



**Høgskulen
på Vestlandet**

MASTEROPPGAVE

**Digital eksamen – mer enn strøm på
gamle eksamensoppgaver?**

**Digital exams - more than digitalizing
old assignments?**

Johanne Marie Trovåg

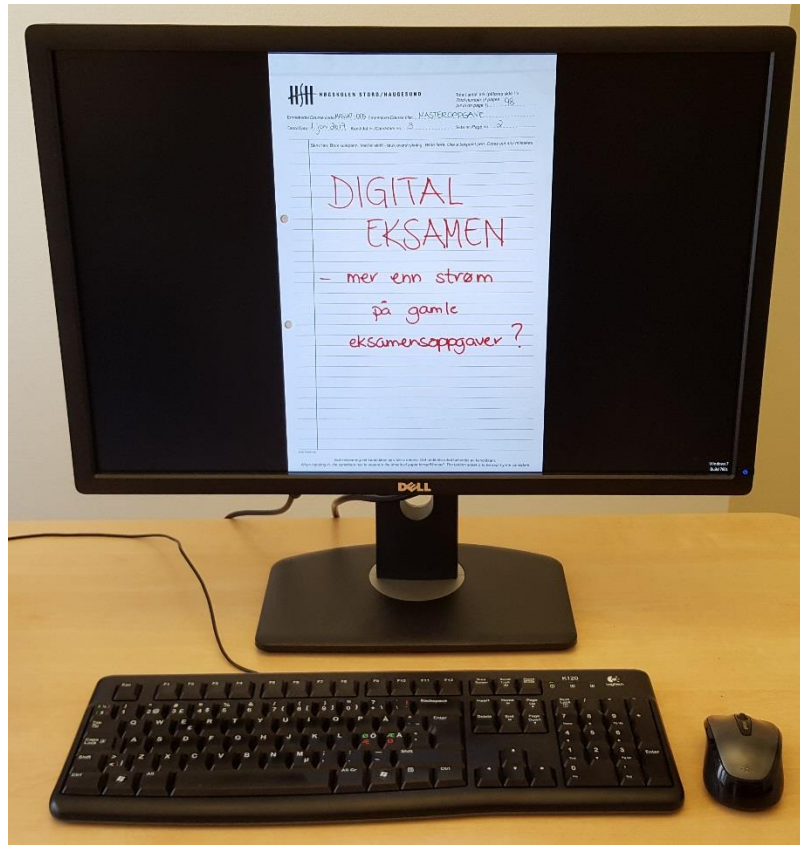
Master i IKT i læring

Avdeling for Lærerutdanning og kulturfag

Veileder: Knut Steinar Engelsen

1. juni 2017

Jeg bekrefter at arbeidet er selvstendig utarbeidet, og at referanser/kildehenvisninger til alle kilder som er brukt i arbeidet er oppgitt, jf. Forskrift om studium og eksamen ved Høgskulen på Vestlandet, § 10.



DIGITAL EKSAMEN – MER ENN STRØM PÅ GAMLE EKSAMENSOPPGAVER?

Høgskulen på Vestlandet, Våren 2017

MASIKT-OPG

Johanne Marie Trovåg

Sammendrag

Innføring av digital eksamen er et aktuelt tema i norsk høyere utdanning, og mange forelesere må i den forbindelse forholde seg til helt nye problemstillinger. Denne masteroppgaven undersøker hvordan tradisjonell skriftlig skoleeksamen på nautikkstudiet kan tilpasses en digital kontekst. Studien søker svar på hvordan foreleserne tenker når de lager eksamensoppgaver, og hva studentene mener om sammenhengen mellom eksamensoppgavene og emnets læringsutbytte.

Prosjektet er gjennomført som en kvalitativ studie med Educational Design Research som overordnet metodikk. Datainnsamlingen er gjort gjennom refleksjonsmøter med foreleserne som er involvert i emnet og fokusgruppeintervjuer med studenter som har tatt emnet. Datagrunnlaget består av synspunktene til to forelesere og totalt tolv studenter, innsamlet gjennom tre refleksjonsmøter og tre fokusgruppeintervjuer.

Teorigrunnlaget for oppgaven tar utgangspunkt i et sosiokulturelt læringssyn, med fokus på praksisfellesskap, appropriering av medierende artefakter, sammenheng mellom læringsutbytte, arbeids- og vurderingsformer og backwash-effekt. I tillegg er forskning på vurdering sentralt, med et hovedfokus på summativ vurdering og validitet i vurderingen. På denne måten søkes problemstillingen belyst ut fra et helhetlig perspektiv.

Studien viser at foreleserne tar utgangspunkt i læringsutbyttet de ønsker å teste, for så å velge hensiktsmessig oppgavetype ut fra dette, og at det derfor er de pedagogiske perspektivene som er hovedfokus for foreleserne når de utarbeider eksamensoppgaver. På denne måten fullender de en helhet som knytter emnets innhold, arbeidsformer og vurdering sammen. Denne helheten gir vurderingen god validitet og bidrar til en positiv backwash-effekt på studentenes læring. Det fremgår av studien at også studentene opplever sammenhengen mellom eksamensoppgavene og emnets læringsutbytte som god. På denne måten har foreleserne lyktes med å tilpasse eksamen til en digital kontekst, og studenter og forelesere har gjennom emnet appropriert digital eksamen som medierende artefakt.

Abstract

In Norwegian higher education of today, the introduction of digital exams is a highly relevant topic, causing many lecturers having to adapt to a new assessment regime. The objective of this master's thesis is to explore how a traditional written exam in one of the courses within the Bachelor of Nautical Science can be adapted to a digital context. The main questions of the thesis are how the lecturers think when they develop the assignments, and the students' perception of the correlation between the exam and the course learning outcomes.

The project is a qualitative study, in which Educational Design Research is used as overall framework. The data collection is carried out through reflection meetings with the involved lecturers and focus group interviews with students. The collected data consist of the thoughts and views of two lecturers and twelve students gathered through three reflection meetings and three focus group interviews.

The theoretical framework of the thesis is placed within a sociocultural perspective on learning, with a focus on collaborative learning, appropriation of mediating artefacts, correlation between learning outcomes, learning activities and assessment, and the backwash effect. In addition, research on assessment is essential, with a main focus on summative assessment and validity in the assessment.

The study shows that it is the pedagogical perspectives, which is the lecturers' starting point in the process of choosing suitable type of task and developing the assessment tasks. This way, they accomplish an alignment between the course content, the learning activities and the assessment within the course, an alignment which ensures good assessment validity and contribute to a positive backwash-effect on teaching and student learning. This corresponds with the students' experience, and the students describe the correlation between the course learning outcomes and the assessment tasks as good. The lecturers have been successful in adapting the assessment of the course to a digital context, and students and lecturers have appropriated digital exam as a mediating artefact.

Forord

«Doing qualitative research is in many respects no different than doing everyday life: it is complex and sometimes downright chaotic»

(Silverman, 2013).

Denne masteroppgaven er et resultat av et spennende og lærerikt prosjekt som ikke hadde vært mulig uten en hel rekke flotte mennesker:

Først og fremst vil jeg rette en stor takk til studenter og forelesere som har stilt opp og latt seg intervju. Ikke bare har dere gitt meg datamateriale til oppgaven, men dere har gitt meg ny innsikt som jeg tar med meg videre i livet og i jobben.

Takk til veilederen min, Knut Steinar Engelsen, som har bidratt med forventninger i tillegg til konstruktive og nyttige tilbakemeldinger underveis.

Takk til min arbeidsgiver – Høgskolen Stord/Haugesund da jeg begynte på studiet og Høgskulen på Vestlandet når jeg nå fullfører – som legger til rette for kompetanseheving.

Takk til familien min som bidrar til å få logistikken til å gå opp i hverdagen, og da spesielt til Mamma; Jeg er deg evig takknemlig!

Takk til mine to gullfine unger, som heier på mamma og som er rågode til å få tankene over på andre ting, enn jobb og studier, når vi er sammen!

Og til slutt en ekstra stor takk til min fantastiske samboer Kjell; Ikke bare er du en superpappa som gjør det mulig å fullføre en mastergrad midt oppi alt annet som til enhver tid skjer i livet vårt, men du gjør i tillegg livet mitt til et utrolig bra sted å være <3

Johanne Marie Trovåg

Haugesund, 31. mai 2017

Innhold

Sammendrag.....	-III-
Abstract.....	-IV-
Forord.....	-V-
Innhold.....	- 1 -
1. Innledning.....	- 4 -
1.1 Bakgrunn for valg av tema.....	- 5 -
1.2 Problemstilling og forskningsspørsmål	- 6 -
1.3 Avgrensninger i oppgaven	- 6 -
1.4 Tidligere forskning	- 7 -
1.4.1 Digital eksamen.....	- 8 -
1.4.2 Effektivisering av evaluering	- 9 -
1.4.3 Evaluering i en digital tidsalder	- 10 -
1.4.4 Pedagogiske perspektiver på digital evaluering	- 11 -
1.4.5 Digital evaluering i maritim utdanning/næring.....	- 13 -
1.4.6 Oppsummering	- 14 -
2. Teoretiske perspektiver	- 15 -
2.1 Sosiokulturell læringsteori	- 15 -
2.1.1 Praksisfellesskap	- 16 -
2.1.2 Medierende artefakter	- 16 -
2.2 Vurdering	- 20 -
2.2.1 Summativ vurdering.....	- 20 -
2.2.2 Sammenheng mellom læringsutbytte, arbeids- og vurderingsformer	- 23 -
2.2.3 Backwash	- 25 -
2.2.4 Reliabilitet og validitet i vurdering	- 26 -
2.3 Vurdering i STCW	- 27 -
3. Metode.....	- 29 -
3.1 Kvalitativ metode.....	- 29 -

3.1.1 Educational Design Research.....	- 29 -
3.1.2 Fenomenologi.....	- 31 -
3.2 Intervju.....	- 31 -
3.2.1 Fokusgruppeintervju.....	- 32 -
3.2.2 Refleksjonsmøter.....	- 34 -
3.2.3 Utvalg.....	- 34 -
3.2.4 Gjennomføring av undersøkelsen.....	- 35 -
3.2.5 Transkribering.....	- 36 -
3.3 Kvalitativ dataanalyse.....	- 37 -
3.4 Metodiske refleksjoner.....	- 38 -
3.4.1 Reliabilitet og validitet.....	- 38 -
3.4.2 Etikk og personvern.....	- 39 -
4. Presentasjon av funn.....	- 41 -
4.1 Digital kontekst.....	- 41 -
4.1.1 Kjennskap til systemet.....	- 42 -
4.2 Vurdering.....	- 46 -
4.2.1 Flere valgmuligheter.....	- 46 -
4.3 Læring.....	- 54 -
4.3.1 Lært av pilotene.....	- 56 -
4.3.2 Behov for støtte i institusjonen.....	- 57 -
4.3.3 Eksamen påvirker læring.....	- 57 -
4.3.4 Videre utvikling.....	- 59 -
5. Drøfting.....	- 60 -
5.1 Digital kontekst.....	- 60 -
5.1.1 Digital eksamen som medierende artefakt.....	- 60 -
5.1.2 Appropriering.....	- 61 -
5.2 Vurdering.....	- 63 -
5.2.1 Meningsskapende helhet.....	- 63 -
5.2.2 Reliabilitet og validitet i vurderingen.....	- 66 -
5.3 Læring.....	- 69 -
5.3.1 Praksisfellesskap.....	- 69 -

5.3.2 Backwash	- 70 -
6. Avslutning	- 73 -
6.1 Oppsummering og konklusjon	- 73 -
6.2 Avsluttende tanker	- 74 -
6.2.1 Fremtidig forskning	- 74 -
Bibliografi.....	- 76 -
Vedlegg 1 – Tabell 1, Litteraturgjennomgang	- 82 -
Vedlegg 2 – Informasjons- og samtykkeskjema	- 84 -
Vedlegg 3 – Godkjenning fra NSD	- 86 -
Vedlegg 4 – Intervjuguide refleksjonsmøter	- 88 -
Vedlegg 5 – Intervjuguide individuelt intervju	- 90 -
Vedlegg 6 – Intervjuguide fokusgruppeintervju	- 92 -

1. Innledning

I tråd med utdanningspolitiske signaler, er digital eksamen på vei inn i norsk høyere utdanning. Stortingsmeldingen *Kultur for kvalitet i høyere utdanning*, også kalt Kvalitetsmeldingen, påpeker det at på tross av en teknologiutvikling som har ført til endrede arbeidsprosesser i samfunnet for øvrig, ikke har vært like store endringer i gjennomføring av eksamen. Det er i følge Kvalitetsmeldingen ennå et stykke igjen før alle eksamener er digitale, og det understrekes at en slik overgang ikke bare dreier seg om innlevering av papireksamen via nettet, men en reell utnyttelse av mulighetene som ligger i de digitale vurderingsformene (Kunnskapsdepartementet, 2017).

Myndighetene har lagt til rette for en overgang til digital eksamen gjennom blant annet *Forskrift om endring i forskrift om egenbetaling ved universiteter og høyskoler* §3-3, 3. ledd, som sier at «*Institusjonen kan kreve at studenten har egen bærbar datamaskin*» (Kunnskapsdepartementet, 2006).

I følge Biggs og Tang (2011) har høyere utdanning endret seg dramatisk siden år 2000, både når det gjelder andelen mennesker som søker seg til høyere utdanning og utdanningene i seg selv. Utdanningssystemene i Europa har gjennom Bologna-prosessen fått et fokus på læringsutbytte. I kjølvannet av denne prosessen ble Nasjonalt kvalifikasjonsrammeverk for livslang læring (NKR) fastsatt (Kunnskapsdepartementet, 2016b). I norsk utdanning er definerte læringsutbytter et nyere fenomen, selv om det lenge har vært fokus på innhold og undervisning (Prøitz, 2015). Gjennom innføringen av det nasjonale kvalifikasjonsrammeverket for livslang læring gjorde læringsutbytte sitt inntog i hele det norske utdanningssystemet. Innføringen av læringsutbytte har ført til at det i undervisning og opplæring ikke lenger er hovedfokus på undervisning, men på læring (Baartman L. , Bastiaens, Kirschner, & van der Vleuten, 2006; Prøitz, 2015).

God undervisning er lagt opp med aktiviteter som legger til rette for at studentene skal oppnå det aktuelle læringsutbyttet. I undervisningen bør derfor det som studentene skal lære, hvordan de lærer det, og hvordan de testes i det oppnådde læringsutbyttet ha fokus. (Biggs & Tang, 2011).

1.1 Bakgrunn for valg av tema

Som studieleder ved en norsk høgskole, er spørsmål rundt innføring av digital eksamen noe jeg jevnlig forholder meg til. Det er mange diskusjoner i utdanningssektoren og ved egen institusjon knyttet til digitalisering og digital eksamen. Slike diskusjoner kretser som illustrert i kapittel 1.4 Tidligere forskning rundt alt fra logistikk og praktiske problemstillinger knyttet til digital eksamen, hvordan digital eksamen skal gjennomføres, til hva digital eksamen faktisk er. Er det egentlig en digital eksamen når studentene skriver besvarelsen sin på en pc? Er det digital eksamen når studentene bruker pc til å levere inn besvarelsen sin i form av et dokument eller andre typer filer? Er det digital eksamen når studentene besvarer eksamen ved hjelp av pc før besvarelsene skrives ut og sendes i posten til sensor? Eller er det digital eksamen når man benytter en bestemt type programvare til å lage oppgaver som man ikke like enkelt kunne skrevet på et ark?

For meg er digital eksamen noe mer enn å sette strøm på papir, og en av diskusjonene jeg til dels har savnet i relasjon til digital eksamen, er diskusjonen rundt hvordan vi kan utnytte nye digitale verktøy til å gjøre gode vurderinger. Dette krever mer av oss, og vi må derfor tenke nytt for å utnytte mulighetene i det digitale verktøyet til å utvikle undervisnings- og vurderingsformene våre.

Læringsutbytte på studieprogram- og emnenivå er også en viktig del av hverdagen i min rolle som faglig ansvarlig for nautikkstudiet, og jeg har utviklet et aktivt forhold det nasjonale kvalifikasjonsrammeverket, NKR (Kunnskapsdepartementet, 2014). I den sammenheng synes jeg at problemstillinger knyttet til hvordan vi kan oppnå god sammenheng mellom læringsutbytte, arbeids- og vurderingsformer er svært interessant.

I perioden 2013-2015 arbeidet med å utvikle ny studieplan på bachelorutdanningen i nautikk, og i den sammenheng fikk vi mulighet til å kombinere de to temaene. Dette førte til at noen av foreleserne ønsket å innføre digital eksamen i sine emner. Som et resultat av dette har vi i forbindelse med innføring av ny studieplan for kull 2016 også innført digital eksamen i ett av emnene på første studieår. Den prosessen har jeg valgt å gjøre til tema for min masteroppgave, og jeg har på denne måten fått mulighet til å grave dypere i et svært interessant felt; digital eksamen i et pedagogisk perspektiv.

1.2 Problemstilling og forskningsspørsmål

Formålet med studien er å undersøke hvordan foreleserne jobber for å gjøre den summative vurderingen i emnet digital. I tillegg ser jeg på hvordan studentene opplever at foreleserne lykkes med å etablere en sammenheng mellom den digitale eksamenen og det som emnet reelt har handlet om.

Problemstillingen er som følger:

Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkstudiet tilpasses en digital kontekst?

For å kunne finne svar på problemstillingen, har jeg formulert to forskningsspørsmål:

- *Hvordan begrunner faglærerne utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i programmet Wiseflow¹?*
- *Hvordan beskriver studentene sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen?*

Problemstilling og forskningsspørsmål utforskes kvalitativt gjennom den overordnede metodikken Educational Design Research, med refleksjonssamtaler og fokusgruppeintervju som datainnsamlingsmetode. Dette beskrives nærmere i kapittel 3 Metode.

1.3 Avgrensninger i oppgaven

Digital eksamen er et vidt område, som i tråd med funnene fra kapittel 1.4 Tidligere forskning har mange ulike definisjoner. Det er også mange ulike innfallsvinkler til digital eksamen, slik kapittel 1.4.1 Digital eksamen viser. Jeg har derfor måttet ta mange valg underveis i arbeidet knyttet til avgrensninger i temaet, siden et masterarbeid er av et begrenset omfang. Dette er årsaken til at jeg når jeg undersøker et emne med to ulike former for vurdering, samt formative og summative elementer, kun har valgt å fokusere på den ene summative vurderingen. På grunn av at jeg har ønsket et helhetlig bilde, vil

¹ Wiseflow – programvare for gjennomføring av digital eksamen. Eies og driftes av selskapet UNIwise.

allikevel den praktiske eksamenen, samt de formative elementene dukke opp i ulike sammenhenger.

Debatten om formative versus summative vurderingsformer har gått over lang tid, og det er mange synspunkter på hva som er den beste formen for vurdering. Dette er et spørsmål som er interessant nok i seg selv, men som ikke belyses i denne oppgaven. Diskusjoner rundt vurderingsform for emnet med tilhørende vurderinger ble gjort i forbindelse med utviklingen av ny studieplan på nautikk. Denne oppgaven forholder seg kun til den vurderingsformen som er valgt i emnet, og som parallelt med denne studien blir innført.

Enhver masterstudie har begrensninger, og i dette tilfellet er en av disse begrensningene at studien er gjennomført på kun ett emne og en enkelt gjennomføring av dette emnet. Det fører til at det er et begrenset utvalg informanter som har bidratt med sine synspunkter knyttet til problemstilling og forskningsspørsmål, og det kan derfor knyttes usikkerhet til hvorvidt deres synspunkter er representative. Av denne grunn er det derfor ikke mulig å generalisere studien på en slik måte at den kan overføres til en hvilken som helst utdanning på universitets- og høyskolenivå. Imidlertid kan det som Postholm (2010) kaller naturalistisk generalisering bidra til at leseren kan trekke paralleller til egen praksis og dra nytte av denne innsikten. På denne måten kan andre som ønsker å starte opp med digital eksamen få nyttig innsikt i en slik prosess.

Det være både en styrke og en svakhet for studien, at jeg som forsker i det daglige har tette koblinger til de jeg forsker på, noe som er nærmere beskrevet i kapittel 3.4 Metodiske refleksjoner. I tillegg vil det faktum at det er første gangs gjennomføring av et emne med nytt innhold og ny vurderingsform påvirke resultatene. Det er ikke sikkert at jeg hadde fått identiske resultater hvis foreleserne hadde gjort dette noen ganger tidligere. På den annen side hadde jeg dermed også mistet innføringsperspektivet og en sentral del av poenget med oppgaven.

1.4 Tidligere forskning

Tidligere forskning ligger til grunn for ethvert forskningsprosjekt, og nye prosjekter forankres gjennom allerede eksisterende forskning på området. Litteraturgjennomgangen er en viktig del av en avhandling og legger grunnlaget for både forskningsdesign og metodisk håndverk (Boote & Beile, 2005; Cohen, Manion, & Morrison, 2011; Krumsvik, 2014). Det er derfor svært interessant å se hva som allerede finnes av forskning relatert

til digital (skole)eksamen, slik at jeg gjennom eksisterende litteratur på området kan forankre mitt prosjekt og på denne måten bygge videre på det andre forskere har gjort før meg. Som American Educational Research Association (1967) sier det; «Intellectual progress is possible only because newcomers can stand on the shoulders of giants» (s. 39).

I arbeidet med litteraturgjennomgangen fulgte jeg Creswells (2014) anbefaling om en femtrinns prosess, hvor de enkelte trinn er definert som følger; 1) Identifisere begreper til litteratursøkene, 2) søke og finne litteratur, 3) lese og sjekke relevans av litteraturen, 4) organisere litteraturen og 5) skrive en litteraturreview (s. 97, min oversettelse). Jeg har tatt utgangspunkt i hvilke temaer som er sentrale i eksisterende litteratur om digital eksamen, hvilke konklusjoner som trekkes, og hva litteraturen peker på som temaer som bør undersøkes, og funnet søkeord ut fra disse. En skjematisk oversikt over søkeord og funn finnes i tabell 1 i vedlegg 1.

Digital eksamen er et relativt nytt fenomen i høyere utdanning, og jeg har derfor tidvis fått et begrenset antall treff på enkelte relevante søkeord. Det har likevel vært et tema lenge nok til at det finnes en del litteratur å ta utgangspunkt i, selv om en klar tendens i litteraturen jeg har funnet på området har et teknologisk fokus sammenlignet med målet for denne studien.

1.4.1 Digital eksamen

Mye av litteraturen jeg har funnet ved søk på *digital eksamen*, *digital exams*, *computer assisted exam* og tilsvarende søkeord har et teknisk preg eller en teknisk orientering (Behrens, Mislevy, DiCerbo, & Levy, 2010; Berggren, Fili, & Nordberg, 2015; Keshavarz, 2011; Pellegrino & Quellmalz, 2010; Sindre & Vegendla, 2015). En del av artiklene som omhandler digital eksamen har også en praktisk tilnærming, hvor rammene rundt den aktuelle eksamenen er fokus for undersøkelsene (Berggren, Fili, & Nordberg, 2015; Hassanien, Al-Hayani, Abu-Kamer, & Almazrooa, 2013; Nielsen & Petersen, 2013). I tillegg er både politikk, autensitet og effektivitet, evalueringssystemer samt muligheter for automatisering av det som i dag er manuelle prosesser omtalt i flere artikler (Bertheussen, 2014; Llamas-Nistal, Fernández-Iglesias, González-Tato, & Mikic-Fonte, 2013; Pellegrino & Quellmalz, 2010; Scalise & Wilson, 2011; Thelwall, 2000; Zoanetti, 2010).

Mange artikler som omhandler digital eksamen tar for seg formativ vurdering (Dochy F., 2001; Dochy & McDowell, 1997; Nix & Wyllie, 2011; Pellegrino & Quellmalz, 2010) eller digital eksamen av typen hjemmeeksamen, mappeeksamen eller lignende (Newhouse, 2010). En rekke artikler omhandler også studentenes holdninger til digital eksamen (Boevé, Meijer, Albers, Beetsma, & Bosker, 2015; Deutsch, Herrmann, Frese, & Sandholzer, 2012; Nielsen & Petersen, 2013) eller selvvurdering/ hverandrevurdering (Cowan, 2010). Dette er ikke direkte relevant for denne oppgavens problemstilling. Artikler som beskriver digitale skoleeksamener omtaler i flere tilfeller flervalgsoppgaver (multiple choice), og utarbeidelse av denne typen oppgaver (Boevé, Meijer, Albers, Beetsma, & Bosker, 2015; Gekara, Bloor, & Sampson, 2011; Thelwall, 2000). Litteratur med fokus på flervalgsoppgaver er ikke så relevant i denne sammenhengen, da multiple choice-eksamener vil være for ensidig til å dekke behovet vi ser for å teste læringsutbyttet på nautikk. Jeg har i mine litteratursøk funnet overraskende lite materiale om digital skoleeksamen relatert til læringsutbytte og hvordan man lager gode spørsmål.

1.4.2 Effektivisering av evaluering

Karaktersetting og automatisering av karaktersettingsprosessen er tema i flere artikler, og Christopher Newhouse er en av forfatterne som har undersøkt dette. Han skriver i artikkelen *Aligning assessment with curriculum and pedagogy in applied information technology* (2010) om digitale mapper og digital skoleeksamen med fokus på karaktersetting og karaktersettingens pålitelighet i disse eksamensformene. Artikkelen *Digital examination in higher education – Experiences from three different perspectives* (Berggren, Fili, & Nordberg, 2015) tar for seg erfaringer med innføring av digital eksamen ved Royal Institute of Technology i Stockholm med fokus på tidsbesparelser og administrative elementer i et karaktersettingsperspektiv. Gjennom artikkelen *Blended e-assessment: Migrating classical exams to the digital world* (Llamas-Nistal, Fernández-Iglesias, González-Tato, & Mikic-Fonte, 2013), beskriver forfatterne et verktøy som de har utviklet, og som skal bidra til digital retting av eksamensoppgaver skrevet på papir. M. Thelwall har i artikkelen *Computer-based assessment: A versatile educational tool* (2000) blant annet beskrevet fordeler ved digital eksamen som for eksempel automatisk karaktersetting som kan føre til tidsbesparelser for foreleser og muligheten for umiddelbar tilbakemelding for studentene. Slike funksjoner øker studentenes muligheter til å gjennomføre selv-tester og til å bruke denne typen tester som øvingsoppgaver i forkant av eksamen (Thelwall, 2000).

Artikkelen *E-exams and exam process improvement* (Sindre & Vegendla, 2015) er også teknisk orientert, og fokuserer på hvordan digitale eksamener kan forbedre gjennomføringen av eksamen, både i relasjon til prosess, eksamensoppgaver og hvilke systemer som bør være på plass for at de antatte forbedringene skal kunne detekteres. Artikkelen har flere forskningsspørsmål, hvorav ett av disse kan virke svært relevant for denne studien; «What didactic improvements in exam questions can be achieved with e-exams?» (s. 2). Spørsmålet besvares i artikkelen svært kursspesifikt for et programmeringsemne, og det er vanskelig å hente ut overføringsverdi til emner hvor læringsutbyttet ikke inkluderer praktisk bruk av dataverktøy og evnen til å programmere. Læringsutbytte står sentralt i artikkelen *Measuring course learning outcomes*, hvor Keshavarz (2011) tar for seg en metode for å kvantifisere læringsutbyttet i et emne, slik at dette kan tilpasses et retteverktøy. Selve retteverktøyet er ikke relevant for min forskning, men artikkelen går også inn på hvordan man tester læringsutbytte og hvordan læringsutbyttet påvirker oppbygningen av et emne, samt sammenhengen mellom læringsutbytte og vurderingskriterier (Keshavarz, 2011). Keshavarz' (2011) forskning er forankret i Biggs og Tang (2011) sine idéer om synkronisering av læringsutbytte, undervisnings- og læringsaktiviteter, samt vurdering.

1.4.3 Evaluering i en digital tidsalder

Selv om en stor andel av litteraturen funnet i mine litteratursøk bygger på tanker og idéer Benjamin Bloom m.fl. forfektet på 1950-tallet (Bloom, Engelhart, Furst, Hill, & Kratwohl, 1956; Bloom, Hastings, & Madaus, 1971), er prinsippene for vurdering fortsatt i stor grad aktuelle. Artikkelen *Rethinking assessment in a digital age: opportunities, challenges and risks* (Timmis, Broadfoot, Sutherland, & Oldfield, 2016) er en grundig litteraturstudie som viser at det ennå er uklart hvor mye digital teknologi vil endre utdanningspraksis og vurdering. De fremhever at litteratur om digital eksamen ofte har fokus på prosedyrer og et teknisk perspektiv, noe som også stemmer med de funnene jeg har gjort i mine litteratursøk. Forfatterne av artikkelen peker på behovet for reevaluering av både formål og prosedyrer for vurdering, og poengterer at det pedagogiske som oftest ikke er fokus i forbindelse med digital eksamen. De hevder at teknologiens rolle i vurdering i litteraturen ofte er definert som multiple choice-oppgaver og automatisk karaktersetning, og mener at fremtidig forskning bør inneholde felles prosjekter mellom forskere og forelesere for å kunne finne opp (rethink) evaluering/vurdering på nytt (Timmis, Broadfoot, Sutherland, & Oldfield, 2016).

Politikk står sentralt i innføringen av et nytt evalueringsregime, og Redecker og Johannessen (2013) skriver i artikkelen *Changing Assessment – Towards a New Assessment Paradigm Using ICT* om vurderingsparadigmer og hvordan politikken må endres for å tilpasses det 21. århundrets vurdering. De fremhever at forskning og utvikling på området blant annet må utvikle nye konsepter for innebygd, autentisk og holistisk vurdering, samt pedagogikk for e-vurdering. Også Bernt Arne Bertheussen fra UiT er opptatt av politiske føringer i forbindelse med digital eksamen, og har i artikkelen *Digital School Examinations: An Educational Note of an Innovative Practice* (2014) undersøkt motsetningene mellom politikk og praksis, og tar for seg effektiviteten av digitale skoleeksamener. I denne artikkelen er noe av fokuset på hvordan det i økonomifag kan utarbeides gode oppgaver, og forfatteren peker på at det bør forskes mer på dette i andre studieretninger for å eliminere kontekstuelle effekter.

1.4.4 Pedagogiske perspektiver på digital evaluering

Flere av artiklene peker på nødvendigheten av studier som er opptatt av de pedagogiske perspektivene på digital eksamen (Bertheussen, 2014; Boevé, Meijer, Albers, Beetsma, & Bosker, 2015; Conole & Warburton, 2005; Fegran & Slettebø, 2012; Nastiti & Campbell, 2010; Redecker & Johannessen, 2013; Timmis, Broadfoot, Sutherland, & Oldfield, 2016), noe som synliggjør behovet for denne studien.

I artikkelen *The nature of Assessment Systems to Support Effective Use of Evidence through Technology* tar Scalise og Wilson (2011) for seg den nasjonale planen *The National Educational Technology Plan 2010* (NETP) og kobler denne mot vurdering og ferdigheter fra det 21. århundret. Artikkelen omhandler i all hovedsak et tilbakemeldings- og rapporteringssystem, men definerer blant annet behovet for mer forskning på nye innfallsvinkler relatert til teknologimediert innhold som kan bidra i evaluering. Forfatterne skriver også at den nasjonale planen viser til at teknologiens omfang ennå ikke har blitt fullt utnyttet til å designe, utvikle og validere nytt evalueringsmateriale, og hvor det kan gjøres fremskritt i relasjon til evaluering. De viser også til at det er viktig med god forståelse av evaluering og evalueringsdesign, og hvorvidt det som evalueres er passende for evalueringens design, karaktersetning og rapportering (Scalise & Wilson, 2011).

Cristopher Newhouse har sammen med en forskergruppe fra Edith Cowan University i Australia gjennomført et større forskningsprosjekt om digital eksamen. Prosjektet har gått

over tre år og involvert fire ulike fag på en videregående skole (Newhouse, 2010; 2012; 2013; 2015). I konferansepaperet *Digital forms of assessment: alignment with pedagogic and curriculum intentions* identifiserer han blant annet et kritisk behov for forskning på bruk av digitale former for vurdering på komplekse oppgaver som er gjennomførbare i skolen (Newhouse, 2012). Newhouse (2012) fremhever også den pedagogiske dimensjonen ved digital evaluering og hvor viktig det er å se på hvordan bruk av digital vurdering kan støtte og berike studentenes læring. Han peker på at digital teknologi kan muliggjøre alternative former for vurdering, slik at man faktisk er i stand til å teste de kunnskaper, ferdigheter og den kompetansen man ønsker og som samsvarer med det aktuelle emnets innhold. Gjennom konferansepaperet *Computer-based exams in schools: Freedom from the limitation of paper?* har Newhouse (2013) fokus på digital eksamen i de tidligere nevnte fagene. Selv om artikkelen ikke omhandler digital skoleeksamen tar den allikevel for seg perspektiver ved digital eksamen som er relevante også for digital skoleeksamen. Newhouse (2013) fokuserer blant annet på hvor viktig det er at eksamensoppgavene og eksamensformen er tilpasset emnets innhold, og viser til metodologiske utfordringer relatert til det å designe passende oppgaver – en utfordring som digital eksamen i svært stor grad har til felles med tradisjonell skriftlig eksamen. Newhouse konkluderer med at det ikke lenger er nødvendig at den tradisjonelle skriftlige eksamenen begrenser vurderingen, og at det bør fokuseres på å sette sammen en best mulig vurdering for å kunne evaluere hvorvidt studentene har oppnådd det forventede læringsutbyttet (Newhouse, 2013).

I artikkelen *Using digital technologies to improve the authenticity of performance assessment for high-stakes purposes* definerer Newhouse (2015) begrepet “construct validity”, et begrep han selv bruker om hvor godt vurderingen i seg selv stemmer med innholdet i det som vurderes. Han viser til Pellegrino og Quellmalz, som sier at hvordan noe vurderes og designet på oppgavene studentene vurderes ut fra er kritiske for vurderingens funksjonelle kvalitet (Pellegrino & Quellmalz, 2010). Målet med studien var i følge Newhouse (2015) å utforske bruken av digitale vurderingsformer som erstatning for penn og papir eller andre ikke-digitale vurderingsformer for å øke autensiteten ved vurderingen, i tillegg til en forbedret kostnadseffektivitet. Han beskriver blant annet at den pedagogiske dimensjonen sier noe om hvilke evalueringsoppgaver som er tilpasset pedagogikken i emnet, og dermed emnets innhold. Newhouse (2010; 2012; 2013; 2015) går ikke i noen av de tidligere nevnte artiklene inn

på de spesifikke vurderingene som ligger til grunn i utviklingen av oppgaver. Han nevner heller ikke i noen særlig grad hvordan eller hvorfor de ulike oppgavene er valgt ut i de ulike emnene, men han er tydelig på at utformingen av oppgaver er viktig, og at videre forskning må sikre at det å ta i bruk ny teknologi fører til at utbyttet blir bedre. Newhouse (2015) peker også på at det er viktig at studentene er trygge på bruken av både pc og software før en viktig eksamen gjennomføres, og at ny teknologi må føre til mer autentiske evalueringer for de ulike emnene.

1.4.5 Digital evaluering i maritim utdanning/næring

Den eneste bransjerelaterte artikkelen jeg fant i litteratursøkene mine som også omhandlet digital eksamen var *Computer-based assessment in safety-critical industries: the case of shipping* av Gekara, Bloor og Sampson (2011). Artikkelen tar i hovedsak for seg evaluering ved hjelp av mutiple choice-oppgaver, og til dels også simulortester, men inkluderes på grunn av at den er bransjerelevant. Målet med studien var å undersøke hvordan pc-baserte vurderingssystemer ble benyttet til sertifikatprøver for sjøfolk og omfattet både litteraturstudier og datainnsamling på tvers av landegrenser. De beskriver hvordan innføringen av pc-basert evaluering i forbindelse med sertifikatprøver har kommet som et resultat av hovedsaklig tre faktorer; økende arbeidsbyrde for sensorer, behov for objektivitet og konsekvens, og det økende behovet for offiserer (Gekara, Bloor, & Sampson, 2011). De viser også til at funnene deres peker i retning av at maritime myndigheter i all hovedsak er motivert av kostnader og praktiske fordeler i motsetning til pedagogikk og problemstillinger relatert til validitet og reliabilitet, noe som forfatterne anser vel så viktig (Gekara, Bloor, & Sampson, 2011). Forfatterne understreker at construct validity, content validity og criterion validity er viktig, hvor construct validity omhandler graden av samsvar mellom vurderingsinstrumentet og målet med evalueringen, content validity omhandler hvordan vurderingen relateres til kjernen av trainingen (læringsutbyttet) og criterion validity til korrelasjonen mellom testresultater sett i lys av de forhåndsdefinerte kriteriene. Talsmenn for autentisk vurdering argumenterer for at vurderingen må være robust og den må simulere virkelighetens arbeidsliv hvor kandidatene skal benytte sine kunnskaper, ferdigheter og kompetanse. De konkluderer med at selv om pc-basert vurdering har mange potensielle fordeler, må bruk og design vurderes nøye før bruk (Gekara, Bloor, & Sampson, 2011).

1.4.6 Oppsummering

Arbeidet med en litteraturgjennomgang bidrar til at man får oversikt over forskningsfeltet, slik at man kan identifisere kunnskapshull og posisjonere en studie (Krumsvik, 2014). Gjennom dette arbeidet har jeg opparbeidet meg en oversikt over forskningsfeltet knyttet til digital eksamen, og det fremstår tydelig for meg at det er behov for kvalitative studier med fokus på de pedagogiske tankene bak utarbeidelse av eksamensoppgaver i et digitalt format, og da spesielt knyttet opp mot læringsutbytte. I og med at jeg ikke har funnet litteratur om digital eksamen relatert til maritim utdanning, vil denne studien dermed også bidra til å fylle et kunnskapshull knyttet til maritim utdanning.

2. Teoretiske perspektiver

«Læring kan sies å være et naturlig og nødvendig aspekt av menneskelig virksomhet» (Säljö, 2001, s. 48). Enhver situasjon som omhandler opplæring og som skal resultere i læring for noen, vil være preget av et læringssyn. Så også en problemstilling knyttet til vurdering av læring, nærmere bestemt hvordan en tradisjonell skriftlig skoleeksamen i nautikkutdanningen kan tilpasses en digital kontekst. I følge Lave og Wenger er «*All theories of learning...based on fundamental assumptions about the person, the world, and their relations...*» (1991, s. 47). Også Vygotsky (1978) fremhever sammenhengen mellom språk, deltagelse i sosial praksis og læring. I tillegg peker Vygotsky med flere på redskapene som benyttes i samhandlingen med omverdenen, såkalte medierende artefakter (Dysthe, 2001; Engeström, 2016; Säljö, 2001; Säljö, 2006; Vygotsky, 1978). Alt dette er sentralt i teorigrunnet som arbeidet med denne oppgaven er fundamentert i, og utdypes nærmere under.

2.1 Sosiokulturell læringsteori

Et sosiokulturelt læringssyn tar utgangspunkt i at kunnskap konstrueres gjennom samhandling i ulike former, heller enn gjennom individuelle prosesser (Dysthe, 1999; Dysthe, 2001; Engeström, 2016; Imsen, 2005; Lave & Wenger, 1991; Säljö, 2001; Vygotsky, 1978).

Sentrale teoretikere som Dewey (1859-1952), Mead (1868-1931), Vygotsky (1886-1934) og Bakhtin (1895-1975) har alle bidratt til utviklingen av det sosiokulturelle perspektivet, ut fra litt ulike tradisjoner. Stikkord som er felles for dem alle er interaksjon og samhandling, og et felles standpunkt er at kunnskap alltid er situert – det vil si at kunnskapen er innvevd i den kulturelle og historiske konteksten (Dysthe, 1999). Hovedfokus blir derfor samspillet mellom kollektiv og individ (Säljö, 2001). Ut fra et slikt perspektiv vil individet i følge Dysthe (1999) ikke kunne forstås uavhengig av en kontekst, og det er derfor interaksjonsprosessene i relasjon til dette prosjektet omfatter samspillet mellom studentene og de digitale hjelpemidlene de bruker – «dei kulturelle reiskapane» (s. 2). Et annet viktig særtrekk ved den sosiokulturelle læringsteorien er at læringen anses å være mediert – enten av språklige eller fysiske verktøy, såkalte

medierende artefakter (Dysthe, 2001; Engeström, 2016; Imsen, 2005; Säljö, 2001; Säljö, 2006; Vygotsky, 1978; Wertsch, 1998).

2.1.1 Praksisfellesskap

Praksisfellesskap står i følge Lave og Wenger (1991) sentralt når situert kunnskap konstrueres gjennom samhandling. Gjennom sine antropologiske studier av læring la Lave og Wenger grunnlaget for en annen forståelse av læring, enn den tradisjonelle psykologiske tradisjonen la til grunn (Dysthe, 2001). Læring er i henhold til Lave og Wengers (1991) og Vygotsky (1978) et grunnleggende sosialt fenomen. På denne måten lærer vi primært gjennom praksisfellesskap.

Engeström (2016) påpeker at praksisfellesskap i organisasjoner ofte er et så velkjent fenomen at disse ikke spesifikt er uttrykt – praksisfellesskapene tas for gitt. Hvor aktivt man deltar og hvilken rolle man spiller i ulike praksisfellesskap varierer, og man kan gå fra å være «mester» i et praksisfellesskap til å være «lærling» i et annet, avhengig kunnskapsnivå og erfaring relatert til det aktuelle praksisfellesskapets aktiviteter. Det som driver et praksisfellesskap frem og som holder det sammen er i all hovedsak vedvarende og gjensidig engasjement, eierskap til praksisfellesskapets mening og kollektiv læring (Engeström, 2016).

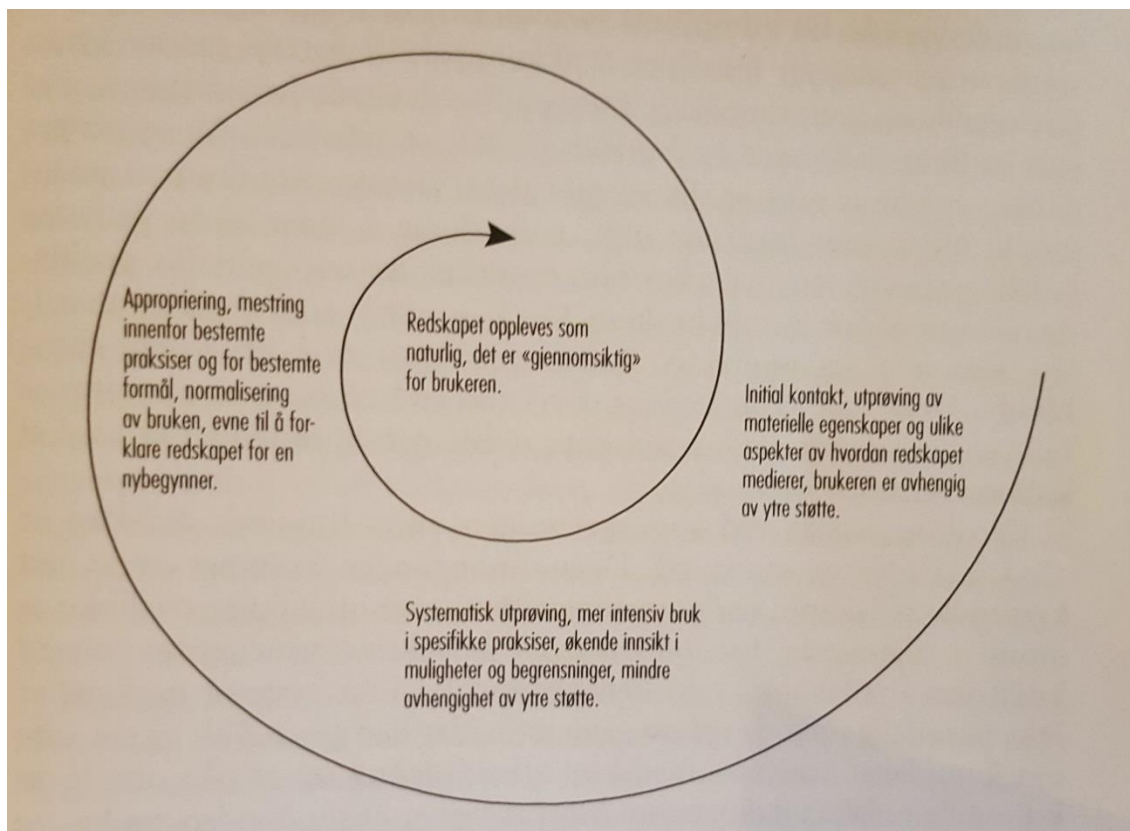
I følge Säljö (2001) utvikles en kollektiv hukommelse gjennom praksisfellesskap, og gjennom slike lærende fellesskap gjenspeiles og forsterkes sosiale holdninger. Lave og Wenger (1991) gir uttrykk for at hierarkiets utforming og mesterrollen vil variere mellom ulike praksisfellesskap, og et praksisfellesskap kan derfor ikke defineres av en bestemt relasjon mellom «mester» og «lærling». Også den enkeltes deltagelse i et praksisfellesskap varierer og individet bidrar på ulike nivå gjennom samspillet med andre «lærlinger» eller «mestere» (Lave & Wenger, 1991). Læringen skjer på ulike måter gjennom deltagelse i et spesifikt praksisfellesskap og hvordan man lærer i det enkelte praksisfellesskap er avhengig av samhandlingen mellom deltagerne i det aktuelle praksisfellesskapet. Læringen situert (Lave & Wenger, 1991).

2.1.2 Medierende artefakter

Medierende artefakter beskrives av Säljö (2001) som redskaper eller verktøy som er «*de ressursene, så vel språklige (eller intellektuelle) som fysiske, som vi har tilgang til, og som vi bruker når vi forstår vår omverden og handler i den*» (s. 21). En artefakt er et

menneskeskapt produkt eller gjenstand, som er utviklet med det formål å bidra i menneskelige handlinger av både fysisk og mer uhandgripelig karakter. Eksempel på mer abstrakte menneskelige handlinger kan være informasjonsbearbeiding (Säljö, 2001). Säljö (2001; 2006) hevder at artefaktene er en materiell del av kulturen vår som er innvevd i våre ideer, holdninger og kunnskaper, og er på denne måten en selvfølgelig del av vår evne til å samle erfaringer og bruke disse hensiktsmessig. I ulike menneskelige virksomheter kan samme artefakt og samme fenomen forstås og benyttes på ulike måter. Det som er avgjørende for tenking og handling er ikke lenger biologisk utvikling, men vår evne til å utnytte og skape de nødvendige intellektuelle og fysiske redskaper (Säljö, 2001; 2006).

Historisk sett har mennesket tatt i bruk flere og flere artefakter. På denne måten har mange ulike funksjoner og kompetanser blitt overført til fysiske redskaper, og bruk av intellektuelle og manuelle artefakter er uløselig knyttet til hverandre. Menneskets bruk av redskaper skiller oss fra andre arter, og våre handlinger er et resultat av kombinasjonen av intellektuell og manuell aktivitet (Dysthe, 2001; Säljö, 2001).



Figur 1 – Appropriering gjennom økende koordinering mellom redskap og bruker (Säljö, 2006, s. 215) – Brukt med tillatelse fra Cappelen Akademisk Forlag.

«Redskaper medierer læring på mange ulike vis» (Dysthe, 2001, s. 46). Et grunnleggende prinsipp for læring sett i et sosiokulturelt perspektiv er derfor at mennesket evner å se nye ting ut fra det som allerede er kjent, for eksempel kan et fysisk redskap tilpasses en ny oppgave som ligner på noe man allerede kjenner. Denne evnen utvikles ved at vi stadig behersker flere intellektuelle redskaper (Säljö, 2001). Dette er en kontinuerlig prosess som resulterer i at mennesket utvikler «redskaper som vi kan ta i bruk for å løse fysiske og intellektuelle problemer» (Säljö, 2001, s. 75). Menneskets utvikling bærer preg av stadig økende abstraksjon, og vi håndterer fysiske og intellektuelle utfordringer gjennom å utvikle nye fysiske redskaper (Säljö, 2001). For å kunne lære og utvikle seg må menneskene utnytte informasjon, prosedyrer, begreper og innsikter som ligger innebygd i artefaktene som allerede er utviklet, de må *appropriere* sosiokulturelle kunnskaper, ferdigheter og erfaringer slik figur 1 illustrerer (Säljö, 2001; 2006).

Figur 1 over viser hvordan Säljö (2006) ser for seg en økende fortrolighet mellom individet og de kulturelle redskapene. Spiralen starter med at mennesket kommer i kontakt med et helt nytt redskap gjennom ulike aktiviteter eller gjennom andre mennesker, f.eks. en datamaskin og programvaren for digital eksamen. I denne fasen oppfattes redskapene som ressurser for bestemte typer aktiviteter, for eksempel digital skoleeksamen. I den neste fasen tester man redskapet og begynner å bruke dette, for eksempel brukes datamaskinen og programvaren nå til digital eksamen. I den andre fasen utforskes ulike sammenhenger hvor redskapet kan brukes og hvordan det brukes i den aktuelle sammenhengen, i tillegg til redskapets begrensninger. Den tredje fasen er fasen hvor mennesket mestrer redskapet – redskapet er appropriert. I denne fasen oppfattes redskapet som en naturlig del av en bestemt handling/praksis, datamaskinen har blitt en naturlig del av planlegging av eksamensoppgavene for foreleserne og for eksamensgjennomføringen til studentene. Selv om man har kommet til denne fasen betyr det ikke at man kjenner til absolutt alle funksjoner ved en datamaskin, eller at man har utforsket alle aspekter ved digital eksamen, men man vet hvordan man skal gjennomføre digital eksamen. Etter hvert går man over i den siste fasen hvor datamaskinen og den digitale eksamenen oppfattes som den naturlige måten å gjennomføre skoleeksamen på, og brukerne er fullt ut koordinerte med datamaskinen og programvaren for digital eksamen. Grensen går ikke lenger ved det vi mennesker i oss selv er i stand til å håndtere, vi har lært oss å bruke nye redskaper for på denne måten å øke våre kunnskaper og ferdigheter (Säljö, 2006).

Säljö (2001; 2006) hevder at teknologiutviklingen bidrar til at læring fremover i større grad blir et spørsmål om å beherske den nye teknologien i aktuelle situasjoner, og at informasjonsteknologi endrer forutsetningene for både læring og undervisning. Ny teknologi endrer samspillet mellom mennesket, artefaktene (her forstått som det aktuelle teknologiske redskapet) og kollektivet, og påvirker på denne måten læringskulturen. Dette fører ikke til at vi slutter å lære, men at vi tvert i mot lærer nye ting og på nye måter. I tillegg kan enkelte kognitive prosesser integreres i artefakten (f.eks. regneregler i kalkulatoren, retterprogram i tekstbehandlingsprogram, etc.). Vi kan ved hjelp av arbeidsformer som problemløsning, veiledning, modellering og lignende få en aktiv kunnskapskonstruksjon (Dysthe, 2001; Säljö, 2001; Säljö, 2006).

2.1.2.1 Artefakter på ulike nivå

Wartofsky (1979) skiller mellom primære, sekundære og tertiære artefakter ut fra et synspunkt hvor artefaktene er en eksternalisering og objektivering av menneskehetens kunnskaper og innsikter.

Primære artefakter er konkrete redskaper som f.eks. datamaskinen, hammer, spade, etc. De primære artefaktene er en forlengelse av menneskekroppen som forenkler og/eller effektiviserer handlingene våre. I et sosikulturelt perspektiv må de primære artefaktene forstås situert i lys av aktiviteten de inngår i (Säljö, 2006; Wartofsky, 1979).

Sekundære artefakter er redskaper som bidrar med modeller for tenkning og handling, og som formidler våre historiske og kollektive begreper og innsikter. Eksempler på sekundære artefakter er for eksempel instruksjonsbøker og tegninger. De sekundære artefaktene gir veiledning om hvordan de primære artefaktene skal brukes, og styrer dermed handlingene våre og forståelsen vår (Säljö, 2006; Wartofsky, 1979).

Tertiære artefakter er en form for forlengelse av de sekundære artefaktene, med tilhørende praksis. De tertiære artefaktene har en løsere og mer hypotetisk kobling til den primære produksjonen, ved at de handler om hvordan man skal framstille, forstå og analysere. Felles for sekundære og tertiære artefakter er at de bygger på og utvikler representasjonssystemer og på denne måten bidrar til å strukturere det vi f.eks. ser (Säljö, 2006; Wartofsky, 1979).

I relasjon til denne oppgaven, vil det primære nivået være datamaskinen som eksamen gjennomføres på og som er det konkrete verktøyet som benyttes. På sekundært nivå er

spørsmålene som beskriver hva studentene skal svare på, og som samtidig fungerer som instruksjon for bruken av artefakten i den situerte konteksten. Det tertiære nivået kan tilsvare koblingen mellom emnets læringsutbytte og de spørsmålene som studentene må besvare under eksamen, det som foreleserne ønsker å teste gjennom eksamen.

2.2 Vurdering

Vurdering deles i all hovedsak inn i formativ og summativ vurdering, hvor formativ vurdering kan beskrives som vurdering *for* læring og summativ som vurdering *av* læring (Bloom, Hastings, & Madaus, 1971; Havnes & Raaheim, 2006; Raaheim, 2016). Skillet mellom summativ og formativ evaluering ble først fremhevet av Scriven (1967), og i henhold til dette er formativ vurdering en evaluering av en foreløpig versjon, mens summativ vurdering er det som er oppnådd til slutt. Det fremheves også at evaluering tjener ulike formål avhengig av sammenhengen den benyttes i (Scriven, 1974). Den digitale skriftlige skoleeksamen som ligger til grunn for arbeidet med denne oppgaven, er en avsluttende vurdering av læring i et emne. Det vil derfor i fortsettelsen fokuseres utelukkende på summativ vurdering, selv om formativ vurdering også er et veldig interessant tema.

2.2.1 Summativ vurdering

Summativ vurdering gjennomføres etter at undervisningen i et emne er avsluttet, med den målsetning å teste hvor godt studentene har lært det læringsutbyttet som er definert for det aktuelle emnet. Det er to viktige funksjoner som dekkes av en summativ vurdering; dokumentasjon og karaktersetting (Biggs & Tang, 2011). Dokumentasjonsbehovet er for nautikkutdanningens del forankret i regelverket i STCW (International Maritime Organization, 2011) og FOR-2011-12-22 nr. 1523 Forskrift om kvalifikasjoner og sertifikater for sjøfolk (Sjøfartsdirektoratet, 2012). Karaktersetting er forankret i skolesystem og utdanningspolitikk, og skal gi en indikasjon på hvor godt studentene har lært det de skal kunne etter endt emne og utdanning.

I litteratursøkene mine har jeg funnet lite nyere teori om summativ vurdering, siden fokus i senere år ser ut til i større grad å ha vært formativ vurdering. Summativ vurdering synes likevel å være nevnt i mange sammenhenger der f.eks. formativ vurdering, vurdering og

digitale verktøy, digitale læringsformer, og lignende tas opp (Fossland, 2015; Havnes & Raaheim, 2006; Raaheim, 2016).

På tross av at det har vært vanskelig å finne nyere teori om summativ vurdering, betyr ikke det at den teorien som etter hvert har blitt noen år ikke lenger er aktuell. Vi ser at for eksempel Bloom m.fl. (Bloom, Engelhart, Furst, Hill, & Kratwohl, 1956; Bloom, Hastings, & Madaus, 1971) tar opp temaer knyttet til vurdering som også er svært aktuelle i dagen vurderingsdebatt; sammenheng mellom det som undervises, hva studentene testes i og hvordan studentene testes. Selv om synet på detaljene rundt selve vurderingen og hensikten med utvikling av læringsobjektiver har endret seg en del fra midten av 1950-tallet til i dag, er grunntankene om at det skal være sammenheng mellom det studentene lærer, hvordan de evalueres, hva de vurderes ut fra og hva de evalueres i fortsatt høyst aktuelle (Bloom et.al. 1956; Bloom, Hastings, & Madaus, 1971). Tanken om en taksonomi hvor man begynner med det enkle for så å bygge opp mot mer avanserte spørsmål i et slags hierarkisk rammeverk, er heller ikke utdatert, selv om gjennomføringen i praksis nok har endret seg i årenes løp og i henhold til nyere teorier om læring. Det ser man ikke bare gjennom Taxonomy of Educational Objectives (Bloom et.al. 1956) fra 1950-tallet, men også gjennom SOLO-taksonomien (Biggs & Tang, 2011) som er utviklet i senere tid.

Historisk sett har vurdering vært sett i lys av to formål; vurdering av studenter opp mot visse kriterier gjennom Criterion Referenced Assessment, CRA, eller rangering av studentene opp mot hverandre gjennom såkalt Norm Referenced Assessment, NRA (Biggs & Tang, 2011; Bloom, Hastings, & Madaus, 1971). NRA har man etter hvert gått bort fra, og fokus ligger nå på å evaluere studentene opp mot en gitt standard i tråd med prinsippene i det som kan kalles utbyttebasert undervisning og læring eller lignende (Biggs & Tang, 2011). En form for utbyttebasert utdanning er for eksempel kompetansebasert utdanning som vi finner i regelverket i STCW (International Maritime Organization, 2011) og FOR-2011-12-22 nr. 1523 Forskrift om kvalifikasjoner og sertifikater for sjøfolk (Sjøfartsdirektoratet, 2012).

Særtrekk ved utbyttebasert utdanning er at det forventede læringsutbyttet er tydelig formulert, at undervisningen er lagt opp for at flest mulig av studentene skal oppnå læringsutbyttet og at studentene prøves i det aktuelle læringsutbyttet (Biggs & Tang,

2011). Dette er også hovedprinsippene i det Biggs og Tang (2011) kaller Constructive Alignment – et begrep som av Raaheim (2016) er oversatt til *meningsskapende helhet*.

I Norge har spørsmål om innhold og undervisning vært sentrale lenge, mens definerte læringsutbytter er et nyere fenomen (Prøitz, 2015). Læringsutbytte har gjennom innføringen av Nasjonalt kvalifikasjonsrammeverk for livslang læring (NKR) (Kunnskapsdepartementet, 2014) fått et sterkere fokus også i det norske utdanningssystemet. NKR er basert på det europeiske kvalifikasjonsrammeverket (EQF) og gjelder for hele det norske utdanningssystemet. Dette har ført til at alle utdanninger skal ha studieplaner som synliggjør læringsutbyttet det forventes at man innehar etter endt utdanning. Det er også krav om læringsutbyttebeskrivelser på emnenivå for å synliggjøre hvilke kunnskaper, ferdigheter og generell kompetanse studenten sitter igjen med etter at et emne er fullført og bestått (Kunnskapsdepartementet, 2014; Prøitz, 2015).

Selv om man i styrende dokumenter ikke har tatt inn læringsutbytte før i senere tid, har dette vært et fokusområde som har eksistert i pedagogikken over lang tid, og da spesielt i tilknytning til vurdering (Biggs & Tang, 2011; Bloom, Engelhart, Furst, Hill, & Kratwohl, 1956; Bloom, Hastings, & Madaus, 1971; Dobson, Eggen, & Smith, 2009; Fosslund, 2015; Havnes & Raaheim, 2006; Prøitz, 2015). Den største endringen etter innføringen av læringsutbytter, er at fokus i undervisning og opplæring nå har dreid fra undervisning til læring (Baartman L. , Bastiaens, Kirschner, & van der Vleuten, 2006; Prøitz, 2015). Med hovedfokus på læring er spørsmålene om hvordan studentene lærer, og hvordan vi best kan legge opp undervisningen sentrale for alle undervisere.

2.2.1.1 Evaluering i en digital tidsalder

I følge Timmis, et. al. (2016) er det ennå uklart hvor mye digital teknologi vil endre utdanningspraksis og vurdering. Forfatterne fremhever et fokus på prosedyrer og et teknisk perspektiv i litteratur som omhandler digital eksamen. Dette stemmer også med egne funn, noe som er nærmere beskrevet i kapittel 1.4 Tidligere forskning. Timmis et al. (2016) viser til at det er behov for en reevaluering av både formål og prosedyrer for vurdering i forbindelse med innføring av digital eksamen, og de poengterer at det i større grad bør fokuseres på de pedagogiske problemstillingene i stedet for teknologiske, praktiske og logistiske utfordringer knyttet til multiple choice-oppgaver, automatisk karaktersetting, og lignende. Dette samsvarer også med Raaheims (2016) tanker om digitalisering av eksamen. Han peker på at vi bør utnytte mulighetene teknologien gir til

å endre hvordan vi underviser og vurderer i stedet for å tilpasse de moderne verktøyene til tradisjonelle tankemønstre og kontekster. Ved å tenke nytt, kan vi i følge Raaheim (2016) oppnå fremskritt knyttet til studentenes læring, samt undervisning, organisering av fysiske omgivelser, etc.

2.2.2 Sammenheng mellom læringsutbytte, arbeids- og vurderingsformer

Siden 2000 har høyere utdanning i følge Biggs og Tang (2011) endret seg dramatisk, både når det gjelder andelen mennesker som søker seg til høyere utdanning og utdanningene i seg selv.

I Europa har utdanningssystemene fått et fokus på læringsutbytte gjennom Bologna-prosessen med tilhørende utarbeidelse av kvalitetsrammeverket for europeisk høyere utdanning (QF-EHEA), som senere også har blitt styrket med Det europeiske kvalifikasjonsrammeverket for livslang læring (EQF) (Biggs & Tang, 2011; European Commission, 2017). EQF ble vedtatt i EU i april 2008 og inkludert i EØS-avtalen i 2009 (European Commission, 2017; Kunnskapsdepartementet, 2016a). I kjølvannet av Bolognaprosessen ble Nasjonalt kvalifikasjonsrammeverk for livslang læring (NKR) utarbeidet for det norske utdanningssystemet. NKR ble fastsatt i 2011 og gjelder for alle deler av det norske utdanningssystemet (Kunnskapsdepartementet, 2016b).

I undervisningen bør det som studentene skal lære og hvordan de lærer dette ha fokus, i stedet for at fokus ligger på hva foreleserne skal forelese. God undervisning er lagt opp med aktiviteter som legger til rette for at studentene skal oppnå det aktuelle læringsutbyttet. Det samme gjelder for utarbeidelse av eksamensoppgaver. Det er viktig at det er sammenheng mellom hva foreleser ønsker å teste med den enkelte oppgaven på eksamen, siden man lett kan teste andre kunnskaper og ferdigheter, enn det som er planen, hvis ikke oppgaven er formulert eller lagt opp riktig. God vurdering får vi først når det er mulig å sjekke at det definerte læringsutbyttet er oppnådd (Biggs & Tang, 2011).

En måte å gjøre dette er ved hjelp av kriteriebasert vurdering (Criterion-referenced assessment, CRA), det vil si vurdering som forholder seg til en oppsatt standard, som oftest konkretisert ved hjelp av spesifikke vurderingskriterier. Dette er for øvrig det de aller fleste utenfor utdanningssystemet gjør når de driver opplæring (Biggs & Tang, 2011). CRA har eksistert siden Thorndike introduserte tanken om kriteriebasert vurdering på tidlig 1900-tall, hvor han la til grunn at studentenes resultater vurderes uavhengig av

hverandre (Biggs & Tang, 2011). Men som Dysthe uttrykker det «Det tar tid å utvikle alternative vurderingsformer som bygger på nye læringsprinsipp» (Dysthe, 1999, s. 3).

Kriteriebasert vurdering tar utgangspunkt i et definert læringsutbytte. Gjennom kriteriebasert vurdering blir studentenes prestasjoner vurdert kvalitativt med fokus på hvor godt de har lært det aktuelle læringsutbyttet. Dette står i motsetning til en mer kvantitativ vurdering av hvor mye studentene har lært, en metode som ikke lar seg forene med målsetningen om en meningssskapende helhet. Ved bruk av kriteriebasert vurdering er kriteriene gitt på forhånd, og det er tydelig for alle parter hvilken standard studentene vurderes opp mot (Biggs & Tang, 2011; Fosslund, 2015; Prøitz, 2015). Målsetningen bør i følge Biggs og Tang (2011) være autentisk vurdering, det vil si at studentene må kunne gjøre mer enn å ramse opp faktakunnskap i situasjoner hvor det ikke er deres kjennskap til fakta som testes. For å vurdere studentene helhetlig, har man behov for et begrepsrammeverk som synliggjør sammenhengene mellom den enkelte bestanddel og helheten, for eksempel SOLO-taksonomien.

SOLO står for Structure of the Observed Learning Outcome og er et begrepsrammeverk som bidrar til å beskrive hvordan studentens mestringsutvikling fra det helt enkle til det svært komplekse. Denne taksonomien kan benyttes både i relasjon til utvikling av læringsutbytte i emner eller et studieprogram, men det kan også benyttes i tilknytning til evaluering av nevnte læringsutbytter (Biggs & Tang, 2011). Denne taksonomien er utviklet på bakgrunn av en utbyttebasert utdanningsfilosofi, med en overraskende enkel logikk; «Say what you want students to be able to do, teach them to do it, try again until they can» (Biggs & Tang, 2011, s. 206). Prinsippet finner man også igjen knyttet til begrepet “mastery learning” hos Bloom et. al. (1971) og “the Keller Plan” (Keller (1968) i Biggs & Tang, 2011).

For å oppnå autentisk vurdering, må det være så stor likhet som mulig mellom læringssituasjonen og virkeligheten i for eksempel arbeidslivet man utdannes til. Dette må også vurderingssituasjonen bære preg av, og vurderingssituasjonen må stå i sammenheng med både det som læres i løpet av emnets/studiets gang og med påfølgende yrkespraksis (Biggs & Tang, 2011; Raaheim, 2016). I praksis vil det si at emner som har et praksiselement, slik som navigasjonsemnene på nautikkstudiet, også må ha en vurderingsform som gjenspeiler dette (Raaheim, 2016). Eksamensformen i det undersøkte emnet er som tidligere nevnt todelt, og består av en praktisk eksamen på

simulator og en digital skoleeksamen. På denne måten ivaretas den nødvendige sammenhengen i dette aktuelle emnet ved at også praksiselementet påvirker studentenes totale resultat i emnet.

2.2.3 Backwash

Backwash er et viktig begrep knyttet til vurdering, siden det er påvist at vurdering i stor grad påvirker hvordan studentene lærer og hvordan foreleserne underviser (Baartman L. K., Bastiaens, Kirschner, & van der Vleuten, 2007a; Biggs & Tang, 2011; Dobson, Eggen, & Smith, 2009; Fosslund, 2015; Havnes & Raaheim, 2006).

Lewis Elton introduserte på 1980-tallet begrepet backwash for å synliggjøre påvirkningen vurdering har på studentenes læring, og hvordan vurderingen påvirker læringen i større grad, enn det læreplanene gjør. Backwash har i mange sammenhenger hatt et litt negativt fokus, men begrepet kan i situasjoner der vurdering står i sterk sammenheng med det studentene skal lære også brukes positivt (Biggs & Tang, 2011).

Biggs (1996) og Havnes (2004) peker på at det er vurderingens utforming og innebygde krav som objektivt danner grunnlaget for backwash-effekten. I realiteten er det ikke dette objektive grunnlaget som avgjør hva studentene lærer, men studentenes egen oppfatning og tolkning av de kravene de tror at vurderingen stiller. Studentenes oppfatning av hva som kreves i en vurdering er ikke nødvendigvis i samsvar med de kravene foreleserne reelt sett stiller. Når studentene forbereder seg til en eksamen prøver de å forstå materialet på en slik måte at de tror kravene i den påfølgende vurderingen dekkes. Hvordan forståelse kan forstås er avhengig av hvordan studentene opplever vurderingen. På denne måten skaper studentenes oppfatning av kravene vurderingen stiller et rammeverk – eller en form for forståelse – som studentene jobber mot (Biggs, 1996; Entwistle & Entwistle, 1992). Vi ser at vurderingen driver læringen i den forstand at det er den forventede standarden på vurderingen, og ikke de reelle kravene eller læringsutbyttet, som avgjør studentenes forståelse av emnets innhold.

Entwistle (1994) beskriver tre ulike læringsstrategier for studenter; en strategi hvor målsetningen er dyp læring integrert i studentenes eget rammeverk, en strategi hvor studentene kun lærer ting utenat og overflattisk og en strategi hvor studentene strategisk trekker veksler på begge de nevnte strategiene ut fra en vurdering av hva som er nødvendig basert på undervisning og vurdering.

Når faktorene beskrevet over er kjent, kan backwash-effekten utnyttes til å synliggjøre hva studentene skal sikte mot, og vi kan på denne måten sørge for at studentenes streben mot læring tar dem i riktig retning. Vurderingen må tydelig reflektere læringsutbyttet, og foreleserne må gi studentene trening i å løse komplekse oppgaver, slik at de er rustet for den vurderingen de skal gjennom. På denne måten fører backwash til en positiv læringseffekt som aktivt kan brukes til å gi studentene retning på læringen i samsvar med kravene som stilles i vurderingen, emnets innhold og det definerte læringsutbyttet (Biggs, 1996; Biggs, 1998; Biggs & Tang, 2011; Entwistle & Entwistle, 1992; Havnes, 2004).

2.2.4 Reliabilitet og validitet i vurdering

I relasjon til vurdering må man sørge for å sikre at resultatene er pålitelige (reliabilitet) og at de måler det de skal måle (validitet) (Biggs & Tang, 2011; Bloom, Hastings, & Madaus, 1971; Eggen, 2009; Raaheim, 2016). I tillegg hadde Bloom, Hastings og Madaus (1971) med en tredje faktor; objektivitet i karaktersetting (scoring objectivity). Denne er hos de andre inkludert som en del av de to andre begrepene.

Reliabilitet relatert til vurdering er et spørsmål om stabilitet knyttet til resultatene, det vil si påliteligheten av vurderingen. Kontekstuelle faktorer som vurderingsform, hvem som vurderer, relevansen av vurderingsformen, etc. påvirker reliabiliteten. Det som er avgjørende for reliabiliteten av en vurdering er at vurderingen er gjort i henhold til de definerte parameterne, og at samme resultat oppnås under identiske forhold (Bloom, Hastings, & Madaus, 1971; Raaheim, 2016). Vurderingen som gjøres må stå i forhold til de definerte kriteriene og referansene (Eggen, 2009), slik Bloom, Hastings og Madaus' (1971) faktor objektivitet i karaktersetting fremhever.

Validitet er et spørsmål om vurderingens gyldighet. I følge Eggen (2009) kan validitet i hovedsak deles inn i innholdsvaliditet, ytre validitet og konsekvensvaliditet. Innholdsvaliditet er knyttet til det faglige innholdet som skal vurderes (Bloom, Hastings, & Madaus, 1971; Eggen, 2009), ytre validitet er knyttet til mulighetene til å generalisere og konsekvensvaliditet er knyttet til vurderingens konsekvenser overfor individet, samfunnet, etc. Raaheim (2016) understreker at det ikke er vurderingsformen som er avgjørende for validiteten, men den typen oppgaver studenten blir prøvd i. Disse må være relevante og autentiske.

I relasjon til utbyttebasert undervisning og læring med en forhåndsdefinert vurderingsstandard avgjøres reliabiliteten av to faktorer; Hvorvidt den samme sensoren gjør den samme vurderingen ved to ulike anledninger og hvorvidt ulike sensorer gjør samme vurdering av samme besvarelse på et gitt tidspunkt. Hvis man bruker dette perspektivet, er reliabilitet ikke et særtrekk ved en bestemt test, men evnen sensorene har til å gjøre konsekvente vurderinger (Biggs & Tang, 2011). For å oppnå konsekvente vurderinger, krever det at læringsutbyttet og vurderingskriteriene er tydelig definert.

Validitet er i kriteriebasert vurdering knyttet til tolkninger og bruk av testresultatene i relasjon til hele undervisningskonteksten. Det er bedømmelsen som fullender helheten som utgjøres av læringsutbyttet, læringsmiljøet og vurderingen i en kvalitativ vurdering. Et mål på validiteten er dermed hvor godt vurderingen samsvarer med det definerte læringsutbyttet, og validiteten avhenger i større og større grad av foreleserens profesjonalitet og ansvar for å bedømme kvaliteten på læringen (Biggs & Tang, 2011).

2.3 Vurdering i STCW

Nautikkutdanningen er underlagt regelverket i International Convention on Standards for Training, Certification and Watchkeeping for Seafarers 78 as Amended, STCW. Dette regelverket legger føringer for både opplæring og vurdering, og sier blant annet i Regulation I/6 Training and assessment at opplæring og vurdering skal være i henhold til standardene satt i regelverket, og at de som skal drive opplæring og vurdering må være kvalifisert i henhold til samme regelverk (International Maritime Organization, 2011). I veiledningsdelen av STCW er det beskrevet under B-II/1 Evaluation of Competence at;

«Assessment of competence should, therefore, encompass more than the immediate technical requirements of the job, the skills and the tasks to be performed, and should reflect the broader aspects needed to meet the full expectations of competent performance as a ship's officer. This includes relevant knowledge, theory, principles and cognitive skills, which, to varying degrees, underpin all levels of competence. It also encompasses proficiency in what to do, how and when to do it, and why it should be done» (International Maritime Organization, 2011, s. 314).

I tillegg bidrar IMO med veiledning til hvilken standard det beskrevne læringsutbyttet fra tabellene (f.eks. tabell A-II/1 og A-II/2) skal vurderes opp mot gjennom kolonne 4 i

tabellene «Criteria for evaluating competence» (International Maritime Organization, 2011).

Veiledningen har fokus på prosessen rundt vurderingen og gir blant annet råd om hvilken dokumentasjon som kreves og hvilke former for vurdering som kan benyttes for å innhente informasjon om hvorvidt kandidatene har oppnådd det spesifiserte læringsutbyttet. STCW-regelverket opp til kompetansebasert opplæring og vurdering, og tabellene spesifiserer;

- Kolonne 1 – den aktuelle kompetansen
- Kolonne 2 – det konkrete læringsutbytte knyttet til kompetansen i kolonne 1, brutt ned i kunnskap, forståelse og ferdigheter (Knowledge, Understanding and Proficiency)
- Kolonne 3 – metoder for å evaluere kompetansen
- Kolonne 4 – Vurderingskriterier.

I tillegg bidrar veiledningen med forslag til ulike måter å teste kompetansen (International Maritime Organization, 2011).

Vurderingskriteriene fastsatt i de obligatorisk tabellene i STCW omfatter ulike kognitive nivåer, for eksempel å identifisere farer i et kart, posisjonsbestemmelse basert på info fra et system i lys av systemets svakheter, kalkulasjon, sjekk av utstyr, loggføring, kommunikasjon, lederskap, beslutninger, tolking og bruk av data, utarbeidelse av planer i tråd med kunnskaper og situasjonsforståelse, bruk av utstyr, evalueringer, med mer (International Maritime Organization, 2011).

3. Metode

Forskning blir av Creswell (2014) beskrevet som en prosess hvor man samler inn og analyserer informasjon for å oppnå økt forståelse for eller innsikt i et tema eller en problemstilling. Metode betyr opprinnelig «veien til målet» (Krumsvik, 2014; Kvale & Brinkmann, 2009), og skal være et verktøy som velges ut fra hva man ønsker å finne svar på. På denne måten bidrar metoden til å gi forskeren svar på forskningsspørsmålene (Creswell, 2014; Krumsvik, 2014; Ringdal, 2013; Silverman, 2013).

I motsetning til kvantitativ metode, hvor forskeren grunngir sin problemstilling ut fra det at noe skjer og at noe kan måles (Creswell, 2014; Krumsvik, 2014), er kvalitativ forskning best egnet til å utforske problemstillinger som vektlegger forståelse og analyse av sammenhenger i forhold til noe som skjer. I kvalitativ forskning er variablene ukjente, og man må lære mer om temaet gjennom utforsking av fenomenet med det formål å forstå deltagerens perspektiv (Creswell, 2014; Krumsvik, 2014; Postholm, 2010).

3.1 Kvalitativ metode

Ringdal (2013) skriver at kvalitativ metode er godt egnet til utforsking av fenomener det i liten grad er forsket på, og påpeker at forskningsspørsmål som begynner med «hva» eller «hvordan», samt studier hvor formålet er å beskrive noe, peker mot kvalitativ metode. Problemstillingen «Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkutdanningen tilpasses en digital kontekst» indikerer således at det riktige verktøyet for å få de svarene man ønsker er å finne ved hjelp av kvalitativ metode. På denne måten har formålet med studien «valgt» metode (Krumsvik, 2014).

3.1.1 Educational Design Research

Metoden brukt i denne undersøkelsen er Educational Design Research (EDR), en metode som ifølge McKenney og Reeves (2012) er en kombinasjon mellom vitenskapelig undersøkelse og systematisk utvikling av praksis. Å velge EDR gir ikke noen forhåndsbestemt fremgangsmåte, men en overordnet strategi for forskningen. Strategien går ut på at man med forskningen ønsker å utvikle praksis og/eller teori, og fører til at man videre må finne de riktige metodiske verktøy egnet til å oppnå en slik målsetning. McKenney og Reeves (2012) definerer EDR som «..a genre of research in which the

iterative development of solutions to practical and complex educational problems also provides the context for empirical investigation... » (s. 7). Dette betyr at kunnskapen som utarbeides gjennom forskningen kan benyttes i praksis. Educational Design Research er i følge van den Akker, Gravemeijer, McKinney og Nieveen (2006) et modigere og mer risikabelt valg, enn konvensjonelle samfunnsvitenskapelige metoder, men «...it stands a much better chance of leading to innovative designs» (van den Akker et.al., 2006, s. 11). Det er nettopp utvikling av eksisterende praksis – innovasjon – vi ønsker å oppnå med innføringen av digital eksamen på nautikkstudiet.

EDR er en forskningsprosess med gjentakende faser, hvor hver gjentakelse bidrar til å spisse målsetningen, øke innsikten og bidra til utvikling. I en EDR-studie følger man en gjentakende syklus som består av tre hovedfaser; analyse og utforskning, design og prototyping og evaluering og refleksjon (McKenney & Reeves, 2012; van den Akker et al., 2006). Både forelesere og forsker har vært involvert i alle fasene. Den første fasen, analyse og utforskning, ble først trigget gjennom diskusjoner i kollegiet om digital skoleeksamen som vurderingsform. Da det var bestemt at digital skoleeksamen var en vurderingsform vi anså egnet som en av to eksamensformer i et nytt emne i nautikks studieplan, ble dette grundig diskutert. Nåsituasjonen ble kartlagt ved at litteratur om digital eksamen ble studert (McKenney & Reeves, 2012; van den Akker et al., 2006). Litteraturen på området avdekket at pedagogiske problemstillinger relatert til digital eksamen er lite utforsket sammenlignet med tekniske, administrative og logistiske sider ved denne typen eksamen. Resultatene fra litteraturgjennomgangen er nærmere beskrevet i kapittel 1.4 Tidligere forskning.

I den andre fasen, design og prototyping, ble en pilot gjennomført i et eksisterende emne med et annet studentkull. Studentene som tok testeksamen, gjorde dette på frivillig basis. I etterkant ble erfaringene fra piloten brukt som grunnlag for planleggingen av det nye emnet. Det ble besluttet å gjennomføre to studietester før studentenes endelige eksamen i emnet, slik at studentene skulle kjenne programvaren når de kom til eksamen. I etterkant av studietestene gjennomførte forelesere og forsker et refleksjonsmøte (gruppeintervju) relatert til deres erfaringen fra disse pilottestene. I dette refleksjonsmøtet var evaluering av studietestene og mulig forbedring av praksis før eksamen tema.

Både forelesere og studenter ble involvert i den siste fasen, evaluering og refleksjon. I fokusgrupper, refleksjonsgruppe og individuelt intervju var eksamen tema. Målsetningen

var å evaluere gjennomføringen av digital eksamen, og hvorvidt læringsutbyttet ble testet i henhold til planen.

Ved å velge Educational Design Research som overordnet metodisk ramme, kan man benytte ulike former for vitenskapelig metode som passer til formålet. I dette arbeidet har jeg valgt å benytte et fenomenologisk perspektiv, for å kunne beskrive den meningen de involverte parter legger i opplevelsen knyttet til erfaringene med fenomenet digital eksamen (Alvesson & Sköldberg, 2008; Krumsvik, 2014; Kvale & Brinkmann, 2009; Moustakas, 1994; Postholm, 2011).

3.1.2 Fenomenologi

Fenomenologien er rotfestet i de filosofiske perspektivene til Edmund Husserl (1859-1938) fra ca. år 1900 og var basert på subjektiv åpenhet, en radikal innfallsvinkel til vitenskapen som den gang møtte mye kritikk (Alvesson & Sköldberg, 2008; Moustakas, 1994; Postholm, 2010). Husserl var blant annet inspirert av Descartes, og fremholdt på tross av kritikk og hån at «Phenomena are the building blocks of human science and the basis for all knowledge» (Moustakas, 1994, s. 26). Husserls transcendentale fenomenologi har senere blitt videreutviklet av Heidegger, Sartre og Merleau-Ponty. Gjennom fenomenologien søker forskeren å beskrive hvordan mennesker opplever et bestemt fenomen i sin livsverden, for på denne måten å forstå fenomenet ut fra aktørens perspektiv (Kvale & Brinkmann, 2009).

I denne studien er det forskningsdeltakernes opplevelse av fenomenet digital eksamen som søkes beskrevet gjennom deres eget perspektiv (Moustakas, 1994; Postholm, 2010; Thagaard, 2013). Dette oppnås gjennom samtaler med forskningsdeltakerne, som for å delta må ha opplevd fenomenet (Postholm, 2010).

3.2 Intervju

«Å intervju mennesker kan dermed bety at en får tak i deler av en annen persons liv som man vanskelig kunne ha fanget opp på andre måter» (Postholm, 2010, s. 68). Forskningsintervjuer deles inn i tre typer; strukturert, semistrukturert og ustrukturert (Krumsvik, 2014; Postholm, 2010). Strukturerte intervju er planlagte og formelle intervjuer, hvor intervjueren har skrevet alle spørsmålene i forkant og det er liten fleksibilitet i forhold til avvik fra planen. Semistrukturerte (halvstrukturerte) intervju er

halvplanlagte og formelle intervjuer hvor intervjueren har en plan for hvilke spørsmål som ønskes besvart, men med fleksibilitet nok til å forfølge intervjupersonens refleksjoner i den retningen disse måtte ta. Ustrukturerte intervjuer er uplanlagte og halvformelle intervjuer, og tar form av et åpent intervju. Grensene mellom et ustrukturert intervju og deltagende observasjon kan være vage (Krumsvik, 2014; Postholm, 2010).

Et forskningsintervju er preget av en asymmetrisk relasjon, hvor forskeren er den som planlegger temaet for intervjuet, og som avgjør hvordan det som kommer frem i intervjuet benyttes inn mot forskningsformålet (Thagaard, 2013). Dette vil i enda større grad være en viktig faktor når intervjueren har en relasjon til intervjupersonene gjennom en maktposisjon, slik som er tilfelle i dette prosjektet. Dette kommer jeg tilbake til i punkt 3.4 Metodiske refleksjoner under.

Datainnsamling ved intervjuer er den mest dominerende metoden innenfor kvalitativ forskning, og individuelle intervjuer er mest vanlig. I tillegg kan gruppeintervjuer benyttes, for eksempel gjennom fokusgrupper (Thagaard, 2013). Ved å stille åpne spørsmål, kan intervjupersonene selv formulere sine erfaringer og tanker rundt et tema, helt uavhengig av forskerens standpunkt eller meninger (Creswell, 2014).

3.2.1 Fokusgruppeintervju

«Fokusgrupper er uformelle gruppeintervjuer eller gruppediskusjoner som ledes av en forsker som fungerer som møteleder eller moderator» (Ringdal, 2013, s. 247). Helt fra 1920-årene har samfunnsforskere benyttet gruppeintervjuer. Fra 1980- tallet ser vi også fokusgruppeintervju i den akademiske samfunnsforskningen. I dag benyttes fokusgruppeintervju til mange ulike formål, blant annet i forbruksforskning (Krueger & Casey, 2015; Kvale & Brinkmann, 2009). Også i utdanningsforskning er bruken av fokusgrupper økende, men i et noe saktere tempo, enn i andre typer forskning (Cohen, Manion, & Morrison, 2011). Fokusgrupper synliggjør et kollektivt standpunkt muliggjort gjennom interaksjon mellom deltagerne i gruppen, og det er deres – ikke forskerens – agenda som dominerer (Cohen, Manion, & Morrison, 2011).

Fokusgrupper varierer i størrelse, og settes sammen basert på formålet med undersøkelsen. Det er mulig å gjennomføre fokusgrupper med alt fra tre til fire og helt opp til tolv deltagere, men slike grupper består vanligvis av mellom fem og åtte medlemmer (Halkier, 2010; Krueger & Casey, 2015; Kvale & Brinkmann, 2009; Postholm, 2010; Silverman, 2013). I denne studien har jeg valgt å gjennomføre tre

fokusgrupper med fire studenter i hver gruppe, i tillegg til intervjuene med foreleserne. Intervjuene med forelesere kommer jeg tilbake til i neste avsnitt, refleksjonsmøter. Årsaken til at jeg har valgt små fokusgrupper er at jeg ønsker å forstå deltagerne erfaringer med digital eksamen, og har derfor behov for mer utdypning av studentenes meninger. I henhold til Krueger & Casey (2015) må man for å oppnå en slik dybdeinnsikt velge små grupper på for eksempel fire deltagere. På denne måten er mulig å gjøre dyperegående analyser (Thagaard, 2013). Mindre fokusgrupper er også mer behagelig for deltagerne (Krueger & Casey, 2015), noe som er viktig når man skal intervjuere studenter. I tillegg er små fokusgrupper lettere å styre (Halkier, 2010), noe som er en fordel i denne sammenhengen, siden forskeren er uerfaren.

Fordeler med fokusgrupper uavhengig av størrelse er at gruppedynamikken kan gi andre svar og innsikt man ikke ville fått gjennom individuelle intervju, i tillegg til at det er tidsbesparende (Ringdal, 2013). Det kan også være en fordel når man skal rekruttere studenter, siden en fokusgruppe vil kunne virke mindre «skummel», siden de deltar sammen med klassekamerater og dermed kjenner de andre relativt godt. Ulemper med fokusgrupper kan være at de hindrer individuelle synspunkter og at en deltager dominerer (Ringdal, 2013). Jeg anser det ikke å være stor fare for at disse fokusgruppene hindrer individuelle synspunkter, siden deltagerne kjenner hverandre godt og dermed er vant til å ytre sine meninger i gruppen. Men hvis så skulle skje, er det moderator sitt ansvar å sørge for at alle meninger kommer frem, og i tilfeller hvor enkeltpersoner dominerer samtalen, må moderator ta jobben med å rette spørsmålene også til de andre deltagerne i gruppen.

Antallet fokusgrupper varierer, men en gylden regel er at man skal fortsette å intervjuere til man ikke får flere nye svar eller ny kunnskap (Halkier, 2010; Krueger & Casey, 2015). Krueger og Casey (2015) anbefaler at det planlegges for minimum tre fokusgrupper for hver deltagergruppe. Etter at disse er gjennomført, må forskeren avgjøre hvorvidt man har oppnådd «metning», og dermed ikke får ny informasjon. I utgangspunktet var det ikke mer enn 12 studenter totalt som meldte seg for deltagelse i studien, så antallet fokusgrupper ble fastsatt av det. Men etter den tredje fokusgruppen opplevde jeg at det heller ikke kom nye momenter frem i intervjuene lenger, og tre fokusgrupper ville derfor vært et naturlig sted og stoppe uansett.

3.2.2 Refleksjonsmøter

I tillegg til fokusgruppeintervjuene med studenter, har foreleserne blitt intervjuet. Intervjuene med foreleserne har blitt gjennomført som refleksjonsmøter (McKenney & Reeves, 2012; van den Akker, Gravemeijer, McKinney, & Nieveen, 2006), i form av to formelle gruppeintervjuer og ett individuelt intervju.

Et særtrekk ved EDR er at forskerens rolle er mindre distansert, enn i ordinær forskning. Forskeren identifiserer «problemet» sammen med de som praktiserer, og forsker og praktiserende lærer av hverandre (McKenney & Reeves, 2012). Dette vil si at både deltagerne, litteraturen og gjennomføringen er det som former resultatet av forskningsprosjektet og dermed også videre praksis (McKenney & Reeves, 2012; van den Akkeret al., 2006). EDR er «...structured to explore, rather than mute, the complex realities of teaching and learning contexts, and respond accordingly» (McKenney & Reeves, 2012, s. 15). Dette utvikles over tid gjentakende faser av undersøkelser, utvikling, testing og forbedring (McKenney & Reeves, 2012).

Noe av det som skiller EDR fra mange andre typer forskning, er at målsetningen ikke bare er å forstå problemet, men faktisk å løse det i praksis (McKenney & Reeves, 2012; van den Akkeret al., 2006).

3.2.3 Utvalg

En kvalitativ studie bygger som regel på et strategisk utvalg av begrenset størrelse, og deltagerne velges ut fra hvilke spørsmål man ønsker besvart (Halkier, 2010; Krueger & Casey, 2015; Kvale & Brinkmann, 2009; Postholm, 2010; Ringdal, 2013; Silverman, 2013; Thagaard, 2013). I noen studier vil man ha behov for homogene grupper, mens man i andre studier vil være mer opptatt av at flere perspektiver og erfaringer er representert i den enkelte fokusgruppe (Halkier, 2010; Krueger & Casey, 2015).

I denne studien er deltagerne valgt ut fra to ulike kriterier; foreleserne er valgt ut fra at de er involvert i planlegging, gjennomføring og retting av digital eksamen, og studentene er valgt ut fra at de har gjennomført den aktuelle digitale skoleeksamenen, samt at de har meldt seg frivillig. Utvalget består derfor av to homogene grupper (Krueger & Casey, 2015). Ved gjennomføring av første fokusgruppe med studenter var det kun nok studenter til en og en halv fokusgruppe som hadde meldt sin interesse for deltagelse i prosjektet, men etter at første gruppe hadde gjennomført var det flere studenter som kom og meldte sin interesse. På denne måten hadde de studentene som deltok i det første intervjuet bidratt

til at flere studenter meldte sin interesse. Utvalget er strategisk valgt siden de er forventet å kunne bidra med innsikt i fenomenet, og består av tre fokusgrupper med totalt fire studenter i hver fokusgruppe, to refleksjonsmøter med begge forelesere hver gang (gruppeintervju) og ett individuelt intervju med en av foreleserne (Halkier, 2010; Krueger & Casey, 2015; Kvale & Brinkmann, 2009; Postholm, 2010; Ringdal, 2013; Silverman, 2013; Thagaard, 2013).

3.2.4 Gjennomføring av undersøkelsen

Undersøkelsen ble gjennomført i siste del av 2016 og i begynnelsen av 2017. Intervjuene med foreleserne ble gjennomført i rekkefølgen refleksjonsmøte om studiekraft, individuelt intervju og refleksjonsmøte om eksamen. Alle intervjuene med forelesere ble gjennomført i 2016. Studentene ble intervjuet i 2017, ca. en måned etter at eksamen var gjennomført. Årsaken til dette var at det er vanskelig å få samlet studentene i eksamensperioden og juleferien. Intervjuene med studenter ble derfor av praktiske årsaker gjennomført når de var tilbake på skolen etter juleferien.

Før jeg skulle gjennomføre intervjuene ble spørsmålene til foreleserne prøvd ut på andre forelesere i avdelingen og spørsmålene til studentene ble testet på andre studenter (Krueger & Casey, 2015; Silverman, 2013; Thagaard, 2013). Noen av studentene som i vårsemesteret gjennomførte pilottesten på digital skoleeksamen i et annet emne på nautikkstudiet sa seg villige til å være med og teste spørsmålene i intervjuguiden min. Resultatet av pilottestingen av spørsmålene var at jeg endret rekkefølgen på noen av spørsmålene, presiserte noen av spørsmålene bedre slik at deltagerne lettere skulle forstå disse, samt formulerte noen mulige oppfølgingsspørsmål. Jeg har prøvd å ha så åpne spørsmål som mulig i fokusgruppene for å få i gang samtalene og for ikke at deltagerne skal «låse» seg i ett spor.

Alle intervjuene ble gjennomført i et møterom i nautikkhuset, et sted som er kjent for både forelesere og studenter. I intervjuene med studentene hadde jeg forberedt litt lett servering, slik at vi spiste sammen før intervjuet begynte (Krueger & Casey, 2015). Dette bidro til å dempe intervjuer sin posisjon som autoritet i gruppen (Halkier, 2010). Før jeg for alvor begynte intervjuet, gjennomgikk jeg informasjon om personvern og anonymitet, at det var frivillig å delta, jeg innhentet samtykke for lydopptak av intervjuene, og informerte om at intervjupersonene på et hvilket som helst tidspunkt kan trekke seg fra studien uten å oppgi grunn (Den nasjonale forskningsetiske komitè for

samfunnsvitenskap og humaniora (NESH), 2016). Jeg oppsummerte også kort hva forskningsprosjektet handler om, i tillegg til å presisere at jeg var ute etter deres erfaringer og opplevelser, og at alle svar var like riktige. Deltagerne ble også informert om at alt blir anonymisert, slik at ingen kjenner igjen deres utsagn eller deres meninger (Den nasjonale forskningsetiske komitè for samfunnsvitenskap og humaniora (NESH), 2016). I tillegg fikk deltagerne praktisk informasjon om hvordan intervjuet ble lagt opp.

Intervjuene forløp greit, og det var en velvillig og åpen atmosfære som preget intervjusituasjonen (Halkier, 2010; Kvale & Brinkmann, 2009). De involverte regulerte seg selv, og sørget for at alles meninger kom frem underveis i diskusjonen. Alle fokusgruppeintervju ble gjennomført som semistrukturerte intervjuer (Halkier, 2010; Krumsvik, 2014; Kvale & Brinkmann, 2009; Postholm, 2010; Thagaard, 2013), hvor jeg hadde forberedt noen tema som medlemmene i fokusgruppene skulle diskutere (se vedlegg 4, 5 og 6 for intervjuguider). I tillegg hadde jeg klargjort oppfølgingsspørsmål, hvis diskusjonen ikke skulle komme i gang av seg selv, eller jeg ikke fikk de svarene jeg hadde behov for (Halkier, 2010). Det individuelle intervjuet ble gjennomført med en noe mer styrt struktur (Halkier, 2010; Krumsvik, 2014; Kvale & Brinkmann, 2009; Postholm, 2010; Thagaard, 2013), hvor jeg hadde forberedt mer detaljerte spørsmål. Spørsmålene var utarbeidet i forkant av intervjuene, men ble justert på bakgrunn av gruppeintervjuene. På denne måten kunne jeg innhente mer detaljert informasjon om ting jeg kanskje ikke hadde klart å grave dypt nok i under gruppeintervjuene. En sannsynlig årsak til dette kan være at jeg ikke er erfaren i intervjusituasjonen, og derfor ikke fanget opp behovet for å grave dypere før jeg transkriberte gruppeintervjuene i ettertid (Thagaard, 2013). Det individuelle intervjuet ga meg også mulighet til å følge sporene fra den første refleksjonssamtalen, slik at jeg fikk en dypere forståelse for noen av svarene derfra.

3.2.5 Transkribering

I etterkant av intervjuene ble lydfilene transkribert. Dette gjorde jeg selv, noe som anbefales i litteraturen siden intervjuer allerede har kjennskap til deltagerne og situasjonen (Halkier, 2010), men som også er praktisk da man allerede i transkriberingsfasen så smått kan begynne å analysere materialet (Kvale & Brinkmann, 2009). Transkriberingen ble gjort i Word ved hjelp av transkriberingsmodulen Olympus Dictation Management System Transcription Module. Dette ble så overført til Hyper Research for analyser. Transkriberingen ble gjort like etter gjennomføringen av intervjuene, slik at det var lett å huske de ulike nyansene i intervjuene, hva ulikt tonefall

indikerte, etc. I tillegg førte jeg inn ting som ((ler)), ((utydelig)) og lignende, slik at sammenhengen skulle fremgå i tråd med Kvale og Brinkmann (2009) sin anbefaling. Når for eksempel deltagerne snakket i munnen på hverandre var det ikke alltid mulig å tyde alt som ble sagt. Bearbeidelsen av transkripsjonene ble gjort så skriftlig som mulig på bokmål, uten at dette skulle påvirke innholdet i det som ble sagt. Fyllord som ble gjentatt, for eksempel «liksom», «på en måte», «vet du», og lignende ble derfor utelatt.

3.3 Kvalitativ dataanalyse

Det er mange måter å analysere kvalitative data (Krueger & Casey, 2015). Kvale og Brinkmann definerer det å analysere som «å dele noe opp i biter eller elementer» (2009, s. 201). Det er forsøket på å få frem deltagerens opplevelse av fenomenet som er hensikten med analysen (Grimen, 2004; Moustakas, 1994).

Proessen med å analysere et intervju kan i følge Kvale og Brinkmann (2009) reduseres til fem trinn; lese helheten i datamaterialet, bryte ned teksten i stikkord for å beskrive meningen, beskrive temaet basert på fortolkningen man har gjort, bruke meningsfortolkningene til å belyse forskningsspørsmålene og sette dette sammen til en beskrivende helhet. Jeg startet derfor med å lese det transkriberte materialet for å få et helhetlig inntrykk av materialet. Deretter benyttet jeg meg av det som Kvale og Brinkmann (2009) kaller meningskodning, og knyttet nøkkelord til avsnitt i teksten. Jeg benyttet såkalt datastyrt koding, og utviklet kodene etter hvert som jeg leste datamaterialet basert på en meningsfortetting. Meningsfortetting er fenomenologisk basert og skjer når man forkorter det intervjupersonene har sagt til helt korte stikkord som gjengir selve meningen med utsagnet. Deretter gjorde jeg en meningsfortolkning, som i følge Kvale og Brinkmann (2009) er en dyp og kritisk fortolkning av teksten som går ut over det å strukturere det som har blitt sagt i intervjuene og over til å finne meninger og betydninger som ikke umiddelbart sees i teksten. Ved vekselvis å lese og tolke datamaterialet, og å gjøre dette flere ganger, fikk jeg gradvis en økt forståelse for hva som var sentralt i intervjuene. Som et resultat av denne prosessen fant jeg også de tre overordnede kategoriene som danner grunnlaget for hvordan presentasjonen av funn er strukturert i kapittel 4, Presentasjon av funn.

3.4 Metodiske refleksjoner

Kvalitativ metode har i henhold til Thagaard (2013) blitt forbundet med «forskning som innebærer nær kontakt mellom forsker og de som studeres» (s. 11). Dette reiser en rekke utfordringer relatert til metodologi og etikk (Thagaard, 2013).

3.4.1 Reliabilitet og validitet

I kvantitativ forskning er reliabilitet og validitet knyttet til hvorvidt man måler det man skal og hvorvidt funnene er generaliserbare (Cohen, Manion, & Morrison, 2011; Creswell, 2014; Ringdal, 2013). I kvalitativ metode er derimot begrepene validitet og reliabilitet ikke like åpenbare. Lincoln og Guba (1985) har brukt begreper som «trustworthiness» for å beskrive intern og ekstern validitet, reliabilitet og objektivitet i kvalitativ metode. Thagaard (2013) sier om reliabilitet at «...forskeren reflekterer over konteksten for innsamlingen av data og hvordan relasjonen til deltakere i prosjektet kan influere den informasjonen forskeren får» (s. 203). Hun siterer Seale (1999) som deler reliabilitet i intern og ekstern reliabilitet, hvor ekstern reliabilitet er knyttet til hvorvidt andre forskere kan komme frem til samme resultat i en annen situasjon. Dette er så godt som umulig i kvalitative studier (Seale i Thagaard, 2013). Intern reliabilitet er knyttet til at forskeren må være spesifikk og konkret i beskrivelsen av datainnsamling og -analyse, slik at fremgangsmåtene kommer tydelig frem og forskningsprosessen dermed er transparent (Thagaard, 2013). Ringdal beskriver reliabilitet i kvalitativ forskning som «forskerens refleksjon over hvordan datainnsamlingen har foregått, med sikte på å bli seg bevisst mulige feilkilder» (2013, s. 248). Dette betyr i praksis at reliabilitet handler om pålitelighet (Krumsvik, 2014; Ringdal, 2013; Thagaard, 2013), og det er viktig for en forsker å sikre denne gjennom klare og tydelige spørsmål i intervjuene. Spørsmålene man stiller under intervjuene må derfor være utvetydige, ikke ledende og utformes i et språk som intervjudeltagerne forstår (Krumsvik, 2014). Dette har jeg jobbet med i utformingen av intervjuguider, med det formål å ha så åpne spørsmål som mulig slik at det er deltagerens egne opplevelser som legger føringene for svarene deres. I tillegg vil reliabilitet også være en utfordring relatert til transkribering og analyse av data, siden to personer som transkriberer eller analyserer intervjuer ikke nødvendigvis kommer frem til det samme (Krumsvik, 2014). Både transkribering og analyse er gjort av samme person som har gjennomført intervjuene, og det er derfor konsistens i hvordan materialet fortolkes.

Min rolle som studieleder er også en faktor i vurderingen av reliabilitet i studien, siden jeg i hverdagen innehar det som kan beskrives som en maktposisjon i relasjon til de som intervjues, noe som kan påvirke svarene som kommer (Kvale & Brinkmann, 2009; Postholm, 2010; Thagaard, 2013). I lys av tilbakemeldinger jeg vanligvis får fra studenter og forelesere på nautikkstudiet gjennom innspill, møter, diskusjoner, evalueringer og lignende, har jeg – uten å ha dokumentasjon på det – et inntrykk av at jeg får ærlige og ufiltrerte tilbakemeldinger fra både studenter og ansatte. Det gjør at jeg har tiltro til at svarene jeg får i intervjuene er studenter og foreleseres reelle meninger om digital eksamen.

Ringdal (2013) beskriver validitet i kvalitativ forskning som noe som kan måles av forskeren selv eller av intervjupersonene. Validitet handler om hvorvidt man har målt det man ønsket å måle og vil i praksis si at det handler om hvorvidt forskerens tolkninger er gyldige eller ikke (Ringdal, 2013; Thagaard, 2013). I EDR benyttes begrepet «ecological validity», og innenfor denne overordnede metodikken er en studie valid hvis metodene, materialene og omgivelsene i studien gjenspeiler den reelle situasjonen som undersøkes. Studiens muligheter for generalisering øker når den gjennomføres så virkelighetsnært som mulig (Brewer i McKenney & Reeves, 2012). Dette opplever jeg at er oppnådd i studien, gjennom samtalene med forelesere og studenter. I et prosjekt hvor EDR benyttes som overordnet metodikk, vil det fordre større involvering fra forskers side, enn hva som er vanlig i andre former for kvalitative studier. Dette fører også til at forsker ikke lenger er en utenforstående, men vil ha en referanseramme for å forstå det som skal tolkes (Kvale & Brinkmann, 2009; McKenney & Reeves, 2012; van den Akker et al., 2006), noe jeg også har opplevd at jeg har.

3.4.2 Etikk og personvern

I et kvalitativt forskningsintervju er maktforholdet mellom intervjueren og de som blir intervjuet asymmetrisk, uavhengig av hvorvidt forskeren utøver makt med overlegg, noe forskeren må være bevisst på (Kvale & Brinkmann, 2009). Dette vil også være en faktor i den videre bruken av informasjonen som er samlet inn, siden det er forskeren som avgjør hvordan de intervjuedes uttalelser bearbeides og benyttes (Thagaard, 2013).

I tillegg til å være oppmerksom på og prøve å utjevne det asymmetriske maktforholdet som intervjuet utgjør (Kvale & Brinkmann, 2009; Thagaard, 2013), er det også flere andre etiske perspektiver som må tas i betraktning. Forskningen må utføres i henhold til

Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi, som blant annet slår fast at forskningen må gjennomføres på en ansvarlig og aktsom måte i samsvar med hensynet til personvern (Den nasjonale forskningsetiske komitè for samfunnsvitenskap og humaniora (NESH), 2016). Dette er ivaretatt gjennom at forskningsdeltagerne er fullt informert om prosjektet, hva dataene som samles inn skal benyttes til, samt hvordan opplysningene lagres og når de slettes. Deltagerne har fått informasjon om formålet med studien, hvordan deres utsagn blir behandlet, hvem som får tilgang til informasjonen, hvordan resultatene skal benyttes og hvilke konsekvenser det kan få for dem å delta i studien. De har også fått informasjon om at deltagelsen er frivillig og at de på et hvilket som helst tidspunkt kan trekke seg fra prosjektet uten å måtte oppgi grunn for det (Den nasjonale forskningsetiske komitè for samfunnsvitenskap og humaniora (NESH), 2016). Deltagerne fikk informasjon om prosjektet både skriftlig (se vedlegg 2 for informasjonsskriv/samtykkeskjema) og muntlig før de meldte sin interesse for deltagelse. I tillegg har de lest og skrevet under på samtykkeskjema før datainnsamlingen startet. De er også blitt anbefalt å ta kontakt hvis de skulle ha spørsmål før, underveis eller i etterkant av datainnsamlingen.

I og med at forskeren i dette tilfellet også er studieleder på nautikkstudiet, må relasjonen i forskningsprosjektet tydeliggjøres ettertrykkelig, slik at det ikke er tvil for deltagerne når jeg opptrer som forsker og når jeg opptrer som studieleder. Studielederrollen vil ikke være en del av forskningsprosjektet, og den informasjonen som kommer frem under intervjuene vil bli brukt i forskningsøyemed. Det er forskers ansvar å fremtre med klarhet og tydeliggjøre grensene mellom forskningen og daglig drift av nautikkstudiet, i tråd med Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi (Den nasjonale forskningsetiske komitè for samfunnsvitenskap og humaniora (NESH), 2016).

Studien ble meldt inn til Norsk Senter for Forskningsdata (NSD) tidlig høsten 2016, og tillatelsen fra NSD forelå i god tid før datainnsamlingen startet (ref. vedlegg 3 Tillatelse fra NSD). I tråd med innmeldt informasjon blir personopplysninger forsvarlig oppbevart adskilt fra annen informasjon fra forskningsprosjektet. Alle uttalelser er anonymisert, slik at ingenting kan spores tilbake til den enkelte forskningsdeltager. Dataene som er samlet inn skal kun brukes til denne studien, og vil bli slettet så snart prosjektet er avsluttet i tråd med NSDs retningslinjer.

4. Presentasjon av funn

I denne delen presenteres funn fra intervjuene. Intervjuene ble transkribert og analysert som beskrevet i metodedelens kapittel 3.2.5 Transkribering og kapittel 3.3 Kvalitativ dataanalyse. Flere ulike kategorier skilte seg ut i analysearbeidet, før det til slutt var mulig å samle disse i tre overordnede kategorier; *digital kontekst*, *vurdering* og *læring*. Funnene er sammensatte og dekker ulike perspektiver. Valg av kategori har derfor ikke til enhver tid vært entydige, og kategoriene kan som resultat av dette tidvis overlappe.

Presentasjon av funn er delt inn i henhold til de tre overordnede kategoriene *digital kontekst*, *vurdering* og *læring*. Hovedfokus er allikevel problemstillingen med tilhørende forskningsspørsmål.

«*Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkstudiet tilpasses en digital kontekst?*»

1. *Hvordan begrunner faglærerne utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i programmet Wiseflow?*
2. *Hvordan beskriver studentene sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen?*

Funnene er presentert som sitater fra intervjudeltagerne eller som oppsummeringer av hva de har sagt. I tillegg er enkelte av mine fortolkninger inkludert i teksten. Jeg har ikke gitt noen av de intervjuede identitetsmarkører for å sikre full anonymitet, og kaller derfor personene som siteres enten «foreleser» eller «student».

4.1 Digital kontekst

Ved innføringen av ny studieplan på nautikkstudiet ble det i prosessen med utvikling av emner gjennomført en grundig gjennomgang for å sikre sammenheng mellom læringsutbytte og arbeids- og vurderingsformer i de ulike emnene. Innføringen av digital eksamen ble som et resultat av denne prosessen igangsatt av fagmiljøet selv. Foreleserne beskriver årsakene til at de ønsket å innføre digital eksamen som flerdelt, men gir uttrykk for at det pedagogiske i all hovedsak har vært drivkraften i prosessen. De involverte foreleserne var i utgangspunktet ikke spesielt interessert i digitale verktøy; «*jeg hopper*

ikke på med en gang det er noe digitalt, fordi det er digitalt, jeg prøver å se hvilken nytteverdi det har», og jeg tolker dem derfor ikke som blendet av den digitale teknologien på noe vis. Mottagelsen i institusjonen beskrives av faglærer som *«det var jo ikke noen sånn hallelujastemning»*, men vedkommende forteller at det likevel ble opprettet en prosjektgruppe som skulle bidra under utprøving og innføring av digital eksamen. Gruppen bestod av medlemmer fra utdanningssenteret, fagavdelingen og IT, og foreleserne forteller at de har inntrykk av at de andre involverte også er fornøyd med gjennomføringen. På denne måten ser vi at ulike deler av organisasjonen har deltatt i det samme praksisfellesskapet, og situert kunnskap om digital eksamen i Wiseflow er konstruert gjennom samhandling.

Vurderingsformen i emnet er todelt, og for å oppnå resultat i emnet, må studentene gjennomføre og bestå både praktisk eksamen på simulator og digital skoleeksamen. Foreleserne synes at denne delingen har fungert bra, og uttaler at de er *«Meget fornøyd med det»*. Jeg tolker det som at de er opptatt av helhet i emnet, og at de opplever å ha lyktes med dette. Foreleserne forteller at de blant annet bruker to studietester til å dokumentere studentenes kunnskaper om sjøveisregler og sjømerker. Studietestene er passeringstester, og regulerer studentenes eksamensrett. De nevnte temaene tas i følge foreleserne ikke inn i eksamen, siden dette ville utgjort en for stor del av eksamen. Foreleserne ser det slik; *«vi tar jo ikke inn sjømerker og sjøveisregler på eksamen, fordi at de har bestått det i studiekravene, det er en del av en helhet»*. Et alternativ hvor man bare tar stikkprøver av denne kunnskapen vil i følge foreleserne ikke dekke regelverkskravene i tilstrekkelig grad; *«Det som er viktig i forhold til studietestene er det at vi kan dokumentere overfor Sjøfartsdirektoratet at de kan det, og det ville vi ikke kunnet hvis vi bare hadde gitt dem stikkprøver gjennom eksamen»*. Studentene er enige med foreleserne, og sier at *«Det hadde blitt så stort tema på eksamen...det hadde ikke vært plass til noe annet»*. Dette viser at foreleserne benytter studiekravene i emnet aktivt som en del av emnets helhet, blant annet for å dokumentere at kravene i regelverket dekkes.

4.1.1 Kjennskap til systemet

Knyttet til utarbeidelse av oppgaver og valg av oppgavetyper i digital eksamen kommer det frem at foreleserne opplever at de har lært mye om hvilke muligheter programvaren gir gjennom pilottestene;

«Jeg vil si at hele opplegget på en måte er en sånn trapp, altså du lærer litt og litt, og så ser du nye behov på nye oppgaver og begynner å lete etter de, og så finner du ut at «jo det kan brukes sånt», eller «det går ikke», og sånn går det fremover».

Foreleserne har påbegynt en prosess hvor de gjennom bruk av programvaren i pilottestene forholder seg aktivt til verktøyet, men de har ennå ikke kommet til det punktet at de kjenner til absolutt alle funksjonene. Hittil har foreleserne opparbeidet seg en viss erfaring med Wiseflow, men de uttrykker at de opplever at det ennå er mange muligheter i programvaren de ikke har prøvd ut, og sier at *«det nok ligger mange muligheter og kanskje også noen begrensinger, men mye av dette er sånt som vi ikke har oppdaget ennå»*. Foreleserne beskriver prosessen som *«en spennende reise»*, men sier samtidig at de ikke synes det er spesielt stor forskjell i prinsippene for utarbeidelse av en digital eksamen kontra en tradisjonell skriftlig eksamen. De omtaler innføring av digital eksamen som tidkrevende; *«Det har gått mye tid i det praktiske rundt det å få det til»*, og de ser rom for forbedringer og muligheter for utvikling i forhold til hvordan de tester studentene; *«vi fokuserer enda mer på hva dette faget består av, som er alt fra arbeid med papirkart, bruk av instrumenter, kursrettelser og sånne ting, og hvordan kan vi teste dem i alt dette med de mulighetene vi har»*. Dette tyder på at foreleserne ønsker å bruke digital eksamen til å forbedre vurderingspraksis.

Både forelesere og studenter er opptatt av at studentene skal ha brukt systemet før de møter det på eksamen, slik at studentene som allerede er i en stressende situasjon ikke opplever ytterligere stress. Foreleserne sier at *«Hvis vi lager en spesiell type oppgave, eller har tenkt å gjøre det til eksamen, så må vi ha en tilsvarende oppgave på studiekravene»*. Jeg tolker det som at det er viktig for foreleserne at artefakten fungerer på det ønskede nivået (primært, sekundært eller tertiært), slik at en digital eksamen tester studentenes kunnskaper, ferdigheter og generelle kompetanse i emnet, ikke deres tekniske innsikt knyttet til hvordan oppgavene rent praktisk skal løses. Studentene uttrykker at de er komfortable med digitale media, og opplever at dette er deres «hjemmebane». Det kan tolkes som at digitale verktøy generelt sett er appropriert hos studentene. Også foreleserne opplever at studentene er mer vant til denne typen verktøy, enn det de selv er; *«det er et språk de skjønner mye bedre enn oss»*. Studentene *«er mer vant til digital eksamen fra før av, fra videregående og sånt»* og synes derfor at det er helt greit. På det jevne fremhever alle fokusgruppene med studenter at de syntes det var positivt at de hadde brukt programvaren før eksamen skulle gjennomføres, slik at de *«vet hvordan programmet*

fungerer, og føler deg litt mer trygg på det». Studentene er enige om at de venner seg til nye typer oppgaver, og en av diskusjonene rundt dette temaet dreide seg om at det *«blir jo fort en vane»*. Slik jeg tolker det approprierer studentene tilsynelatende det digitale verktøyet relativt greit, og eksamen bør derfor kunne fungere på tertiært nivå i henhold til forelesernes intensjoner.

4.1.1.1 Digital eksamen som verktøy

Digital skoleeksamen beskrives av foreleserne som en god måte å teste studentene på, ut fra et helhetlig perspektiv. De sier at *«det aller viktigste det er ikke akkurat det tekniske, men hvilke muligheter det gir for å lage enda bedre oppgaver»*, noe som illustrerer en kontinuerlig utvikling av redskaper som kan benyttes for å løse nye oppgaver. Foreleserne opplever at de befinner seg i en prosess som gir stadig økende fortrolighet med programvaren, og de sier blant annet at *«Vi kommer helt sikkert til å gå litt videre for hver gang»*. Slik jeg ser det vil denne utviklingen fortsette helt til foreleserne opplever at programvaren for digital eksamen er fullstendig appropiert, og dermed integrert som en naturlig og usynlig del av deres arbeidshverdag.

Digital eksamen beskrives av studentene som bedre og mer ryddig, enn en skriftlig eksamen. De fremhever også at det er en fordel at besvarelsene kan sendes *«til sensor etterpå, så går det mye fortere, så de kan rette mye raskere»*, og at de som sensurerer *«slipper å tyde dårlig skrift»*. Dette er også foreleserne enige i. Det kom opp i alle fokusgruppene med studenter at mulighetene til å redigere svaret sitt uten at dette fører til at de må skrive alt på nytt, eller lignende, er en fordel; *«Hvis du skriver en side og så er det noe du ikke skulle hatt med på midten, så må du eventuelt stryke det eller føre inn på et nytt ark (på tradisjonell skriftlig, min anm.)»*. De opplevde det også som en fordel at de fikk bedre tid til å tenke på svarene sine; *«Du klarer gjerne å skrive litt fortere så man får litt bedre tid til å tenke på»*.

Studentene mener at tidsutnyttelsen på digital eksamen gjør at det kan være flere oppgaver på en slik eksamen, enn på en tradisjonell skriftlig eksamen; *«jeg tenker nivåmessig, hvis du skal ha samme vanskelighetsgrad, bør du kanskje legge til et par ekstra oppgaver på en digital i forhold til en skriftlig»*. En av fordelene med denne tidsutnyttelsen er i følge studentene at *«du kan teste litt bredere med en digital eksamen»*. På denne måten kan eksamenens validitet styrkes sammenlignet med tradisjonell skriftlig eksamen, siden den

effektive tidsutnyttelsen gir rom for flere oppgaver knyttet til emnets læringsutbytte og dermed et bredere grunnlag for vurdering.

Autentisk læring og vurdering fremheves av Biggs og Tang (2011) og Raaheim (2016) som viktig. Dette er også helt sentralt for en profesjonsutdanning som nautikk. I fokusgruppeintervjuene kommer det frem at studentene har et positivt inntrykk av gjennomføringen av digital eksamen, og de gir uttrykk for at de har opplevd Wiseflow som stabilt og trygt i forbindelse med avvikling av studiekraav og eksamen; *«jeg har inntrykk av at Wiseflow er ganske stabil...det virker i alle fall veldig sikkert»*. Det positive inntrykket er felles for alle studentene, og en av studentene oppsummerer det slik; *«Samlet så har jeg et veldig godt inntrykk av den digitale eksamen. Det gikk veldig greit»*. Studentene ser fordelene digitale verktøy gir, også for visualisering i undervisningen;

«Vi snakker for eksempel mye om fyrkarakteristikker, og i stedet for at vi pugger f.eks. Oc28s...får et bilde hvor du ser en fyrlykt som blinker, da må du faktisk kunne det, slik som du må kunne det når du kommer ut».

Vedkommende student ser at en slik form for visualisering ikke bare kan gi en autentisk læresituasjon, men også en mer autentisk vurdering, siden situasjonen under eksamen vil ligne mer på den virkeligheten kunnskapen skal brukes i når studentene er ferdig utdannet.

Hvis oppgavene ikke er relevante og autentiske, kan dette utfordre vurderingens validitet (jfr. Biggs og Tang, 2011). Foreleserne er ikke like godt fornøyde med hvordan regneoppgaver fungerer i Wiseflow, sammenlignet med andre oppgavetyper. De ser ikke for seg at dette er det best egnede verktøyet i regnetunge emner, men beskriver det som en treningssak å venne seg til et slikt verktøy; *«det krever trening. Men det er ikke mer trening som trengs, enn å bruke en kalkulator. Bruker du forskjellige kalkulatorer, må du øve. Tilsvarende må de øve på digitale verktøy»*. Studentene er også samstemte om at regneoppgaver er mer tungvint, enn andre oppgaver. Alle fokusgruppene kommenterte dette, og sa blant annet at *«Det ble litt knotete når man skulle inn i det egne tastaturet»*. De tror at arbeidet med regneoppgavene *«tok litt lenger tid, enn hvis du skulle regnet på ark»*. Alle studentene er ikke overbevist om at dette vil bli bedre med trening, men de fleste tror det. En student uttalte at *«alt er jo en treningssak»*. Vurderingens validitet vil slik jeg ser det kunne utfordres hvis oppgavene ikke er relevante og autentiske, og hvis de i større grad tester hvor godt studentene mestrer det digitale verktøyet rent teknologisk, heller enn emnets læringsutbytte.

4.1.1.2 Fokus på regelverket

En regelverksstyrt utdanning som nautikk stiller store krav til kvalitetssikring og dokumentasjon. Foreleserne gir uttrykk for at de er bevisste på ansvaret de har relatert til regelverket, og de har fokus på at «vi oppfyller det kravet vi har fra Sjøfart til å dokumentere at de faktisk kan det de skal...». Dette gjøres på flere måter, og «Noe er jo dokumentert i form av studiekraft, selvfølgelig, men at vi også da kan vise at vi har sjekket at studentene kan det de skal». Todelt eksamen i emnet sikrer at «vi enda tydeligere si at vi har testet ferdighetene deres, helt spesifikt ned på detaljnivå. Så det er veldig bra». På denne måten jobber foreleserne i det enkelte emnet aktivt for å sikre samsvar med det særskilte regelverket som gjelder for nautikkutdanningen.

4.2 Vurdering

Sammenhengen mellom læringsutbytte og emnets innhold, arbeids- og vurderingsformer er viktig for foreleserne, og de uttaler at «hele den pedagogiske biten, sammenhengen mellom læringsutbytte og det vi holder på med i emnene, det er kjempeviktig». De uttaler også at de til stadighet prøver å «jobbe med å bli bedre». I den sammenheng mener de at det er viktig å ha en klar plan bak det man gjør; «vi har på en måte en plan på det vi holder på med». Foreleserne fremhever det som viktig at det er ønskelig å teste så bredt som mulig innenfor emnet; «For det er jo viktig uansett eksamen at du får testet bredden også, at ikke du bare tar en stikkprøve». De anser vurderingsformen som helhetlig, og forteller at studentene også vurderes ut fra de praktiske ferdighetene sine gjennom den praktiske eksamenen. Dette illustrerer slik jeg tolker det en helhet i emnet som er basert på nær sammenheng mellom emnets læringsutbytte og de arbeids- og vurderingsformene som benyttes.

4.2.1 Flere valgmuligheter

Foreleserne har erfart at den digitale programvaren fører med seg flere valgmuligheter for hvilken oppgavetype som er best egnet til å teste et bestemt læringsutbytte. I praksis betyr det at de bruker mer tid på å vurdere mulighetene før de tar valget om hvilken oppgavetype de skal bruke for en spesifikk oppgave;

«all den tid du har flere muligheter i et digitalt system...så må du gjerne bruke litt mer tid på å tenke gjennom akkurat det. Jo flere muligheter du har, jo mer må du tenke gjennom hvordan det er mest hensiktsmessig å teste akkurat dette».

Teknologiutviklingen bidrar dermed til å endre forutsetninger for læring og utvikling, ved at samspillet mellom foreleserne og programvaren for digital eksamen påvirker hvordan for eksempel studiekravene gjennomføres. Denne utviklingen bidrar også til læring i seg selv og fører til at foreleserne har økt sin kunnskap og sine ferdigheter – de har lært nye ting på nye måter. Dette er en naturlig utvikling i prosessen for appropriering. Foreleserne uttaler at; *«så vil jo faktisk denne digitale eksamensformen få oss til å prøve studentene på nye måter»*, noe som viser at de er bevisste på at de i økende grad vil kunne benytte seg av flere funksjoner, etter hvert som de blir kjent med mulighetene som ligger i verktøyet. Utarbeidelse av eksamensoppgaver tar i følge foreleserne mye mer tid i starten, siden *«du bruker en del tid på å sette deg inn i det og finne ut hvordan det fungerer, og hvilken av malene som er den mest hensiktsmessige til denne oppgaven»*. Denne første gangen har foreleserne hatt størst fokus på en god og trygg gjennomføring for studentenes del, og for egen del; *«mer fokus på at nå må vi få gjennomført dette på en anstendig måte, på en best mulig måte, uten å ta absolutt alle metodene i bruk»*. Slik jeg tolker det har pilotgjennomføringene gitt foreleserne situert kunnskap, som de kan bygge videre på og utvikle fremover, uten at de har satt gjennomføring i fare med for mye eksperimentering.

Valg av oppgavetype ble i følge foreleserne gjort på bakgrunn av hva de ønsket å teste med den aktuelle oppgaven, og deres viktigste prioritet var *«det er det svaret jeg er på jakt etter...Og så ser jeg på hvilken oppgavetype jeg kan bruke for å få svar på dette»*. Dette er et eksempel på en målsetning om å prøve studentene i det spesifikke læringsutbyttet som er definert for emnet, noe som bidrar til en mer autentisk vurdering og øker vurderingens validitet (jfr. Biggs & Tang, 2011). At oppgaven skal sensureres er for foreleserne også en viktig faktor i valg av oppgavetype og utforming av spørsmål; *«i bakhodet ligger det hele tiden at – i alle fall på sånne tekstoppgaver og sånt – at det må være håndterlig å sensurere, og at studentene har mulighet til å vise at de skjønner dette»*. På denne måten har foreleserne fokus på at eksamensoppgavene skal gi studentene en reell mulighet til å vise hva de kan gjennom oppgavene de testes i.

Foreleserne har gjennom studietestene og eksamen fått erfaring med bruk av forskjellige oppgavetyper, og brukte blant annet *«ordinære tekstoppgaver der de kunne skrive inn svar, multiple choice, en del matematikkoppgaver, en sånn drag and drop-funksjon...»* på eksamen. Disse ble valgt for å fylle en bestemt hensikt, det vil si for å teste et konkret læringsutbytte, for eksempel; *«flervalgsoppgaver, for å få testet spesifikke kunnskaper»*.

I tillegg ble tekstoppgaver benyttet for i større grad å sjekke forståelse, gjennom for eksempel «et ark om radar med vanlige tekstoppgaver. Da prøvde jeg å lage oppgavene slik at man ser om de forstår sammenhengene. Når de svarer med tekst, ser du lettere om de har forståelse, kan resonnere og lignende». Oppgavene laget med ulik vanskelighetsgrad «prøver å lage litt forskjellig vanskelighetsgrad på tekstoppgavene», slik at de fikk testet det som i teorien beskrives som ulike kognitive nivåer. Foreleserne opplever det som viktig at «Det man skal lære bort må være tilpasset det man holder på med. I noen tilfeller vil tavleundervisning og i andre tilfeller mer fancy ting passe best. Verktøyet må tilpasses formålet». På denne måten fungerer oppgavene som artefakt på et sekundært nivå i den formen at spørsmålene gir instruksjoner om bruk av artefakten i den situerte konteksten. Studentene vil ikke kunne bruke artefakten uten oppgavene, siden disse beskriver hva studentene skal gjøre.

4.2.1.1 Utforming av oppgaver

Foreleserne opplever at «Jeg har hatt de samme tankene, altså dette med at det skal være enkelt å sensurere og at vi vil teste dem i det som jeg vil at de skal sitte igjen med. Sånt sett har det vel vært likt». Når oppgavene skal utarbeides, sier foreleserne at det viktig å tenke «dette temaet, hvordan kan jeg teste for eksempel kunnskapen til studentene?», siden det finnes mange oppgavetyper de kan velge. Spørsmålet er i praksis «hvordan er det mest hensiktsmessig å bruke dette for å teste akkurat dette?», noe som illustrerer forelesernes fokus på å utarbeide gode eksamensoppgaver som står i sammenheng med emnets læringsutbytte. Foreleserne er klare på at de må ta bevisste valg, og reelt utnytte mulighetene som ligger i programvaren. Det er ikke bare valg av oppgavetype og formulering av oppgaver som har stått i fokus hos foreleserne, de har også vært opptatt av at dette skal være oversiktlig, ryddig og greit for studentene;

«i tillegg til det som jeg nevnte nå med det pedagogiske, så har jeg tenkt mye på hvordan det blir ryddig for studentene, hvordan de lettest mulig har en oversikt, som er minst like god som når de får det på papir. Det tror jeg er viktig».

Dette illustrerer forelesernes ønske om at digital eksamen som medierende artefakt skal fungere på et ønsket nivå, slik at de tester studentenes kunnskaper, ferdigheter og generelle kompetanse i emnet, ikke hvor gode de er til å finne frem i eksamensoppgavene.

Hvordan oppgavene utformes og hva de krever av den som skal svare er avhengig av hvilke nivåer av kunnskap, ferdigheter og generell kompetanse som skal vurderes. I

refleksjonssamtalene kommer det frem at foreleserne har fokus på hvordan de kan bygge opp oppgavene, og hvordan de kan teste ulike former for og ulike nivåer av kunnskaper og ferdigheter;

«der er det jo veldig lett å gjøre, for da ligger disse «blokkene» hulter til bulter nederst, og så er det bare å ta tak i de og sette de på rett plass...og så kan jeg ha tekstopp-gaver i etterkant av en slik oppgave; «hvilken funksjon har triggeren?», for eksempel, og jeg kan stille enda vanskeligere spørsmål, for eksempel «når du justerer gainen på en radar, hvilken av modulene i blokkdiagrammet, osv.». Du kan utvikle en sånn oppgave».

Relatert til utvikling av oppgaver og progresjonen i vanskelighetsgrad, kommer det frem at foreleserne når de lager oppgavene *«har den taksonomiske trappen, og at vi kan se om de er helt blanke. At ikke vi bare har ett spørsmål innen et tema som gjerne er litt vanskelig, og så er det akkurat det de ikke kan».* Dette er en problemstilling de fremhever at gjelder for utarbeidelse av eksamensoppgaver generelt. På tross av fokus på problemstillingen, fremgår det at det være utfordrende å lykkes i praksis;

«det har kanskje litt med fagenes karakter også, at det ofte er mange tema i et fag, og da kan du jo ta en stikkprøve og gå dypt inn i det, men du må liksom ha med litt av hvert tema, og da blir det en god del spørsmål, så da er den taksonomiske trappen vi snakket om ikke alltid like enkelt å få til».

Foreleserne forteller at studentenes resultater i kullet *«fordelte seg sånt som de pleier».* Dette indikerer at foreleserne har klart å utarbeide eksamen på en slik måte at den digitale gjennomføringen tilsynelatende ikke har påvirket resultatene i noen særlig grad. På tross av at det ikke har vært nevneverdige resultatmessige avvik knyttet til den nye digitale eksamenen, påpeker studentene i den ene fokusgruppen at de opplever det som en svakhet med digital eksamen at det ikke ligger et tegneverktøy inne i programvaren. De fremhever at et slikt tegneverktøy kunne bidratt til å synliggjøre deres forståelse, for eksempel ved at de i enkelte tilfeller kunne tegnet en skisse for å illustrere besvarelsen sin. En av studentene sier at *«Det er jo en fordel med skriftlig i forhold til digital, at da kan man tegne og illustrere mye lettere»*, noe de andre i fokusgruppen er enige i. En av de andre utdyper at *«Og så gir det forklaring til det du skriver, selv om du ikke har ordene til å forklare det».* Studenter uttrykker seg på mange ulike måter, og behovet for et tegneverktøy er forståelig ut fra et slikt perspektiv.

Foreleserne legger vekt på kvalitetssikring av eksamensoppgavene, blant annet for å sikre «at vi ikke har feil i spørsmålsstillinger». De anser sammenheng mellom emnets innhold, arbeids- og vurderingsformer som svært viktig og er opptatt av tilbakemeldingene studiekravene gir underveis i semesteret. Foreleserne opplever at de gjennom digital eksamen «får testet det vi har prøvd å lære dem», noe som gir vurderingen god validitet (jfr. Biggs og Tang, 2011). Foreleserne har en helhetlig plan bak emnets oppbygging og gjennomføring, noe som understrekes ytterligere av at; «vi tester ikke alt læringsutbyttet på den digitale skriftlige eksamenen, fordi ferdighetene og en del av budskapet blir testet på den praktiske og i studiekravene». Et eksempel på dette er at «Vi tar ikke inn sjømerker og sjøveisregler på eksamen, fordi de har bestått studiekravene», noe som bidrar til å styrke sammenhengen i emnet gjennom en helhetlig utarbeidet plan for oppbygging og gjennomføring.

4.2.1.2 Rettferdig sensur

Foreleserne er opptatt av objektivitet i vurderingen, noe som kommer frem i flere sammenhenger. Målsetningen er at studentene skal få en objektiv og rettferdig sensur basert på hvordan vurderingskriteriene er oppfylt og læringsutbyttet som testes er dokumentert. Havnes og Raaheim (2006) understreker at kriteriebasert vurdering gir vurderingen økt reliabilitet, og for å oppnå en objektiv sensur, må sensuren ligge i bakhodet helt fra utarbeidelsen av oppgaver. Dette har foreleserne et bevisst forhold til; «altså når jeg lager en eksamensoppgave...jeg må si at jeg tenker mye på hvordan den er å sensurere...den skal være håndterlig å sensurere på et mest mulig objektivt vis».

I tilknytning til vurdering er objektiv sensur et sentralt punkt, og for å sikre dette kan ulike tiltak iverksettes. Et eksempel på slike tiltak er å rette eksamen spørsmål for spørsmål på tvers av kandidater, heller enn å rette hele eksamen for hver kandidat. Det kommer frem i refleksjonsmøtene med foreleserne at programvaren er lagt opp slik at de har opplevd utfordringer med å sensurere eksamen på en hensiktsmessig måte. De mener at det er «lettere hvis det hadde vært en funksjon der vi kunne velge å rette spørsmål for spørsmål, for da har du mer gangen i spørsmålet, hva forventninger vi har, fremfor å ta det fra gang til gang». Den ene foreleseren sier at ideelt sett så retter vedkommende «en og en deloppgave for alle kandidatene, så retter du neste deloppgave for alle, for å få det mest mulig rettferdig», noe begge foreleserne fremhever som standard praksis. Det kommer frem at dette var noe de lærte av mer erfarne forelesere når de selv begynte å jobbe på

høgskolen. På den digitale eksamenen opplevde foreleserne at det var *«mye mer tungvint å ta spørsmål en, på alle...så det kan være en svakhet...du må være veldig bevisst på hvor opplagt du er når du retter»*. I og med at dette var en svakhet som var kjent, hadde de satt i verk tiltak for å forebygge at dette skulle få innvirkning på objektiviteten i sensuren; *«Jeg var ikke så redd for det, nettopp fordi at jeg hadde laget spørsmålene litt kortere og alle (tekstoppgavene, min anm.) var vektet likt...Det gjorde at jeg var litt mer komfortabel når jeg rettet»*. Slik jeg ser det har begrensninger i programvaren ført til at foreleserne har måttet gå på akkord med sine egne prinsipper for hvordan de retter eksamen, men i og med at de visste om dette som et resultat av erfaringene fra pilottestene, har de gjennom utviklingen av eksamensoppgavene kunnet sette i verk tiltak for å sikre objektiviteten.

4.2.1.3 Standard for vurdering

I vurderingslitteratur er vurderingsstandard et sentralt element (jfr. Bloom, Hastings og Madaus, 1971; Havnes og Raaheim, 2006; Raaheim, 2016). Gjennom intervjuene kom det frem at dette er et viktig punkt for foreleserne, og de peker på at standarden for vurderingen er etablert tidnok og opp mot kriteriene for vurderingen (læringsutbyttet). I tillegg er regelverket² styrende i forhold til hva som er minimumsstandard for godkjenning. Sensorene må i forbindelse med utarbeiding av kriteriene for eksamen derfor gjøre vurderinger relatert til hvorvidt studentene har *«så store hull i helt sentrale områder som er regelverksstyrt at de faktisk stryker»*.

Spørsmålene på eksamen er i følge foreleserne utarbeidet for å teste læringsutbyttet i emnet. Dette fordrer at kriteriene for vurderingen er definert og at vektingen av spørsmål opp mot hverandre gir mening. Vekting er derfor noe foreleserne forteller at de har jobbet mye med. En funksjon de savner i den sammenheng er *«negative poeng. Nå tenker jeg ikke på vanlige tekstoppgaver, men på for eksempel et blokkdiagram, savner jeg den funksjonen, og jeg har snakket med Wiseflow...den virker ikke som den skal»*. Vekting av oppgavene er et sentralt punkt for foreleserne i forbindelse med utarbeiding av oppgavene, og dette kommer opp gjentatte ganger i alle intervjuene. De er tydelig opptatt av hvor viktig samsvar i omfang er relatert til uttelling;

«når du bruker forskjellige oppgavetyper, så blir det feil hvis du lager for eksempel en mcq med tre alternativer, og så får de 1 poeng hvis de svarer rett på

² Regelverket foreleserne viser til er STCW og FOR-2011-12-22 nr. 1523 *Forskrift om forskrift om kvalifikasjoner og sertifikater for sjøfolk* (International Maritime Organization, 2011; Sjøfartsdirektoratet, 2012)

det. Og så lager du gjerne noe sånt drag and drop...Så får de 1 poeng på det også, og så har du en tekstoppgave som er litt omfattende og så får de 1 poeng på det».

Gjennom summativ vurdering er målsetningen kvalitativt å vurdere hvor godt læringsutbyttet i et emne er oppnådd. Dette stiller store krav til utarbeidelse av oppgaver. En og samme oppgave kan teste flere ulike læringsutbytter, og samme læringsutbytte kan testes i flere ulike oppgaver. Gjennom refleksjonsmøtene med foreleserne kommer det frem at dette er noe de er bevisste på. Med tanke på spørsmålene de stilte på eksamen, er forelesernes tilbakemelding at disse *«jevnt over så ga de de svarene du forventet»*. De uttaler at de *«føler at vi får testet det vi har prøvd å lære dem, ja»*. Basert på samtalene med foreleserne er min oppfatning at de har et svært bevisst forhold til prosessen rundt utvikling av oppgaver, og at de søker å teste både ulike nivå av og ulike former for kunnskap. På denne måten har de et bredt grunnlag for å gjøre en god kvalitativ vurdering hver enkelt besvarelse.

Et bredt grunnlag for vurdering som nevnt over, får man ikke uten at man tar i bruk flere ulike typer oppgaver. I en av fokusgruppene fremhevet studentene at de syntes det var positivt med ulike oppgavetyper; *«Kjekt å ha litt av alt, egentlig, at ikke det bare er skriveoppgaver»*. De opplevde at oppgavene og oppgavetyperne hadde ulike formål, og at noen av oppgavene testet faktakunnskap, mens andre testet forståelse. En student sier at *«Det var sånne spørsmål der det bare er en måte, eller ett rett svar. Du kunne ikke snakke rundt det»* og en annen sier om noen av de andre oppgavene at når du skal svare *«Så du må liksom forstå det da»*. På denne måten får foreleserne testet studentene helhetlig ut fra et bredt grunnlag av oppgaver, noe som påvirker vurderingens validitet positivt.

I summativ vurdering er validitet viktig, siden dette er et mål på vurderingens gyldighet. For å oppnå god validitet må vurderingen blant annet være tett koblet til det faglige innholdet i emnet (jfr. Eggen, 2009). I tillegg må oppgavene studenten prøves i være relevante og autentiske (Raaheim, 2016). Studentene gir uttrykk for at de tror de kan bli testet i flere temaer innenfor et emne på digital eksamen, siden de skriver fortere. De opplever at det er lettere å rette hvis de har skrevet noe feil; *«det var enkelt å rette hvis du hadde gjort noe feil, du kom deg litt fortere gjennom spørsmålene. Det at du kunne gå tilbake igjen var også veldig greit»*, at de kommer raskere gjennom spørsmålene og at den digitale eksamenen gir en grei oversikt; *«Når du har mulighet til å gå inn i et sånn*

ikt-program, så får du mye bedre oversikt over det du driver med». Studentene fremhever det som positivt at de slipper tenke på skriften og dermed kan ha mer fokus på selve besvarelsen; *«du får på en måte tenkt gjennom og svart på spørsmålene, i stedet for å måtte jobbe med skriveingen i tillegg»* Dette fører til at studentene ikke føler seg sliten på samme måte som de gjør under en tradisjonell skriftlig eksamen og at det dermed er lettere å holde fokus;

«...spesielt det der med å skrive...det er mye lettere, du vil jo sørge for at du skriver pent nok til at det blir forstått, men det er ikke der når du skriver på pc. Så det er et par ting mindre å tenke på, og fokuset blir større på å besvare spørsmålene»

Det kan tolkes som at studentene opplever at digital eksamen som medierende artefakt fungerer på det ønskede nivået og tester studentenes læringsutbytte, i stedet for deres evne til å skrive pent nok til at sensorene forstår, hardt nok til at det går gjennom tre ark, og lignende. Dette indikerer at vurderingens validitet er god.

4.2.1.4 Sammenheng mellom læringsutbytte og eksamensoppgaver

Studentene i alle fokusgruppene er tydelige på at de opplever en klar sammenheng mellom emnet, pensum, studiekraft og eksamen, og sier blant annet at *«det trekker liksom en linje»* når de snakker om innleveringene og eksamensoppgavene. I den ene fokusgruppen kommer det frem at de *«Klarer ikke å se noe som konkret kan bli bedre for at sammenhengen skal bli bedre mellom undervisningen og eksamen»*. Studentene er omforente om at de ser sammenhengen mellom de to eksamenene i emnets helhet, og i den ene fokusgruppen ble dette spesifikt nevnt; *«så har vi en praktisk del også, for det er jo mye praktisk innenfor dette temaet. Så der får du vist andre ting igjen»*. Studentene omtaler digital skoleeksamen som en *«helt flott og balansert eksamen»* og at foreleserne da de laget eksamen *«tok med det viktigste»*. De uttrykker at de opplevde den digitale eksamenen som mer omfattende, enn en ordinær skriftlig eksamen kunne vært; *«Jeg syntes at vi fikk flere spørsmål enn hva vi ville fått på en skriftlig eksamen, så vi hadde et bredere spekter å svare på. Det er veldig bra, for da får vi dekt mer»*. Alle studentgruppene var enige om at sammenhengen mellom emnet, pensum og eksamen var tydelig og god, og en av studentene sa at *«spørsmålene var relevante til faget»*. De fremhever at de ble testet både dypt og overflatisk, og at det var flere oppgaver innenfor samme tema. En av studentene uttaler at *«det var litt av alt, og den hadde spørsmål som du gikk litt dypere enn andre på»*. En annen student uttrykker at vedkommende *«følte*

eksamen dekket ganske godt det som var forelest i og det som vi hadde lest på», mens en tredje student supplerer med at eksamen «dekket alt unntatt de to prøvene vi hadde». Dette var også formålet med de to testene. Flere studenter på tvers av fokusgruppene trekker frem at eksamen var «veldig godt oppbygd, rett og slett» og en student sier til og med at det var «den beste prøven jeg har hatt noen gang». Studentene trekker flere ganger frem sammenhengen mellom innleveringer og eksamen; «Det at eksamen har sammenheng med innleveringene er også veldig bra». Slik jeg tolker det viser studentenes tilbakemeldinger at foreleserne har oppnådd den ønskede sammenhengen mellom emnets innhold og de benyttede arbeids- og vurderingsformene, og at vurderingens validitet basert på dette er svært god.

4.3 Læring

Vurdering står i sammenheng med læresituasjonen i emnet. Det kommer tydelig frem under intervjuene at foreleserne er opptatt av studentenes læring, og at de aktivt benytter studentaktive arbeidsformer i emnet. De sier blant annet at «Uansett hva vi gjør og ikke gjør, må de bli aktive på et eller annet vis...at ikke de bare sitter og tar imot». Dagens digitale verktøy gjør det i følge foreleserne enklere med studentaktivitet, for eksempel gjennom digital quiz som er «en enkel måte å aktivisere studentene på, nå når alle har sånne smarttelefoner». Slike interaktive quizer utarbeidet spesifikt for undervisningen er i tråd med tanken om en sammenheng mellom læringsutbytte, arbeids- og vurderingsformer. Quizene gir i følge foreleserne både studenter og dem selv en umiddelbar tilbakemelding på hvordan henholdsvis de selv og studentene i klassen har forstått det som er gjennomgått. Slik tilbakemelding underveis i et emne er et eksempel på formativ vurdering, vurdering for læring, og gir studentene informasjon om hvordan de ligger an og hva de må jobbe mer med. I følge foreleserne er dette en av fordelene med digitale quizer; «da kan hver enkelt sitte og tenke på «hva var det?», og så får de umiddelbar tilbakemelding; Er det feil, er det rett?», og de ser at denne typen tilbakemelding underveis i et emne gir studentene nyttig informasjon om hvordan de ligger an og hva de må jobbe mer med. I den sammenhengen opplever foreleserne også at studentene er mer mottagelige for repetisjon i tidsrommet like etter at de har fått denne tilbakemeldingen;

«da er de varme liksom, hvis du ser at her har en god del svart feil, så kan du ta opp temaet...ikke bare si at det er rett eller feil, men...forklare hvorfor det er rett.

Da tror jeg at de er mer innstilt på å motta det du sier...og da tror jeg at de lærer mye».

På denne måten benytter foreleserne arbeidsformene i emnet til å rette studentenes fokus mot repetisjon av sentralt stoff fra undervisningen, samtidig som de sørger for at studentene får en tilbakemelding på hvorvidt de har forstått det som er gjennomgått.

Både studenter og forelesere opptatt av mulighetene til læring som ligger i bruk av programvaren Wiseflow både i tilknytning til og ut over gjennomføringen av digital eksamen. I de aller fleste intervjuene kom det frem at bruk av denne typen programvare på eksamen og i studiekrav gir muligheter for økt bruk av multimedia i undervisnings- og læresituasjonen. Dette kan for eksempel bidra til økt visualisering. En av foreleserne uttaler at *«spesielt for disse studentene som aldri har vært ute (til sjøs, min anm.), så er det viktig for innlæringen også. For forståelsen»*. Bruk av visuelle hjelpemidler bidrar til at læringen er situert (også når studenter og forelesere befinner seg i en simulert situasjon), og fører til mer autentisk opplæring og vurdering ved hjelp av medierende redskaper på flere nivåer.

Studentene anser multimedia som svært positivt i læresituasjonen, og de fremhever at det kan bidra til å belyse perspektiver og gi mer realistiske situasjoner;

«Jeg synes det høres interessant ut å bruke det i undervisningen, for da har du forståelsen for det. Det er enkelt nok å si at den båten skal gå der og den båten skal gå der i et kart...men når du står på broen og ser derfra...Da blir det fort noe annet».

I tilknytning til eksamen vil denne typen visualisering kunne bidra til at foreleserne kan lage oppgaver i det de beskriver som et hittil ubenyttet format. Det er for eksempel mulig å lage video av situasjoner på brosimulatorene som igjen kan brukes til oppgaver som tester studentenes forståelse for blant annet bruken av sjøveisreglene. Foreleserne ser muligheter for å bruke slike elementer i eksamensoppgaver; *«vi kan bruke mer visuelle ting i oppgavene...kan hjelpe oss og teste studentene på en bedre måte»*. Etter min vurdering ser foreleserne mulighet for å utnytte simulatorene, redskaper som allerede er innvevd i utdanningens kulturelle og historiske kontekst, på en hensiktsmessig måte i samspillet mellom individ og kollektiv, både i opplærings- og vurderingssituasjon.

Studentene på sin side skjønner at de har behov for ulike former for kunnskaper og ferdigheter, og har fokus på at disse skal være relevante for yrkeslivet de skal ut i etter endt utdanning. De er opptatt av at det de lærer på høghskolen skal danne grunnlag for «*en forståelse for hva du gjør når du kommer ut*». Denne grunnleggende forståelsen skal dekke «*alt av oppgaver, alt av jobber som du kan komme borti, slik at du har noe å bygge videre på*». Dette betyr at studentene ser at de trenger en forståelse for «*hvordan alt fungerer og ha en grunnleggende forståelse hvordan radaren fungerer, men du skal ikke kunne gå opp og skru på den, for eksempel*», slik at de «*skal forstå hva som skjer når du trykker på de forskjellige knappene*». Jeg tolker det slik at studentene er opptatt av en autentisk og helhetlig opplæring, ikke bare i emnet som denne studien omhandler, men også i nautikkutdanningen som helhet.

4.3.1 Lært av pilotene

Erfaringene som ble gjort i forbindelse med den første piloten har i følge foreleserne blitt innarbeidet i både digitale studiekravtester og digitale eksamensoppgaver. Også gjennomføringen av studiekravene ga mange gode svar, og foreleserne forteller at de har ervervet omfattende kunnskap gjennom pilottestingen. Denne kunnskapen kan forstås som situert kunnskap. Samtidig ser foreleserne at det å inkludere studiekravene som i denne omgang ikke ble gjennomført i Wiseflow, vil gi større rom for å teste ut nye oppgavetyper; «*Det tror jeg faktisk kommer mer når vi får inn oppgavene. Der er det mer varierte spørsmål*». Dette viser at foreleserne opplever økende mestring knyttet til Wiseflow, de approprierer artefakten. Videre læring og utvikling er viktig for foreleserne. De ønsker hele tiden å bli bedre til å lage eksamensoppgaver, og til å utnytte stadig flere av de tilgjengelige verktøyene som de finner hensiktsmessige «*For at vi ikke skal stå for mye i ro når vi skal gå videre*». Pilottestene bidrar til at læringen er situert, den er et samspill mellom foreleserne og testingen, mediert gjennom selve programvaren for digital eksamen.

Foreleserne gir uttrykk for at de har lært mye gjennom disse første erfaringene med digital eksamen og programvaren Wiseflow, og i et sosiokulturelt perspektiv kan de beskrives å være inne i en prosess kalt appropriering. Dette betyr at de videreutvikler sin kjennskap til og sin bruk av redskapet Wiseflow i takt med økende koordinering mellom redskapet og dem selv som brukere. Foreleserne sier at «*vi har jo begynt fra scratch...og egentlig synes jeg at vi har lært veldig mye på relativt kort tid*». De beskriver hvordan

læreprosessen har blitt påvirket av det de opplever som god brukerstøtte fra UNIwise; *«hvis de ikke hadde vært så kjappe og svare, så hadde det vært et irritasjonsmoment som gjerne hadde vippet meg over til at «nei, dette gidder jeg ikke, det er jo ikke noe...» det har veldig mye å si, responsen deres»*. På denne måten tolker jeg det som at læringen er situert – den er avhengig av rask tilbakemelding og god brukerstøtte fra programvarens leverandør der og da når behovet oppstår.

4.3.2 Behov for støtte i institusjonen

I sosiokulturell læringsteori ligger det til grunn at læring er et resultat av samspill mellom individ og kollektiv, for eksempel gjennom praksisfellesskap. I refleksjonsmøtene med foreleserne kommer det frem at de føler et behov for et slikt praksisfellesskap knyttet til digital eksamen. De forteller at de opplever mange kollegaer som skeptiske til bruk av digital eksamen på tross av at de ikke har testet programvaren; *«Jeg har oppfordret mange, men det er jo ingen som har gjort det, utenom oss»*. I prosessen rundt innføring av digital eksamen og knyttet til utarbeidelse av oppgaver til en digital eksamen, uttrykker foreleserne behov for støtte;

«det hadde vært veldig greit og hatt en slags superbruker i nærmiljøet, som du kunne spurt om både tekniske og andre spørsmål....det tror jeg er viktig, for vi er veldig forskjellige. Noen er veldig flinke med tekniske ting, og andre er flinke i det pedagogiske og noen er flinke i begge deler...så det å kunne overføre erfaring når sånne ting blir innført, det er svært viktig».

Jeg tolker dette som at foreleserne i realiteten opplever et behov for et praksisfellesskap hvor man kan lære av hverandres erfaringer i en utvidet og mer systematisert form, et behov som nok vil foreligge gjennom hele approprieringen og frem til en større gruppe av kollegiet i høgskolen er etablerte brukere av digital eksamen og dermed naturlig kan overta denne rollen gjennom et kollegialt praksisfellesskap (jfr. Säljö, 2006).

4.3.3 Eksamen påvirker læring

Innleveringsoppgavene ga i følge studentene god støtte før eksamen, og de uttrykker at det var tette koblinger mellom emnet som helhet og eksamen. På denne måten var alt gjennomgått på forhånd, og studentene fikk ingen ubehagelige overraskelser på eksamensdagen. En av studentene sier at *«Jeg synes det var veldig relatert til pensum og det var veldig vidt spekter på eksamen, så vi måtte være forberedt på så å si alt»*. Alle fokusgruppene er enige om at de hadde *«bruk for det som vi har hatt i disse*

innleveringsoppgavene». Dette tolker jeg som at studentene opplever eksamen som autentisk, og at det er god sammenheng mellom arbeidsoppgavene i emnet og eksamen.

Vurderingen påvirker studentenes læring, siden studenter innretter studiene sine mot den kommende vurderingen i emnet (Entwistle & Entwistle, 1992). Dette kan utnyttes positivt når det er en sterk sammenheng mellom emnets læringsutbytte, arbeidsformene i emnet og vurderingen, såfremt koblingen mellom arbeidet og vurderingen tidlig er synliggjort (Biggs, 1996). Foreleserne forteller at de opplever at innføringen av en praktisk eksamen som tester studentenes ferdigheter bidrar til at studentene tar inn det de lærer på simulatorøvelsene på en annen måte, enn tidligere; *«Det at de visste at de fikk en eksamen i det, det tror jeg påvirket at de la mer arbeid ned»*. Dette illustrerer hvordan innføringen av den praktiske eksamenen tilsynelatende har påvirket studentenes innsats og læring knyttet til gjennomføringen av de praktiske simulatorøvelsene. I refleksjonsmøtene kommer det også frem at foreleserne kobler læringsutbyttet og eksamensoppgavene sammen blant annet gjennom de *«skriftlige obligatoriske øvelsene som de har i løpet av semesteret. Det er essensen de får oppgaver i...det som jeg vil at de skal sitte igjen med»*. De obligatoriske øvingene krever mer av studentene, enn eksamen gjør, siden de har tilgang til alle hjelpemidler og mye tid til rådighet. Dette fører til at de obligatoriske øvingene inneholder spørsmål; *«som er litt på siden, fordi jeg gjerne vil at de skal søke informasjon andre steder, lese læreboken litt grundigere, eller gå ut og søke informasjon»*, selv om hovedtyngden i de obligatoriske øvingene ligger på *«kjernestoff, det jeg vil at de skal sitte igjen med, egentlig»*. Foreleser er også tydelig på *«at de må jobbe godt med de obligatoriske øvingene, siden hoveddelen av eksamen blir plukket fra disse»*. På denne måten synliggjør foreleserne at de obligatoriske øvingene er sentrale i forberedelsene til eksamen, og dette bidrar til å gi studentenes læring i emnet retning.

Studentene selv er klare på at emnet inneholder mange ulike tema, og har et variert innhold. De opplevde allikevel at *«eksamen tok spørsmålene fra alle delene»*. I tillegg fremhever studentene at det var første gang dette emnet gikk, og at dette ga usikkerhet med tanke på hva som kom til å bli gitt på eksamen; *«det var kanskje litt vanskelig å tenke hva man skulle få til eksamen, siden vi ikke har hatt noen tidligere eksamener å se på»*, men de opplevde at foreleserne bidro til å dempe denne usikkerheten ved å være tydelige på hva som er viktig i emnet; *«Han har jo sagt noen ganger at «dette må dere kunne»»*. Studentene fortalte også at arbeidet med innleveringene ga dem muligheten til å jobbe

jevnt med emnet gjennom hele semesteret, og uttrykte at en av fordelene med gruppeinnleveringene var; «*Når du jobbet i grupper så fikk du en diskusjon på de forskjellige områdene, og det hjalp når du kom til eksamen*». På denne måten har sammenhengen i emnet og arbeidsformene som foreleserne har lagt opp til vært sentralt i forsøket på å oppnå en hensiktsmessig læringseffekt.

4.3.4 Videre utvikling

Studentene ser muligheter for videre utvikling og bruk av andre typer oppgaver og omtaler bruk av multimedia i undervisnings- og vurderingssammenheng som «*spennende*» og «*positivt*». De fremhever likevel at det er en fordel å ha øvd på slike oppgaver i undervisningen og på innleveringer før det eventuelt gis på eksamen. Slik jeg tolker det, er dette et eksempel på at det ikke bare er foreleserne som gjennomgår en appropriering knyttet til digital eksamen, men viser at studentene også må gjennom en tilsvarende prosess. Studentene mener at multimedia kan bidra til mer realistisk opplæring og vurdering. En student uttrykker det slik;

«Hvis du viser fra simulatoren, og liksom står på broen og det kommer en båt fra styrbord, og du ser det på filmen, så spør de deg hva du skal gjøre, eller hvem som har vikeplikt...da handler det jo om at vi faktisk skal stå på en bro på en båt og vurdere situasjonen».

Dette viser at studentene er opptatt av hvordan digitale verktøy i vurderingssituasjonen kan bidra til en autentisk vurdering, noe som bør være målsetningen i enhver vurdering (Biggs & Tang, 2011). Flere studenter uttaler at det kunne vært greit å ha slike oppgaver på eksamen, «*hvis de hadde laget en quiz med mange forskjellige situasjoner med sånne videosnutter og du hadde blitt litt kjent med det*». En student mener at «*det er jo akkurat som spørsmål du ikke kjenner*». Enkelte andre studenter er litt mer skeptiske, og en av disse studentene uttalte at; «*i en eksamenssituasjon så er du jo litt stresset, og hvis det kommer noe du ikke har sett før...*». Slik jeg tolker det, befinner studentene seg på ulike stadier i prosessen for appropriering av multimedia i læring og vurdering, og det kan virke som at noen av studentene befinner seg på et nivå hvor de opplever at de mestrer verktøyet, mens andre fortsatt er på et lavere nivå og dermed er noe mer usikre.

5. Drøfting

De empiriske funnene fra studien ble presentert i kapittel 4 Presentasjon av funn. I denne delen drøfter jeg disse funnene i lys av de teoretiske perspektivene presentert i kapittel 2 Teoretiske perspektiver, for på denne måten å kunne besvare problemstillingen: «Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkstudiet tilpasses en digital kontekst?». Sentralt i diskusjonen står de tre overordnede kategoriene som kom til uttrykk gjennom dataanalysen; *digital kontekst*, *vurdering* og *læring*. Disse tre begrepene bidrar til å strukturere diskusjonen presentert under.

5.1 Digital kontekst

Store deler av samtalene i fokusgruppene med studenter og refleksjonssamtalene med forelesere omhandlet i en eller annen form den digitale konteksten rundt digital eksamen. Denne konteksten er vanskelig å komme utenom, i og med at all læring er situert, og at læringen derfor påvirkes av konteksten (Dysthe, 1999).

5.1.1 Digital eksamen som medierende artefakt

Medierende artefakter er i følge Säljö (2001; 2006) en integrert del av kulturen vår og på denne måten innvevd i våre handlinger, idéer og holdninger. Samme artefakt kan forstås og benyttes på ulike måter, og en allerede eksisterende artefakt kan tilpasses nye oppgaver. Dette er en kontinuerlig prosess, og for å kunne utnytte artefaktene til fulle, må vi mennesker i følge Säljö (2001; 2006) appropriere sosiokulturelle kunnskaper, ferdigheter og erfaringer. Funn fra refleksjonssamtaler og fokusgruppeintervjuer viser at forelesere og studenter gjennom innføringen av digital eksamen tilpasser en eksisterende artefakt til nye oppgaver. Datamaskinen som medierende artefakt er integrert i kulturen vår på mange ulike vis, blant annet til å lese aviser, skrive dokumenter og lage presentasjoner. Det er i dag mange situasjoner hvor datamaskinen benyttes med den største selvfølgelighet. Gjennom innføringen av digital eksamen og bruk av ny programvare blir datamaskinen benyttet og forstått på en helt ny måte for både forelesere og studenter. De må derfor utvikle en stadig økende koordinering mellom seg selv og datamaskinen som redskap for digital eksamen, slik at digital eksamen approprieres.

5.1.2 Appropriering

Säljö (2006) peker på at en appropriering av et nytt redskap er en prosess med fire faser, hvor man i den innledende fasen prøver ut verktøy og egenskaper. I denne fasen er man fortsatt avhengig av ekstern støtte, og gjennom refleksjonssamtalene kom det frem at foreleserne har gått gjennom denne fasen. De viser flere ganger til at de har testet ulike oppgavetyper, men at det fortsatt er mye de ikke har prøvd ut. Dette kom spesielt tydelig frem i den første refleksjonssamtalen. Behovet for ekstern støtte har også vært en faktor i prosessen fra piloter til eksamen, og det fremgår av refleksjonssamtalene at kontakten med brukerstøtten fra UNIwise har vært svært viktig for den første fasen av approprieringen. I tillegg har god brukerstøtte og raske tilbakemeldinger spilt en stor rolle for forelesernes holdninger til programvaren, og de hevder at dette er en medvirkende årsak til at de fortsatt ønsker å bruke programvaren Wiseflow på tross av enkelte feil og mangler.

I følge Säljö (2006) er neste fase i approprieringen uttesting og bruk av redskapet, noe som i dette tilfellet er perioden hvor programvaren for digital eksamen har blitt benyttet på studiekraft og forberedelse til eksamen. Også i denne fasen utforskes bruken av redskapet i den konkrete sammenhengen, og muligheter og begrensninger oppdages. Foreleserne og studentene får økende innsikt i bruken av programvaren. Det fremgår av refleksjonssamtalene at ikke alle funksjonene i programvaren har blitt testet i denne første gjennomføringen, men foreleserne har fått en viss oversikt over tilgjengelige funksjoner med tilhørende fortrinn og begrensninger. Siden brukerne i store deler av semesteret befant seg på et lavere nivå i approprieringen, hadde de fokus på en trygg gjennomføring av eksamen, uten å ta inn for mange nye elementer. I denne fasen av approprieringen har fordeler med programvaren blitt avdekket, og både forelesere og studenter peker på at tidsutnyttelsen ved digital eksamen er bedre, enn ved en tradisjonell skriftlig eksamen. Dette gjør at digital eksamen kan inneholde flere oppgaver, og på denne måten kan dermed læringsutbyttet testes bredere. Verken forelesere eller studenter ga uttrykk for at de ser veldig mange ulemper med digital eksamen, bortsett fra i regnetunge emner, hvor begge grupper opplever programvaren som noe tungvint.

Gjennom den tredje fasen i approprieringen oppnår forelesere og studenter mestring innenfor bestemte praksiser, og bruken av digital eksamen begynner å bli normalisert (Säljö, 2006). I denne fasen kan brukerne forklare bruk av programvaren til en nybegynner, selv om de fortsatt ikke kjenner til alle funksjoner. Basert på funn fra studien

tyder det på at det per i dag er på dette stadiet foreleserne befinner seg. De viser at de kan forklare hvordan programvaren brukes, men fremhever samtidig at de tror det er mange funksjoner som de ennå ikke er kjent med. Dette medfører at de etter kun ett semester med bruk av Wiseflow ennå ikke har kommet til den siste fasen i approprieringen, fasen hvor man betrakter Wiseflow som den naturlige måten å gjennomføre eksamen på. I denne fasen har programvaren blitt en «usynlig» del av arbeidshverdagen. Foreleserne er foreløpig ikke fullt ut koordinerte med digital eksamen. Programvaren er dermed fortsatt synlig for dem, selv om de stadig oppnår økende koordinering mellom seg selv som bruker og Wiseflow som medierende redskap (Säljö, 2006).

5.1.2.1 Artefakter på ulike nivå

I lys av Säljö (2006) og Wartofskys (1979) beskrivelser av artefakter på ulike nivåer, indikerer funn fra studien at digital eksamen som medierende artefakt i dette tilfellet har en funksjon på alle tre nivå; primært, sekundært og tertiært. Det primære nivået er selve datamaskinen som eksamen gjennomføres på, det konkrete verktøyet som benyttes. Spørsmålene som stilles fungerer som instruksjon for bruk av artefakten i den situerte konteksten på et sekundært nivå. Sammenhengen mellom emnets læringsutbytte og spørsmålene som studentene må besvare under eksamen utgjør det tertiære nivået.

Gjennom refleksjonssamtaler og fokusgrupper kommer det frem at både forelesere og studenter anser det nødvendig at digital eksamen fungerer i henhold til intensjonen på det tertiære nivået – som en test av studentenes kunnskaper, ferdigheter og generelle kompetanse relatert til det aktuelle emnet. En fare med digital eksamen er at det kan bli en test på det som Säljö (2006) og Wartofsky (1979) beskriver som primært nivå, slik at det i realiteten er studentenes teknologiske ferdigheter knyttet til selve verktøyet som testes. Dette har foreleserne vært bevisste på, og de har derfor benyttet Wiseflow på studiekraft i løpet av semesteret, slik at studentene er kjent med programvaren når de kommer til eksamen. På denne måten sørger foreleserne for å sikre at eksamen tester studentene på det tertiære nivået, den tester læringsutbyttet.

Behovet for familiarisering med oppgavetyper som benyttes oppleves ulikt av forelesere og studenter, og det kan ut fra dette tolkes dithen at studentene opplever å befinne seg på et høyere nivå av appropriering, enn hva foreleserne ser for seg at studentene gjør (Säljö, 2006). Også mellom studentene er det forskjeller i hvordan de oppfatter behovet for familiarisering med oppgaver i forkant av eksamen. Det er allikevel ingenting som tilsier

at studentene som er opptatt av fordelene med familiariseringen befinner seg på et lavere nivå, enn approprieringens nest siste trinn. De har dermed i stor grad appropriert den digitale artefakten. Resten av studentene ser ut til å ha fullstendig koordinasjon med redskapet, og gir uttrykk for at de anser dette som en naturlig del av eksamensavviklingen. Et flertall av studentene påpeker for eksempel at digital eksamen er deres hjemmebane, at pc er høgskolestudenters eget medium, og at de var overrasket over at digital eksamen ikke var standard på høgskolen. I motsetning til foreleserne gir studentene uttrykk for at de finner digital eksamen mer ryddig og oversiktlig, enn tradisjonell skriftlig eksamen. Studentene i samtlige fokusgrupper fremhever at det var tungvint med regneoppgaver på eksamen, en oppgavetype de ikke var kjent med i forkant av eksamen. Allikevel gir de uttrykk for at det gikk greit, og det er det ingen av dem som på noe tidspunkt nevnte at de ikke syntes dette var håndterbart, selv om de ikke hadde benyttet oppgavetypen i forkant. Det er derfor sannsynlig at studentene er i stand til å benytte artefakten på et tertiært nivå (Säljö, 2006; Wartofsky, 1979) både med og uten familiarisering i forkant av eksamen.

5.2 Vurdering

Summativ vurdering, vurdering *av* læring, har som formål å teste hvor godt studentene har oppnådd emnets definerte læringsutbytte (Bloom, Hastings, & Madaus, 1971; Havnes & Raaheim, 2006; Raaheim, 2016). En summativ vurdering skal dekke to formål; dokumentasjon og karaktersetting. For å kunne oppfylle de to formålene, må en kvalitativ vurdering gjennomføres basert på de forhåndsbestemte kriteriene som studentene skal vurderes opp mot (Biggs & Tang, 2011; Raaheim, 2016). Refleksjonssamtaler og fokusgruppeintervjuer viser at både studenter og forelesere uttrykker tilfredshet med den gjennomførte vurderingen, og i følge foreleserne var det ingen nevneverdige resultatmessige avvik sammenlignet med tilsvarende emner tidligere år.

5.2.1 Meningsskapende helhet

Helhet i et emne – constructive alignment – er i følge Biggs og Tang (2011) oppnådd når emnets læringsaktiviteter, vurdering og læringsutbytte systematisk står i sammenheng, en tankegang som har røtter tilbake til Bloom, Hastings og Madaus. Allerede tidlig på 70-tallet fremhevet de viktigheten av samsvar mellom undervisning og vurdering knyttet til læringsutbytte (Bloom, Hastings, & Madaus, 1971). Raaheim (2016) har på norsk oversatt constructive alignment til meningsskapende helhet, og jeg kommer i fortsettelsen

til å benytte det norske begrepet. Det skal allikevel ikke være tvil om at det er Biggs og Tangs (2011) prinsipper for constructive alignment som ligger til grunn når meningsskapende helhet omtales i den videre teksten.

For å oppnå meningsskapende helhet, er det nødvendig med helhetlig fokus helt fra emnet planlegges og til vurderingen er fullført og emnet avsluttet (Biggs & Tang, 2011). Det fremgår av refleksjonssamtalene at foreleserne i forbindelse med utvikling av nye emner tok initiativ til innføring av digital eksamen, siden dette var den vurderingsformen som de i kombinasjon med praktisk eksamen fant mest hensiktsmessig for deres emne. Det kommer også frem i studien at deres valg av digital eksamen ikke består i å sette strøm på papir, men faktisk å bruke funksjonene som ligger i programvaren til å teste studentene på nye og bedre måter. Forelesernes beskrivelser viser at de har hatt fokus på de pedagogiske hensiktene med bruken av digital eksamen, noe Newhouse (2012) peker på at kan støtte og berike studentenes læring. Ved innføring av digital eksamen har foreleserne valgt det Raaheim (2016) beskriver som en høyterskelrute til digitalisering. Dette går ut på at de benytter mulighetene teknologien fører med seg til å gjøre endringer i undervisningsopplegget, i stedet for å trekke det digitale verktøyet inn i eksisterende strukturer og etablerte arbeidsformer. Foreleserne har tenkt nytt om undervisning og vurdering. De beveger seg i retning av det Redecker og Johannessen (2013) kaller det 21. århundrets evalueringsstrategier, og som bidrar til at teknologiske fremskritt støtter pedagogisk utvikling. Gjennom innføringen av en helhet som omfatter ulike arbeidsformer og ulike typer vurdering, viser funn fra studien at foreleserne også benytter formative elementer. Dette er i samsvar med det Fosslund (2015) sier om at digital teknologi sannsynligvis kan muliggjøre en mer integrert prosess hvor summativ vurdering står i sammenheng med formativ vurdering i et helhetlig læringsforløp.

I en meningsskapende helhet bidrar teknologi i tilknytning til vurdering og læring. Biggs og Tang (2011) og Raaheim (2016) viser hvordan teknologien kan bidra til utvikling i undervisningssituasjoner, for eksempel gjennom studentaktive læringsformer. Funn fra studien viser at foreleserne blant annet benytter digitale quizer. Disse quizene er utarbeidet spesifikt for undervisningen og gir studenter og forelesere en umiddelbar tilbakemelding om hvorvidt studentene har forstått det som er gjennomgått i undervisningen. Tilbakemelding underveis i et emne er et eksempel på formativ vurdering, vurdering *for* læring, og gir studentene informasjon om hvordan de ligger an og hva de må jobbe mer med (Bloom, Hastings, & Madaus, 1971; Havnes & Raaheim,

2006; Raaheim, 2016). Inntrykket som sitter igjen etter intervjuene med foreleserne er at det er lite som er overlatt til tilfeldighetene i planleggingen av dette emnet. Alt har en hensikt, både faglig og knyttet til hvordan foreleserne ønsker at studentene skal jobbe med de ulike arbeidsoppgavene. Dette gir en meningsskapende helhet og funn fra fokusgruppeintervjuene med studenter bekrefter at foreleserne i svært stor grad har lyktes med å etablere den ønskede sammenhengen mellom emnets innhold og de benyttede arbeids- og vurderingsformene.

Som en del av den meningsskapende helheten utviklet Biggs og Collis SOLO-taksonomien. SOLO (Structure of the Observed Learning Outcome) er en generell modell for læringskvalitet, og kan tilpasses et bestemt faglig innhold. Gjennom taksonomien får man et rammeverk for å beskrive hvordan en students mestring blir mer kompleks etter hvert som studenten lærer og mestrer mer (Biggs & Tang, 2011). For å kunne teste kunnskaper, ferdigheter og generell kompetanse på ulike nivåer i henhold til en taksonomi som for eksempel SOLO er det nødvendig med tette koblinger mellom spørsmålene på eksamen og læringsutbyttet som skal testes. Gjennom refleksjonssamtalene kom det frem at foreleserne målrettet jobber ut fra hva de ønsker å teste med de ulike oppgavene når de utarbeider eksamensoppgaver. De er også opptatt av progresjonen i oppgavene, og at ingen oppgaver skal føre til «blindveier». På denne måten unngår de at studentene ikke kommer videre, hvis de har feilet på en deloppgave. Raaheim (2016) beskriver hvordan taksonomier som for eksempel SOLO-taksonomien former studenters forståelse av kunnskap, slik at studentene oppfatter at multiple choice-oppgaver tester lavere kognitive funksjoner, mens tekstopp-gaver tester høyere kognitive funksjoner som forståelse og refleksjonsevne. Funn fra fokusgruppeintervjuene med studenter viser at dette også stemmer for denne studentgruppen. Det kommer frem gjennom intervjuene at studentenes oppfatning er at ulike typer oppgaver benyttes for å teste ulike kognitive nivåer av kunnskap. Dette viser at foreleserne ved utarbeidelse av oppgavene har vært bevisste på hva de vil at studentene skal oppnå, i tråd med de grunnleggende tankene bak enhver taksonomi (Biggs & Tang, 2011; Raaheim, 2016).

Også gjennom vurdering som del av en meningsskapende helhet er det behov for å sjekke at det definerte læringsutbyttet er oppnådd. Biggs og Tang (2011), Fossland (2015) og Prøitz (2015) fremhever i denne sammenheng kriteriebasert vurdering som et godt egnet redskap. I kriteriebasert vurdering gjennomføres vurderingen i henhold til en oppsatt standard, og studentene vurderes ut fra forhåndsdefinerte vurderingskriterier.

Kriteriebasert vurdering tar ikke høyde for hvordan studentene gjør det i forhold til hverandre, slik som normbasert vurdering gjør, men forholder seg til angitte vurderingskriterier. Disse kriteriene er kvalitative og forhåndsbestemte, slik at det er tydelig for alle parter hvilken standard studentene vurderes opp mot (Biggs & Tang, 2011; Fossland, 2015; Prøitz, 2015). Dataene viser at det for foreleserne er viktig å etablere standarden for vurderingen som skal gjennomføres, ikke minst fordi de opplever et ekstra ansvar knyttet til regelverket i STCW og FOR-2011-12-22 nr. 1523 *Forskrift om forskrift om kvalifikasjoner og sertifikater for sjøfolk* (International Maritime Organization, 2011; Sjøfartsdirektoratet, 2012). I refleksjonssamtalene med foreleserne kommer det frem at de har jobbet mye med problemstillinger knyttet til vekting, noe som er mer aktuelt i forbindelse med en digital eksamen, enn på en tradisjonell skriftlig eksamen. På digital eksamen kan mange ulike oppgavetyper benyttes, og det vil ikke være samsvar mellom omfanget på oppgavene hvis for eksempel multiple choice-oppgaver med fem svaralternativer og dermed 20% sjanse for at man kan gjette seg til riktig svar gir samme uttelling som en tekstoppgave som krever både kunnskap og forståelse. For at vurderingskriteriene skal være gode, må vektingen av spørsmål opp mot hverandre også være passende (Biggs & Tang, 2011; Bloom, Hastings, & Madaus, 1971; Havnes & Raaheim, 2006; Raaheim, 2016). Det kom frem under samtalene at vekting er en svakhet med programvaren, og at foreleserne håper på utbedringer knyttet til dette. Gjennom refleksjonsmøtene kom det frem at foreleserne er bevisste på viktigheten av samsvar mellom oppgavens omfang og uttellingen den gir. For å sørge for passende uttelling for ulike oppgavetyper satte de derfor i verk tiltak. På denne måten oppnådde de på tross av begrensninger i programvaren en pålitelig og god kriteriebasert vurdering.

5.2.2 Reliabilitet og validitet i vurderingen

Reliabilitet og validitet i vurdering beskriver henholdsvis hvorvidt resultatene er pålitelige og om de måler det de skal måle (Biggs & Tang, 2011; Bloom, Hastings, & Madaus, 1971; Eggen, 2009; Raaheim, 2016).

5.2.2.1 Vurderingens reliabilitet

For at en vurdering skal ha god reliabilitet, må vurderingene som gjøres være stabile, og de må stå i forhold til de definerte kriteriene (Eggen, 2009; Havnes & Raaheim, 2006). Funn fra studien viser at foreleserne har fokus på å sikre god reliabilitet i vurderingen gjennom å sørge for tydelig definerte kriterier for den aktuelle vurderingen. I tillegg er de

opptatt av objektivitet i vurderingen, for på denne måten å sikre en stabil, konsekvent og rettferdig vurderingsstandard. I refleksjonssamtalene kom det frem at foreleserne vanligvis retter eksamensbesvarelser deloppgave for deloppgave, og at de stikker om på bunken mellom hvert spørsmål, slik at de ikke retter de samme besvarelsene etter hverandre i de ulike deloppgavene. Dette er i tråd med blant annet Biggs og Tang (2011) sine anbefalinger for hvordan man kan oppnå en mest mulig rettferdig sensur. Foreleserne gir uttrykk for at programvaren i dette tilfellet har gjort det vanskelig for dem å sensurere deloppgave for deloppgave, og at de derfor har måttet forebygge bias i sensuren på annen måte. Siden dette var kjent etter pilottesting, har foreleserne blant annet hatt samme omfang og lik vektning på alle tekstoppavene på eksamen. Foreleserne har dermed prøvd å forebygge partiskhet i sensuren gjennom oppgavens oppbygning, i tillegg til at de har benyttet ekstern sensor.

Havnes og Raaheim (2006) understreker at kriteriebasert vurdering gir vurderingen økt reliabilitet, og for å oppnå en objektiv sensur, må sensuren ha en sentral rolle allerede under utarbeidelse av oppgaver. Funn fra studien viser at den fremtidige sensuren ligger i bakhodet til foreleserne helt fra de begynner å utarbeide oppgaver. De ønsker ikke bare å sikre at oppgavene er mulige å sensurere på en objektiv og god måte, men at oppgavene også gir studentene en reell mulighet til å demonstrere hva de kan.

Kvalitetssikring av oppgaver i forkant av eksamen bør i følge Biggs og Tang (2011) gjennomføres for å unngå at oppgavene er tvetydige eller uklare. Studien viser at foreleserne gjennom en grundig kvalitetssikring av oppgaver sørger for at oppgavene fungerer i tråd med intensjonen før studentene skal gjennomføre eksamen. Denne kvalitetssikringen gjennomføres i samarbeid med ekstern sensor eller kollegaer.

5.2.2.2 Vurderingens validitet

Validitet er et spørsmål om en vurderings gyldighet. Biggs og Tang (2011) beskriver validitet som samsvar mellom vurderingen og det definerte læringsutbyttet. I følge Eggen (2009) kan validitet i undervisningssammenheng deles inn i innholdsvaliditet, ytre validitet og konsekvensvaliditet. Ytre validitet og konsekvensvaliditet handler om konsekvens av vurderingen, og beskrives av hvorvidt vurderingene som gjøres er gyldige uavhengig av sammenheng (Eggen, 2009). Dette kan knyttes til Regelverket i STCW (International Maritime Organization, 2011) og FOR-2011-12-22 nr. 1523 *Forskrift om kvalifikasjoner og sertifikater for sjøfolk* (Sjøfartsdirektoratet, 2012), som er styrende for

undervisning og vurdering i nautikkutdanningen. Vurderinger gjort opp mot regelverket må være konsekvente på tvers av studenter og kull. Det fremgår av intervjuene at foreleserne er bevisste på regelverkskravene og har fokus på hvordan de ivaretar disse.

Innholdsvaliditet er i følge Eggen (2009) knyttet til kriteriene for vurdering, og er på denne måten tett koblet til vurderingens reliabilitet. Samtlige fokusgruppeintervjuer viser at studentene opplever en klar og helhetlig sammenheng mellom emnet, pensum, studiekraav og eksamen. Det fremgår i flere av fokusgruppene at studentene ikke kan se hvordan sammenhengen mellom eksamensoppgavene og emnets innhold kunne vært bedre. Rett nok er det førsteårsstudenter som har begrenset erfaring med eksamen på høgskolenivå som intervjues, men alle har i lavere skoleslag gjennomgått en rekke eksamener. Funn fra studien viser at studentene er samstemte i at de ser en sammenheng mellom emnets to eksamensformer, og at de anser praktisk eksamen som en del av en helhetlig test på emnets læringsutbytte. Validiteten i den digitale eksamen er dermed god i henhold til Biggs og Tang (2011), Bloom, Hastings og Madaus (1971), Eggen (2009) og Raaheims (2016) definisjoner.

Valg av oppgavetype er en viktig faktor knyttet til vurderingens validitet, og Raaheim (2016) fremhever at det ikke er vurderingsformen i seg selv som er avgjørende for validiteten, men typen oppgaver som benyttes. Gjennom refleksjonsmøtene kom det frem at valg av oppgavetype ble gjort på bakgrunn av formålet med den aktuelle oppgaven, det vil si hvilket læringsutbytte oppgaven skulle teste. Biggs og Tang (2011) og Raaheim (2016) fremhever viktigheten av at studentene prøves i og vurderes ut fra læringsutbyttet. Dette bidrar blant annet til autentisk vurdering og øker vurderingens validitet. Vurderingens validitet vil derfor kunne utfordres, hvis oppgavene ikke er relevante og autentiske (Biggs & Tang, 2011; Eggen, 2009; Raaheim, 2016). Det kommer frem i undersøkelsen at foreleserne frykter at dette kan skje under eksamen, hvis studentene ikke har kjennskap til oppgavetyperne de benytter. Foreleserne gir gjennom refleksjonssamtalene uttrykk for at de ved hjelp av digital eksamen får testet læringsutbyttet studentene skal ha oppnådd (Biggs & Tang, 2011; Eggen, 2009; Raaheim, 2016). En medvirkende årsak til dette er at studentene gjennom studiekraavene har blitt kjent med Wiseflow, og dermed mestrer verktøyet rent teknisk når de benytter det på eksamen.

5.3 Læring

Sosiokulturell læring er situert og et resultat av samspillet mellom individ og kollektiv (Dysthe, 1999; Säljö, 2001). Dette betyr at læring i et sosiokulturelt perspektiv foregår i bestemte situasjoner og primært skjer i praksisfellesskap (Dysthe, 1999; Dysthe, 2001; Engeström, 2016; Imsen, 2005; Lave & Wenger, 1991; Säljö, 2001; Vygotsky, 1978). Funn fra studien viser at foreleserne i prosessen med innføring av digital eksamen har opplevd et udekket behov for støtte fra institusjonens side. De har hatt – og har fremdeles – behov for et praksisfellesskap knyttet til digital eksamen.

5.3.1 Praksisfellesskap

Da foreleserne fremsatte ønsket om å innføre digital eksamen var i følge funn fra studien mottagelsen i sentrale deler av institusjonen noe lunken. Foreleserne gir uttrykk for at de har opplevd mange kollegaer som skeptiske til bruk av digital eksamen på tross av at de som uttrykker skepsis ikke har testet programvaren, og dermed ikke eksakt vet hva de snakker om. I forbindelse med oppstarten og det å skulle bli kjent med et nytt verktøy, samt å bruke dette som del av et emne, har ført til at foreleserne ser behov for en ressursperson som kan bidra med faglig og pedagogisk veiledning i tilknytning til arbeidet med programvaren Wiseflow. Dette illustrerer behovet foreleserne har for et praksisfellesskap hvor de kan lære av hverandres erfaringer i en utvidet og mer systematisert form, enn hva som har vært tilfelle frem til nå. Et behov som for eksempel kan dekkes gjennom en ressursperson som fungerer som «mester» i et læringsfellesskap (Lave & Wenger, 1991). I tråd med sosiokulturell læringsteori er det sannsynlig at behovet for en slik ressursperson vil bestå gjennom hele approprieringen (Säljö, 2006) frem til en større del av kollegiet i høgskolen er etablerte brukere av digital eksamen, og kollegiet dermed naturlig kan overta denne rollen gjennom et større praksisfellesskap (Dysthe, 1999; Dysthe, 2001; Engeström, 2016; Imsen, 2005; Lave & Wenger, 1991; Säljö, 2001; Vygotsky, 1978).

På tross av at foreleserne gir uttrykk for at de savner et praksisfellesskap knyttet til arbeidet med digital eksamen, kan det argumenteres for at de faktisk har vært en del av et praksisfellesskap i denne prosessen, men at dette tilsynelatende ikke har dekket alle behov. Lave og Wenger (1991) gir uttrykk for at praksisfellesskap finnes i mange ulike former, og at disse derfor fungerer på ulikt vis. I prosessen med innføring av eksamen viser funn fra studien at foreleserne har inngått i et praksisfellesskap med ansatte i

utdanningscenteret og på IT knyttet til innføringen av digital eksamen. Siden de andre som har vært en del av dette praksisfellesskapet heller ikke hadde erfaring med digital eksamen, fremstår det tydelig at behovene som omtales over ikke har kunnet dekkes gjennom denne gruppen. De andre deltagerne i gruppen har ikke hatt ren faglig eller pedagogisk funksjon, og det er av den grunn ikke naturlig at de kan bidra i slike sammenhenger. Vi ser likevel at ulike deler av organisasjonen har deltatt i det samme lærende fellesskapet, noe som gjør at situert kunnskap om digital eksamen i Wiseflow har blitt konstruert gjennom samhandling. Samarbeidet har i ettertid ført til at det i påfølgende semester gjennomføres digitale eksamener i flere emner på nautikkstudiet, at flere ulike typer eksamen gjennomføres i Wiseflow, samt et felles ønske om at ytterligere emner på sikt skal ha digital eksamen.

Engeström (2016) og Lave og Wenger (1991) peker på at det er vedvarende gjensidig engasjement, eierskap til praksisfellesskapets mening og kollektiv læring som holder praksisfellesskap sammen og driver disse frem. Foreleserne gir uttrykk for at de ønsker utvikling og videre læring relatert til digital eksamen, men at de savner et større faglig og pedagogisk fellesskap knyttet til denne læreprosessen, derav også det tidligere nevnte ønsket om en ressursperson på institusjonen. Foreleserne har en forventning om at et hensiktsmessig og engasjerende praksisfellesskap kan oppstå hvis flere fagansatte er engasjert i digital eksamen, og at deltagerne i et slikt fellesskap på denne måten kan drive frem læring og utvikling i fellesskap, basert på egne og hverandres erfaringer. At foreleserne har vært del av slike praksisfellesskap i andre sammenhenger kommer frem i et av refleksjonsmøtene, hvor de gir uttrykk for at de har lært hvordan de bør sensurere av forelesere som var erfarne da de selv var nye i jobben og uerfarne med sensur. Dette viser at foreleserne i forbindelse med sin aller første sensur gikk inn i en rolle som «lærling» i et praksisfellesskap med en mer erfaren kollega. På denne måten ervervet foreleserne situert kunnskap om sensurering gjennom praksisfellesskapet, noe som er i tråd med Lave og Wengers (1991) beskrivelser av praksisfellesskap.

5.3.2 Backwash

Vurdering ser i stor grad ut til å påvirke både hvordan studenter lærer og hvordan forelesere underviser. Dette kalles *backwash*, og er en effekt som er mest kjent som en negativ påvirkning, men som også kan brukes formålstjenlig knyttet til studentenes læring (Biggs, 1996; Biggs & Tang, 2011; Havnes, 2004).

Raaheim (2016) fremhever at emner med sterke praksiselementer bør synliggjøre dette gjennom vurderingsformen. Funn fra studien viser at en positiv backwash-effekt i dette tilfellet synes å oppstå som et resultat av at det aktuelle emnet har todelt eksamensform. På denne måten testes ikke studentene utelukkende teoretisk, men også deres praktiske ferdigheter testes gjennom en praktisk eksamen. Dette fremstår som et resultat av at foreleserne er bevisste på behovet for en sterk sammenheng mellom vurderingen og emnets definerte læringsutbytte i henhold til Biggs og Tangs (2011) prinsipper for meningsskapende helhet. Gjennom refleksjonssamtalene kommer det frem at foreleserne opplever at innføringen av praktisk eksamen bidrar til at studentene tar inn det de lærer på simulatorøvelsene på en annen måte, enn tidligere, noe som har påvirket gjennomføringen av de praktiske simulatorøvelsene positivt.

I henhold til Entwistle og Entwistle (1992) påvirker emner studentene opplever at har en manglende struktur og logikk deres søken etter forståelse i emnet negativt. Dette står i sammenheng med Entwistle (1994) sine beskrivelser av tre læringsstrategier for studenter; en strategi med mål om dyp læring, en overflatisk strategi og en strategi hvor studentene strategisk benytter en kombinasjon av de to førstnevnte strategiene. Samtlige studenter gir under intervjuene uttrykk for at de anser emnet å ha god sammenheng mellom læringsutbytte, arbeids- og vurderingsformer, og dermed god struktur og logikk. I fokusgruppeintervjuene kom det frem at enkelte studenter var litt usikre på hvordan eksamen kom til å bli, siden dette var første gangs gjennomføring av emnet og de av den grunn ikke hadde tidligere eksamensoppgaver å øve på. Studien viser at arbeidsformene i emnet, forelesernes opplegg og informasjon sørget for at studentene jobbet med emnet alene og i grupper, og at dette bidro til innsikt i standarden på eksamen. Funn fra fokusgruppeintervjuene viser at for eksempel arbeidet med innleveringsoppgaver i grupper ga studentene trygghet knyttet til eksamen, siden de visste at de hadde jobbet godt med stoffet og diskutert det i gruppene. Dette stemmer med Entwistle og Entwistles (1992) beskrivelser av studenters måte å jobbe mot vurdering, og at arbeidsformene og sammenhengen i emnet har bidratt til at studentene kunne jobbe hensiktsmessig mot vurderingen. Studentene har på denne måten jobbet med det spesifiserte læringsutbyttet inn mot vurderingen.

I tillegg til å ha en innvirkning på studentenes læring, er vurderingen i følge Havnes (2004) en manifestasjon på kravene i et emne. Vurderingen definerer læringsobjektivene i større grad, enn emnets læringsutbytte. Dette medfører at backwash-effekten også har

en innvirkning på undervisning, pensum og annet læringsmateriale (Biggs, 1998). Funn fra studien viser at foreleserne gjennom å etablere en planlagt og tydelig sammenheng mellom arbeidsformer, vurdering og læringsutbytte har utnyttet backwash-effekten positivt. Det har ført til at læremateriell, innleveringsoppgaver og lignende har vært gjennomført med vurderingen i bakhodet, og har resultert i det studentene beskriver som et godt grunnlag for å kunne lykkes på eksamen. Når studentenes oppfatning av kravene som stilles på eksamen stemmer overens med kravene foreleserne faktisk stiller, påvirkes studentenes læring hensiktsmessig (Biggs, 1996). Funn fra studien viser at foreleserne har vært tydelige på hvordan de kommer til å legge opp eksamen, og at studentene også opplever det slik. Dette betyr ikke at studentene har fått vite hvilke spørsmål som kom på eksamen, men at foreleserne og deres helhetlige opplegg førte til at studentene visste hvordan de skulle forberede seg til eksamen og hvilke krav som ble stilt.

6. Avslutning

Under presenteres konklusjoner fra studien basert på problemstilling og forskningsspørsmål.

Studien har blitt gjennomført med problemstillingen: «*Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkstudiet tilpasses en digital kontekst?*», med tilhørende forskningsspørsmål:

1. *Hvordan begrunner faglærerne utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i programmet Wiseflow?*
2. *Hvordan beskriver studentene sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen?*

6.1 Oppsummering og konklusjon

Det første forskningsspørsmålet tar for seg hvordan faglærerne begrunner utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i Wiseflow. Analysene viser at de pedagogiske aspektene er avgjørende for foreleserne når de utarbeider eksamensoppgaver, og at programvaren har en viktig rolle i denne prosessen, siden denne legger føringer for hvilke oppgavetyper som er tilgjengelige. Foreleserne tar utgangspunkt i emnets læringsutbytte når de planlegger læringsaktiviteter og vurdering, og oppnår på denne måten en meningssskapende helhet i tråd med prinsippene beskrevet i kapittel 2 Teoretiske perspektiver.

Dataene gir klare indikasjoner på at både forelesere og studenter gjennomgår en appropriering av digital eksamen som medierende artefakt. Digitale artefakter blir et naturlig aspekt ved tenkningen og samtalen om læring og vurdering. Studentene synes i større grad å være kommet til siste fase i approprieringen, enn foreleserne. De fleste studentene regner digital eksamen som en selvfølge ved en moderne norsk høgskole. Foreleserne synes fortsatt å befinne seg i den tredje og nest siste fasen av approprieringen. Digital eksamen har ennå ikke blitt en selvfølgelig del av hverdagen deres. Analysene viser likevel at faglærerne gjennom semesteret har oppnådd en stadig økende koordinering mellom dem selv og programvaren for digital eksamen (Säljö, 2006).

Det andre forskningsspørsmålet tar for seg hvordan studentene beskriver sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen. En slik sammenheng er i tillegg til å bidra til en meningssskapende helhet i det aktuelle emnet blant annet med på å gi vurderingen god validitet (Biggs & Tang, 2011; Raaheim, 2016), noe analysene indikerer at er tilfelle. Analysene viser at også studentene opplever en tett og god sammenheng mellom emnets innhold, pensum, studiekraft, eksamen og læringsutbytte. Dette indikerer at man har oppnådd en formålstjenlig utnyttelse av backwash-effekten, og at eksamensformen også har bidratt positivt på studentenes læring.

6.2 Avsluttende tanker

Gjennom studien har jeg funnet at dyktige og bevisste forelesere har klart å tilpasse den tradisjonelle skriftlige skoleeksamenen i et av emnene på nautikk til en digital kontekst ved hjelp av en grunnleggende tanke om helhet og en sterk sammenheng mellom læringsutbytte, arbeids- og vurderingsformer. I denne prosessen gir foreleserne uttrykk for at de har savnet et praksisfellesskap i institusjonen, noe som kan være verdt å merke seg i forbindelse med videre overgang til digital eksamen. På tross av et udekket behov for et institusjonelt praksisfellesskap knyttet til digital eksamen, viser studien at foreleserne har gjennomført en vellykket innføring av digital eksamen. I tillegg har både studenter og forelesere kommet langt i prosessen for appropriering av digital eksamen, noe som gir et godt grunnlag for fremtidig bruk av digital eksamen som vurderingsverktøy i større målestokk.

6.2.1 Fremtidig forskning

I arbeidet med studien har jeg som litteraturgjennomgangen i kapittel 1.4 Tidligere forskning og kapittel 2 Teoretiske perspektiver viser lest en hel del om digital eksamen. Det som har slått meg underveis er at det i den litteraturen er relativt lite informasjon om digital skoleeksamen i et pedagogisk perspektiv. Det har denne studien forsøkt å gjøre noe med, men studien er et masterarbeid og dermed av begrenset omfang. Det kunne derfor vært interessant i en annen sammenheng og gått nærmere inn på helheten i emnet med læringsaktiviteter, og formative og summative elementer, for å undersøke hvordan programvaren best kan benyttes i en helhetlig sammenheng.

Tanken om en oppfølgingsstudie har også slått meg. I et av intervjuene sier foreleserne i sammenheng med hvilke funksjoner de har brukt for å teste de ulike formene for læringsutbytte i emnet at dette «*et spørsmål vi egentlig kunne svart mye bedre på om et års tid, når vi har kommet enda lenger, på en måte*». Det kunne det vært veldig interessant å undersøke nærmere; tenker de på samme måten når de har fått mer erfaring og verktøyet er fullstendig appropriert?

Bibliografi

- Alvesson, M., & Sköldberg, K. (2008). *Tolkning och reflektion*. Lund: Studentlitteratur AB.
- American Educational Research Association. (1967). *Perspectives of Curriculum Evaluation. American Educational Research Association Monograph Series on Curriculum Evaluation* (Vol. 1). (R. Tyler, R. Gagne, & M. Scriven, Red.) Chicago, USA: Rand McNally & Company.
- Baartman, L. K., Bastiaens, T. J., Kirschner, P. A., & van der Vleuten, C. P. (2007a). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review*, 2, ss. 114-129.
- Baartman, L., Bastiaens, T., Kirschner, P., & van der Vleuten, C. (2006). The wheel of competency assessment: Presenting quality criteria for competency assessment programs. *Studies in Educational Evaluation*(32), ss. 153-170.
- Behrens, J., Mislevy, R., DiCerbo, K., & Levy, R. (2010). *An Evidence Centered Design for Learning and Assessment in the Digital World*. UCLA. Los Angeles: The National Research Center for Research on Evaluation, Standards, and Student Testing.
- Berggren, B., Fili, A., & Nordberg, O. (2015). Digital examination in higher education - Experiences from three different perspectives. *International journal of Education and Development using Information and Communication Technology (IJEDICT)*, Vol. 11(3), ss. 100-108.
- Bertheussen, B. (2014). Digital School Examinations: An Educational Note of an Innovative Practice. *International Business Research*, Vol 7(6), ss. 129-139.
- Biggs, J. (1996, March). Assessing learning quality: Reconciling institutional, staff and educational demands. *Assessment & Evaluation in Higher Education*, 21(1), ss. 5-15.
- Biggs, J. (1998, July). Assessment and Classroom Learning: a role for summative assessment? *Assessment in Education*, 5(1), ss. 103-110.
- Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University* (4. utg.). Berkshire, England: Open University Press.
- Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University* (4. utg.). Berkshire, England: Open University Press.
- Bloom, B., Engelhart, M., Furst, E., Hill, W., & Kratwohl, D. (1956). *Taxonomy of Educational Objectives. Handbook I: Cognitive Domain* (16th printing, March 1971. utg.). (B. S. Bloom, Red.) New York: David McKay Company, Inc.
- Bloom, B., Hastings, J., & Madaus, G. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill Book Company.

- Boevé, A., Meijer, R., Albers, C., Beetsma, Y., & Bosker, R. (2015, December). Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment. *PLOS ONE*, ss. 1-13.
- Boote, D., & Beile, P. (2005). Scholars before researchers: On the Centrality of the Dissertation Literature Review in Research Preparation. *Educational Researcher*, 34(6), ss. 3-15.
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7. utg.). New York, NY, USA: Routledge.
- Conole, G., & Warburton, B. (2005, March). A review of computer-assisted assessment. *Research in Learning Technology*, 13(1), ss. 17-31.
- Cowan, J. (2010, June). Developing the ability for making evaluative judgements. *Teaching in Higher Education*, 15(3), ss. 323-334.
- Creswell, J. W. (2014). *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research* (4. utg.). Essex: Pearson Education Limited.
- Den nasjonale forskningsetiske komitè for samfunnsvitenskap og humaniora (NESH). (2016, April). Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi. Oslo: De nasjonale forskningsetiske komiteene.
- Deutsch, T., Herrmann, K., Frese, T., & Sandholzer, H. (2012). Implementing computer bases assessment - A web-based mock examination changes attitudes. *Computers & Education*, 58, ss. 1068-1075.
- Dobson, S., Eggen, A., & Smith, K. (Red.). (2009). *Vurdering, prinsipper og praksis*. Oslo: Gyldendal Akademisk.
- Dochy, F. (2001). A new assessment era: different needs, new challenges. *Research Dialogue in Learning and Instruction*, 2001(2), ss. 11-20.
- Dochy, F., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), ss. 279-298.
- Dysthe, O. (1999). *Ulike teoriperspektiv på kunnskap og læring*. Hentet April 11, 2017 fra Stiftelsen Hvasser: <http://www.stiftelsen-hvasser.no/documents/Teoriperspektivpaakunnskapoglering.pdf>
- Dysthe, O. (2001). *Dialog, samspel og læring*. Oslo: Abstrakt Forlag AS.
- Eggen, A. B. (2009). Vurdering og validitet. I S. Dobson, A. Eggen, & K. Smith (Red.), *Vurdering, prinsipper og praksis*. Oslo: Gyldendal Akademisk.
- Engeström, Y. (2016). *Studies in Expansive Learning*. New York, USA: Cambridge University Press.
- Entwistle, A., & Entwistle, N. (1992). Experiences of understanding in revising for degree examinations. *Learning and Instruction* (Vol. 2), ss. 1-22.

- Entwistle, N. (1994). *Experiences of Understanding and Strategic Studying*. American Educational Research Association.
- European Commission. (2017, April 29). *Learning Opportunities and Qualifications in Europe*. Hentet April 29, 2017 fra European Commission: <https://ec.europa.eu/ploteus/content/how-does-efq-work>
- Fegran, L., & Slettebø, Å. (2012). Læringsutbytte, læringsaktiviteter og vurderingsformer. *Nordisk Sygeplejeforskning*, 2(1), ss. 64-73.
- Fossland, T. (2015). *Digitale Læringsformer i høyere utdanning*. Oslo: Universitetsforlaget.
- Gekara, V., Bloor, M., & Sampson, H. (2011, March 17.). Computer-based assessment in safety-critical industries: the case of shipping. *Journal of Vocational Education and Training*, 63(1), ss. 87-100.
- Grimen, H. (2004). *Samfunnsvitenskapelige tenkemåter* (3.. utg.). Oslo: Universitetsforlaget.
- Halkier, B. (2010). *Fokusgrupper*. Oslo: Gyldendal Norsk Forlag AS.
- Hassanien, M., Al-Hayani, A., Abu-Kamer, R., & Almazrooa, A. (2013). A six step approach for developing computer based assessment in medical education. *Medical Teacher*, 35, ss. 15-19.
- Havnes, A. (2004, April). Examination and Learning: an activity-theoretical analysis of the relationship between assessment and educational practice. *Assessment & Evaluation in Higher Education*, 29(2), ss. 159-176.
- Havnes, A., & Raaheim, A. (2006). Vurdering og eksamen. I H. Strømsø, K. Lycke, & P. Lauvås (Red.), *Når læring er det viktigste - Undervisning i høyere utdanning*. Oslo: Cappelen Akademisk Forlag.
- Imsen, G. (2005). *Elevers Verden*. Oslo: Universitetsforlaget.
- International Maritime Organization. (2011). *International Convention on Standards for Training, Certification and Watchkeeping for Seafarers*. London: International Maritime Organization.
- Keshavarz, M. (2011). Measuring Course Learning Outcomes. *Journal of Learning Design*, 4(4), ss. 1-9.
- Krueger, R., & Casey, M. (2015). *Focus Groups: A practical guide for applied research* (5. utg.). Thousand Oaks, California, USA: SAGE Publications, Inc.
- Krumsvik, R. J. (2014). *Forskningsdesign og kvalitativ metode*. Bergen: Fagbokforlaget.
- Kunnskapsdepartementet. (2006). *FOR-2005-12-15-1506 Forskrift om egenbetaling ved universiteter og høyskoler*. Hentet Mai 21, 2017 fra Lovdata: <https://lovdata.no/dokument/SF/forskrift/2005-12-15-1506?q=FOR-2016-06-27-821>

- Kunnskapsdepartementet. (2014, Oktober 13). *Regjeringen.no*. Hentet April 27., 2017 fra Nasjonalt kvalifikasjonsrammeverk for livslang læring (NKR): <https://www.regjeringen.no/contentassets/e579f913fa1d45c2bf2219afc726670b/nkr.pdf>
- Kunnskapsdepartementet. (2016a, Juni 13). *Regjeringen.no*. Hentet April 29, 2017 fra Voksnes læring og kompetanse: <https://www.regjeringen.no/no/tema/utdanning/hoyere-utdanning/nasjonalt-kvalifikasjonsrammeverk/id564809/>
- Kunnskapsdepartementet. (2016b, Juni 13). *Regjeringen.no*. Hentet April 29, 2017 fra Voksnes læring og kompetanse: https://www.regjeringen.no/no/tema/utdanning/voksnes_laering_og_kompetanse/artikler/nasjonalt-kvalifikasjonsrammeverk/id601327/
- Kunnskapsdepartementet. (2017). *Meld. St. 16 (2016–2017) Kultur for kvalitet i høyere utdanning*. Hentet fra *Regjeringen.no*: <https://www.regjeringen.no/contentassets/aee30e4b7d3241d5bd89db69fe38f7ba/no/pdfs/stm201620170016000dddpdfs.pdf>
- Kvale, S., & Brinkmann, S. (2009). *Det kvalitative forskningsintervju* (2.. utg.). Gyldendal Norsk Forlag.
- Lave, J., & Wenger, E. (1991). *Situated learning; Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic Inquiry*. Newbury Park, California, USA: SAGE Publications, Inc.
- Llamas-Nistal, M., Fernández-Iglesias, M., González-Tato, J., & Mikic-Fonte, F. (2013). Blended e-assessment: Migrating classical exams to the digital world. *Computers & Education*, 62, ss. 72-87.
- McKenney, S., & Reeves, T. (2012). *Conducting Educational Design Research*. New York, NY, USA: Routledge.
- Moustakas, C. (1994). *Phenomenological Research Methods*. Thousand Oaks, California, USA: SAGE Publications, Inc.
- Nastiti, H., & Campbell, A. (2010). *Digital assessment: assessing what counts, the performance*. Hentet August 25., 2016 fra International Association for Educational Assessment: http://www.iaea.info/documents/paper_4d52f56.pdf
- Newhouse, C. P. (2010). Aligning assessment with Curriculum and Pedagogy in Applied Information Technology. *Australian Educational Computing*, Vol 24(2), ss. 4-11.
- Newhouse, C. P. (2012). Digital forms of assessment: aligning with pedagogic and curriculum intentions. *Proceedings of Australian Computers in Education Conference (ACEC)*, (ss. 1-7). Perth, Australia.

- Newhouse, C. P. (2013). Computer-based exams in schools: Freedom from the limitations of paper? *Research and Practice in Technology Enhanced Learning*, 8(3), ss. 431-447.
- Newhouse, C. P. (2015). Using digital technologies to improve the authenticity of performance assessment for high-stakes purposes. *Technology, Pedegogy and Education*, 1, ss. 17-33.
- Nielsen, K., & Petersen, L. (2013). Digital Assessment. *19th EUNIS Congress "ICT Role for Next Generation Universities"*. Riga: EUNIS. DOI: 10.7250/eunis.2013.016.
- Nix, I., & Wyllie, A. (2011). Exploring design features to enhance computer-based assessment: Learners' view on using. *British Journal of Educational Technology*, 42(1), ss. 101-112.
- Pellegrino, J., & Quellmalz, E. (2010). Perspectives on the Integration of Technology in Assessment. *Journal of Research and Technology in Education*, 43(2), ss. 119-134.
- Postholm, M. B. (2010). *Kvalitativ metode: En innføring med fokus på fenomenologi, etnografi og kasusstudier* (2. utg.). Oslo: Universitetsforlaget.
- Prøitz, T. S. (2015). *Læringsutbytte*. Oslo: Universitetsforlaget.
- Raaheim, A. (2016). *Eksamensrevolusjonen*. Oslo: Gyldendal Norsk Forlag AS.
- Redecker, C., & Johannessen, Ø. (2013). Changing Assessment - Towards a New Assessment Paradigm Using ICT. *European Journal of Education*, Vol. 48(No.1), ss. 79-96.
- Ringdal, K. (2013). *Enhet og mangfold* (3. utg.). Bergen: Fagbokforlaget Vigmostad & Bjørke AS.
- Scalise, K., & Wilson, M. (2011). The Nature of Assessment Systems to Support Effective Use of Evidence through Technology. *E-Learning and digital Media*, 8(2), ss. 121-132.
- Scriven, M. (1967). The Methodology of Evaluation. I R. W. Tyler, R. M. Gagné, & M. Scriven (Red.), *Area Monograph Series on Curriculum Evaluation: Perspectives of Curriculum Evaluation* (Vol. 1, ss. 39-84). Chicago, USA: Rand McNally & Company.
- Scriven, M. (1974). *Evaluation: A study guide for educational administrators*. Fort Lauderdale, Florida, USA: Nova University.
- Silverman, D. (2013). *Doing Qualitative Research* (4. utg.). London: Sage Publications Ltd.
- Sindre, G., & Vegendla, A. (2015). E-exams and exam process improvement. *UDIT/NIK 2015 Conference*.
- Sjøfartsdirektoratet. (2012, Januar 1). *Forskrift om kvalifikasjoner og sertifikater for sjøfolk*. Hentet April 29, 2017 fra Lovdata: <https://lovdata.no/dokument/SF/forskrift/2011-12-22-1523>
- Säljö, R. (2001). *Læring i praksis*. Oslo: J. W. Cappelens forlag a.s.
- Säljö, R. (2006). *Læring og kulturelle redskaper*. Oslo: Cappelen Akademisk Forlag AS.

- Thagaard, T. (2013). *Systematikk og innlevelse: En innføring i kvalitativ metode* (4. utg.). Bergen: Fagbokforlaget Vigmostad & Bjørke AS.
- Thelwall, M. (2000). Computer-based assessment: A versatile educational tool. *Computers & Education*, 2000, ss. 37-49.
- Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016, June). Rethinking assessment in a digital age: opportunities, challenges and risks. *British Educational Research Journal*, Vol. 42(No.3), ss. 454-476.
- van den Akker, J., Gravemeijer, K., McKinney, S., & Nieveen, N. (2006). *Educational Design Research*. New York, NY, USA: Routledge.
- Vygotsky, L. S. (1978). *Mind In Society*. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Red.) Cambridge, Massachusetts, USA: Harvard University Press.
- Wartofsky, M. W. (1979). *Models - Representation and the Scientific Understanding*. Dordrecht, Holland: D. Reidel Publishing Company.
- Wertsch, J. V. (1998). *Mind As Action*. New York, New York, USA: Oxford University Press.
- Zoanetti, N. (2010). Interactive computer based assessment tasks: How problem-solving process data can inform instruction. *Australasian Journal of Educational Technology*, 26(5), ss. 585-606.

Vedlegg 1 - Tabell 1, Litteraturgjennomgang

Tema	Inkludert	Ekskludert
Database	Academic Search Premier, ERIC, Teacher Reference Center (ARC/ERIC/TRC) Idunn (Id) Science Direct (SD) Google Scholar (GS)	Alle andre databaser
Tid	All litteratur etter 2010 Manuelle søk etter konkrete artikler/forfattere foretatt uten tidsbegrensinger.	All litteratur eldre enn fra 2010.
Fokus	Maritim utdanning/utdanning, digital skoleeksamen, digital vurdering, summativ vurdering, læringsutbytte i sammenheng med vurdering.	Utdanningsnivå under høyskole/universitet, formativ vurdering, digitale praktiske eksamener, digitale mapper, digital hjemmeeksamen Alle sitater og patenter utelatt fra søk.
Type aktivitet	Gjennomføring og evaluering av digital eksamen og digital vurdering, utvikling av eksamensspørsmål, forankring av eksamensspørsmål i læringsutbytte/innhold	Artikler som omhandler teknologien for digital eksamen, andre former for eksamen, multiple choice-oppgaver.
Språk	Norsk, svensk, dansk, engelsk	Alle andre språk
Søkeord	<ol style="list-style-type: none"> 1. digital assessment 2. digital exam 3. computer based assessment 4. maritime education + digital assessment 5. maritime education + assessment 6. digital assessment + good questions 7. summative evaluation + learning outcomes in higher education 8. digital eksamen (ingen tidsavgrensning I Idunn) 9. digital vurdering 10. maritim utdanning + digital vurdering 11. digital vurdering + gode spørsmål 12. digital eksamen + gode spørsmål 13. computer supported assessment 	

Metode	Kvalitativ og kvantitativ	Ingen
Resultat	<ol style="list-style-type: none"> 1. ASE/ERIC/TRC; 8.168 treff, sjekket 250 → 6 relevante (5 etter abstract) SD; 4 treff, sjekket alle → 0 relevante 2. ASE/ERIC/TRC; 492 treff, sjekket alle → 23 relevante (9 etter abstract) SD; 3 treff, sjekket alle → 0 relevante GS; 496 treff, sjekket alle → 15 relevante (5 etter abstract) 3. ASE/ERIC/TRC; 183 treff, sjekket alle → 21 relevante (13 etter abstract) 4. ASE/ERIC/TRC; 0 treff 5. ASE/ERIC/TRC; 66 treff, sjekket alle → 5 relevante (2 etter abstract) 6. ASE/ERIC/TRC; 23 treff, sjekket alle → 1 relevant 7. ASE/ERIC/TRC; 95 treff, sjekket alle → 7 relevante (2 etter abstract) 8. Id; 34 treff, sjekket alle → 4 relevante (3 etter abstract) GS; 46 treff, sjekket alle → 10 relevante (4 etter abstract) 9. Id; 138 treff, sjekket alle → 8 relevante (6 etter abstract) 10. Id; 0 treff 11. Id; 71 treff, sjekket alle → 5 relevante (samme treff som på søk nr. 9) 12. Id; 14 treff, sjekket alle → 4 relevante (samme treff som på søk nr. 9) 13. ASE/ERIC/TRC; 2 treff, sjekket begge → 0 relevante 	

Vedlegg 2 – Informasjons- og samtykkeskjema

Forespørsel om deltakelse i forskningsprosjektet «Digital skoleeksamen – mer enn strøm på gamle eksamensoppgaver?»

Bakgrunn og formål

Jeg er masterstudent på master i IKT i læring ved Høgskolen Stord/Haugesund, og skal i studieåret 2016/2017 gjennomføre et forskningsprosjekt for å undersøke hvordan man ved innføringen av digital skoleeksamen i ett emne på nautikkstudiet ved HSH kan bruke de mulighetene som en digital eksamen gir for å lage gode eksamensoppgaver, og på denne måten teste studentenes kunnskaper, ferdigheter og generelle kompetanse ut fra emnets definerte læringsutbytte. Dette gjøres gjennom problemstillingen «Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkutdanningen tilpasses en digital kontekst?» med de tilhørende forskningsspørsmålene «Hvordan begrunner faglærerne utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i programmet Wiseflow?» og «Hvordan beskriver studentene sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen?».

Hva innebærer deltakelse i studien?

Utvalget blir satt sammen av frivillige studenter og forelesere fra ett spesifikt emne ved HSHs nautikkstudium, hvor digital eksamen gjennomføres, og som har deltatt på studiekrav og digital eksamen. Etter godkjenning fra Norsk Samfunnsvitenskapelig Datatjeneste (NSD) retter jeg derfor denne forespørselen om deltagelse til deg.

Datainnsamlingen i studien vil skje gjennom fokusgruppeintervju med studenter og fokusgruppe-intervju med forelesere. Intervjuene gjennomføres i perioden oktober 2016 til februar 2017, og spørsmålene vil omhandle sammenhengen mellom oppgavene på den digitale skoleeksamenen og emnets innhold.

Hva skjer med informasjonen om deg?

Jeg ønsker å ta lydopptak av intervjuene, som senere transkriberes til tekst. Intervjuene er konfidensielle, og svarene anonymiseres. Det vil ikke være mulig å vite hvem som har svart hva i masteroppgaven. Alle personopplysninger vil bli behandlet konfidensielt, og det er kun jeg som vil ha tilgang til opplysningene underveis i prosjektet. Personopplysningene jeg registrerer om dere omfatter kun fornavn, og koblingsnøkkelen vil være lagret adskilt fra øvrige forskningsdata. Prosjektet skal etter planen avsluttes innen 1. juli 2017, men det er ønskelig å lagre dataene til 1. desember 2017 i tilfelle det skjer noe uventet. Alle data slettes umiddelbart etter dette.

Frivillig deltakelse

Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert.

Med vennlig hilsen

Johanne Marie Trovåg

Telefon 52 70 27 29/ 91 37 26 53 eller e-post johanne.trovag@hsh.no.

Kontaktopplysninger til veileder:

Professor Knut Steinar Engelsen, Høgskolen Stord/Haugesund; knut.engelsen@hsh.no.

Studien er meldt til Personvernombudet for forskning, NSD - Norsk senter for forskningsdata AS.

Samtykke til deltakelse i studien

Jeg har mottatt informasjon om studien, og er villig til å delta

(Signert av prosjektdeltaker, dato)

Vedlegg 3 – Godkjenning fra NSD



Knut Steinar Engelsen
Avdeling for lærerutdanning og kulturfag Høgskolen Stord/Haugesund
Klingenbergvegen 8
5414 STORD

Vår dato: 22.09.2016

Vår ref: 49659 / 3 / ASF

Deres dato:

Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 30.08.2016. Meldingen gjelder prosjektet:

49659	<i>Digital skoleeksamen - mer enn strøm på gamle eksamensoppgaver?</i>
Behandlingsansvarlig	<i>Høgskolen Stord/Haugesund, ved institusjonens øverste leder</i>
Daglig ansvarlig	<i>Knut Steinar Engelsen</i>
Student	<i>Johanne Marie Trovåg</i>

Personvernombudet har vurdert prosjektet og finner at behandlingen av personopplysninger er meldepliktig i henhold til personopplysningsloven § 31. Behandlingen tilfredsstiller kravene i personopplysningsloven.

Personvernombudets vurdering forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, ombudets kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, <http://www.nsd.uib.no/personvern/meldeplikt/skjema.html>. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, <http://pvo.nsd.no/prosjekt>.

Personvernombudet vil ved prosjektets avslutning, 01.12.2017, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

Kjersti Haugstvedt

Amalie Statland Fantoft

Kontaktperson: Amalie Statland Fantoft tlf: 55 58 36 41

Vedlegg: Prosjektvurdering

Dokumentet er elektronisk produsert og godkjent ved NSDs rutiner for elektronisk godkjenning.

NSD – Norsk senter for forskningsdata AS
NSD – Norwegian Centre for Research Data

Harald Hårfagres gate 29
NO-5007 Bergen, NORWAY

Tel: +47-55 58 21 17
Faks: +47-55 58 96 50

nsd@nsd.no
www.nsd.no

Org.nr. 985 321 884

Personvernombudet for forskning



Prosjektvurdering - Kommentar

Prosjektnr: 49659

I følge meldeskjemaet skal deltakerne i studien informeres skriftlig og muntlig om prosjektet og samtykke til deltakelse. Informasjonsskrivet er godt utformet.

Personvernombudet legger til grunn at dere behandler alle data og personopplysninger i tråd med Høgskolen Stord/Haugesund sine retningslinjer for innsamling og videre behandling av forskningsdata og personopplysninger.

I meldeskjemaet har dere informert om at forventet prosjektslutt er 01.12.2017. Ifølge prosjektmeldingen skal dere da anonymisere innsamlede opplysninger. Anonymisering innebærer at dere bearbeider datamaterialet slik at ingen enkeltpersoner kan gjenkjennes. Det gjør dere ved å slette direkte personopplysninger, slette eller omskrive indirekte personopplysninger og slette digitale lydopptak.

Vedlegg 4 – Intervjuguide refleksjonsmøter

Tema for refleksjonsmøte med forelesere

Informasjon

Forskningsprosjektets tema er å undersøke hvordan man ved innføringen av digital skoleeksamen i ett emne på nautikkstudiet ved HSH kan bruke de mulighetene som en digital eksamen gir for å lage gode eksamensoppgaver, og på denne måten teste studentenes kunnskaper, ferdigheter og generelle kompetanse ut fra emnets definerte læringsutbytte.

Problemstillingen jeg har valgt er «*Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkstudiet tilpasses en digital kontekst?*», med forskningsspørsmålene;

1. *Hvordan begrunner faglærerne utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i programmet Wiseflow?*
2. *Hvordan beskriver studentene sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen?*

Introduksjonsrunde

Alle kjenner hverandre, så det er ikke nødvendig med en introduksjonsrunde.

Info om anonymitet, samt bekreftelse på at intervjuet kan tas opp. Eventuelle spørsmål. Info om at man kan trekke seg uten grunn når som helst både under intervjuet og i den videre prosessen.

Intervjuet

Intervjuet i dag dreier seg om deres opplevelser av den digitale skoleeksamenen, og hvordan dere utformer eksamensoppgavene til digital skoleeksamen i emnet (navn på emnet).

Intervjuet varer i 1-1 ½ time, og det blir tatt opp med lydopptaker. Ingen andre enn jeg og eventuelt min veileder hører på opptakene eller ser utskriftene.

Det som skal foregå her er;

- Dette intervjuet er annerledes enn man normalt forbinder med et intervju, hvor intervjueren stiller masse spørsmål hele tiden.
- Nå er det mest dere som skal snakke og diskutere med hverandre.
- Jeg har ulike emner som jeg gir dere en om gangen til å snakke om. I tillegg har vi en helt åpen diskusjon avslutningsvis («annet»).

- Dere leder selv diskusjonen. Hvis den sporer av, eller hvis dere går tomme for noe å si, eller hvis ikke alle blir hørt, vil jeg gripe inn, såfremt ingen av dere gjør det.
- Forestill dere at det er som å sitte sammen og diskutere fag over en kaffe i lunsjrommet.
- Jeg er først og fremst interessert i hvordan dere tenker og jobber når dere skal utforme eksamensoppgavene til den digitale skoleeksamenen i emnet (navn på emnet).
- Alles meninger og opplevelser er like viktige.
- Det finnes ingen rette eller gale svar, alle deres oppfatninger er like riktige.

Tema – fokusgruppe etter gjennomført pilot/studiekraft:

- Kan dere si litt om hvordan dere synes at piloten på digital eksamen (studiekraftgjennomføringen) har fungert?
- Kan dere si noe om hvordan dere utarbeidet oppgavene?
- Kan dere si noe om hvorvidt spørsmålene ga de svarene dere forventet, og mulige årsaker til dette?
- Er det noe dere vil gjøre annerledes i prosessen med utforming av eksamensoppgavene?
- Annet?

Tema – fokusgruppe etter gjennomføring av eksamen og fullført sensur:

- Kan dere si litt om hva dere har gjort annerledes enn i piloten?
- Kan dere si noe om hvordan dere utarbeidet oppgavene?
- Kan dere si noe om hvorvidt spørsmålene ga de svarene dere forventet, og mulige årsaker til dette?
- Har dere noen tanker om programvaren?
- Kan dere si noe om hvordan eksamensresultatene samsvarte med deres forventninger?
- Annet?

Vedlegg 5 – Intervjuguide individuelt intervju

Tema for individuelt refleksjonsmøte

Informasjon

Forskningsprosjektets tema er å undersøke hvordan man ved innføringen av digital skoleeksamen i ett emne på nautikkstudiet ved HSH kan bruke de mulighetene som en digital eksamen gir for å lage gode eksamensoppgaver, og på denne måten teste studentenes kunnskaper, ferdigheter og generelle kompetanse ut fra emnets definerte læringsutbytte.

Problemstillingen jeg har valgt er «*Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkstudiet tilpasses en digital kontekst?*», med forskningsspørsmålene;

1. *Hvordan begrunner faglærerne utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i programmet Wiseflow?*
2. *Hvordan beskriver studentene sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen?*

Tema	Spørsmål
1. Introduksjon/ Bakgrunnsinformasjon	Informasjon om prosjektet, tema og mål.
	Informasjon om lydopptak, samtykke, anonymitet, eventuelle spørsmål. Gjenta info om at de som intervjues har mulighet til å trekke seg uten grunn på et hvilket som helst tidspunkt.
2. Utarbeidelse av oppgaver	Bakgrunn: <ul style="list-style-type: none">• Hvilken erfaring har du med utarbeidelse av eksamensoppgaver?• Hvilken erfaring har du med digitale verktøy?• Hvilke funksjoner i Wiseflow har du brukt i utarbeidelsen av oppgavene?• Hvilke funksjoner i Wiseflow har du valgt bort?• Var det noen funksjoner du savnet i programvaren? Hvilke? Utforming av oppgaver: <ul style="list-style-type: none">• Hvordan er eksamensoppgavene utformet?• Hva er fokuset når du utarbeider eksamensoppgaver?

	<ul style="list-style-type: none"> • Hva er driveren i prosessen ved utarbeidelse av oppgaver (teknologi eller pedagogikk)? Hvorfor? • Hvordan jobber du for å knytte eksamensoppgavene til emnets læringsutbytte? • Hva kan du få testet ved bruk av digital eksamen, som du ikke tidligere kunne teste på en ordinær skriftlig eksamen? • Hva opplever du som forskjellen(e) mellom å utvikle oppgaver til tradisjonell skriftlig eksamen og digital skoleeksamen? • Hva har du ikke fått testet på digital skoleeksamen som du kunne teste på en vanlig skriftlig eksamen?
<p>3. Oppsummering/avrunding</p>	<ul style="list-style-type: none"> • Hvordan synes du at prosessen med innføring av digital eksamen har gått? • Hvilke utfordringer har du støtt på underveis i prosessen med utvikling av oppgaver til digital skoleeksamen? • Beskriv hvorvidt du er fornøyd med oppgavene sett i lys av emnets innhold og studentenes resultater? Hvorfor/hvorfor ikke? • Er det noe annet du har tenkt på i denne sammenhengen?

Vedlegg 6 – Intervjuguide fokusgruppeintervju

Tema for fokusgruppeintervju med studenter

Informasjon

Forskningsprosjektets tema er å undersøke hvordan man ved innføringen av digital skoleeksamen i ett emne på nautikkstudiet ved HSH kan bruke de mulighetene som en digital eksamen gir for å lage gode eksamensoppgaver, og på denne måten teste studentenes kunnskaper, ferdigheter og generelle kompetanse ut fra emnets definerte læringsutbytte.

Problemstillingen jeg har valgt er «*Hvordan kan tradisjonell skriftlig skoleeksamen i nautikkstudiet tilpasses en digital kontekst?*», med forskningsspørsmålene;

1. *Hvordan begrunner faglærerne utformingen av eksamensoppgaver i lys av emnets innhold og tilgjengelige funksjoner i programmet Wiseflow?*
2. *Hvordan beskriver studentene sammenhengen mellom emnets innhold og spørsmålene på digital skoleeksamen?*

Introduksjonsrunde

Alle presenterer seg med fornavn, før båndopptakeren startes.

Info om anonymitet, samt bekreftelse på at intervjuet kan tas opp. Eventuelle spørsmål. Info om at man kan trekke seg uten grunn når som helst både under intervjuet og i den videre prosessen.

Intervjuet

Intervjuet i dag dreier seg om deres opplevelser av den digitale skoleeksamenen, og hvordan dere opplever sammenhengen mellom innholdet i emnet (navn på emnet) og eksamensoppgavene.

Intervjuet varer i 1-1 ½ time, og det blir tatt opp med lydopptaker. Ingen andre enn jeg hører på opptakene eller ser utskriftene.

Det som skal foregå her er;

- Dette intervjuet er annerledes enn man normalt forbinder med et intervju, hvor intervjueren stiller masse spørsmål hele tiden. Nå er det mest dere som skal snakke og diskutere med hverandre.

- Jeg har fem emner som jeg gir dere en om gangen til å snakke om. I tillegg har vi en helt åpen diskusjon avslutningsvis («annet»).
- Dere leder selv diskusjonen. Hvis den sporer av, eller hvis dere går tomme for noe å si, eller hvis ikke alle blir hørt, vil jeg gripe inn, såfremt ingen av dere gjør det.
- Forestill dere at det er som å sitte sammen og diskutere fag over en kaffe i kaffebaren.
- Jeg er først og fremst interessert i hvordan dere opplever sammenhengen mellom eksamensoppgavene og innholdet i emnet (navn på emnet), altså om dere mener at dere blir testet på det fagstoffet som dere opplever at emnet handler om, og eventuelt hvorfor/hvorfor ikke?
- Alles meninger og opplevelser er like viktige.
- Det finnes ingen rette eller gale svar, alle deres oppfatninger er like riktige.

Tema:

- Kan dere si litt om hvordan dere opplevde det å gjennomføre en digital skoleeksamen? Kan dere si noe om sammenhengen mellom det faglige innholdet i eksamensoppgavene og (navn på emnet)?
- Kunne noe vært gjort annerledes for å tydeliggjøre sammenhengen mellom emnets innhold og eksamensoppgavene?
- Er det noen spørsmål dere ser for dere at ikke kunne blitt gitt på en tradisjonell skriftlig eksamen?
- Er det noen spørsmål dere ser for dere at dere kunne fått på en tradisjonell skriftlig eksamen, som ikke kan gis på digital eksamen?
- Får dere bedre uttrykt det dere kan på en digital skoleeksamen i forhold til en tradisjonell skriftlig eksamen?
- Annet?