

Dypevaluering av studentenes læring - Er fokusgrupper svaret?

Welie Annett Schaathun

Hans Georg Schaathun

Høgskolen i Ålesund, Avdeling for ingeniør- og realfag

Sammendrag

Med utgangspunkt i en dypevaluering av et emne ved dataingeniørutdanningen ved Høgskolen i Ålesund, diskuterer vi i denne artikkelen hvordan samfunnsvitenskapelige, kvalitative metoder er egnet til å studere og evaluere studentenes læring på en helt annen måte enn de spørreskjemaene som vanligvis brukes ved emneevaluering. Artikkelen hevder at dette gir en formativ evaluering heller enn en retrospektiv, summativ evaluering. Dermed blir fokus på "quality enhancement" heller enn "quality assurance". Vi viser hvordan vi gjennom fokusgrupper med studenter har fått data som er svært nyttige for kvalitetsutviklingsarbeidet både med dette emnet og andre emner i vårt studium.

Innledning

Alle som underviser har vært gjennom evalueringer, av enkeltemner og av studieprogrammer. Institusjonene er pålagt ved lov å ha systemer for kvalitetssikring og kvalitetsutvikling, og evalueringene er en del av disse systemene. Hvorvidt det er de fagansatte eller administrasjonen som er ansvarlig for utformingen og gjennomføringen kan variere fra institusjon til institusjon. Likeså varierer metodene; spørreskjema er typisk men ikke universelt. Det eneste som er sikkert er at resultatene skal rapporteres til og arkiveres av administrasjonen, slik at gjester fra Riksrevisjonen og fra NOKUT kan druknes i data.

Mange som underviser har antagelig studert evalueringsresultatene nøye i håp om å finne idéer til forbedring. Med et genuint ønske om å bli en bedre foreleser og skape et bedre emne, er det naturlig å forvente nyttig informasjon i evalueringsmaterialet.

Som regel blir vi skuffet. Svarene vi får er som regel overfladiske, og skulle vi finne et svar som kan synes nyttig vet vi aldri om det er representativt for studentmassen eller om det er en statistisk utligger. Mange forelesere har opplevd å endre emnet basert på konstruktive tilbakemeldinger fra en gruppe studenter, bare for å møte like konstruktive tilbakemeldinger fra en annen gruppe det påfølgende året med bønn om å reversere endringen.

Kildene til problemer er mange. Lav svarprosent fra studentene er bare én av dem. Spørreskjemaene gir som regel overfladiske svar, gjerne med vage spørsmål og tvetydige svar. Fritekstsvarene kaster gjerne frem gode innspill, men ingen spør de andre studentene om å bekrefte eller problematisere de gode innspillene. I de fleste tilfeller er hele evalueringsprosessen utformet for å tilfredsstille et dokumentasjonskrav og ikke for å lære noe eller forbedre noe. Det er viktigere at evalueringen dekker alle tenkelige sider ved undervisningsopplegget, enn at det gir konstruktiv og nyttig informasjon med sikte på konkret forbedring.

Utforming av spørreskjema er en vitenskapsdisiplin i seg selv. Det er ikke lært i en håndvending, og det er en grunn for at samfunnsviterne har det på pensum og bruker tid på det. Sjelden eller aldri finner vi spørsmålsformuleringene problematisert i et informatikk- eller annet realfagsstudium. Hvor skal foreleserne ha kunnet lære god evalueringsmetodikk?

Ved Høgskolen i Ålesund har vi gjort et forsøk med dypevaluering av et emne som bl.a. inngår i dataingeniørstudiet. For å sikre både dybde og objektivitet har vi brukt samfunnsvitenskapelig, kvalitativ forskningsmetode, og investert tid både i datainnsamlingen og ikke minst i analysen. Evalueringen ble gjennomført i samarbeid mellom to fagansatte og en samfunnsviter innleid for formålet. Bare én av de tre hadde undervist på emnet, og da bare for to år siden.

Vi har i all hovedsak meget gode erfaringer fra dette forsøket. De som har vært med på prosjektet har lært mye, både om det konkrete emnet og generelt om undervisning og læring. I denne artikkelen drøfter vi erfaringene våre og sammenligner dem med de typiske og rutinemessige evalueringsmetodene som vi er vant med. Vi vil også se det hele i sammenheng med eksisterende litteratur om evaluering og kvalitetssikring generelt og i høyere utdanning spesielt. Tidligere arbeider som bygger på data fra denne studien har fokusert på erfaringer med aktiv læring og studentenes læring i lys av tradisjonelle læringsteorier (se Schaathun, Schaathun & Bye, 2015 og Schaathun & Schaathun, 2015).

Litteraturstudie: Evaluering og kvalitetssikring

Evaluering: Hva, hvem og hvordan?

Evaluering er et viktig og utbredt styringsinstrument i offentlig virksomhet. Som Vedung (2010, s. 263) beskriver: "If you carefully examine and assess the results of what you have done and the paths toward them, you will be better able to orient forward."

Evaluering kan defineres som nøye retrospektiv vurdering av intervensjoner i offentlig sektor, deres organisering, innhold, implementering og utfall hvor intensjonen er at disse vurderingene skal spille en rolle i framtidige praktiske situasjoner (Vedung, 2010, s. 264). En kortere definisjon finner vi hos Tornes (2012) hvor evaluering defineres som det å vurdere gjennomføringen av et tiltak eller et program. Disse evalueringene settes i gang når vi ønsker å gjøre oss opp en mening om reformene/programmene/tiltakene har gått bra eller mindre bra. Hun mener derfor at evalueringer kan karakteriseres som en form for tilbakemelding. Denne definisjonen er nok dekkende for hvordan begrepet brukes i dagligtale.

Evaluering har dermed å gjøre med å sikre kvalitet. Biggs (2003) snakker om to typer av "quality assurance". Den første typen er den retrospektive evalueringen som Vedung beskriver, som ser tilbake på det som har vært gjort og gjør summative vurderinger opp mot eksterne standarder. Biggs beskriver den som administrativ og byråkratisk, med ovenfra-og- ned prosedyrer for å sikre konformitet til eksternt oppsatte standarder. Mot dette setter han opp en annen type, en prospektiv quality assurance som kontinuerlig streber mot forbedringer, kalt "quality enhancement". Evaluering kan dermed være både summativ og orientert mot beslutninger og formativ og orientert mot utvikling (Henkel, 1998). Selv om både Biggs og Henkel er særlig opptatt av evaluering i høyere utdanning, vil deres innspill også være gyldig i forhold til evaluering generelt.

Vedung (2002) skiller mellom to hovedgrupper av evalueringsmodeller: substansmodeller ser på aspekter ved intervensjonen, virksomheten eller saken, mens de økonomiske modellene er opptatt av produktivitet og effektivitet. Innenfor substansmodellene finner vi både modeller som er opptatt av måloppnåelse og brukermodeller som retter seg mot møtet mellom virksomheten og dens klienter og spørsmålet om virksomheten leverer den kvaliteten som mottakeren krever eller behøver.

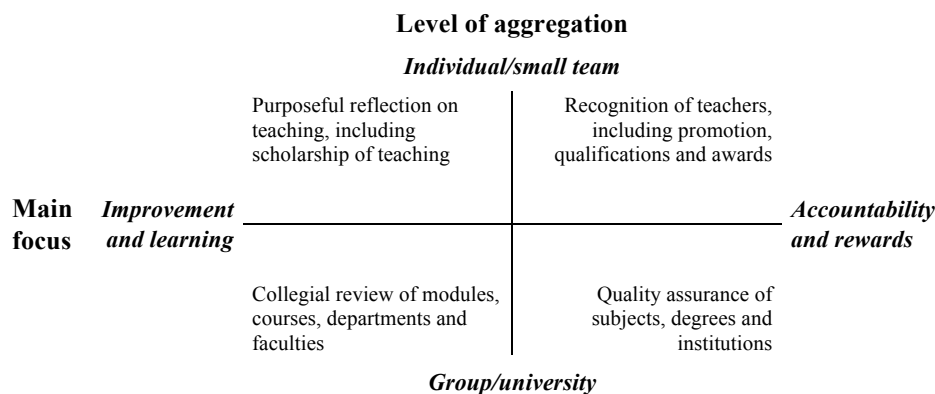
Ulike metoder for evaluering gjenspeiles i spørsmål rundt hvem som evaluerer, hva som evalueres, hvordan det evalueres, og når og hvor ofte det evalueres (Brennan & Shah, 2000; Handal, 1996). Et grunnleggende skille går også mellom eksterne og interne evalueringer (Brennan & Shah, 2000). En del evalueringer gjennomføres av eksterne organer, en del tar institusjonene selv initiativ til, men bruker eksterne evaluere, og en del skjer både på internt initiativ og ved interne krefter (Handal & Sandvik, 2008). Her vil noen argumentere for at evalueringer best gjøres av andre enn ens egne kollegaer, mens andre ser fordelene ved å dra nytte av den dype kunnskapen om området som finnes i kollegiet (Henkel, 1998).

Evaluering i høyere utdanning

Om evaluering i høyere utdanning sier Ramsden (2003, s. 209):

Evaluation is a way of understanding the effects of our teaching on student's learning. It implies collecting information about our work, interpreting the information and making judgements about which actions we should take to improve practice. To reflect on what helps students to understand a concept or argument, and to apply the results to teaching – that is to engage in evaluation.

Ramsden legger her vekt på at evalueringene må ta utgangspunkt i studentenes opplevelse av læring. Han konseptualiserer evaluering langs to dimensjoner. Den ene dimensjonen strekker seg fra evaluering av en enkelt forelesers undervisning via evaluering av emner til evaluering av hele institusjonen, mens den andre dimensjonen tar for seg hva som er hovedfokus for evalueringen, se figur 1 (Ramsden, 2003, s. 214).



Figur 1. To dimensjoner i evaluering i høyere utdanning

Undersøkelser blant studenter kan omfatte et bredt spekter av temaer. De kan spørre om årsaker og motiver for valg av utdanning og lærested, om atferd og holdninger til studieinnsats og læringsstrategier, om faktiske forhold som demografi, helse og levekår, om læringsutbytte eller om vurderinger av ulike aspekter ved utdanning og lærested (Wiers-Jenssen, 2014). Samtidig som undersøkelsene kan omfatte ulike tema, kan de også ha ulike formål. De kan gi studentene en mulighet til å bli hørt, ønske å forbedre studiekvaliteten, øke tilfredsheten eller sammenligne studietilbud. Disse undersøkelsene inngår gjerne som en del av studiestedenes kvalitetssikringssystemer. Siden studentundersøkelsene kan gi informasjon om forhold som man ikke får gjennom andre foreliggende data, er de et viktig supplement til tradisjonelle kvalitets- og styringsindikatorer (Wiers-Jenssen, 2014).

Handal (1996) beskriver en rekke former for studentevaluering. Tradisjonelt er spørreskjema mye brukt, men Handal kommer også med andre eksempler på skriftlige evalueringer som kan brukes, f.eks. notater, loggskrivning eller "brev til læreren". I tillegg til skriftlige evalueringsformer kan undervisningsevalueringen også foregå i samtaleform, både med bruk av smågrupper, referansegrupper, samtaler med kritiske studenter og intervjuer.

Evaluering og kvalitet

Gynnild (2014) beskriver kvalitetsbegrepet som prinsipielt åpent, og det gir rom for ulike tilnærminger eller perspektiver. Hvilket perspektiv som blir valgt, får betydning for både fokusering og progresjon av arbeidet. Det valgte perspektivet retter oppmerksomheten mot et aspekt ved kvalitet, noe som samtidig betyr at andre perspektiver havner i bakgrunnen. En tydeliggjøring må derfor til for at man skal kunne måle eller sammenligne kvalitet.

Hva kan så avgjøre kvaliteten på undervisningen? Avgjøres den av lærernes evne til å presentere faget, av undervisningens synliggjøring av fagområdets prosedyrer, handle- og tenkemåter, av programmets potensiale for å forberede studentene til et bestemt yrke eller bør kvalitet bli vurdert ut fra hvordan studieopplegget fremmer kognitiv endring hos studentene (Wittek & Habib, 2012)? Her vil svaret kunne variere ut fra hvem man spør. Politikere, akademikere, studenter, arbeidsgivere og andre vil ha ulike synspunkt på dette (Wittek & Kvernbekk, 2011). Til tross for at begrepet har vært diskutert i lange tider, er det fremdeles ingen enighet om en enkelt definisjon. Dette har igjen betydning for evalueringen i høyere utdanning, hva som evalueres, hvordan, når og av hvem.

Læringsutbytte og grad av oppnåelse av læringsmål har den senere tiden fått hovedfokus når utdanningenes innhold og nivå skal defineres, og det er da en naturlig følge at kvalitetsvurderinger av studier og institusjoner gir økt oppmerksomhet til studienes læringsresultat (Haakstad, 2011). Læringsutbytte retter fokuset på studentens læring og hva de skal kunne etter fullført studium, i motsetning til hva som skal undervises/læres bort. Det deles vanligvis inn i to hovedkategorier: fagspesifikt læringsutbytte som er de kunnskapene og ferdighetene som hører til faget og generelle (generiske) ferdigheter, også kalt nøkkelferdigheter, som er overførbare til andre områder (Aamodt, Prøitz, Hovdhaugen & Stensaker, 2007). Imidlertid har det vist seg vanskelig å finne en god operasjonell definisjon på læringsutbytte, og det er dermed vanskelig å måle. Karakterer, strykprosent, studiepoengsproduksjon og andel studenter som fullfører på normert tid, er alle mål på effektivitet i utdanningen, snarere enn mål på læringsutbytte (Aamodt et al., 2007). Det kan også stilles spørsmål ved om det er ønskelig eller i det hele tatt mulig å bryte ned og beskrive all kunnskap i målbare atferds- og mestringsstermer. Karlsen (2010) diskuterer hvordan det er stor fare for at registrering og måling av læringsutbytte blir en teknisk og byråkratisk øvelse. Dette fører til lite bruk av skjønn og liten forståelse for at dannelse, som også er en del av utdanningen, er et langsiktig prosjekt.

Kvalitetssikringssystemer i norsk høyere utdanning

Forskrift om kvalitetssikring og kvalitetsutvikling i høyere utdanning og fagskoleutdanning (Kunnskapsdepartementet, 2010) stiller i §2-1 krav om at

1. Universiteter og høyskoler skal ha et system for sitt kvalitetssikringsarbeid som sikrer kontinuerlige forbedringer, gir tilfredsstillende dokumentasjon av arbeidet og avdekker sviktende kvalitet.
2. Kvalitetssikringssystemet skal sikre og bidra til å utvikle kvaliteten i hele utdanningen, inkludert praksisstudier. Systemet skal omfatte alle forhold som har betydning for studiekvaliteten, fra informasjon overfor mulige søkere til avslutning av studiet inklusive studiets relevans for arbeidslivet.

En casestudie fra Høgskolen i Ålesund

Kvalitetssikringssystemet

Kvalitetssikringssystemet ved Høgskolen i Ålesund er beskrevet i Kvalitetshåndboken (Høgskolen i Ålesund, 2014). Kvalitetssystemet skal ikke bare sikre kvalitet, men også bidra til å utvikle høgskolen, og man har derfor valgt å benytte begrepet kvalitetsutvikling heller enn kvalitetssikring.

PA10 i prosedyrehåndboken (Kvalitetshåndbokens del 2) gir retningslinjer for emneevalueringer. Formålet med evalueringen er å sikre god kvalitet i undervisningen i tråd med læringsutbytte i alle emner på lavere og høyere grads nivå. Alle emner skal evalueres minst hvert femte år, og i tillegg etter behov ved bl.a. etablering eller endring av emne eller

ansvar, eller ved høy strykprosent eller klager fra studenter. Evalueringene foretas muntlig eller skriftlig i tråd med avdelingens retningslinjer, og skal fokusere på undervisningskvalitet, informasjon om faget, studentdeltakelse, andre aktiviteter som arbeid på laboratorier og praksis og andre tema som er relevant for faget. Funnene fra evalueringene rapporteres så oppover i systemet gjennom de rapporteringsskjemaene som er utarbeidet.

Dypevaluering i emnet *Mikrokontrollere*

Høsten 2014 gjennomførte vi en grundig og systematisk evaluering av undervisningsmetodene og studentenes læring i førsteårsemnet *Mikrokontrollere*, som er felles for programområdene data og elektro ved ingeniørutdanningen. Dette emnet ble startet opp som ledd i omleggingen da den nye rammeplanen for ingeniørutdanningene i Norge kom i 2012. Aktiv læring var et grunnprinsipp for emnet helt fra idéstadiet, og emnet inneholder mye praktisk labarbeid.

Da emnet ble kjørt for tredje gang høsten 2014, ønsket man en grundigere evaluering av emnet sammen med pedagogisk forankring i litteraturen. En uavhengig prosjektgruppe gjennomførte denne studien som søkte å gi svar på tre spørsmål ut fra studentenes synspunkt og deres erfaringer med emnet: *Hvilke aktiviteter skaper læring? Hvilke læringsaktiviteter skaper trivsel? Og hvordan bør vi utvikle læringsaktivitetene videre?*

For å finne svar på disse spørsmålene brukte vi kvalitativ, samfunnsvitenskapelig forskningsmetode, som gir en langt dypere analyse enn det som er mulig gjennom det ordinære kvalitetssikringsarbeidet ved høyskolen. Det var også viktig å avvike fra de tradisjonelle spørreskjemaene som mange studenter er lei av. Vi holdt derfor fokusgrupper med studenter fra emnet. Fokusgrupper er en metode som fungerer med relativt få informanter, og som ofte gir dype svar og mye data per informant (Krueger & Casey, 2009). Gjennom fokusgruppene kan forskerne utvikle en forståelse av *hvorfor* folk føler og mener det de gjør, og deltakerne i gruppen kan utfordre hverandre, diskutere, oppfordre til utdyping av meninger, eller få nye tanker og ideer (Bryman, 2008). Metoden er derfor godt egnet til å få fram ulike synspunkt i forhold til et spesifikt tema, til å finne ut hvor det er enighet og hvor deltakerne har ulike syn, og ikke minst hvorfor det er enighet eller uenighet. I tillegg til fokusgruppene gjorde vi også intervju med de involverte lærerne.

Invitasjon til å delta i prosjektet ble sendt til alle studentene på alle de tre årskullene på de to studieprogrammene som bruker emnet. Vi ønsket å ta med studenter både fra inneværende og tidligere år, for å kunne se endringer i emnet fra år til år, og for å se emnet bedre i sammenheng med resten av studiet. Det var også et håp om at å blande studenter fra ulike årskull ville bidra positivt til de samtalene og diskusjonene fokusgrupper som metode legger opp til.

Tre fokusgrupper med til sammen 19 studenter ble gjennomført. Den opprinnelige planen var at hver fokusgruppe skulle ha studenter fra alle årskullene fra hvert studieprogram. På grunn av vansker med rekrutteringen, lot ikke dette seg gjøre. Samlet sett er likevel alle årskullrepresentert fra begge studieprogrammene. Vanskene med rekrutteringen gjorde også at vi måtte rekruttere strategisk og spørre utvalgte studenter direkte om de ville være med. Dette har trolig ført til en overrepresentasjon av studenter som mestrer emnet godt og en underrepresentasjon av dem som sliter. Samtidig var det et par studenter som åpent fortalte om de vanskene de opplevde med emnet slik at også dette perspektivet er godt representert.

Fokusgruppene ble gjennomført av en moderator som ikke jobber ved høyskolen til daglig, og som ble hentet inn spesielt til dette prosjektet. Ekstern moderator så ut til å virke positivt, ved at studentene trolig var mer åpne og torde komme med negative kommentarer i større grad enn de ville ha gjort med en intern moderator. Også andre studier har gjort tilsvarende erfaringer (Gibbings & Bowtell, 2013). I tillegg kom studentene gjerne med

utfyllende og forklarende svar for å opplyse en moderator som ikke selv kjenner studiet.

Transkripsjonene fra fokusgruppene og lærerintervjuene ble først kodet og analysert av forskergruppen i samarbeid (Krueger & Casey, 2009), før kodeboken ble revidert og kodet opp igjen ved hjelp av dataprogram (RQDA) for videre analyse. Funnene knyttet til forskningsspørsmålene er tidligere presentert på to ulike konferanser i 2015, hvor fokuset på den første konferansen var aktiv læring, mens den andre så ingeniørstudenters læring i et utviklingspsykologisk perspektiv (Schaathun, Schaathun & Bye, 2015; Schaathun & Schaathun, 2015). I tillegg er det utarbeidet en mer omfattende intern rapport til bruk ved utvikling av emnet.

Drøfting

Spørreskjema eller intervju?

Ved evaluering av emner er det vanlig å bruke en eller annen form for spørreskjema, gjerne på papir som fylles ut i klasserommet eller elektroniske spørreskjema som studentene fyller ut hjemme. Spørreskjema har fordelen av at de er billige og raske å administrere og de kan nå ut til et stort antall personer. I tillegg er spørreskjema ofte praktiske og bekvemme for respondentene fordi de selv kan bestemme når og hvor de ønsker å besvare dem, og fordi spørsmålene ofte er enkle å svare på (Bryman, 2008). Bruk av spørreskjema er derfor velegnet når man ønsker å nå et stort antall studenter, som f.eks. NOKUTs nasjonale studentundersøkelse om studiekvalitet - Studentbarometeret.

Det er ofte en holdning til at spørreskjema er en enkel og grei evalueringsmåte som verken er spesielt tidkrevende eller vanskelig. Realiteten er imidlertid at det er krevende å lage gode spørreskjema som virkelig gir deg svar på de spørsmålene du har og som gir data du kan bruke videre. Lukkede spørsmål har den fordelen at de er lette både å svare på og å analysere, og det er lett å sammenligne svarene man får. Ulempen er, som vi skal se nedenfor, at de fort kan bli for generelle og upresise, og dermed gi svar som egentlig ikke sier noe som helst, eller at spørsmålene er tvetydige eller forsøker å spørre om flere ting samtidig (Bryman, 2008). Vi må innrømme at vi selv også har gått i disse fellene.

Et større problem når spørreskjema nærmest automatisk velges når emner skal evalueres, er at det sjelden stilles spørsmål ved om dette faktisk er den metoden som er best egnet til å nå det som er målet med evalueringen. Man begynner på en måte i feil ende, ved å først velge metoden som skal brukes, for så å bestemme formålet med undersøkelsen. Som Kvale og Brinkman (2015, s. 151) sier, er ”forskjellige metoder forskjellige verktøy for å svare på forskjellige spørsmål”. Problemet er, for å si det med Edström (2008, s. 99), at ”audit and development need rather different investigations”, og samme type data passer sjelden begge formål. Studier viser da også at foreleserne ofte ikke finner evalueringene nyttige med tanke på videre utvikling av emnet, og derfor gjerne ignorerer resultatene (Spooren, Brockx & Mortelmans, 2013).

Kvantitative metoder og spørreskjema er velegnet til å finne utbredelsen av noe, og til å si hvor fornøyde studentene er ved ulike aspekter ved undervisningen. Fokus kan da lett bli på de praktiske sidene ved undervisningen, som en test på hvor effektiv undervisningsmateriellet eller undervisningsmetodene er eller på foreleserens prestasjoner (Edström, 2008). Dette er også informasjon som kan være nyttig for å vurdere et undervisningsopplegg, og dersom det i hovedsak er denne type data man er ute etter, så er spørreskjema den metoden som bør velges. Problemet er at disse undersøkelsene ikke spør hvorfor studentene er fornøyde eller misfornøyde (Stålhane, Bratsberg & Midtstraum, 2012). Dersom fokus for evalueringen skal være læring, må man derfor tenke nytt. Det er nettopp denne erkjennelsen som ligger til grunn for vårt prosjekt.

Sammenligning av data fra evalueringene i *Mikrokontrollere*

I emnet *Mikrokontrollere* ble det gjennomført skriftlige evalueringer i form av elektroniske spørreskjema til studentene høsten 2012 og høsten 2013. Det første året ble evalueringen gjennomført etter at emnet var avsluttet, mens det året etter ble gjennomført midtsemesterevaluering. Høsten 2014 ble det ikke gjennomført skriftlig evaluering. Det er interessant å sammenligne data fra de skriftlige evalueringene med data vi fikk gjennom fokusgruppene, for å vurdere omfang, kvalitet og verdi for videre utvikling av faget.

Begge de skriftlige evalueringene hadde lav deltakelse fra studentenes side. Det første året kom det inn 12 svar etter flere runder med purringer, mens det året etter kom inn 22 svar. Dette er lite tatt i betraktning at emnet har rundt 90 studenter, og vi vet ikke om de studentene som har svart er representative for studentgruppen som helhet. Kanskje er det bare de mest kritiske og misfornøyde studentene som har svart, eller kanskje tvert i mot de studentene som har vært særlig positive?

Som nevnt er det å lage gode spørreskjema utfordrende. Gjennom vår studie har vi blitt bevisste på hvordan spørreskjemaene som tidligere har blitt brukt i emneevalueringene nok lider under et ønske om å både måle hvor fornøyde studentene er med undervisningen og få data for å forbedre emnet. Kombinasjonen av få svar og varierende kvalitet og presisjonsnivå på spørsmålene som ble stilt, gjør at det etter vår mening ikke er så mye nyttig informasjon å hente med tanke på utvikling av emnet. Dette kan vi vise gjennom å sammenligne noen av dataene fra spørreskjemaene og fokusgruppene.

Når studentene skal gradere hvor vidt de er enige eller uenige i en påstand som ”Måten faget var lagt opp var passende”, kan man spørre seg om hva man egentlig får svar på. Eller påstanden ”Arbeidsmengden gjennom semesteret var passelig”. For begge disse påstandene er det lett å svare ”jo, greit nok”, noe studentene i stor grad gjorde: 58% av studentene var delvis eller ganske enig i at faget var lagt opp på en passende måte, mens 75% var ganske eller helt enig i at arbeidsmengden var passelig. Men hva betyr det egentlig at noe er ”passende”?

I fokusgruppene ble både måten faget er lagt opp på og arbeidsmengden inngående diskutert. Emnet har vært bygget på en ide om at studentene først skal få prøve seg på egen hånd på labben før de så får forklart teorien bak det de har gjort. Studentene stiller seg mer tvilende til denne metoden: ”...nyttan av forelesningen forsvinner litte grann, føler jeg, hvis du allerede har lært deg gjennom å plages, og så kommer du på forelesning og så sier de alt du har lært deg gjennom å plages...” sa en student, mens en annen påpekte at ”Men vi ville liksom ha en liten sånn, et lite sånn programmerings baby-step kurs, da, før vi begynte med labber og sånt.”

Arbeidsmengden hadde studentene ulike oppfatninger om. ”...i alle fall vi som tok det, tok det ganske fort da, skulle jeg til å si, mikrokontroller-faget. Så for vår del så var det, jeg skal selyfølgelig ikke si for lite, men ja, vi arbeidet ikke ræva av oss...” sa en student, mens en av de andre hadde en annen opplevelse og sa ”Vi burde brukt mye, mye mer tid på å lære hva, hva som faktisk skjer, hva vi faktisk skal gjøre, i stedet for å bare, det skal være en sånn konkurranse med å ploge gjennom mest mulig oppgaver på kortest mulig tid.” Dette ble ytterligere nyansert av en tredje student; ”...med tanke på arbeidsmengde, da, så følte jeg av og til at det var veldig mye, men veldig lite utfordrende.” Samtidig kom det også fram hvor positivt det er at studentene må bruke mye tid på labben: ”Sånn du sitter borte på labben en hel masse folk, og kaster lysdioder mot hverandre og prøver å få dette her til å virke og. Der er en god stemning, og folk har det gøy, ikke sant, og du sitter jo i mange timer og knoter og leker deg og...”

En vanlig feil ved spørreskjema er også at man prøver å stille flere spørsmål i ett spørsmål. I spørsmålet ”Hva synes du om forelesningene (innhold, metode, mengde)?” hvor svarene skal graderes fra 1 dårlig til 5 meget bra er det jo faktisk tre spørsmål i ett, noe som gjør det umulig å svare hvis studenten for eksempel mener at innholdet av forelesningene er

bra, men at det er alt for få av dem. Over 70% av studentene svarte 2 eller 3, så noe var de misfornøyd med, men hva, er vanskelig å si. Tilsvarende får vi når studentene skal gradere spørsmålet ”Veiledning, tilbakemelding og oppfølging fra faglærerne?”, hvor veiledningen studenten faktisk får kanskje er bra, men det er for lite av den, noe som var akkurat det mange studenter sa i fokusgruppene. På dette spørsmålet fordeler svarene seg på alle svarkategorier.

Gjennom fokusgruppene fikk vi data med en helt annen dybde med tanke på videre utvikling av emnet. Her fikk studentene gå inn på hva, hvorfor, hvordan – alle de spørsmålene som lett melder seg når vi leser resultatene fra en spørreundersøkelse. Ta spørsmålet ”Hva synes du om forelesningene (innhold, metode, mengde)?”. I fokusgruppene snakket studentene som vist over om hva de syntes var bra og hva som var mindre bra med innholdet og metoden i forelesningene. De kom også med forslag til hvordan de syntes opplegget rundt forelesningene skulle være, og forholdet mellom lab og forelesning: ” ... hvis en hadde hatt forelesning på mandag, der en gikk igjennom litt av oppgaven, da, og så for eksempel seinere i uken, da, så kunne en hatt en spørretime eller noe sånn”.

Tilsvarende fikk vi lange og utfyllende diskusjoner rundt oppgavene og oppgaveutformingene og forslag til konkrete endringer for hvordan dette kunne bli bedre. ”...følte jeg, jeg satt veldig mange timer der nede og gjorde ting som jeg like godt kunne droppet å gjøre på disse innleveringene. Ville heller ha kunnet bryne meg på noe vanskelig.” sa en student, mens en annen utdypet med ”Gjentagende spørsmål, da, utfylling av kommentarer på... Først skal du kommentere og forklare på hele koden, så kommer det spørsmål etterpå der skal du forklare enkeltvis på det du har forklart i oppgaven før.”

Studentene trakk også inn andre emner for sammenligning, både positivt og negativt, og kom også med ideer fra andre emner som de mente kunne være aktuelle i *Mikrokontrollere*. En student forteller for eksempel om foreleseren i et annet emne at ”...han har videosnutter på, på undervisning hvor han forklarer enkelte deler av oppgaver eller emner vi har i faget. Så du kan når som helst gå tilbake til videoene og se dersom det er noe du lurer på, det hadde kanskje vært bedre om det hadde vært noe sånn i *Mikrokontrollere* også.”

Det ligger i de kvalitative metodene at dataene som samles inn er lite generaliserbare (Bryman, 2008). Tvert imot ligger styrken ved kvalitative data i dybde og bredde, noe som også viser nyanser i studentenes oppfatninger. Vi får et godt inntrykk av generaliserbarheten innenfor de aktuelle studentkullene når fokusgruppen viser konsensus eller uenighet, og også når de viser til hva andre studenter gjør og mener. Selv om vi kun intervjuet et ganske lite antall studenter kan vi trekke mange robuste slutninger som må regnes som gyldige for en signifikant andel av studentene. Samtidig får vi et solid inntrykk av konkurrerende perspektiver og synspunkter, der vi nok må ta hensyn til alle. Generaliseringen kan derimot ikke kvantifiseres, og vi kan ikke bedømme hvor stor andel av studentene som har ens oppfatning.

Vi fant at dataene vi fikk gjennom fokusgruppene er overførbare til andre emner, og prosjektet ble dermed noe mer enn bare en evaluering av dette ene emnet. Studentene snakket bl.a. mye om samarbeid med andre studenter: ”Men det er mye samarbeid, det vil jeg si, med de som, skulle til å si, klikker med. Vi sitter og jobber og går og spør hverandre og... og noen har gjort det mer enn andre og noen var flink på, noen har hatt elektronikk før og noen har programmert mye før, og så mingler man rundt og, ja... Og det fungerer bra.” Som en følge av dette er det planlagt et forsøksprosjekt med samarbeidslæring i fellesemnet i matematikk for alle førsteårsstudenter på ingeniørutdanningene fra høsten 2016.

Det må likevel påpekes at på samme måte som kvaliteten på dataene som samles inn gjennom spørreskjema avhenger av godt utformede spørreskjema med gode spørsmål, er også kvaliteten på dataene fra fokusgruppene avhengig av at det stilles gode spørsmål. Vi brukte mye tid på å utarbeide en intervjuguide med både nøkkelspørsmål og en rekke oppfølgingsspørsmål og det som engelskspråklig litteratur kaller ”probes”. I dette arbeidet var

det sentralt at en av deltakerne i gruppen hadde god kjennskap til samfunnsvitenskapelige forskningsmetoder.

Hva med andre evalueringsmetoder?

Det finnes mange ulike metoder for evaluering som ulike aktører har prøvd med vekslende hell. Handal (1996) beskriver f.eks. bruk av referansegrupper hvor en liten gruppe studenter jevnlig møter med faglærer og tar opp ulike tema knyttet til undervisningen som studentgruppen som helhet er opptatt av. I praksis varierer det imidlertid hvor godt dette fungerer. Metoden er prøvd med noe hell ved NTNU (Stålhane et al., 2012), men det viser seg å være vanskelig å engasjere hele studentgruppen og det som diskuteres i referansegruppen blir derfor ofte meningene til de studentene som sitter i gruppen. Tilsvarende har en av artikkelforfatterne erfaring med at studentene i slike referansegrupper gjerne har lite tid til rådighet både til å snakke med de andre studentene og med faglærer, og at det derfor ofte bare er praktiske detaljer ved undervisningen som blir tema. I teorien kunne en referansegruppe gi samme type data som vi fikk gjennom fokusgruppene, men det ville kreve at referansegruppen fikk bedre tid og gode spørsmål de skulle diskutere, og at fokus dermed ble dreid bort fra de praktiske sidene ved undervisningen og over på diskusjoner om vilkårene for læring.

Stålhane et al. (2012) beskriver hvordan de med stor suksess bruker en metode hvor studentene skal oppgi de 3 tingene de liker ved undervisningen og de 3 tingene de hater, og hvordan disse svarene videre brukes i utviklingsarbeidet. Sentralt her er at studentene blir fortalt at foreleser ønsker å forbedre undervisningen og derfor trenger deres hjelp, og at studentene raskt får tilbakemelding på hva foreleser vil styrke eller fjerne. En tilsvarende metode, kalt ABC¹, er brukt med hell av en av artikkelforfatterne. Her får studentene i oppgave å skrive en ting de vil "Alter", en de vil "Begin" og en de vil "Continue", og dette blir så brukt for å forbedre undervisningen.

Felles for disse metodene er imidlertid at de lett fokuserer på praktiske aspekter ved undervisningen, at de har fokus på "teaching" heller enn "learning" (Edström, 2008). Nettopp her mener vi fokusgruppen har et fortrinn, ved at det er tid til de dype diskusjonene og at de nøye forberedte spørsmålene lettere leder diskusjonen inn på læringsaspektet. På den andre siden finnes det også eksempler hvor evaluering ved bruk av fokusgrupper har vært mer undervisningsfokuserende og også målt studentenes tilfredshet med emnene (Gibbins & Bowtell, 2013). Dette viser hvor viktig det er å vite hva man ønsker med evalueringen og lager gode spørsmål som passer til dette formålet uansett hvilken metode man velger å bruke.

Dypevaluering – en kostbar affære?

I denne studien ble de innsamlede dataene grundig analysert ved hjelp av samfunnsvitenskapelige forskningsmetoder. Alle fokusgruppene ble fullstendig transkribert, noe som tar mye tid, og dataene ble analysert både manuelt og ved hjelp av dataverktøy av en forskergruppe på tre personer. Denne grundige og systematiske analysen førte til vi har fått god forståelse av studentenes læring både generelt og i dette emnet spesielt.

Fokusgruppeintervjuene ble gjennomført av en ekstern moderator. Dette fant vi å være svært vellykket, og en medvirkende årsak til at vi fikk den dybden i dataene som vi fikk. Det var tydelig at studentene torde være mer åpne og også kritiske siden de ikke snakket med en av avdelingens lærere som de kanskje måtte forholde seg til ved en senere anledning. Studentene visste også på forhånd at de i transkripsjonene av fokusgruppeintervjuene var

¹ Andreforfatteren møtte denne metoden da Dr. Andrew Comrie brukte den i kurset Postgraduate Certificate of Academic Practice ved Universitetet i Surrey 2006/07.

anonymisert, og at kun moderator ville ha tilgang til opptakene av intervjuene. Det at moderator ikke var godt kjent med høgskolen, avdelingen, studiet og emnet gjorde også at studentene var mer utdypende/utfyllende i svarene og villig fortalte og beskrev ting de nok ikke ville ha gjort med en moderator som var kjent ”på huset”. Dette ga etter vår mening bedre og dypere data.

Det ble lagt ned anslagsvis 400 timeverk på prosjektet, halvparten på ekstern samfunnsviter og halvparten på de interne fagansatte. Timeantallet ble ekstra høyt siden vi fikk en del førstegangskostnader ved at dette var en ny metode for evaluering for alle parter. I tillegg ble mye arbeid lagt i analysen, en kostnad som ved bruk av fokusgrupper alltid vil være høyere enn når vi evaluerer med spørreskjema. Så føler vi også at vi sitter igjen med kunnskap som er langt ut over det vi har fått gjennom tidligere evalueringer. Å gå i dybden på studentenes læring har gitt innsikt som har potensielt stor betydning for utformingen av undervisningen både i dette og andre emner. Vi opplever derfor at denne dypevalueringen har vært verdifull, og at et slikt dypdykk med jevne mellomrom kan være viktig for kvalitetsutviklingen av studiene våre.

I forhold til å ha fokusgrupper som ordinær emneevaluering, vil det imidlertid ikke være mulig å bruke like mye tid, penger og personell. Det er likevel måter å tilpasse metoden på for å gjøre den mer håndterlig og mindre ressurskrevende: Det er ikke nødvendigvis behov for å ha flere fokusgrupper; en fokusgruppe vil kunne gi mye nyttig informasjon. I tillegg er det ikke nødvendig med en fullstendig transkripsjon av intervjuene. Det blir i løpet av fokusgruppeintervjuene alltid noe prat om ting som er på siden av tema, og ikke alt som sies er like interessant og relevant. Ved kun å transkribere de relevante delene av intervjuet, vil man kunne spare mye tid. Det er også mulig å la være å transkribere, og heller basere seg på å notere ned interessante poeng. I en ordinær evaluering vil man også ha en mindre omfattende analyse, hvor fokuset ikke er så mye på å få en fullstendig forståelse av materialet, men mer leter etter det som er viktig og nyttig for den videre utviklingen av emnet. Vi erfarte at en ekstern moderator var verdifullt, men noe av den samme effekten kan trolig oppnås ved at fokusgruppen gjennomføres av en person fra høgskolen som studentene ikke har jevnlig kontakt med.

Quality assurance eller quality enhancement?

Biggs (2003) beskriver to ulike typer av evalueringer, ”quality assurance” og ”quality enhancement”. Ut fra det som er beskrevet over om de dataene vi får ved hhv. spørreskjema og fokusgrupper kan vi si at spørreskjema i hovedsak representerer den første typen og fokusgruppe den andre.

Emneevaluering ved bruk av spørreskjema kan lett bli en retrospektiv, summativ vurdering av det som er gjort. Ofte er det også sentralt utformede spørreskjema som alle skal bruke, selv om dette ikke er tilfelle ved Høgskolen i Ålesund. De rapportskjemaene som finnes legger likevel føringer for evalueringene. Samtidig kan evaluering som en del av en administrativ rutine hvor evalueringen skjer til bestemte tider også passe dårlig inn tidsmessig i forhold til når det passer med tid for revisjon av et emne. Evalueringen kan dermed lett bli en administrativ og byråkratisk øvelse.

Hva som faktisk skal gjøres med resultatene fra en emneevaluering er heller ikke alltid like klart. På generelt grunnlag er vel erfaringene at de ofte havner i en skuff uten at de brukes til videre utviklingsarbeid (Spooren et al., 2013). Dette kan skyldes at evalueringsarbeidet er noe som er pålagt ovenfra, og som de som faktisk underviser i emnene ikke selv ser nytten av. Det kan imidlertid også skyldes at kvaliteten på dataene man får gjennom spørreskjema som vist ikke alltid er like god, og at det rett og slett er vanskelig å få noe ut av resultatene som kan brukes til utvikling av emnet (Edström, 2008).

De dataene vi samlet inn gjennom fokusgruppene, er etter vår mening bedre egnet til å utvikle emnet videre. Her er fokus på hva som kan forbedres og på studentenes læring og læringsutbytte. Dette harmonerer godt med intensjonene for Kvalitetssikringsystemene for høyere utdanning som gradvis har gått over til et økt fokus på læringsutbytte. Med økt fokus på læringsutbytte må fokuset også økes på studentenes læring, på hva som skaper læring og på vilkår for at denne læringen skal finne sted. Dette er informasjon man vanskelig kan få gjennom skriftlige spørreskjemaevalueringer. Her må man snakke med studentene. Fokusgruppene viste seg å være en særlig godt egnet metode for å ta opp denne typen tema med studentene, med diskusjon og refleksjon i fellesskap. Studentene i vår studie uttrykte da også at de likte denne typen av evaluering svært godt, og de ville gjerne ha fokusgrupper som så på hele studieprogram.

Konklusjon

Med bakgrunn i en dypevaluering i et førsteårsemne, har vi i denne artikkelen diskutert hvordan vi best kan evaluere undervisningen med fokus på studentenes læring. En sammenligning av de dataene vi har fått gjennom fokusgrupper og data fra tidligere spørreundersøkelser, viser at dersom målet er å forbedre undervisningstilbudet, heve kvaliteten på studiet og bedre studentenes læring, vil en mer kvalitativ tilnærming på evalueringen gi et bedre grunnlag for det videre utviklingsarbeidet. Gjennom bruk av fokusgrupper opplevde vi at målet ble "quality enhancement" heller enn "quality assurance", for å bruke Biggs (2003) sine begrep.

Selv om fokusgrupper er en ganske ressurskrevende metode, er det likevel mulig å tilpasse metoden slik at det er praktisk gjennomførbart å bruke den til emneevaluering. Siden fokusgrupper gir en unik dybdeinnsikt i undervisning og læring som også er overførbart til andre emner, vil man kunne få mye igjen for å ta en slik dypevaluering. Aller best er det kanskje dersom man klarer å kombinere det beste fra flere metoder: fokusgrupper for å gi dybde og retning for bedre læring, gode spørreskjema for å måle hvor fornøyde studentene er med ulike aspekter ved undervisningen (som kan fylles ut i klasserommet/labben for å få en høy svarprosent) og i tillegg metoder for rask tilbakemelding i klasserommet på hvordan det som foregår der fungerer akkurat nå.

Referanser

- Biggs, J. (2003). *Teaching for Quality Learning at University* (2. utg.). Maidenhead: The Society for Research into Higher Education & Open University Press.
- Brennan, J., & Shah, T. (2000). *Managing quality in higher education: an international perspective on institutional assessment and change*. Buckingham: Open University Press.
- Bryman, A. (2008). *Social Research Methods* (3. utg.). Oxford: Oxford University Press.
- Edström, K. (2008). Doing course evaluation as if learning matters most. *Higher Education Research & Development*, 27(2), 95-106.
- Gibbings, P., & Bowtell, L. (2013). *Student focus groups*. Paper presented at AAEE Conference, Queensland, Australia
- Gynnild, V. (2014). Kvalitetssystemet i praksis: Ressurser på avveie? *Uniped*, 37(3), 4-22.
- Handal, G. (1996). *Studentevaluering av undervisning. Håndbok for lærere og studenter i høyere utdanning*. Oslo: Cappelen Akademisk Forlag.
- Handal, G., & Sandvik, L. (2008). Internevalueringens legitimitet - et eksempel fra Universitetet i Oslo. *Norsk pedagogisk tidsskrift*, 92(6), 428-442.

- Henkel, M. (1998). Evaluation in Higher Education: conceptual and epistemological foundations. *European Journal of Education*, 33(3), 285-297.
- Høgskolen i Ålesund. (2014). *Håndbok for kvalitetsutvikling for Høgskolen i Ålesund*. Ålesund: Høgskolen i Ålesund Hentet fra <http://kvalitet.hials.no/khb>.
- Haakstad, J. (2011). Læringsutbytte: Begrepets anvendelighet i kvalitetsvurdering av høyere utdanning. *Uniped*, 34(4), 72-81.
- Karlsen, G. E. (2010). Kvalifikasjonsrammeverk - virkemiddel for kvalitet eller ensretting? *Uniped*, 33(3), 5-17.
- Krueger, R. A., & Casey, M. A. (2009). *Focus groups: A practical guide for applied research* (4. utg.). Los Angeles, Calif: Sage.
- Kunnskapsdepartementet. (2010). *Forskrift om kvalitetssikring og kvalitetsutvikling i høyere utdanning og fagskoleutdanning*. Oslo: Departementet.
- Kvale, S., & Brinkman, S. (2015). *Det kvalitative forskningsintervju* (3. utg.). Oslo: Gyldendal Akademisk.
- Ramsden, P. (2003). *Learning to Teach in Higher Education*. London: RoutledgeFalmer.
- Schaathun, W. A., & Schaathun, H. G. (2015). *Mellom klassisk betingning og støttende stillas: Et utviklingspsykologisk perspektiv på ingeniørstudenters læring*. Paper presentert på Fjordkonferansen 2015, Loen.
- Schaathun, W. A., Schaathun, H. G., & Bye, R. T. (2015). Aktiv Læring i Mikrokontrollarar. *Uniped*, 38(4), 381-389.
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research*, 83(4), 598-642.
- Stålhane, T., Bratsberg, S. E., & Midtstraum, R. (2012). Course improvement the TQM way. I: Norsk informatikkonferanse (Red.), *Norsk informatikkonferanse (NIK 2012 : Bodø)*. Trondheim: Akademika.
- Tornes, K. (2012). Evaluering - definisjon og fokus. I: K. Tornes (Red.), *Evaluering i teori og praksis*. Oslo: Akademia forlag.
- Vedung, E. (2002). Utvärderingsmodeller. *Socialvetenskaplig tidskrift*(2-3), 118-143.
- Vedung, E. (2010). Four Waves of Evaluation Diffusion. *Evaluation*, 16(3), 263-277.
- Wiers-Jenssen, J. (2014). Studentenes vurderinger av utdanning og lærested. I: N. Frølich, E. Hovdhaugen & L. I. Terum (Red.), *Kvalitet, kapasitet & relevans. Utviklingstrekk i norsk høyere utdanning*. Oslo: Cappelen Damm Akademisk.
- Wittek, L., & Habib, L. (2012). Undervisningskvalitet i praksis. *Norsk pedagogisk tidsskrift*, 96(3), 223-236.
- Wittek, L., & Kvernbekk, T. (2011). On the Problems of Asking for a Definition of Quality in Education. *Scandinavian Journal of Educational Research*, 55(6), 671-684.
- Aamodt, P. O., Prøitz, T. S., Hovdhaugen, E. & Stensaker, B. (2007). Læringsutbytte i høyere utdanning. En drøfting av definisjoner, utviklingstrekk og måleproblemer. Oslo: NIFU STEP.