretisk perspektivmangfold. Datainnsamling må sees i lys av valgt perspektiv, med vekt på dypere innsikt i sammenhenger bak tendenser i studieatferd og læring. Oppmerksomheten rettes mot årsaks- og virkningsforhold. Dersom vi skal forklare høy strykporsent i et emne, ligger den neppe i «dårlig undervisning» i klassisk betydning, men i et sett av mulige variabler der studenten selv er den viktigste aktøren.

Fra individbasert til systembasert tilnærming

Undersøkelsen har dokumentert et misforhold mellom intensjoner om kvalitet i utdanningen og institusjonens dokumentasjon med forslag til tiltak. Dersom *læring* er målet, trenger partene innsikt i hvilke mekanismer som hemmer eller fremmer tilsiktet læring. Å ta et læringsperspektiv på utdanningen er med andre ord nødvendig, og spørsmålet blir til sist hva studentene gjør, ikke hva læreren gjør (Biggs, 1999). Dette åpner for et handlingsteoretisk perspektiv på utdanning – fra et instruksjonsparadigme til et læringsparadigme med ulike typer design som grunnleggende utfordringer (Barr & Tagg, 1995). Selv om mekanismene i læringsprosessene ikke er direkte tilgjengelige, kan vi studere effektene i læringsresultatene.

En utdanning kan oppfattes som et delvis lukket sosialt system der vesentlige rammer fastholdes over tid, mens andre kan manipuleres i større eller mindre grad. Strukturer gir føringer for handling, og handling kan modifisere struktur, men dette tar tid fordi mange strukturer som oftest fastholdes. Hvis en begrepsmessig setter likhetstrekk mellom intensjoner i undervisningen og måloppnåelse, fortales kompleksiteten i systemet og dermed muligheten til aktivt å påvirke resultatet. Intensjonsutsagn angir retningen for læringsarbeidet, men realiseringen av mål skjer ved aktivering av de mekanismene som bidrar til måloppnåelsen. Årsaksforklaringer blir dermed systemisk og ikke individuelt forankret (Biggs, 1999).

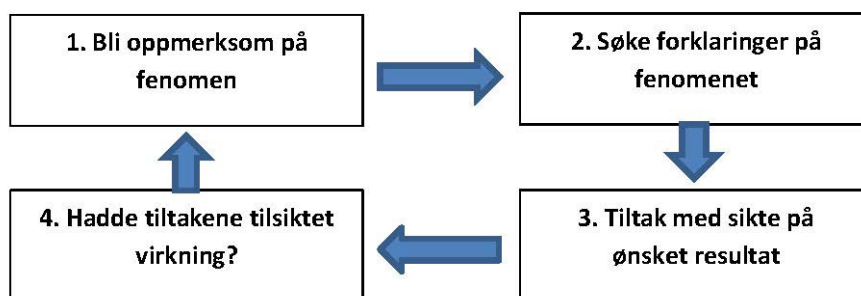
Fra beskrivelse til analyse og teoriarbeid

De sakkyndige evalueringsrapportene konkluderer med at kvalitetssystemene har vesentlige mangler: «Komiteens hovedinntrykk er at det i for liten grad fremkommer hvordan kunnskap generert av kvalitetsarbeid blir brukt til kvalitetsforbedring» (Tabell 1). Selvrapporteringen beskrives som deskriptiv og usystematisk der analyser, vurderinger og konklusjoner etterlyses spesielt. Ikke mindre tankevekkende er ekspertpanelenes råd for å bøte på situasjonen. Blant disse er forslag om å «tydeliggjøre kvalitetssystemet og utarbeide årsrapport»; «gjennomgå dagens kvalitetssystem»; «systembeskrivelsen er mangelfull» og «analyser, vurderinger og forslag til tiltak for forbedringer av kvaliteten i utdanningen må gjennomgående gis større oppmerksomhet og dokumenteres bedre» (fra Tabell 2). Som det framgår av eksemplene, er forslagene fri for konseptuelle modeller og teori til hjelp for institusjonene. Dermed framstår jobben med å forbedre kvalitetssystemene som en nesten uoverkommelig oppgave.

De omtalte kvalitetssystemene sliter med ulike utfordringer. Dersom hensikten med kvalitetsarbeidet er å sikre eller å forbedre «undervisningen», er det teoretisk sett mulig, men vurderingsgrunnlaget framstår som usikkert. Hva er kriteriene for god undervisning,

og hvem kan bedømme kvaliteten? Hvis hensikten er å forbedre læring, er datatilfanget rikere og arbeidsmåtene bedre utviklet. Figur 3 beskriver en fenomenologisk, resultatbasert tilnærming, som ikke benytter evaluering som metode. Ufordringen består i å undersøke distribusjon og omfang av et gitt fenomen, og hvilke mekanismer som skaper og opprettholder fenomener.

Figur 3. Resultatbasert modell for utviklingsarbeid



Arbeidsmåten krever systematikk og disiplin, men når tilnærmingen er kjent, slipper aktørene å streve med de foran omtalte utfordringene knyttet til alle former for evaluering. Teoriarbeid handler om å etablere sammenhenger mellom variabler for å forstå hvordan fenomen henger sammen i et system. Innsikt i mekanismene gir grunnlag for målrettede tiltak til endring.

KONKLUSJON

Et fellestrekk i selvevalueringene er deskriptive tilnærminger og fravær av gode analyser. Det blir mye rapportering og få troverdige konklusjoner for hvordan systemene kan forbedres. Datainnsamlingen skjer vesentlig ved bruk av studentevalueringer, uten bidrag fra andre kilder, som dokumentstudier, observasjon, intervju, logg eller vurdering. Datagrunnlaget blir dermed noe smalt, og det virker som evaluering langt på veg har blitt et mål i seg selv, men der institusjonene ikke evner å dra tilstrekkelig nytte av sitt eget prosjekt. Dette gir konturene av byråkratiske, toppstyrte, prosedyreorienterte og faglig svakt forankrete kvalitetssystemer, som blir lite produktive for det oppdraget som er gitt.

Hvilke råd gir så de sakkyndige komiteene til forbedring av utdanningskvaliteten? Et karakteristisk bidrag er det jeg har omtalt som «fine tuning», altså justeringer innen etablert praksis. Forslagene er velmente, men det gis liten grunn til å tro at tiltakene vil gjøre en stor forskjell. Evalueringspanelene bidrar ikke konseptuelt i den grad at institusjonene evner å se verden med nye øyne. I stedet blir de fortalt at «systembeskrivelsen er mangelfull» og at «rapportering om kvaliteten må styrkes». Kunnskapsgrunnlaget virker ikke tilstrekkelig hos aktørene. Samlet oppstår det et misforhold mellom innsats og resultater, slik at ressursbruken trolig tapper institusjonene mer enn den beriker. I tillegg er det fare for at kvalitetsarbeidet kommer i miskreditt, slik at bidragsviljen blant tilsatte reduseres. Selv om denne studien ble utført med utgangspunkt i fem institusjoner, rører trolig undersøkelsen ved tema av interesse for langt flere universiteter og høyskoler, som strever med å omsette intensjoner til praksis.

LITTERATUR

- Atherton, J., Hadfield, P., & Meyers, R. (2008). *Threshold Concepts in the Wild*. Paper presentert på Threshold Concepts: from Theory to Practice conference, Queen's University, Kingston Ontario.
- Barr, R. B., & Tagg, J. (1995). From Teaching to Learning: A New Paradigm for Undergraduate Education. *Change*, 27(6), 12–25.
- Beerens, M. (2015). Agencification Challenges in Higher Education Quality Assurance. I E. Reale & E. Primeri (Red.), *The Transformation of University Institutional and Organizational Boundaries* (Vol. 8, s. 43–61). Rotterdam: Sense Publishers.
- Biggs, J. (1987). *Student Approaches to Learning and Studying*. Newcastle, Australia: Australian Council for Educational Research Ltd.
- Biggs, J. (1996). *Enhancing teaching through constructive alignment*. *Higher Education*, 32(3), 347–364. DOI: <http://dx.doi.org/10.1007/bf00138871>.
- Biggs, J. (1999). What the Student Does: teaching for enhanced learning. *Higher Education Research & Development*, 18(1), 57–75.
- Biggs, J. (2001). The Reflective Institution: Assuring and enhancing the quality of teaching and learning. *Higher Education*, 41(3), 221–238.
- Case, J., & Marshall, D. (2009). Approaches to learning. I M. Tight, J. Huisman, K. H. Mok, & C. Morphey (Red.), *The Routledge International Handbook of Higher Education* (s. 9–21). London and New York: Routledge Falmer.
- Dahler-Larsen, P. (2013). Kvalitetens beskaffenhet og den svage tænkning. I T. Fosslund, G. Nettelund, & Y. Nordkvelle (Red.), *Kvalitet i fleksibel høyere utdanning – nordiske perspektiver* (s. 29–41). Oslo: Akademika forlag.
- Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, 12(2), 219–245.
- Forskrift om tilsyn med utdanningskvaliteten i høyere utdanning (studietilsynsforskriften) (2011).
- Goodyear, P. (2015). Teaching as design. *Herdsa Review of Higher Education*, 2, 27–50.
- Gynnild, V. (2014a, 25. August). From 'quality assurance' to 'quality enhancement': Addressing the transition from 'teaching' to 'learning' in engineering education. Paper presentert på iCEER2014 – International Conference on Engineering Education and Research, Hamilton, Ontario, Canada.
- Gynnild, V. (2014b). Kvalitetssystemet i praksis: Ressurser på avveie? *Uniped*, 37(3), 4–22.
- Gynnild, V., & Myrhaug, D. (2012). Revisiting approaches to learning in science and engineering: a case study. *European Journal of Engineering Education*, 1–13. DOI: <http://dx.doi.org/10.1080/03043797.2012.709483>.
- Hussey, T., & Smith, P. (2003). The Uses of Learning Outcomes. *Teaching in Higher Education*, 8(3), 357–368.
- Marton, F. (1994). Phenomenography. I T. Husén & T. N. Postlethwaite (Red.), *The International Encyclopedia of Education*, 8, 4424–4429). Oxford: Pergamon.
- Newton, J. (2002). *From policy to reality: enhancing quality is a messy business*, 16. Hentet fra http://www.heacademy.ac.uk/resources/detail/resource_database/id158_from_policy_to_reality.
- Sadler, D. R. (2009). Grade integrity and the representation of academic achievement. *Studies in Higher Education*, 34(7), 807–826.
- Schein, E. (1996). Kurt Lewin's change theory in the field and in the classroom: Notes toward a model of managed learning. *Systems practice*, 9(1), 27–47. DOI: <http://dx.doi.org/10.1007/BF02173417>.
- Sierpinska, A. (1994). *Understanding in Mathematics*. London: Falmer.
- Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG)*. (2015). Hentet fra Brüssels, Belgia: http://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf
- Swan, M. (2014). *Designing tasks and lessons that develop conceptual understanding, strategic competence and critical awareness*. Paper presentert på Encontro de Investigação em Educação Matemática (EIEEM 2014), Sesimbra.
- Yin, R. K. (1994). *Case study research: design and methods*. Thousand Oaks, California: Sage.