

Kreativitet og ambivalens

**En undersøkelse av variasjoner i vurdering
og kjennetegn ved sprikvurderte tekster
fra eksamen i norsk hovedmål 2015**

Fakultet for humaniora og utdanningsvitenskap

Masterprogram i norskdidaktikk

Liv Cathrine Krogh

Juni 2016

Sammendrag

Undersøkelsen av eksamen i norsk hovedmål 2015 er tredelt. For å finne svar på spørsmål om mulig sammenheng mellom sprikvurdering og oppgaveformuleringer, er både eksamensoppgavene med tilhørende eksamensveiledning og karakterer gitt før fellessensuren analysert. Videre ble et utvalg eksamensbesvarelser analysert for å kunne forklare sprikvurderingen. Datamaterialet består av dokumenter, tall og tekst, og undersøkelsen har derfor kombinert kvalitative og kvantitative metoder.

Det er funnet én oppgave som skiller seg ut, både når den analyseres kvalitativt og i undersøkelsen av vurderingene. Dette er den kreative oppgaven der eksaminandene skal reflektere. Den skiller seg ut ved sin åpenhet, noe som kanskje er nødvendig fordi den skal føre til nettopp kreativ skriving. Likevel kan en spørre seg om den bør være fullt så åpen, alternativt om vurderingsveiledningen bør utvikles. Dette er spørsmål som får tyngde av tallmaterialet, som viser at oppgaven har trukket andel sprikvurdering opp. En av fem kreative tekster i utvalget er sprikvurdert, og korrelasjonen sensorene imellom på 0.46. Dette kan bety at flere og motstridende normer i sensorcorpset er mer fremtredende i møter med kreative refleksjoner enn andre oppgaver. Informativ, argumenterende og resonnerende oppgave skiller seg også ut. Her korrelerer sensorenes karakterforslag i større grad: 0.69.

Utvalget av kreative tekster er mangfoldig, noe som kan ha komplisert vurderingen. Det er funnet sammenheng mellom sprikvurdering og ulik vektning av kortsvaret, manglende svar på langsvaret og måten dette er besvart. Det er sannsynlig at kortsvarets potensiale til at eleven får vist mer av sin fagkunnskap på eksamen ikke er fullt ut realisert. Langsvaroppgaven ber om refleksjon om tematikk i to tekstvedlegg, og det kreative kan vises gjennom språk og innfallsvinkler. Bortimot halvparten av tekstene i ekstensivutvalget bærer preg av at de ikke innfrir hele denne kontrakten, og sprikvurdering avdekker usikkerhet hos sensorene i å håndtere dette. Sprikvurdert er også eksaminandene som på ulike måter innfrir. Intensivanalysen konkluderer med at det mest sannsynlig er måten de realiserer nettopp refleksjon, kreativt språk og innfallsvinkler som har sammenheng med sprikvurderingen. Dette betyr at tolkningsfellesskapet av sensorer sannsynligvis rommer flere og motstridende normer til eksempelvis subjektiv skriving, metakommunikasjon, språklige og litterære virkemidler.

Uenighet mellom sensorer ser ikke ut til å ha rammet sluttvurderingen av de sprikvurderte tekstene. En kan si at sprikvurderingen har gitt usikkerhet som stort sett er kommet eksaminanden til gode. Dette bekreftes av at kvantitativ undersøkelse ikke viser sammenheng mellom sprikvurdering og klagefrekvens. Imidlertid gjør funn av generell lav konsistens og konsensus at det kan knyttes usikkerhet til vurderinger det er enighet om. Analyse av tekster det var klaget på bekrefter språklige og innholdsmessige tendenser i sprikteksene, men her har sensorene stort sett vært enige i sine karakterforslag. Det er sannsynlig at et sensorpar med motstridende normer blir enige om et karakterforslag som ikke er urimelig strengt. Det er også sannsynlig at når to med mer tilsvarende normer møtes, øker risiko for at tekstene vurderes for strengt eller mildt.

Forord

Det er mai og eksamensdagene nærmer seg igjen. Klassen jeg er norsklærer for, 3STf på Horten videregående skole, skal vise kompetanse i hovedmålet sitt. I likhet med om lag 40.000 andre unge som jakter studiekompetanse. Selv etter at jeg nå har jobbet mye med eksamen i norsk hovedmål gjennom det siste året, er jeg fortsatt veldig glad i eksamen. Jeg gleder meg til å se oppgavene, ser frem til å møte dyktige kolleger på sensorskolering og ikke minst er jeg spent på om «mine» elever får vist hva de er gode for 31. mai. Og at de i midten av juni får vurderingen de skal. Karakteren de kommer til å huske. Hva skaper deres eksamenstekst? Det er blant annet språklig bakgrunn og erfaringer, fra hjemmet, skolen og andre steder, men også såpass umiddelbare vilkår som nattesøvn og oppgaveformuleringer. Rimelig eller ikke, så er også svaret, uttrykt i sluttkarakteren, med på å skape teksten. Både fordi eksamensteksten foregriper svaret, men også fordi det er dette svaret som blir stående.

Eksamen i norsk er en viktig institusjon, som vi trenger å ha tillit til. I den sammenheng er også åpenhet viktig. Også derfor vil jeg begynne med å takke Tone Vindegg i Utdanningsdirektoratet. Uten hennes tro på prosjektet, ville jeg ikke hatt empiri til denne undersøkelsen. I den anledning har jeg også fått god hjelp av Sissel Endung Berger hos Fylkesmannen i Buskerud. Det har også vært godt å bli veiledet. Jannike Ohrem Bakke og Dagrun Skjelbred skal ha stor takk for dytt, råd, avgrensninger, faglige påminnelser, støtte og pirk. Det har vært litt sårt å se teksten min endre seg til vitenskapelig struktur, men utholdenhetsprøven har likevel gjort meg godt. Takk skal også Ottar Hellevik ha, som tok seg tid til å lese det jeg har skrevet om kvantitativ metode og analyse. Det ga nødvendig trygghet. Fordi det er så lenge siden forelesningene nå, kunne jeg nesten ha glemt det jeg har satt aller mest pris på som masterstudent: Superflinke forelesere og erfarne medstudenter. Det var et privilegium å få gå inn i det faglige rommet – for så å kunne diskutere det opp mot vår egen praksis i pausen etterpå. Ingen nevnt, ingen glemt. Nevner bare Kjersti Heldaas, fordi du sa ordet sprikvurdering i en avgjørende fase. Takk.

Horten 7. mai 2016

Liv Cathrine Krogh

Innhold

1	Innledning: Bakgrunn for undersøkelse	6
1.1	Problemstilling og forskningsspørsmål	7
2	Teori og tidligere forskning	7
2.1	Teori om tekst og skrivekompetanse	7
2.1.1	Tekst i kontekst.....	7
2.1.2	Kategorisering av tekster.....	10
2.1.3	Literacy og eksplisitt skriveundervisning.....	11
2.2	Tidligere forskning om skriveoppgaver og vurdering av skrivekompetanse	14
2.2.1	Tidligere forskning om skriveoppgaver	14
2.2.2	Nasjonale skriveprøver og Normprosjektet.....	17
2.2.3	Tidligere forskning om elevens eksamenstekst	19
2.2.4	Normfellesskap og <i>true score</i> i vurdering av eksamenstekster	22
3	Eksamensoppgavenes kontekst: Fra sjanger til teksttype	24
3.1	Endringer av eksamensoppgavene i 2014	24
3.2	Reaksjoner på endringer av eksamensoppgavene	25
3.3	Reaksjoner på oppgavetyper kreativ tekst	26
3.3.1	Kreativ tekst som eksamensoppgave: Reflektere	29
4	Metode og materiale	32
4.1	Forskningsdesign.....	32
4.2	Svar på spørsmål 1: Dokumentanalyse og kvantitativ undersøkelse	33
4.2.1	Eksamenskonteksten: Oppgaver og grunnlag for vurdering	33
4.2.2	Innhenting, utvelging og arbeid med kvantitativt materiale.....	34
4.3	Svar på spørsmål 2: Tekstanalyse av eksamensbesvarelser	37
4.3.1	Ekstensivutvalg av eksamensbesvarelser	37
4.3.2	Intensivutvalg av sprikvurderte eksamensbesvarelser	38
4.4	Drøfting av metode og materiale	40
4.4.1	Materialets nøyaktighet og pålitelighet.....	40
4.4.2	Studentens pålitelighet: Deltakelse, roller og fortolkning	40
4.4.3	Utvalgsusikkerhet og validitet.....	42
4.4.4	Forskningsetikk.....	47
5	Analyse og resultat	49
5.1	Eksamen i norsk hovedmål 2015: Skriveoppgaver i kontekst	49
5.1.1	Eksamensoppgavene i norsk hovedmål 2015.....	49
5.1.2	Kontekst: Grunnlag for vurdering av eksamensoppgavene	54
5.1.3	Oppsummering av dokumentanalyse av eksamensoppgaver i kontekst.....	63
5.2	Kvantitativ analyse	64
5.2.1	Samlet sprikvurdering, korrelasjon og vurderingssamsvar i bruttoutvalget	64
5.2.2	Eksaminanders oppgavevalg	65

5.2.3	Sprikvurdering og oppgaver	66
5.2.4	Karakterfordeling i brutto- og oppgaveutvalg	67
5.2.5	Korrelasjon og samsvar i oppgaveutvalget	68
5.2.6	Klageutvalget og oppgaver	69
5.2.7	Oppsummering av kvantitativ analyse	70
5.3	Tekstanalyse av eksamensbesvarelser	71
5.3.1	Skrivehandlinger i ekstensivutvalget	73
5.3.2	Ekstensivutvalget og svar på oppgaven	73
5.3.3	Ekstensivutvalget og vurdering	76
5.3.4	Ekstensivutvalget oppsummert	77
5.3.5	Intensivutvalget: Situasjonsteksten	78
5.3.6	Analyser av tekster i intensivutvalget som blander teksthandlinger	79
5.3.7	Analyser av tekster i intensivutvalget som i hovedsak forteller	89
5.3.8	Analyser av tekster i intensivutvalget som i hovedsak reflekterer	100
5.3.9	Intensivtekstene oppsummert	109
6	Drøffing av resultater	110
6.1	Svar på spørsmål om sammenheng mellom oppgaver og vurdering	110
6.1.1	Sammenfatning av dokumentanalyse og kvantitativ undersøkelse	110
6.1.2	Drøffing av kvalitative funn i eksamensoppgaver og rammer for vurdering	111
6.1.3	Drøffing av kvantitative funn	115
6.2	Svar på spørsmål om sammenheng mellom tekster og vurdering	118
6.2.1	Sammenfatning av tekstanalysene	118
6.2.2	Drøffing av kvalitative funn i eksamenstekster	119
6.3	Avsluttende refleksjoner	124
7	Litteratur	126

Vedlegg

128

1 INNLEDNING: BAKGRUNN FOR UNDERSØKELSE

Eksamen i norsk hovedmål angår drøye 40.000 elever hvert år. Det er også mange som mener noe om den, i og utenfor fagmiljøet. Vurdering som spriker og varierer har potensiale til å svekke tilliten til eksamensordningen. En undersøkelse av eksamensoppgaver og -tekster som kan forklare variasjon, ble ytterligere aktualisert våren 2014. Da ble eksamensoppgavene revidert i forlengelse av læreplanrevisjonen i norsk høsten 2013. Der noen oppgaver tidligere hadde spesifisert sjanger, ble skriveoppdraget endret til ulike skrivehandlinger med tilhørende instruksjonsverb. Heretter skulle bruk av skrivehandlinger realisere oppgavens formål.

Norskfagets og skrivekompetansens posisjon blir hvert år stadfestet når hele kullet på VG3 studieforbereende har eksamen i hovedmålet, den eneste eksamenen som er obligatorisk for alle, i tillegg til at halve kullet testes i sidemålet. Det gir føringer for undervisning i norskfaget. Det er også avgjørende at eksamen gjennomføres slik at eksaminandene og samfunnet har tillit til vurderingene som blir stående. Mange tekster får heldigvis samsvar i vurdering. Andre blir dessverre sprikvurdert. Jeg vil se nærmere på tekster som skiller seg ut og får ulik vurdering. I hvilken grad normfellesskapet fungerer, kan avdekkes ved for eksempel å se på vurderingssamsvar og klagefrekvens, men også via de såkalte spriktekstene; besvarelser som to medsensorer vurderer med en ulikhet på minst to karakterer.

Sprikvurdering kan gi usikkerhet rundt vurderingen som de to sensorene skal bli enige om, noe som igjen kan svekke tillit til at eksaminanden har fått en reliabel og valid karakter. I tillegg til konsekvenser dette kan ha for eksamensinstitusjonen, vil sprikvurdering eller motstridende vurderingsnormer potensielt ha implikasjoner for den enkelte elev.

Våren 2014 fulgte jeg en avgangsklasse til eksamen etter de nye oppgavetyperne, og jeg skulle gjerne hatt lenger tid til å forberede dem på det som var nytt. I en overgangsfase begynte jeg også å tvile på om sjangerundervisningen var hensiktsmessig. Dette er noe av grunnen til at jeg ønsket å se nærmere på dette. Ønsket ble forsterket da jeg ved skolestart høsten 2015 startet norskundervisningen i nåværende VG3-kull med å se på vårens eksamensoppgaver, og det slo meg at begreper som jeg selv tar for gitt ikke nødvendigvis ikke er like gitt for elevene. Dette kom blant annet til syne da vi kom til den kreative teksten, og jeg spurte klassen om hva de forbandt med en kreativ tekst. De fleste svarte ulike varianter av fiksjonstekster. Da jeg pekte på oppgavetekstens «(...) der du *reflekterer* over tematikken», og foreslo at en kreativ tekst også kunne være en sakprosatext, stusset de fleste, mens noen umiddelbart gikk inn i sjangerrepertoaret sitt for å foreslå praktiske løsninger på hvordan en kunne løse oppgaven.

Det er ikke bare i skrive- og norskfeltet at eksamensoppgaver diskuteres. Nylig ble det hevdet at eksamensoppgavene til grunnskoleeksamen i matematikk hadde negativ innvirkning på elevenes resultater (Skårderund, 2015). Klasseroms-eksempelet over kan for så vidt illustrere hvor viktig det er å jobbe med eksamensoppgavene og -formuleringene i undervisningen, herunder avklaring av forholdet mellom sjanger og skrivehandlinger. Disse erfaringene ga retning. I denne masteroppgaven vil jeg konkret undersøke sprik i sammenheng med oppgave- og teksttyper, og hvordan sprik kan forklares ut fra elevenes eksamenstekster.

1.1 PROBLEMSTILLING OG FORSKNINGSSPØRSMÅL

Problemstilling:

I hvilken grad er det sammenheng mellom sprikvurdering og annen variasjon i vurdering, og trekk ved eksamensoppgaver og -besvarelser?

Forskningsspørsmål:

1. Er det sammenheng mellom oppgavetyper og variasjon i vurdering?
2. Er det trekk som er utpreget i eksamenstekster som har utfordret normfellesskapet spesielt?

2 TEORI OG TIDLIGERE FORSKNING

2.1 TEORI OM TEKST OG SKRIVEKOMPETANSE

Tekster blir til i en historisk og sosial kontekst. Undersøkelsen handler også om tekster i kontekst, og jeg presenterer derfor noe teori om dette. Sosiokulturell læringsteori retter søkelyset mot funksjon. For skrivingens del handler det om funksjonen skriving har for individet, og at det posisjonerer seg til andre via tekst (Smidt, 2009, s. 42), slik eksaminanden til sensor via eksamensteksten. Bakhtins syn på tekst som ytring i dialog med andre, Hallidays grammatikk for språk i bruk og retorikk er viktige teoretiske perspektiv i skriveopplæring (Smidt, Tønnessen, & Aamotsbakken, 2011, s. 11). Når eleven skriver, er dennes bakgrunn og literacy også del av tekstens kontekst. Derfor anvender jeg også skriveteori fra dette feltet.

2.1.1 Tekst i kontekst

Mikhail Bakhtin fremhever språkets rolle som kommunikasjonsmiddel. I våre dager kan det virke selvsagt at tekster forholder seg til ytringer før og etter seg, men ikke da *Spørsmålet om talegenrane* ble utgitt i 1953-4. Her knytter Bakhtin ytringen til situasjonen den oppstår i.

Dagens funksjonelle språksyn har røtter her, og fordi ytringen oppstår i en kontekst, finnes det alltid førende sjangrer. «Ytringar og ytringstypar, dvs. talegenrar, er overføringsreimene som fører frå samfunnshistoria til språkhistoria» (Bachtin, 2005, s. 7). Eksamensoppgavene i norsk har lenge inneholdt konkrete sjangerinstruksjoner. Disse forsvant tilsynelatende ved endring til teksttyper i 2014, men sjanger er fortsatt viktig i skolens språk- og skriveundervisning og på eksamen. På Bakhtins tid var det vanlig at kommunikasjonsteori ga mottaker en passivt lyttende rolle, men Bakhtins mottaker svarer. Enten umiddelbart, eller ved «(...) ei svarande forståing med tidsforseinking» (op.cit, s. 11). Dette er avsender innstilt på, for i ytringskjeden er denne også mottaker. Språket blir dialogisk (op.cit, s. 14). I dag praktiserer vi et slikt språksyn når vi er opptatt av at skriveoppgaven skal ha et klart formål, gjerne operasjonalisert gjennom en mottaker. En kan også forvente at eksamensteksters avsendere er denne dialogen mer eller mindre bevisst. Uavhengig av om adressaten er ubestemt eller deltar i situasjonen, er ytringens adressivitet det sentrale prinsippet i Bakhtins fortolkningsteori (Slaatelid i op.cit, s. 53). Komposisjon og stil bestemmes av adressaten: «Når eg lagar mi ytring freistar eg å gisse det kommande svaret, og samtidig freistar eg å føregripe det» (op.cit, s. 40). Dette kommer tydelig frem når en jobber med hensiktsmessig språk, struktur og innhold i eksamensskrivning, i tillegg til hvordan eksaminandene synes å tilpasse sine tekster til mottakerne.

I 1984 bidro Carolyn Miller ytterligere til å plassere sjangerstudiet innenfor kulturstudiet. I en innledning til sin oversettelse av Miller, skriver Kjell Lars Berge at hun foreslo å se sjanger som en *samfunnskonstituent*, en av institusjonenes maktstrukturer. Det er nærliggende å se eksamensoppgavene som en slik samfunnskonstituent. Miller kaller sjanger en handling som får mening fra situasjonen og dens sosiale kontekst, og som gjentatte mønstre for språkbruk, bidrar sjangrene til å skape kulturelt innhold (Miller, 2001, s. 32). En sjanger er et retorisk redskap for å formidle private intensjoner og sosiale påtrengende problemer. Den kopler det private med det offentlige, enkelttilfellet med det som gjentas (op.cit). En kan si at sjangeren er det offentlige mønsteret en tar i bruk for å formidle sitt eget, men fordi det individuelle hvis det gjentas kan bidra til at mønsteret endres, vil det i sin tur endre sjangerreglene. Dette er ikke helt ulikt det Bakhtin sier når han kaller sjangerformene mer plastiske og friere enn språkformene, og at jo bedre en mestrer sjangrene, jo friere kan de brukes (Bachtin, 2005, s. 25). I sin tur er det slik at når en i skriveundervisningen jobber med sjangerrepertoar, er noe av formålet at elevene skal ha mulighet til å velge hensiktsmessige sjangere eller trekk for å realisere ulike skriveoppgaver eller -handlinger, samtidig som adressaten, her sensor, bør ta høyde for plastisk bruk av eller brudd med dem. Endringene av eksamensoppgavene i norsk i

2014 fra sjangre til oppgavetyper, overlater eksaminandene til selv å velge sjanger ut fra oppgavens formål, noe jeg kommer tilbake til senere i kapittelet.

Eksamenstekstens formål innebærer blant annet å skape mening. Michael Hallidays funksjonelle grammatikk deler meningssskaping i tekst inn i *ideational*, *interpersonal* og *textual metafunction* (Halliday, 2014, s. 30-31). Med norske begreper er dette å se muligheter for både *ideasjonell*, *mellompersonlig* og *tekstuell* mening (Maagerø, 2005, s. 100). En kan identifisere flere aspekter ved mening i språk, ikke ved å velge ut et ord eller en frase, men ved å *ta hele ytringen opp igjen* og nok en gang analysere helheten fra et nytt perspektiv:

Hver setning i en tekst er multifunksjonell. De forskjellige typer mening er vevd sammen i en veldig tett struktur og på en slik måte at vi ikke kan se separat på de forskjellige delene for å forstå dem. Isteden ser vi på helheten samlet fra forskjellige vinkler, og hvert perspektiv bidrar til den samlede tolkningen. Dette er et avgjørende trekk ved en funksjonell tilnærming (Halliday i Berge et al., 1998, s. 89).

Bakhtin var også motstander av oppstykket analyse, fordi ytringen må ses under ett og som del av ytringskjeden, for ikke å miste spor av adressiviteten og det foregrepne svaret (Bakhtin, 2005, s. 43). Halliday kaller dette forholdet mellom teksten og situasjonskonteksten. Kontekst er et viktig begrep hos Halliday, og med faglige røtter via J.R. Firth til Bronislaw Malinowski (Halliday, 2014), deler han begrepet i to, *kulturkontekst* og ulike *situasjonskontekster* innenfor det kulturelle domenet (op.cit, s. 33). Meningen blir utviklet og forstått innenfor en kulturkontekst, og realiseres i konkrete situasjonskontekster der den fyller en funksjon. Eksamenskonteksten er eksempel på en slik situasjonskontekst. Her står systemisk funksjonell grammatikk i relasjon til translingvistikken. I sin lesning av Bakhtin, bidro Julia Kristeva med begrepet *intertekstualitet* (Kristeva, 1969) for å uttrykke tekstens relasjon til konteksten, hos Atle Kittang forklart som: «(...) korleis diktverket alltid er vove saman med dei historiske og sosialt betinga meiningssamanhengane det er omslutta av» (Kittang, 1975, s. 52). Tekstens forhold til situasjonskonteksten kan via Halliday utforskes gjennom kategoriene *field*, *tenor* og *mode* (Halliday 2014), gjerne oversatt til *felt*, *relasjon* og *mediering* (Maagerø, 2005, s. 44). Feltet i diskursen uttrykkes gjennom den erfaringsbaserte funksjonen i semantikken, relasjon gjennom den mellompersonlige og mediering gjennom den tekstuelle (Halliday i Berge et al., 1998, s. 90-91). Dette handler om mye av det samme i møte med en ny situasjon:

Du tildeler konteksten et felt idet du legger merke til hva som foregår, en relasjon ved at du gjenkjenner de personlige forbindelsene som er involvert, og en mediering ved å se hva som oppnås gjennom språket (...) Ellers ville det være umulig å forklare hvordan vi i det virkelige liv så raskt klarer å slutte oss til og delta i en situasjon som vi tidligere ikke visste noe om (op.cit, s. 94).

Læring skjer altså sammen med andre, i og mellom individer og omgivelser. Språk og tekst kommuniserer, og måtene vi handler med språket på kan forklares ut fra funksjon. I analyse

finnes ideasjonell mening ved å se på innholdet i ytringen. Språket representerer verden gjennom prosesser, deltakere og omstendigheter (Maagerø 2005, s. 102). Prosesser kan indikere skrivehandlinger, og virkelighetsfremstillinger kan være mer eller mindre relevante som oppgavesvar. Relasjonell mening ses i språket som samhandling med andre, avsenders og mottakers taleroller, og gjennom modus og modalitet. Taleren kan tilby, konstatere eller kreve, og dette kan adressaten reagere på gjennom å godta eller motsi, svare eller avvise (op.cit, s. 137). Eksaminanden kan skrive på en kategorisk eller nyansert måte, noe sensor på sin side kan honorere eller ikke. Tilslutt er tekstuell metafunksjon meningen som skaper sammenheng i teksten, og gjør den effektiv ut fra hensikten og omgivelsene (op.cit, s. 159). Tekstuell mening realiseres gjennom temamønstre og kohesjon, og det er snakk om å finne sammenheng både til verbale omgivelser (koteksten), som her vil være oppgaveteksten og tekstvedlegg, og til ikke-språklige forhold utenfor teksten; konteksten. *Metafunksjonell* tekstanalyse kan utdype både form- og innholdsmessige trekk i for eksempel eksamenstekster. Eksempelvis kan den mellompersonlige metafunksjonen brukes til å analysere hvordan teksten realiserer mening som skal vurderes av en adressat, og relasjonen denne har til situasjons- eller kulturkonteksten. Når teksten leses mellompersonlig, kan en for eksempel identifisere mening og sannsynliggjøre respons, noe jeg kommer tilbake til i analysen.

2.1.2 Kategorisering av tekster

Forhold mellom tekst og kontekst kan vise seg i måten teksten kategoriseres, og her er flere innfallsvinkler. Skillet mellom sakprosa og skjønnlitteratur er en vanlig todeling «(...) som gjelder forholdet til virkeligheten» (Tønnesson, 2012, s. 17). Det er sammenhengen som danner rammen for adressatens forventninger. Det er ikke slik at romanen må være fiktiv, men vår tekstkultur har en uskreven kontrakt som fastsetter at den er det. Sjangerbetegnelsene hjelper å plassere ytringene på hver sin side «(...) av skillet som gjelder forholdet til virkeligheten» (ibid.). Sakprosa kan også være litterær og uansett hvor objektiv en fremstilling er, vil den alltid foregripe svaret (op.cit, s. 22). Her viser Tønnesson til Bakhtin.

En annen innfallsvinkel er tekst- eller skrivehandlinger. *Skrivehjulet* kategoriserer i skrivehandlingene i å *overbevise, samhandle, reflektere, beskrive, utforske og forestille seg* (Skrivesenteret, 2013). *Skrivehjulet* ble utarbeidet som ressurs for lærere i arbeidet med nasjonale skriveprøver, som blir presentert i kapittel 2.2.2. Modellen understreker skrivingens formål, og ved å stille opp nevnte skrivehandlinger «(...) unngikk modellskaperne å ta utgangspunkt i sjangrer» (Smidt, 2014, s. 210). Dagens læreplan i norsk uttrykker et slikt funksjonelt tekstsyn, og da gjennom *fortellende, beskrivende, forklarende, argumenterende* og

reflekterende skrivehandlinger, som fra 2014 også er brukt i eksamensoppgavene. Ulike teksthandlinger utelukker ikke hverandre (Skjelbred, 2014, s. 36). Hensikten med å betegne dem er å poengtere at tekster brukes til ulike formål (ibid). Imidlertid kan skillet som angår forholdet til virkeligheten bli uklart ved slik typologisering, og jeg vil komme tilbake til hvordan dette kan utfordre adressatens forventninger og lese måte.

Sjanger skiller seg fra skrivehandlinger ved at de både er tekstnormer og -mønstre som allerede finnes, som er perspektiv fra blant annet Bakhtin og Miller. Dagrun Skjelbred skriver at teksthandlingene er på et overordnet nivå fordi en kan bruke flere ulike sjangre til å realisere dem (op.cit, s. 38). Revidert læreplan i norsk toner sjangerarbeid ned og vektlegger det funksjonelle. «Det må imidlertid også sies at sjangerkunnskap og tekstmønstre gir god hjelp til å utforme tekster som fungerer» (op.cit, s. 40), skriver Skjelbred. Det som kan forstås som en motsetning mellom skrivehandlinger og sjangere vil bli omtalt nærmere i kapittel 3. Undersøkelsen vil etter hvert handle om tekster der én av skrivehandlingene skal realiseres, og denne vil jeg derfor undersøke nærmere i kapittel 3.3.1.

2.1.3 Literacy og eksplisitt skriveundervisning

Kunnskapsløftet er «(...) fundert på teori om literacy og språklige nøkkelkompetanser» (Solheim & Matre, 2014, s. 76). Også gjennom skriving skal eleven myndiggjøres til å delta i offentligheten. Internasjonalt brukes begrepet *literacy* om evnen til å forstå, bruke og produsere tekster på meningsfull måte (Smidt et al., 2011), og det er blitt vanlig å bruke begrepet uoversatt i norsk (Skaftun, 2009, s. 33). Kunnskapsløftets vektlegging av grunnleggende ferdigheter i alle fag gjør den til en literacy-reform (Berge, 2005b). Det kan også ses i sammenheng med OECDs identifisering av nøkkelkompetanser, *DeSeCo* (OECD, 2005), med det formål at unge mennesker skal mestre samfunnets utfordringer og delta i en kritisk offentlighet (Smidt et al., 2011 s. 11). Dette henger også sammen med dannelsesbegrepet, som defineres slik av Laila Aase: «En sosial prosess som fører til at man forstår, behersker og kan delta i vanlige, oppvurderte kulturformer. Dette innebærer både tenkemåter, handlingspotensial og kunnskaper innenfor et bredt og variert felt» (Aase, 2005).

David Barton og *The New London Group* har vært sentrale i forståelsen og utviklingen av literacybegrepet. Dagrun Skjelbred og Aslaug Veum sammenfatter literacy ved å forstå det som «(...) en kyndighet som gjør den enkelte i stand til å tolke og selv benytte en rekke semiotiske ressurser for å gjenskape og produsere tekster, slik at han eller hun kan leve og utvikle seg i et komplekst samfunn». Begrepet rommer tilgang til tekster, og til samfunnet gjennom tekster, i tillegg til kritisk tekstrefleksjon (Skjelbred & Veum, 2013, s. 19). Jeg har

referert lingvistisk teori som forstår tekster som produkt av konteksten de blir til i, og følgelig avhenger det å takle tekster også av sosial og kulturell kunnskap. Dette handler også om literacy (Solheim, Sjøhelle, & Matre, 2012), fordi elevenes bakgrunn virker inn på skrivingen deres. Mary Macken-Horarik (1996 og 1998, her lest hos Tønnessen og Maagerø 2015) har beskrevet tre domener for læring, også av språk: *Dagliglivets* som omfatter hjem og nærmiljø, det *spesialiserte* som angår fagområdene på skolen, og det *kritiske* eller *refleksjonsdomenet* i og utenfor utdanningsinstitusjonene. I dagliglivets domene lærer en primært gjennom muntlig kommunikasjon. Hvilke tekster unge møter her, avhenger av aktivitetene de driver med og motiveres til. Noen har sammenfall mellom dagliglivets og skolens tekster, og for disse er ikke skolen fremmed, mens mange vil oppleve kulturforskjell (Tønnessen & Maagerø, 2015, s. 27). Dagliglivets språk er muntlig, konkret og uformelt, ordforrådet enkelt og dagligdags med mange forsterkende ord. I det spesialiserte domenet systematiseres erfaringer, objekter og fenomener i over- og undergrupper, og en finner fagterminologi, abstraksjon, formelle uttrykksmåter, generalisering av enkelterfaringer, detaljkunnskaper og nyanseringer. Det kritiske eller refleksjonsdomenet er der en drøfter og evaluerer kunnskapen fra de to andre domene. Språket blir brukt til å reflektere. Slik kritisk literacy utvikles i skolen og hjem der det stilles spørsmål til absolutte sannheter (op.cit, s. 28). I tekstanalysen vil jeg også undersøke om ovennevnte spor av språklige domener kan bidra til å forklare ulik vurdering.

Literacy er å beherske diskurs, sier James Paul Gee (2003), som også har vært sentral i forståelse av tekstkultur. Gee, her lest hos Atle Skaftun (Skaftun, 2009, s. 37) sier at literacy er diskursive felt vi beveger oss inn og ut av, eller står utenfor. Dagliglivets domene kan sies å ha en del felles med Gees begrep *primærdiskurs*, der domenet refererer til hva som påvirker språket som individet tar i bruk der, og fordi de respektive domene er forskjellige, er det ikke samme primærdiskurs ulike folk praktiserer. Sylvi Penne skriver at årsak til forskjeller i skolen er forskjell i primærdiskurser (Penne, 2010, s. 35), fordi skolens diskurs likner primærdiskurs i lesende middelklasse. I tillegg møter middelklassebarn tidligere *sekundærdiskurser* gjennom hyppige møter med skriftspråket, og det er literacy. Skolen handler om sekundærdiskurser, og igjen har de som kjenner dette hjemmefra et fortrinn. Penne problematiserer skolens toleranse for alle språklige ytringer, for eksempel det narrative og subjektive, fordi det bidrar til primærdiskursens utvidete virkeområde, noe svakere elever taper på (op.cit, s. 42). Tidligere forskning i grunnskolen har også vist at skolsk sekundærdiskurs har tapt terreng. I et literacy- og myndighetsperspektiv ses det som bekymringsfullt at store forskningsprosjekter som SKRIV og KAL, som jeg kommer nærmere

inn på i 2.2, har avdekket at vekten har ligget på litterær og personlig skriving, samtidig som skriving i skolen fortsatt preges av fragmenterte skriveoppgaver som i liten grad er rettet mot mottakere og formål (Smidt et al., 2011). Didaktiske konsekvenser er at utdanning må utvikle refleksjonskompetanse, noe som «(...) forutsetter en literacy-praksis som gjør det å snakke om og med tekster interessant og meningsfylt» (op.cit, s. 28).

Skillet mellom skolens praksis og hva nylig forskning har funnet hensiktsmessig, reflekterer uenighet i fagmiljøet mellom prioriteringen (Smidt, 2014, s. 211) av elevens prosjekt og stemme – og samfunnets sjangrer og krav. Et bakteppe er 1980-tallets dikotomi i skrivefeltet mellom prosessorientert skriving og eksplisitt sjangerundervisning (ibid.). Miller kalte sjanger en handling som kopler det private med det offentlige, så en skulle tro at perspektivene ikke er uforenelige. Et spørsmål som melder seg er hvilket domene som preger skolens literacyarbeid, eller snarere tekstene elevene skriver til eksamen etter avsluttet skolegang. Bakhtin sa at adressaten styrer ytringens innhold og form, men i en skole der skolsk sekundærdiskurs blir mindre tydelig, er det kanskje ikke lenger gitt at samme skriveoppgave får samme adressat. Med det mener jeg at noen elever i eksamenskonteksten skriver for og til sensor og slik underordner seg samfunnets sjangrer og krav, mens andre kan vektlegge andre relasjoner. Tekstanalysen kommer tilbake til dette, og i det følgende beskriver jeg skrivepedagogisk teori som henger sammen med literacyperspektivet.

Jon Smidt skriver at eksplisitt skriveundervisning er en viktig trend i skrivepedagogikken (op.cit, s. 203). Teoretisk grunnlag er langt på vei hentet fra australsk sjangerskole. Frøydís Hertzberg skriver at av fire sentrale skrivepedagogiske forskningsposisjoner, tillegger sjangerskolen formkunnskap mest vekt. Det går tilsynelatende mot strømmen av all empiri som går i disfavør av grammatikkundervisning som middel til skrivekompetanse (Hertzberg, 2001, s. 94). Men sjangerskolen integrerer funksjonell grammatikk i «(...) allsidig og sterkt kontekstualisert» sjangerundervisning, for at elevene skal beherske en rekke sjangrer som samfunnet krever (op.cit, s. 96). Metodikken har et sosialt perspektiv og hevdes å møte barn fra hjem der det ikke leses og skrives mye. Dette handler om å jevne ut ovennevnte literacy-forskjeller, blant annet fordi metodikken er basert på studier av hvordan barn lærer sammen med voksne, og fordi en bruker modelltekster tilpasset elevens nivå (op.cit, s. 101).

Det hersker enighet i sentrale forskningsmiljøer i Norge om at blant annet modelltekster er viktig i lese- (Tønnessen & Maagerø, 2015) og skriveopplæringen. Både ved at faste tekstmønstre fungerer som læringsstøtte (Helstad og Hertzberg i Skjelbred & Veum, 2013), og ved eksplisitt undervisning i skrivestrategier. Skrivestrategier kan være fremgangsmåter

skrivere bruker for å gjennomføre et skrivearbeid (Hertzberg, 2006). Modellering kan brukes til å utvikle elevene til strategiske skrivere. Begrepet er her forstått som lesing, vurdering og diskusjon av eksempeltekster, pluss modellering og veiledning i skriveprosessen (Kringstad & Kvithyld, 2014, s. 15). Nylig publiserte også Skrivesenteret flere eksamensbesvarelser som modelltekster til bruk i opplæringen (Metliaas & Solbu, 2015). På bakgrunn av perspektivet over vil elevene få best utbytte av ressursen om den brukes og diskuteres i klasserommet.

En kan altså si at mye teori betoner viktigheten av et funksjonelt tekstsyn. Formålet med skrivingen er også vektlagt i læreplanen i norsk, men det knyttes tvil til i hvilken grad dette er implementert i skolens skrivepraksis. Oppgavene i sentralgitt skriftlig eksamen i norsk kan leses som myndighetenes tolkning av egen læreplan. Det blir interessant å se hvordan eksaminandene balanserer egen stemme med samfunnets krav til skriving, og om vurderingen av oppgavene deres kan avdekke eventuell usikkerhet eller ambivalens til dette i skolen.

2.2 TIDLIGERE FORSKNING OM SKRIVEOPPGAVER OG VURDERING AV SKRIVEKOMPETANSE

Gyldighet er en sentral utfordring når det kommer til vurdering (Evensen, 2010). Utfordringer knyttet til lavt sensorsamsvar drøftes ikke bare innenfor grunn- og videregående skoler i Norge, men også innenfor høyere utdanning (Bjølseth, Havnes, & Lauvås, 2011) og internasjonal forskning: IEA (International Studies in Educational Achievement) satte seg i 1988 fore å utvikle skriveprøver som kunne måle elevers skriveferdigheter på tvers av skoleslag og 14 deltakerland (Gorman, Purves, & Degenhart, 1988). Forskerteamet fant fellesnasjonale og brukbare ferdighetskriterier, men blant annet vektingen av ferdighetene var så ulike mellom land at man kom frem til at teksters kvaliteter må vurderes lokalt, eller «(...) det man i IEA-studien kalte *rhetorical communities*» (Berge, 2009, s. 44). Dette kan tjene som begrunnelse for at jeg i det følgende vil konsentrere meg om forskning rundt norske forhold. Først ser jeg på aktuelle utfordringer knyttet til utarbeiding av skriveoppgaver og vurdering av skrivekompetanse, før jeg ser på funn knyttet til hvordan eksamen er vurdert.

2.2.1 Tidligere forskning om skriveoppgaver

Skriftlig eksamen i norsk hovedmål inneholder skriveoppgaver. I *Å invitere elever til skriving. Ulike perspektiver på skriveoppgaver* (Otnes, 2015), tar flere til orde for det mulige paradokset at det er forsket såpass lite på en så viktig del av elevenes læring. Jorun Smemo og Hanne Solem skriver blant annet at det fortsatt er lite forskning på skriveoppgaver og deres «(...) betydning for elevenes besvarelser og vurderingen av disse» (Smemo & Solem, 2015, s. 63). Det er

likevel sannsynlig at selve skriveoppgavene har innvirkning på vurderingen av dem, og at valid og reliabel eksamensavvikling også kan avhenge av oppgaveformuleringer.

Egil Børre Johnsen, som i sin doktoravhandling fra 1994 foretok en historisk gjennomgang av artiumsstilens eksamensoppgaver, skriver at «(...) formuleringsproblemet blir sjelden gjenstand for generell drøfting fra rådets side i sensorrapportene» (Johnsen, 1994, s. 161), men at sensorrapporten i 1955 er alene om å ha et prinsipielt kapittel om oppgaveformulering, der Johnsen siterer følgende utsagn (og etterlysning) fra en sensorgruppe:

Jo klarere oppgaven er formet, jo mindre blir sensorene bebyrdet under vurderingen av besvarelsene og justeringen av karakterene. Hvis oppgavens form ikke er helt klar, vil de forskjellige sensorers og sensorgruppers vurdering av besvarelsene kunne avvike mer eller mindre fra hverandre. Og dette kan føre til at motiveringen for karakterene kan bli mer uensartet enn den bør være. Dette kan igjen svekke tilliten til en stabil og pålitelig sensur (Johnsen, 1994).

I våre dager er det avdekket at formålet ved skrivingen er lite vektlagt i oppgaver. SKRIV-prosjektets hovedfunn var at formål sjelden får nok oppmerksomhet i skoleskriving (Aasen, Solheim, & Smidt, 2011, s. 14). Formål er imidlertid så sentralt at når skriving skal vurderes, bør vurderingskriterier innrettes for å vurdere både innhold, form og bruk ut fra nettopp formål. Dette er også sentralt i Normprosjektets utvikling av felles forståelse av forventninger til skriving (op.cit, s. 21). Formålsrettet skriving henger også sammen med sjanger. Formål og sammenheng skaper sjangeren, så «(...) sjanger handler like mye om *innhold* og *bruk* som om form» (op.cit, s. 27). Her viser Smidt til blant annet Bakhtins translingvistikk, og han oppfordrer til å jobbe med modeller for å bevisstgjøre om sjangertrekk, men uten sjangerformalisme (op.cit, s. 28). Eksamensoppgavene har vært uten eksplisitte sjangerinstruksjoner siden 2014, og jeg vil i kapittel 3 komme tilbake til en del av reaksjonene denne endringen fikk.

I en gjennomgang av empiri, skriver Hildegunn Otnes at *skriveoppgavens* hensikt er å dokumentere hva elevene har lært, eller at de skal utvikle seg som skrivere (Otnes, 2015, s. 11). På spørsmålet om hva oppgaven bør inneholde, svarer mange at «(...) det er avgjørende å ha med kontekstuelle beskrivelser i oppgaveteksten. Særlig er formålet med skrivingen fremhevet» (Otnes, 2015, s. 22). I tillegg er det mest markante skillet i forskningslitteraturen mellom dem som fremhever formål som avgjørende og dem som ikke nevner det (ibid.).

Erfaringer fra *Normprosjektet*, som blir beskrevet nærmere i 2.2.2, har vist at «(...) skriveoppgavene la avgjørende føringer for elevenes prestasjoner» (Solheim & Matre, 2014, s. 77-78). Jo mer upresist oppgaven er formulert, jo vanskeligere er den å vurdere. Det var også sensorgruppen fra 1955 opptatt av. Men hva kjennetegner en tilstrekkelig presist

formulert oppgave? Anne Holten Kvistad og Jorun Smemo (Kvistad & Smemo, 2015) har funnet to fellestrekk ved oppgaveformuleringer der de fleste har produsert gode tekster: Det første var at oppgavene inngikk i et større undervisningsopplegg, noe som ga elevene felles referanser og noe relevant å skrive om. Det andre var at oppgavene ga detaljerte krav til innhold og struktur. Elever profiterer altså på eksplisitte forventninger (op.cit, s. 239). Dette funnet tar jeg med til analysen av eksamensoppgavene fra 2015, der jeg undersøker i hvilken grad de viser til felles referanser og hvor tydelige innholds- og strukturkravene er. Forskerne fant også at slike oppgaver ga like tekster. I tillegg ble skillet mellom de beste og de andre skriverne visket ut, noe som tyder på at slike oppgaver ikke gir nok utfordringer til de beste (ibid.). Dette kommer jeg tilbake til i drøftingen, der jeg også vurderer relevans disse funnene fra grunnskolen kan ha i møte med eksamensoppgaver etter videregående.

Da jeg skrev om kategorisering av tekster over, var jeg inne på hvordan sjanger som et nivå under skrivehandlinger i større grad kommuniserer skillet mellom sakprosa og litteratur. I en undersøkelse av tekster fra utvalgsprøvene er det avdekket hvordan vurdering blir vanskelig når sjanger ikke er oppgitt (Skar & Aasen, 2015). Dette gjelder spesielt for de tekstene som tar i bruk uvanlige sjangere for å realisere skrivehandlingen (op.cit, s. 2), og at valg av lese måte bestemmer om teksten leses som sakprosa eller litterært, noe som igjen fører til sprikvurdering (op.cit, s. 1). Dette funnet tilsier at oppgaveformuleringer også kan ha sammenheng med sprikvurdering, noe som kommer til syne i analysen.

Det er altså funnet sammenheng mellom skriveoppgaver og prestasjoner, og mellom skriveoppgaver og sprikvurdering. I den forbindelse kan det være interessant å ta med at Berge i 1994 konkluderte med at sensorenes normfellesskap varierte med oppgavetypen (Berge, 1994, s. 254). Fellesskapet var best utviklet (korrelasjonsnivå 0.64) for såkalt tradisjonelle skolestiler og saksorienterte oppgaver, og dårligst for tekstanalyse (0.51) eller sjangerskriving (0.43). Denne siste store undersøkelsen av eksamen etter videregående skole konkluderte med at prøven ikke var pålitelig og hadde for dårlig begrepsvaliditet. «I det hele tatt er bedømmermønsteret så ulikt fra oppgavetype til oppgavetype at det er rimelig grunn til å hevde at de ulike oppgavene egentlig er kvalitativt ulike prøver (...). Skal den avsluttende prøven i norsk skriftlig bli pålitelig, forutsetter det en betydelig bedre definert begrepsvaliditet og at bedømmerne skoleres i denne begrepsvaliditeten» (op.cit, s. 254). KAL-undersøkelsen, som utdypes i 2.2.3, fant også høyest korrelasjon i oppgaver de færreste valgte, nemlig de som var innrettet mot sak (Vagle og Evensen 2005, s. 173). Dette funnet kan støtte behovet for et pensum- eller saksrelatert skriveoppdrag, noe jeg kommer tilbake til i analysen av

eksamensoppgavene. Når det kommer til eksamen som skriveprøve, ble det også innført nye begreper våren 2014, noe som igjen har utfordret nevnte begrepsvaliditet blant lærere og sensorer. Det er uenighet om det har vært hensiktsmessig å erstatte sjangerinstruksjer i oppgavene med *tekst-* eller *oppgavetyper*, i tillegg til at nyskapningen *kreativ tekst* er omdiskutert. Jeg vil i kapittel 3 presentere endringene og fagmiljøets diskusjoner. I analysen undersøker jeg også hvordan normfellesskapet i 2015 håndterte de ulike oppgavetyperne.

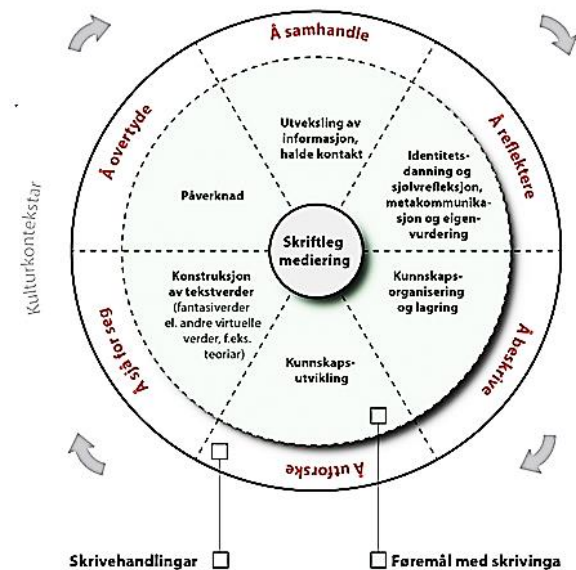
2.2.2 Nasjonale skriveprøver og Normprosjektet

Berge setter arbeidet med de nasjonale skriveprøvene i sammenheng med *Kunnskapsløftets* grunnleggende ferdigheter i alle fag: «For første gang i norsk skolehistorie var det slått fast at det å forstå, lære og utøve et fag ikke kan ses uavhengig av det å skape mening med språket», ergo at elevene settes i stand til å «(...) utøve fagrelevant skriving, lesing og muntlighet» (Berge, 2005, s. 2), men også egenutvikling og reflekterende selvkommunikasjon (ibid.). Berge deltok i planarbeidet og forskergruppen som jobbet med å utvikle prøvene, og erfarte at det er vanskelig å få valide og reliable resultater av skriveprøver.

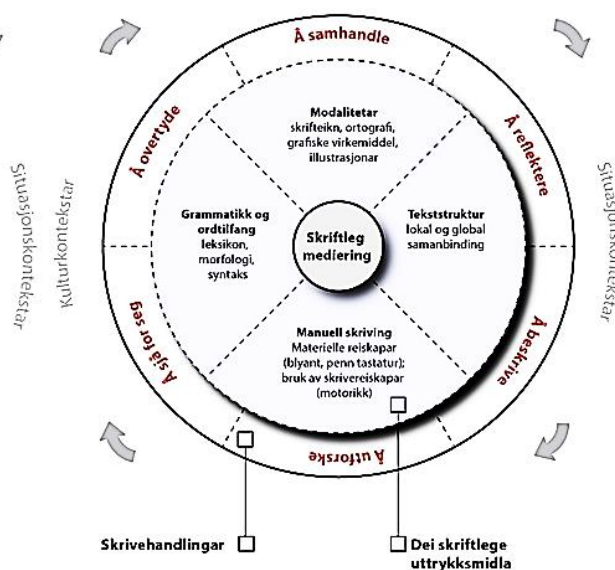
Berges hovedpoeng er at skrivekompetanse er kompleks og krever store vurderingsressurser. Det kan ikke generaliseres ut fra én type skriving, og vurderingskriterier må også være innrettet mot kompleksiteten i antall skrivemåter og del av en vurderingsveiledning med eksempeltekster. Tolkningsfellesskap må utvikles systematisk ved at lærerne får tilgang til kvalitetssikret dokumentasjon av elevens kompetanse innenfor de ulike delferdighetene, ut fra om de er mye bedre eller dårligere enn en kan forvente, eller som en kan forvente, ut fra et normativt syn på hvilke former for skriving samfunnet vil ha (Berge, 2005b, s. 16). Det er her samfunnets behov som prioriteres, og det allerede i grunnskolens skriving. Det henger sammen med *Kunnskapsløftet* som literacyreform med prioritert plass til fagskriving i læreplanen. Berge skriver videre at det har brakt norsk skriveopplæring tettere mot tidligere omtalt australsk sjangerskoles vekt på eksplisitt læring. «En genreorientert skriveopplæring vil tvinge elevene til å lære å hanske med faglige kunnskaper og egne erfaringer gjennom skrivemåter de må tilegne seg på skolen. Muntlig fortelling vil ikke lenger være en privilegert normressurs for elevenes skriftlige tekster» (op.cit, s. 20).

Skriveprøvene i 2005 viste svak pålitelighet på 0.42 (Berge, 2014, s. 8). Forklaringen var mangelfullt tolkningsfellesskap (Evensen, 2014, s. 250), noe som igjen ble forklart med at det ikke har vært prioritert i lærerutdanningen (op.cit, s. 254). Når den nye lederen, Lars Sigfred Evensen, presenterer det som heter *utvalgsprøvene* noen år senere, er det bestemt at skrivekompetanse som skal vurderes er håndskrift, ortografi, tegnsetting, ordforråd,

grammatikk og tekststruktur (Evensen, 2010, s. 14). Erfaringer fra utviklingsarbeidet fra 2003-2005 gjorde også at det sosiale perspektivet ble lagt til grunn i prøvene, for at det skulle bli klart for elevene hva de skal bruke kompetansen til. Hjulet kan ses som «(...) en slags generalisering over hva vi kan bruke skriving til i klasserommet og samfunnet» (Bakke, 2014, s. 73). Kanskje også derfor har de kulturelle skriveformålene fått en sentral plass i *Skrivehjulet*, som en kan se av Figur 1. Skriftlig mediering er viktig og utgjør navet i hjulet:



Figur 1: Framheving av funksjonelle sider ved skriving. (Skrivesenteret, 2013).



Figur 2: Framheving av semiotiske medieringsressurser

I en annen operasjonalisering av *Skrivehjulet* (Figur 2), deles semiotiske uttrykksredskaper i fire; modaliteter, tekststruktur, manuell skriving og grammatikk og ordtilfang. Evensen foreskriver en lese måte som går utenfra og innover for å lese semiotiske ressursar i lys av situerte skrivehandlingar eller ytringer. En slik lese måte er av «(...) vesentlig betydning for vurdering», ved at ulike skrivehandlingar medfører ulike relevans og vekt på dei ulike ressursane (op.cit, s. 22). Denne måten å dele opp skrivekompetansen på, kalles en primærtrekksrelatert strategi, og den er videreført i *Normprosjektet*, et vurderingsprosjekt rettet mot grunnskolen der Evensen også deltar på forskningsiden. Vurdering ut fra primærtrekk begrunnes med at bedre presiserte læringsmål gir mer pålitelighet, og at «IEA-studien viste at mål- eller nivåformuleringar må følges opp av gode eksemplar som viser oppnådd nivå» (op.cit, s. 23). I tillegg kommer at både en intuitiv/holistisk og analytisk vurderingsstrategi vil kunne undervurdere skrivehandlingenes betydning.

Normprosjektet skal gi både kunnskapsgrunnlag for skriveprøvene og utvikle tolkningsfellesskap ved hjelp av eksplisitte forventningsnormer til syv *vurderingsområder*, hvorav de to første er overordnet og knyttet til skrivehandling og formål (Solheim & Matre, 2014, s. 81). Forventningsnormene som skal utvikles er eksplisitte, fordi manglende tolkningsfellesskap relateres til tradisjonen med holistisk vurdering, der lærernes erfaring går fremfor gitte kategorier (ibid.). Sentrale mål i dette vurderingsprosjektet er å utvikle og prøve ut nasjonale normer for skriving, finne ut hvilke konsekvenser bruken av normene har for utviklingen av kvalitet i elevtekster og for lærernes vurderingsarbeid. Gyldig vurdering henger blant annet sammen med at normfellesskapet blant skriveledere er så godt utviklet som mulig, ved at det finnes felles normer og forventninger til tekstkvalitet.

2.2.3 Tidligere forskning om elevens eksamenstekst

Eksamenstekster fra VG3 er i liten grad studert. Jeg har omtalt to doktoravhandlinger fra 1990-tallet om eksamen i videregående skole. Den ene rettet seg i hovedsak mot oppgavene (Johnsen, 1994) som ble gitt, og den andre mot vurderingsnormer (Berge, 1996). Berge hadde da også gjennomført en undersøkelse som gikk spesifikt på sensuren (Berge, 1994), som ble referert i 2.2.1. Nylig foretok NIFU en gjennomgang av ulike intervjuundersøkelser av vurdering i norskfaget, og oppsummerte med at kriterier og kollegasamarbeid fremheves som viktige for reliabel og rettferdig vurdering, i tillegg til at «(...) det er behov for norskdiraktisk utvikling i vurdering i norskfaget» og at behovet er størst i de øverste trinnene (Seland, Lødding, & Prøitz, 2015, s. 64). Selv om det seneste forskningsprosjektet som har undersøkt eksamen er fra grunnskolen, vil jeg referere funn som kan være relevante herfra. Det kan gi viktige perspektiv til analysen av undersøkelsens spriktekster fra eksamen etter videregående.

2.2.3.1 KAL, sprikvurdering og risikotrekk

Samarbeidsprosjektet «Kvalitetssikring av læringsutbyttet i norsk skriftlig» (KAL-prosjektet) undersøkte sensur, oppgaver og eksamenstekster i et utvalg fra perioden 1998-2001.

Prosjektets leder skriver at den statistiske påliteligheten ikke var «(...) fullt ut tilfredsstillende i teoretisk forstand» (Evensen, 2005, s. 206), men den viste stigning og var relativt høy i forhold til internasjonal forskning. Den var også høyere enn da Berge et tiår tidligere hadde undersøkt dette i videregående skole (Berge, 1994). Blant annet derfor ble resultatene fra KAL-prosjektet av mange mottatt som en friskmelding av normfellesskapet blant landets norsklærere og -sensorer, og positiv utvikling ble forklart av godt utviklet samtalekultur i grunnskolen (Evensen, 2005). KAL-prosjektet er oppsummert i ti delstudier. Studiene som er spesielt relevante for min undersøkelse er studiene av såkalte *risikotrekk*, blant annet i

spriktekster, og hva sensor sa om disse (studie 9 og 4), studiene av oppgaver og oppgavevalg (studie 5 og 6) og av undersøkelsens pålitelighet (studie 3).

I «Studie 9: Å ta en risk – «Utfordrerne» blant eksamensskriverne», undersøkte Frøydis Hertzberg og Sissel Andersen trekk i blant annet spriktekster. Trekkene ble sett som symptomer på eksaminandens tillit overfor sensor, men ble kalt *risikotrekk* (Hertzberg & Andersen, 2005, s. 276). I et intensivmateriale på 171 tekster, ble det i 40 tekster funnet ett eller flere risikotrekk knyttet til språk, innhold, sjanger og oppgavekontrakten. Et par trekk som ble funnet, men valgt bort, var uklar bruk av ironi og personlig utleverende tekster. Begge deler var det mye av, men sensorkommentarene tydet ikke på at ironi ga problemer. Det personlig utleverende ble en uhåndterbar kategori fordi det var umulig å ha noen oppfatning om tekstene var fiktive eller ikke (ibid.).

Risikotrekkene knyttet til språk var mange setningsemner, muntlig språk, tabuord, trivialpreg eller «høystemthet». De tre første er selvforklarende, men med trivielt og høystemt menes henholdsvis skrivestil som hører hjemme i triviallitteraturen og skrivestil som er svulstig, pompøs eller gir mistanke om for høyt ambisjonsnivå (op.cit, s. 283). Ikke alt ble funnet risikabelt, men trivielt og høystemt språk ble det. I de sterke tekstene ble det trivielle i liten grad kommentert og forbundet med dårligere vurdering, men det førte også til sprikvurdering. Det høystemte ble det reagert mer ulikt på. Når det kom til innhold, så forskerne for seg at både provoserende holdninger, vold og brudd på forventninger om seriøsitet kunne innebære risiko. Trekkene ble funnet i få tekster, og kun sistnevnte ble kommentert av sensorene. Når det kom til sjanger, var brudd på sjangerforventninger risikabelt, i tillegg til valg av fortelling der oppgaven la opp til noe annet (op.cit, s. 289). Her ble det også funnet sprikende vurderinger. Det forskerne kalte gambling med kontrakten, kunne være å skrive kort, ikke svare på oppgaven, gamble bevisst eller metakommunisere; «(...) å tre ut av den forventede kommunikasjonssituasjonen og gå i dialog med sensor» (op.cit, s. 298), også indirekte ved å referere til eksamenssituasjonen. Det viste seg at de som ikke svarte på oppgaven utgjorde en stor kategori som kunne skilles mellom de med manglende evne eller vilje til å svare, der det gjerne ble sprik i vurdering av sistnevnte. Bevisst gambling gjaldt sterke elever, og flere ble sprikvurdert på grunn av det. Det ble også sprikvurderinger av nevnte metakommunikasjon.

Når det kom til hvordan sensorene møtte risikotekstene, fant ikke forskerne noe enkeltstående trekk som så ut til å være avgjørende for karakteren, heller ikke der det var sprik. Imidlertid var det ett trekk som oftest ble diskutert, og det var i hvilken grad teksten svarte på oppgaven. Sensorene var likevel rause her. Forskerne mente at skolsk skrijving ser ut til å tape terreng for

mer kreative eksperimenter, i tillegg til at sensorene virker godt skolert og lite normative (op.cit, s. 300). I teorikapittelet ble problematisering av denne rausheten referert, fordi en i et literacyperspektiv er opptatt av at utydeliggjøring av det skolske spesielt kan ramme elever med mindre teksttrening hjemmefra.

KAL-sensorene var også gjenstand for et eget delstudium. I «Studie 4: Hva sier sensor? Sensor kommentarer om tekstkvalitet», utdypet Harald Nilsen spriktekstene i KAL-utvalget ved å analysere sensor kommentarene (Nilsen, 2005). Utvalget viste relativt høy korrelasjon i vurderingen: 0.69 for hele utvalget (Berge, 2005a, s. 104), herunder 0.62 for personlig skriving, 0.75 for fortellende og 0.85 for resonnerende. Likevel er sprik og outsiders interessant, for «(...) individuelle vurderingssyn, som skiller seg tydelig ut fra et flertallssyn, kan gi interessante funn om verdsettingssignifikante aspekter» (op.cit, s. 128).

Bedømmersignifikante trekk ved spriktekster var *Fokus, Relevans og Svar på oppgaven*: Disse aspektene var mest representert og førte til stor betydning for karakterene, særlig de lave. I tillegg ble det sprik når det kom til kreativitet. Autentisitet eller troverdighet er også sprikgrunn, i sjangrene personlig brev og dagbok, altså når det skrives for intimsfæren. Nilsen skriver at dette er kontroversielle sjangrer som eksamensskrivning (Nilsen, 2005, s. 136). Nilsen fant videre at mangel på overordnet struktur ikke var høyfrekvent i kommentarene til spriktekstene, med unntak av at tvil om treff på sjanger i noen tilfeller ga utslag på karakter. Språk og formell ferdighet var heller ikke kvaliteter som ofte og separat pekte seg ut. Konklusjonen ble at sprik sier mer om sensor enn om tekstene (Nilsen, 2005, s. 159), men at sensorene har forhandlingsvilje, noe Berge også fant i 1994 (Berge 1994, s. 255).

Nilsen fant altså at sprik kan knyttes til kreativitet og personlig skriving, og kalte sjangrene som åpnet for dette kontroversielle. Det kan henge sammen med omtalte faglige uenighet rundt prioritering av elevens prosjekt versus samfunnets sjangrer og krav. Bakhtin kalte ytringen individuell, men at de litterære er best egnet til å reflektere talerens individualitet (Bachtin, 2005, s. 5). En kan altså si at individet eller elevens stemme har best vilkår i sjangre en ikke forbinder med samfunnets behov, om en ser bort fra sjangre fra kultursfæren. Det er en norskfaglig uenighet knyttet til om det da er riktig å åpne for disse sjangrene på eksamen, noe jeg kommer tilbake til i analysen. Likevel kan en ut fra KAL-funn si at elevens prosjekt hadde rimelig gode vilkår ved grunnskoleeksamen i norsk ved årtusenskiftet, og problematisering av dette ble omtalt i 2.1.3 (Smidt og Tønnessen 2011).

Sprik i grunnskoletekster fra eksamen 2009 er også undersøkt i en masteroppgave skrevet av Mette Thoresen (2010), med utgangspunkt i et utvalg av sensorers kommentarer. Analysen

viste at det for hovedmålseksamen særlig finnes tre årsaker til sprik (Thoresen, 2010, s. 62): I hvilken grad besvarelsen *inneholder* det oppgaven ber om, i hvilken grad eleven mestrer valgt *sjanger* og sensorenes oppfatning av *språk og stil*. Overordnet er det samme kategorier som ble funnet i analysene av KAL-utvalgets tekster. Thoresen leste også sine funn opp mot KAL, særlig når det gjelder språk og stil, og fant mye av det samme (Thoresen, 2010, s. 79-80).

KAL-undersøkelsen gjaldt eksamen etter grunnskolen, og det har vært og er til dels fortsatt ulike tekstkulturer i de to skoleslagene. Det er også «(...) markant strengere bedømming» av eksamen etter videregående enn grunnskolen (Vagle, 2005, s. 96). I videregående skole er det spesielt ulike analysesjangrer som får mye tid og oppmerksomhet, mens ungdomstrinnet utvikler selvstendige skrivere ved å be om erfaringer, opplevelser og meninger (Smidt i Haugaløkken, Evensen, Hertzberg, & Otnes, 2009, s. 70). Dette understøttes av KAL, som avdekket dominans av fortellende, ekspressiv skrivning, mens «(...) oppgaver med temaer innrettet mot sak tiltrekker seg relativt få skrivere» (Vagle & Evensen, 2005, s. 173). På tross av stor valgfrihet skrev mellom 65 og 75% av de studerte årskullene fortellende, i tillegg til at åtte av ni mest populære oppgaver angikk nære relasjoner eller eksistensielle spørsmål (op.cit, s. 198). Sistnevnte funn i KAL ble beskrevet som et problem i opplæringspraksisen i grunnskolen, i tillegg til Nilsens funn av ekspressiv og kreativ skrivning som kontroversielt på eksamen. Skjelbred (2014) skriver at det var uheldig med fravær av saklig, argumenterende og upersonlig skrivemåte, fordi elever senere får bruk for akkurat dette. Et svar på problemet har vært økt vektlegging av sakprosa og funksjonell skrivning, samt omtale av teksthandlinger og grunnleggende ferdigheter i LK06 (Skjelbred, 2014, s. 37). I tillegg bidro revisjonen av læreplanen i norsk høsten 2013 til større grad av sammenheng gjennom hele skoleløpet. Etter hvert kan man derfor forvente større mulighet for å sammenlikne tekster fra de ulike skoleslagene, og at elevene fra grunnskolen er mer forberedt til tekstkulturen i videregående.

2.2.4 Normfellesskap og true score i vurdering av eksamenstekster

Lavt vurderingssamsvar og sprikvurdering kan være en indikasjon på at normfellesskapet er for lite utviklet. Jeg har tidligere referert Berges funn av lavt vurderingssamsvar i eksamen etter videregående. Svensk skriveforskning har lenge problematisert spørsmål rundt vurderingssamsvar, likeverdig vurdering og sprikvurdering. I en doktoravhandling fra 2002 ble det blant annet avdekket et veldig lavt vurderingssamsvar av svenskeksamen (Östlund-Stjärnegårdh, 2002). Mer nylige arbeider har avdekket at det opereres med ulike vurderingsnormer på ulike skoler og at tilsvarende erfaringer som er gjort i de norske

nasjonale prøvene, tilsier at det også må jobbes med utvikling av normfelleskap i svensk skole (Gustafsson, Cliffordson, & Erickson, 2014).

Eric Borgström og Per Ledin setter sprikvurdering i sammenheng med skrivekompetansens kompleksitet. Likevel må vi forvente samsvar, skriver de, fordi eksamen er en såkalt *high stakes* vurderingssituasjon, en prøve som påvirker individets fremtid (Borgström & Ledin, 2014, s. 33). De resonnerer ut fra en beskrivelse av vurdering som en paradoksal praksis fordi det forener ytterpunktene teknisk og hermeneutisk rasjonalitet, der førstnevnte er forutsigbar og mindre menneskelig, mens sistnevnte tilsvarer profesjonelt skjønn (op.cit, s. 34). I en metastudie av 16 studier av vurdering, foreskriver de likevel balansegangen, fordi ingen av disse tenkemåtene bør trekkes for langt. Vurderingskriteriene i Utdanningsdirektoratets eksamensveiledning kan sies å uttrykke både teknisk og hermeneutisk rasjonalitet, der oppdelingen i ulike områder og kriterier for disse er av teknisk art, mens presiseringen av at helheten skal vurderes, ber om hermeneutisk rasjonalitet (Utdanningsdirektoratet 2015). Det er i denne sammenheng interessant at *Normprosjektet* bryter med helhetstradisjonen, noe som nevnt tidligere er begrunnet i at den utfordrer tolkningsfelleskapet (Solheim & Matre, 2014).

Når intervurderingsreliabilitet studeres, utgår man fra en *true score*; en middelvei som er teoretisk basert på alle vurdererne innen faget, og som uttrykker kvaliteter i teksten som ikke er påvirket av hvem som vurderer den. I svenskdidaktikk er Björnsson (1960) et pionerverk. Hans studie, her lest hos Borgström og Ledin, omfattet 104 lærere som vurderte de samme tekstene. På denne måten fant han at noen tekster trengte hele 10-20 sensorer for å nå *true score* (Borgström & Ledin, 2014, s. 8), mens det i mange tilfeller holdt med fem. Ulike forskere har foreskrevet ulikt antall vurderere til én tekst for at sensur skal uttrykke *true score*. Borgström og Ledin skriver at for de aller fleste tekster er «(...) tre bedömare en möjlig rekommendation» (op.cit, s. 156). Dette fordi tre reduserer variasjonen mer enn to gjør, og at for eksempel seks bedømmere for tre av fire tekster ikke reduserer ytterligere. Det blir fortsatt noe variasjon med tre bedømmere, men en får sjaltet ut eventuelle avvikeres vurderinger (ibid.). Forskerne kaller avvikerne *hauker*, som er strengere enn de andre, og *duer*, som er mindre strenge. Dette tilsvarer to av avvikergruppene som Berge fant i sin studie av sensuren, men han avdekket også en avvikergruppe som vurderte annerledes på vilkårlig måte (Berge, 1994). Variasjon mellom medsensorers vurderinger kommer jeg tilbake til i analysen.

Oppsummert viser forskning på både skriveoppgaver og -vurdering at en må ta høyde for skrivekompetansens kompleksitet. Sprikvurdering settes i sammenheng med både denne kompleksiteten (Borgström og Ledin, 2014) og oppgaveformulering (Skar og Aasen, 2015).

Til sistnevnte er klarhet et sentralt stikkord. Presisjon av formål (Otnes, 2015), ønsket innhold og struktur (Smemo og Solem, 2015) ser ut til å gi bedre tekster og enklere vurdering. En kan heller ikke generalisere skrivekompetanse ut fra én type skriving (Berge 2005), og tolkningsfellesskap innebærer systematisk jobbing med kvalitetssikret dokumentasjon av eksplisitte forventningsnormer (Solheim og Matre, 2014), gjerne gjennom eksempeltekster. KAL-analysene av risikotrekk og sprikvurdering viser at det som oftest ble diskutert var svar på oppgaven. Dette kan henge sammen med mindre utviklet normfellesskap på dette området, men også oppgavens utforming. I tillegg ble det funnet uenighet i vurdering av høystemt og trivielt språk, metakommunikasjon og kreativ og personlig skriving. Dette kan tale for både diskusjon av oppgaver som åpner for dette, men også behov for utvikling av normfellesskapet knyttet til vurderinger av slike trekk.

3 EKSAMENSOPPGAVENES KONTEKST: FRA SJANGER TIL TEKSTTYPE

3.1 ENDRINGER AV EKSAMENSOPPGAVENE I 2014

Eksamensoppgavene fikk en endret form våren 2014, og jeg vil i dette kapittelet se på hva endringene innebar og reaksjoner de fikk. Et dokument fra Utdanningsdirektoratet fra 28.10.2013 beskrev endringene (Utdanningsdirektoratet, 2013), og våren 2014 ble eksamensveiledningen oppdatert (Utdanningsdirektoratet, 2014). Direktoratet begrunnet endringene med at eksamensoppgavene skulle være utformet i samsvar med revisjonen av norskfagets læreplan i 2013. Kortsvaroppgaven ble gjort obligatorisk i alle oppgavesettene, og sjangerinstruksene i langsvarene ble erstattet med *teksttyper*. Til sistnevnte har jeg tidligere vist at læreplan og eksamensoppgaver er forankret i skriveideologi og -forskning som vektlegger *skrivehandlinger* som kan realisere skrivingens ulike formål.

Endring i kortsvaroppgaven: Fra våren 2014 ble kortsvaret obligatorisk, og det for at «Eksamen i norsk skal kunne dekke en større bredde av teksttyper og skrivemåter enn når elevene bare skriver én tekst. Slik får eleven mulighet til å vise mer av fagkompetansen sin til eksamen» (Utdanningsdirektoratet, 2013, s. 2). Før dette byttet eksamenssettene i hoved- og sidemål på kortsvaret, så strengt tatt er ikke denne endringen mer i samsvar med revidert læreplan enn den forrige. Likevel er det uttrykk for at Utdanningsdirektoratet implementerer kunnskap som blant annet er jobbet frem i forbindelse med utvalgsprøvene, for eksempel at skriveferdighet ikke kan «(...) generaliseres ut fra en type skriving» (Berge, 2005b, s. 4).

Endring i langvaroppgaven: Der elevene tidligere fikk konkret sjangerinstruks i noen av langvaroppgavene, innførte direktoratet samlebegrepet *teksttyper* og delte inn i *informative*, *argumenterende*, *resonnerende* tekster som noe for seg, og *kreative* tekster som noe annet (Utdanningsdirektoratet, 2013, s. 2-3), samtidig som *litterære tolkninger* og *retoriske analyser* er beholdt (op.cit, s. 4). Jeg vil i analysen komme tilbake til hvordan de ulike oppgavetyper nå er beskrevet. Direktoratet skriver om overgangen at «Fordi sjangrer endrer seg over tid og kan variere mye, har læreplanen mer overordnede begreper for ulike typer tekster» (op.cit, s.2), og det vises til kompetansemål fra revidert læreplan. Der sjangere som kåseri, artikkel og essay før revisjonen ble brukt i kompetansemål, finner en nå formuleringer som at elevene skal «(...) skrive kreative, informative og argumenterende tekster» (ibid.).

3.2 REAKSJONER PÅ ENDRINGER AV EKSAMENSOPPGAVENE

En kan lese teoretiske perspektiv inn i resepsjonen til endringene av eksamensoppgavene. I artikkelen «Fra sjangerformalisme til sjangeranarki» spør Hertzberg og Marte Blikstad-Balas om hva som skjer med sjangerforventninger på eksamen når sjangerkrav fjernes. De leser endringene inn i konteksten av funn fra SKRIV-prosjektet, som fant mye bruk av fragmenterte oppgaver uten eksplisitt formål i skolen, men lurer likevel på om det «(...) av den grunn nødvendig å fjerne sjangrene fra eksamensoppgavene?» (Blikstad-Balas & Hertzberg, 2015, s. 48). Forskerne finner både fordeler og ulemper ved å tone ned sjanger. På sitt beste kan de nye oppgavene tvinge frem større bevissthet rundt skrivingens formål, men en større gevinst er at risiko for sjangerformalisme minkes. Tapet er at elever nå kan «gjøre hva som helst», så lenge formålet blir oppfylt, men «(...) selv om alt er lov, betyr det vel ikke at alt er lurt?» (op.cit, s. 49). Artikkelforfatterne ser ikke formåls- og sjangerskriving som motsetninger: «Vår påstand er at en elev som kjenner til ulike sjangrer, alltid vil være bedre rustet til å oppfylle et formål» (ibid.). De skriver videre at striden mellom prosesskriving og eksplisitt sjangerundervisning er historie, og at vi nå kan kombinere det beste fra begge. De sier med andre ord ja til formål, men tviler på at det er tilstrekkelig. Det må kombineres med sjanger, «(...) fordi det vil gjøre elevene bedre i stand til å oppfylle de ulike formålene de møter både på eksamen og senere» (op.cit, s. 50). Hvis ikke formål kombineres med sjanger, blir «(...) alt overlatt til elevenes implisitte sjangerkompetanse, og den er som vi vet, avhengig av elevenes hjemmebakgrunn og lesererfaringer» (ibid.), ref. omtale av literacyforskjeller i 2.3.

Tittelen til omtalte artikkel ga navn til et seminar som *Landslaget for norskundervisning* (LNU) arrangerte 10. september 2015. Her var både Utdanningsdirektoratet og

forskningsfeltet representert, i tillegg til en representant for norsklærerne. Blikstad-Balas holdt en presentasjon av artikkelen som hadde gitt seminaret navn (Blikstad-Balas, 2015a). Jeg har sett opptak på LNUs hjemmeside, og jeg siterer fra presentasjonene som ble holdt og en oppsummerende artikkel i *Norsklæraren* (Botheim, 2015).

Lærer- og sensorrepresentanten Celia Berg pekte på en problemstilling fra klasserommet, nemlig at det er mange begreper i norskfaget (Berg, 2015, s. 2). Hun stilte spørsmålsteget ved om eksamensoppgavens nye begreper er intuitive for elevene, og at dette krever klare oppgaveinstruksjoner. Fordelen ved den kreative teksten er at den er mindre sjangerbundet, samtidig som friheten ikke nødvendigvis bare er en fordel (op.cit, s. 10). Skrivesenterets presentasjon ble holdt av leder Arne Johannes Aasen. Han var opptatt av at både sjanger og teksttyper er begreper som ikke har entydig mening, og at det er viktig å vektlegge at teksthandlinger er overordnede kategorier for sjanger (Aasen, 2015, s. 12), ref. kapittel 2.1.2. Teksttyper må heller ikke bli gjenstand for formalistisk operasjonalisering (Botheim, 2015).

Utdanningsdirektoratets Bente Heian poengterte at sjangerne ikke er ute av læreplanen. Hun viste til konkrete kompetansemål som beskriver arbeid med sjangere. Men spesifiserte sjangere er erstattet med formål, skritesituasjon og «norskfaglige sjangere» (Heian, 2015, s. 3). Når det gjelder endringene i læreplanen som angår skrivingen, viser direktoratet til forskning som viser behov for mer funksjonell skriveopplæring, og læreplaner i andre land (op.cit, s. 6). Både teksttyper og sjangere skal være med i skriveopplæringen, men elevene skal kunne velge selv, og elevene bør få modeller for tekstene de skal skrive (Botheim 2015). Heian påpekte at oppgaveformuleringene bruker begrep fra læreplanen, underforstått at om lærerne følger planen, vil elevene være forberedt til eksamen. Likevel innrømmet direktoratet at den kreative oppgaven gir «svært krevende» sensur, og viser her til tilbakemeldingene som kom etter vårens eksamen 2015 (Heian, 2015, s. 10). Utfordringen møtes ved å bygge felles forståelse, og det ved å følge direktoratets eksamensveiledning, veiledning til læreplanen i norsk og Skrivesenterets ressurser (op.cit, s. 12). Undersøkelsen kommer tilbake til hvordan denne forståelsen fremkommer i omtalte dokumenter fra Utdanningsdirektoratet.

3.3 REAKSJONER PÅ OPPGAVETYPEN KREATIV TEKST

Reaksjonene på de nye oppgavetyperne gikk oppsummert ut på at det ikke bør være motsetninger mellom skrivehandlinger og sjangrer, men at fravær av rammer kan gjøre det vanskeligere å skrive. Mest debattert i så måte er nok den kreative teksten. Oppgavetypen er kritisert i seg selv, men også for å løsrive det kreative fra annen skriving. I tillegg viser en

fersk gjennomgang av sensorrapporter (Bach 2015) at sensorene rapporterer at oppgavene risikerer å bli vanskelige å skrive og vurdere. Før jeg ser på dette, tar jeg et historisk tilbakeblikk for å vise at kreativ skriving også tidligere har vært omdiskutert.

Den kreative oppgaven erstattet i 2014 eksamenssjangrene novelle, essay og kåseri, der de to førstnevnte hadde vært på læreplanen for videregående skole siden 1967. Johnsen (1994) forteller historien om hvordan novelle også hadde en kort karriere som oppgavesjanger på eksamen etter videregående. Den dukket opp på eksamen i sidemål i 1991 og både på hoved- og sidemål året etter. I tilbakemeldingen fra sensorene var dommen entydig: Ingen oppgave fikk så lav gjennomsnittskarakter og «Kreativ skriving av denne typen egner seg dårlig i eksamenssammenheng» (Johnsen, 1994, s. 160). Novellen forsvant som eksamensoppgave året etter, begrunnet i «(...) sensorenes vansker med å sette en rettferdig karakter» (op.cit, s. 161), og fra 1993 skulle eksamenssjangerne være artikkel, kåseri og personlig formet essay. Johnsens gjennomgang slutter her, men novellen kom tilbake under LK06 og er fra tid til annen, før revisjonen, gitt som oppgave. Berges analyse av skriftlig eksamen fra samme år Johnsen refererte fra, fant at sjangerskrivingen hadde lavest bedømmersamsvar, 0.43 (Berge, 1994). Tilbakeblikket viser at kreativ skriving også tidligere har vært kontroversielt, også innenfor rammene av sjangerskriving på eksamen.

I et foredrag på årskonferansen til *Landslaget for norskundervisning* i mars 2015, som i sin helhet var viet kreative tekster, understreket Jon Smidt det estetiske som grunnleggende i norskfaget. Likevel er han ikke udelt positiv til kreativ teksttype, selv om han forstår at det skal hindre sjangerformalisme (Smidt, 2015). Smidt kaller de kreative eksamensoppgavene spennende, men de passer best for klasseromsarbeidet (ibid.). Han problematiserer måten oppgavetyper beskrives i direktoratets eksamensveiledninger. Veiledningen for 10. klasse skriver at den ikke er en teksttype eller skrivehandling, men en skrivemåte. Den «(...) settes opp som noe annet enn formålsrettede skrivehandlinger» (Smidt, 2015, s. 19). Overgangen til teksttyper og en egen kreativ sådan, åpner for et skille mellom praktisk og kreativ skriving. Det er problematisk, fordi: «(...) tanken om en «fri» og «autentisk» skriving, som liksom skulle være løsrevet fra sjanger og sosial sammenheng (bruk/handling)» (Smidt, 2015, s. 23), gir en oppdeling som skyver det kreative til side når ferdigheter skal måles.

Eksamensveiledningen fra 2015, som ble publisert samme vår som Smidts foredrag, skriver nå at sensorene skal honorere kreativitet i alle oppgavene (Utdanningsdirektoratet, 2015a). Det er usikkert om dette er tilstrekkelig til at Smidts poeng mister gyldighet, fordi det kan gi et sterkere signal at en av oppgavene heter nettopp kreativ.

Blikstad-Balas presenterte sitt syn på kreativ tekst i *Norsklæreren* i 2015. Læreplanendringen har «solid faglig forankring» (Blikstad-Balas, 2015b, s. 30), men spesielt kreativ tekst er en nøtt for mange norsklærere. Hun ønsker heller ikke å løsrive det kreative: «Kreativitet kan være en vesentlig del av skriveprosessen også innenfor de andre teksttypene» (ibid). I refererte artikkel til Blikstad-Balas og Hertzberg, kalles dette også et brudd med tankegangen om at formålet skal være avgjørende, med mindre formålet med oppgaven er kreativitet i seg selv (Blikstad-Balas & Hertzberg, 2015, s. 50). Det er også viktig at de kreative oppgavene ikke blir vage, men angir retning for skrivningen. Er de vage, gjør de vurderingsarbeidet «nærmest umulig» (Blikstad-Balas, 2015b, s. 31), og videre: «Gode skriveoppgaver har et tydelig formål og en tydelig mottaker», noe som ikke står i motsetning til sjangerskriving: «God skriveopplæring kommer ikke utenom systematisk opplæring i sjangere» (ibid). Om det er slik at en oppgave som ber om handlingen *reflektere*, implisitt ber om et essay, slik *fortelle* egentlig ber om en narrativ fiksjonstekst, kan det være en fordel for de skriverne som ser dette, og kanskje i tillegg behersker valg av kreative innfallsvinkler så godt at de også evner å reflektere gjennom novellen eller fortelle gjennom rapporten. Samtidig er det en tilsvarende ulempe for de som ikke har dette repertoaret, ei heller mulighet til å lese egnet sjanger ut fra skrivehandling eller formål. For disse elevene kan fravær av konkret sjanger øke usikkerhet og risiko for å velge feil.

Nevnte antakelse kan finne støtte i en analyse av sensorrapportene fra 2012 og 2014, der Cecilie Bach, som også er medlem av *Fagnemnda i norsk*, finner to aspekt som utmerker seg. Det ene, som gjelder begge rapportene, er hva sensorene oppfatter som en såkalt norskfaglig oppgave, og at de her ønsker at kunnskapsmål fra hovedområdet *Språk, litteratur og kultur* blir etterspurt i oppgavene. Det andre er nettopp reaksjoner i 2014 på nye oppgaveinstruksjoner (Bach, 2015, s. 29). Våren 2014 var første eksamensavvikling med nye oppgavetyper, og etter sensuren dette året, er det spesielt den kreative oppgaven sensorene kommenterer. Det viser seg at fraværet av sjangerforventning i skriveoppgaven ga problemer. «Elever, lærere og sensorer innenfor norskfaget har tilhørt et diskursfellesskap og en tekstkultur der sjangrer som artikkel, essay, kåseri og novelle har vært gjengangere i eksamenssettene, og der sjangeropplæring har stått sterkt i skriveundervisninga. Revisjonen av læreplanen høsten 2013 markerte derfor et veiskille der fokus dreide seg vekk fra sjangrer over til mer overordna teksttyper» (op.cit, s. 38). Dette hadde altså konsekvenser for eksamen, og svært mange sensorer kommenterte den nye instruksjonen «Skriv en kreativ tekst der du reflekterer...» (op.cit, s. 39). Sensorene rapporterte usikkerhet til hvordan de skulle tolke

instruksen og vurdere oppgaven. Koplingen mellom kreativ tekst og reflektere er av flere sensorer tolket som et essayoppdrag, men det er uklart og kunne føre til at eleven misforsto, i tillegg til at det utløste stor usikkerhet på sensorskoleringen. Både elever og sensorer var altså usikre på hvilke tekstnormer som skal gjelde. Dette er en usikkerhet Bach knytter til det Berge kaller *begrepsvaliditeten*, som jeg tidligere har vært inne på at Berge fant svakheter i allerede i 1994. Når usikkerheten blir så stor som den fremstilles her, utfordrer det gyldigheten som eksamen i norsk har som skriveprøve. Utdanningsdirektoratet skriver i den årlige undersøkelsen som danner grunnlaget for sensorrapportene at sensorene er en svært viktig gruppe i arbeidet med eksamen (Bach, 2015). På tross av sensorenes tilbakemelding på formuleringen av den kreative oppgaven, er det en relativt likelydende oppgave som ble gitt til eksamen i hovedmålet våren 2015. Om usikkerheten knyttet til vurdering av kreativ tekst som reflekterer, koples med den andre tendensen Bach fant, nemlig at sensorene i 2012 hevdet at det ikke kreves særlig norskfaglig kompetanse i to av oppgavene dette året, fordi de «(...) åpnet for egne refleksjoner» (op.cit, s. 35), kan en slutte at det er skepsis i sensorkorpset til å skulle sluttvurdere med utgangspunkt i denne teksthandlingen.

For at kreativ oppgavetype skal åpne for større mangfold og variasjon i elevsvar, betyr det i eksamenssammenheng at «Den som vurderer, må se på teksten som en språklig handling mer enn som en variasjon over en gitt sjanger» (op.cit, s. 40), men sjanger og formålsskriving trenger ikke være en motsetning, som jeg tidligere har vært inne på (Aasen, 2015). «Sett i lys av den funksjonen eksamen i norsk skal ha som reliabel og valid vurdering, kan fraværet av eksplisitte sjangerkrav se ut til å ha vært en utfordring for både skrivere og sensorer våren 2014» (op.cit, s. 42). Den kreative teksten er i et normvakuum, noe som fører til mer risikabel skriving og vurdering. Dermed kan det «(...) se ut som om vurderinga av elevsvarene ikke har vært tilstrekkelig valid, fordi det ikke har vært god nok felles forståelse av hva som skulle vurderes» (ibid.). Det er med andre ord en del som gjenstår før elever og sensorer får klarere for seg hvordan den kreative teksten kan løses.

3.3.1 Kreativ tekst som eksamensoppgave: Reflektere

Over har jeg referert nylige betraktninger rundt nykommeren kreativ tekst, men hvordan er den formulert som oppgave på eksamen i hovedmålet? På våreksamen i 2014 og 2015 er den operasjonalisert i skrivehandlingen *reflektere*, som vist under, og derfor presenterer jeg både oppgavene og spesielt skrivehandlingen nærmere.

Vedlegg:

- «For trykkefriheten» av Henrik Wergeland, 1840
- «Du må ikke sove» av Arnulf Øverland, 1937
- «Lagnad sa vi» av Arnljot Eggen, 1969

Skriv en kreativ tekst der du reflekterer over tema som frihet, likegyldighet og engasjement. Bruk ett eller flere av tekstvedleggene.

Kommentar:

Denne oppgaven åpner for ulike løsninger. Du kan velge hvilken tekst og hvilket tema du vil legge mest vekt på. Du skal bruke kunnskap om sjangrer og språklige virkemidler når du utformer teksten din, og du kan vise kreativitet i måten du bruker språket på, i hvordan du strukturerer teksten, og i de perspektivene og tilnærmingene du velger.

VÅR 2015: Oppgave 2

Vedlegg:

- Tarjei Vesaas: «Kvart menneske er ei øy»
- John Donne: «Intet menneske er en øy»

Ta utgangspunkt i de to tekstvedleggene, og skriv en kreativ tekst der du reflekterer over tematikken i tekstene. Lag overskrift selv.

Kommentar:

Denne oppgaven åpner for ulike løsninger, men det skal komme tydelig fram i svaret ditt at du tar utgangspunkt i vedleggene. Du kan vise kreativitet gjennom de innfallsvinklene og perspektivene du velger, og gjennom bruken av språklige virkemidler.

De to oppgaveformuleringene er relativt likelydende. Eksaminanden skal skrive en kreativ og reflekterende tekst, men 2014-oppgaven verbaliserer refleksjonens tematikk. I tillegg sier kommentaren i 2014 at eksaminanden «(...) skal bruke kunnskap om sjangrer og språklige virkemidler», mens 2015-kommentaren sier «(...) kan vise kreativitet gjennom ...». Begge disse momentene gjør 2014-oppgaven mindre åpen enn den i 2015. 2015-oppgaven blir analysert nærmere i kapittel 5.

Jeg har vært inne på oppfatninger av kreativ oppgavetype. I det følgende viser jeg ulike tilnærminger til reflekterende skriving, fordi det gir kontekst for analysen av eksamenstekstene. Sindre Dagsland (2015) drøfter Skrivehjulets refleksjonsbegrep med *Normprosjektet* som kontekst. Skrivehandlingen *reflektere* gjør refleksjon til en utvendig ytring, ikke innvendig tenkning, som er en vanlig konnotasjon (Dagsland, 2015, s. 202). Flere har funnet at refleksjonsbegrepet er utbredt, men utilstrekkelig definert (op.cit, s. 203), og ved

å gjøre det til en skrivehandling, distanserer «(...) refleksjonsbegrepet i Skrivehjulet seg fra en del av forskningslandskapet om refleksjon», der det altså omtales som tenkning (op.cit, s. 204). Det er med andre ord «noen potensielle utfordringer» ved skriving og oppgavelaging ut fra hjulets å *reflektere* (op.cit, s. 218). To masteroppgaver fra 2013 har undersøkt refleksjon som skrivehandling i fire elevtekster fra utvalgsprøvene. De fant at skrivehandlingen er kompleks og befinner seg mellom det private og faglige, noe som også kompliserer vurdering (Buffon & Marti, 2014, s. 70). Analysen vil vise om dette også kan gjelde eksamenstekster.

Torlaug L. Hoel er opptatt av læringsstøttende formål ved reflekterende skriving. Den «(...) ligg nær opp til talespråket. Den er ofte springande, fragmentarisk, assosiasjonsrik», men kan bli lesbar om den bearbeides ut fra leserens behov (Hoel, 2008, s. 39). Her er det blant annet snakk om å gi teksten hensiktsmessig struktur, jobbe med uttrykksmåte og rettskriving (ibid.), og da endres teksten fra tenke- til presentasjonsskriving (Dysthe, Hoel, & Hertzberg, 2010). Reflekterende eksamenstekster kan befinne seg i grenseland. Dysthe og Hertzberg plasserer for eksempel det utprøvende og dialogiske i tenkeskriving, mens en skriftlig eksamensoppgave åpenbart tilhører presentasjonsskriving (Dysthe & Hertzberg, 2014, s. 22)

Ruth Grütters kaller refleksjon et begrep det er innforstått doxaforståelse, men ikke eksplisitte beskrivelser av (Grütters, 2011, s. 161). Hun utforsker skjæringspunktet mellom denne forforståelsen og selve refleksjonsteksten, her realisert i bloggformatet. Refleksjonspotensialet i tekstene forklarer hun via ulike tekststrategier hos henholdsvis dialogiske og narsissistiske modellsendere. En dialogisk modellsender er på jakt etter en faglig dialogpartner, mens den narsissistiske er på leting etter seg selv gjennom selvframstilling (op.cit, s. 169). Samtidig sier Grütters at refleksjon «(...) alltid innebærer selvbetragtning og dermed selvframstilling» (op.cit, s. 162). Refleksjonsmarkører kan finnes ved å studere verbbruk, som hun kaller metaforisk når refleksjon fremmer faglighet og dialog, og metonymisk når refleksjonen ikke gjenspeiles i sin helhet, men kun i relasjon til en påstått erkjennelse (op.cit, s. 171). En dialogisk sender speiler *tanken*, mens den narsissistiske speiler *seg* (ibid.). Intensivanalysen av reflekterende tekster vil kunne finne begge varianter, kanskje også i en og samme tekst.

Aase undersøker forholdet mellom skriving og dannelse (Aase, 2012). Selvrefleksjon kan utvikle selvet, men utfordrer spenningsforholdet mellom individ og kollektiv. For at selvdanningen ikke skal bli ekstrem subjektivitet, er interaksjon med verdier utenfor en selv nødvendig. Selvet dannes ved at «(...) samfunnsmessige praksiser blir transformert og gjort til ens eget» (Aase, 2012, s. 52). Ved å lese Aase sammen med Grütters, kan en se tilsvarende problematisering av at refleksjon nødvendigvis er subjektiv, men at det ikke må overta

teksten. En kan si at omtalte dikotomi mellom elevens stemme og samfunnets krav, hos Aase løses gjennom individuell refleksjon som går via det kollektive, noe som også kan være en praktisk tilnærming for å finne ovennevnte modellsendere i analysetekstene.

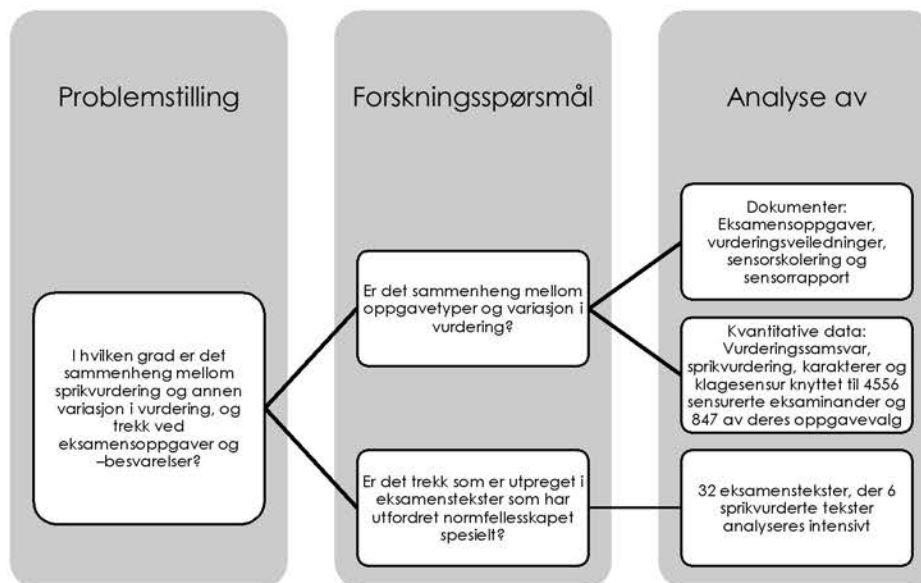
Skjelbred skriver også at reflekterende teksthandlinger i størst grad er vendt mot skriveren selv, at vi bruker det for å klargjøre egne tanker og at formålet er metakommunikasjon, selvrefleksjon og egenvurdering (Skjelbred, 2014, s. 75). Tekstene retter oppmerksomheten mot nåsituasjonen eller et punkt på tidslinjen, men de kan inngå i presentasjonsskriving, og her trekker Skjelbred frem essayet som en velegnet sjanger å reflektere i (op.cit, s. 83).

Skrivehandlingen *reflektere* kan altså knyttes til både innoverrettet tenke- og læringsstøttende skriving, men også selvrefleksjon og dialogisk kunnskapsutvikling. Skrivehandlinger er beskrevet som et nivå over sjanger. «Valg av sjanger er dermed et viktig valg for at vi skal oppnå det vi vil at teksten skal gjøre for oss. I arbeidet med elevtekster er spørsmålet om hensiktsmessig sjangervalg ut fra tekstens formål viktig å stille» (Skjelbred, 2014, s. 38). Tradisjonelt har essayet vært en litterær sakprosasjanger (Tønnesson 2012, s. 34), som gir plass til refleksjon rundt både faglige og hverdagslige emner. Ved å se på essaybeskrivelser, finnes likhetstrekk med ovennevnte tilnærminger til refleksjon. Litteraturvitenskapelig leksikon skriver blant annet at essayet er drøftende, men uten samme krav til logisk struktur som vitenskapelig drøfting (Lothe, Refsum, & Solberg, 1997, s. 69), noe som gir rom for assosiasjoner og digresjoner. Videre er essayet dialogisk og krever mottakerens forståelse av det underforståtte og bidrag til resonnementet (op.cit, s. 70). Det finnes ulike essaytradisjoner, men hovedretninger er didaktiske og subjektive (ibid.), retninger som kan stå til Grüters modellsendere. Landets felles digitale læreverk, NDLA, oppgir flere av de samme trekkene, selv om det subjektive essayet er vektlagt (Federl, 2009). Tønnesson vektlegger essayistisk *skrivemåte*; god språklig utforming fremfor sjangermønstre (Tønnesson, 2012). Dette er en tilnærming som står til den kreative oppgaven. Mot denne bakgrunnen blir det spennende å se hvordan refleksjon er realisert i intensivutvalget, både med tanke på sjanger, men også hvordan eksaminandene balanserer forholdet mellom selvet og omgivelsene når de reflekterer.

4 METODE OG MATERIALE

4.1 FORSKNINGSDESIGN

Problemstillingen og de to forskningsspørsmålene er operasjonalisert i figuren under:



Undersøkelsen er beskrivende og forklarende, og jeg kombinerer kvalitative og kvantitative data fordi de gjensidig kan supplere og komplementere hverandre (Holter & Kalleberg, 1996). Et mål med undersøkelsen er å oppnå forståelse av fenomenene over, og derfor vil også fortolkning bli av betydning (Thagaard, 2013). I det følgende redegjør jeg for metoder og utvalg, og i kapittel 4.4 drøfter jeg metodologiske utfordringer til fortolkning, utvalg og etikk.

4.2 SVAR PÅ SPØRSMÅL 1: DOKUMENTANALYSE OG KVANTITATIV UNDERSØKELSE

4.2.1 Eksamenskonteksten: Oppgaver og grunnlag for vurdering

Jeg har foretatt en dokumentanalyse av eksamensoppgavene som ble gitt i norsk hovedmål 2015, i tillegg til relevante styringsdokumenter som er publisert av Utdanningsdirektoratet, herunder veiledninger og vurderingskriterier. Dette foreliggende datamaterialet er supplert av Utdanningsdirektoratets presentasjon og egne notater fra sensorskoleringen 3. juni 2015, samt sensorenes tilbakemeldinger i *Sensorrapporten* fra 2015. Dette er beskrevet nærmere under.

4.2.1.1 Rammebetingelser fra Utdanningsdirektoratet og tilbakemeldinger på disse

4.2.1.1.1 Skriftlige rammebetingelser

Dokumentene under omhandler endringene av eksamensoppgavene og vurderingsveiledning til eksamen 2015:

Revidert læreplan og eksamen. Eksempeloppgaver (Utdanningsdirektoratet 2013)

Eksamensveiledning – om vurdering av eksamenssvar (Utdanningsdirektoratet 2014)

Eksamensveiledning – om vurdering av eksamenssvar (Utdanningsdirektoratet 2015)

4.2.1.1.2 Sensorskolering

Sensorskoleringen foregikk i seks parallelle sesjoner, der jeg deltok i Drammen 3. juni 2015. Når jeg refererer dette møtet, er det hovedsakelig powerpoint-presentasjon med kommentarer, som var felles for alle sesjonene. I tillegg ble det gitt muntlige avklaringer på spørsmål i møtet. Når jeg gjengir dette, er det basert på egne notater og fortolkninger fra møtet.

4.2.1.1.3 Sensorrapport

Sensorrapporten oppsummerer en frivillig undersøkelse i PAS som Utdanningsdirektoratet bruker som et internt arbeidsverktøy. Rapporten publiseres ikke, men er ikke unntatt offentlighet. Jeg mottok den som et word-dokument på 54 sider fra seniorrådgiver Tone Vindegg i direktoratet. Undersøkelsen er todelt. Den ene besvares av sensorene i fagkodene 1211 og 1218, altså hovedmål og sidemål, og den andre gjelder 1231, påbygg. Undersøkelsen består av 12 spørsmål. De 189 sensorene som har deltatt har svart ut fra en firedelt skala, i tillegg til alternativene «vet ikke» og «ikke besvart». Alle spørsmålene er fulgt av mulighet til å gi en kvalitativ tilbakemelding i form av *forslag til forbedringer* eller *kommentar*. Det er gitt mellom 15 og 90 tekstsvaer til hvert av spørsmålene. En del av disse kommentarene, som innspill om hotellovernatting og frustrasjon over medsensorer, er ikke relevante for undersøkelsen. Det som er relevant og derfor blir omtalt, er hva sensorene mener om de enkelte oppgavene, vurderingskriterier og avklaringer på sensorskolering. Det er vanskelig å presentere svarene systematisk (for eksempel gjennom telling), fordi noen er konstaterende og andre både positive og negative innenfor samme kommentar. Når jeg refererer fra rapporten, vektlegger jeg derfor tendenser, uten at det betyr at det ikke finnes motstemmer.

4.2.2 Innhenting, utvelgning og arbeid med kvantitativt materiale

4.2.2.1 Populasjon

Undersøkelsens populasjon er alle eksaminanders besvarelser i norsk hovedmål, studieforberedende utdanningsprogram (fagkode 1211), fra 29. mai 2015. Våren 2015 var det snakk om 30.823. I tillegg kom 11.106 eksaminander på påbygg, som svarer på samme oppgaver, men sorterer under en annen fagkode (1231). Tallene er hentet fra *Skoleporten* (Utdanningsdirektoratet). Dette gir en populasjon på 30.823 sensurerte besvarelser som er fordelt på om lag 450 sensorer. Avvikling og sensur av eksamen koordineres av seks fylkesmannsembeter; Sør-Trøndelag, Oslo og Akershus, Buskerud, Rogaland, Hordaland og Tromsø. Det er gjort flere utvalg. Utvalgsusikkerhet blir drøftet i kapittel 4.4.3.

4.2.2.2 Bruttoutvalg

Undersøkelsen bygger på foreliggende data i et utvalg av samtlige eksaminander koordinert av Fylkesmannen i Buskerud, myndigheten med ansvar for avvikling av eksamen i Buskerud, Telemark og Vestfold. Det omfatter 4556 eksaminander og rundt 50 sensorer. Bruttoutvalget utgjør 14,8 % av populasjonen. 4556 sensurerte besvarelser ble mottatt i en passordbelagt excel-tabell fra Utdanningsdirektoratet. I tabellen fremgikk eksaminandnummer, sensor-id til sensor 1 og 2, karakterforslag fra sensor 1 og 2 og karakter etter fellessensur.

Først regnet jeg ut korrelasjon, vurderingssamsvar og karakterfordeling, før jeg beregnet andel sprikvurderinger. Det er flere utregninger en kan bruke for å måle konsistens mellom to vurderere, men Pearsons produktmomentkorrelasjonskoeffisient (Pearsons r) er blant de vanlige (Stemler, 2004, s. 6). Den ble også brukt i de to refererte undersøkelsene av eksamen i norsk: Berges undersøkelse fra 1990-tallet og KAL-studien. Berge brukte også Spearmans ρ , men det ble ikke gjort i KAL, og derfor er Pearsons r også valgt for å kunne sammenlikne resultater. Korrelasjonskoeffisienten forteller hvor sterk tendensen er til at bestemte verdier på variablene opptrer sammen (Hellevik, 2002, s. 260). Det har vært vanlig å beskrive vurderingsreliabilitet gjennom ett konsept, men dette er blitt beskrevet som upresist og i verste fall villedende (Stemler, 2004, s. 1). Derfor valgte jeg å supplere konsistensmål med mål for konsensus. For å beskrive i hvilken grad det er nettopp de identiske verdiene på de to variablene som kombineres, brukes en koeffisient for samsvar. Det finnes også flere konsepter for på denne måten å beskrive konsensus. Hellevik omtaler Scotts π (Hellevik, 2002, s. 261), mens Stemler beskriver Cohens κ og prosentvis sammenfall (Stemler, 2004, s. 4). Jeg har derfor valgt alle for å beskrive konsensus i brutto- og oppgaveutvalg.

4.2.2.3 Utvalg 1: Sprikvurderte eksamensbesvarelser i tre fylker

I bruttoutvalget fant jeg hvor mange av eksaminandene som hadde blitt vurdert med en forskjell på to karakterer eller mer. I dette utvalget fremgår forholdet mellom karakterene satt av sensor 1 og sensor 2, i tillegg til karakteren de to ble enige om. Enhetene er vurderte eksamensbesvarelser, og variablene er karakterene satt av de to sensorene. Hver eksaminand utgjør én rad i tabellen, og for å finne hvilke som har fått sprikvurdering før fellessensuren, gikk jeg gjennom tabellen og kopierte ut relevante rader til det såkalte sprikutvalget.

4.2.2.4 Utvalg 2: Eksaminanders valgte langvaroppgaver

Da omfanget av sprikvurderinger var kartlagt, ønsket jeg å se dette opp mot hvilket valg av langvaroppgave eksaminanden hadde gjort. Denne informasjonen foreligger ikke, så her trengte jeg egne data. For å bli kjent med eksamensoppgaver, tekster og sensur, deltok jeg

selv som sensor. Dette hang også sammen med å få tilgang til informasjon om oppgavevalg. Jeg fikk tillatelse fra Fylkesmannen i Buskerud, ved Sissel Endung Berger, til å informere sensorene muntlig og skriftlig under sensorskoleringen 3. juni 2015, samt be om hjelp til å fylle ut et skjema (vedlegg 1). På skjemaet skulle sensorene føre eksaminandnumre til besvarelsene de hadde sensurert og hvilke langvaroppgaver disse hadde valgt. Da ikke alle sensorene deltar på skoleringen og dette ekstraarbeidet ikke kunne pålegges sensorene, forventet jeg et betydelig bortfall, som blir drøftet i kapittel 4.4.3. Videre samlet jeg inn utfylte skjema i forkant av fellessensuren 18. juni 2015. Dette ga en oversikt som jeg kaller oppgaveutvalget. Oppgaveutvalget ble på 1230 stykker, men dette måtte jeg redusere. Én hadde levert blankt og en del eksaminandnumre var duplikater. Den største reduksjonen, på 341 oppgavevalg, kom av at sensorene sensurerer både fagkodene 1211 (studiespesialiserende studieretning) og 1231 (påbygging til studiekompetanse), mens jeg bare skulle undersøke 1211. Til slutt ble oppgaveutvalget på 847 eksaminanders besvarelser, og det utgjorde 18,6% av eksaminandene i de tre fylkene. Oppgavevalget sammenstilte jeg med sprikutvalget, der det jeg fant ga meg muligheten til å finne sammenheng mellom oppgavevalg og vurdering. Dette ble gjort ved å regne ut vurderingssamsvar, karakterfordeling og andel sprikvurdering innenfor hver av de fire oppgavene, og det i tre separate operasjoner:

Først sammenholdt jeg oppgaveutvalget med bruttoutvalget på 4556. Dette gjorde jeg for å se hvilke karakterer eksaminandene jeg kjente oppgavevalget til hadde fått. Dette ble også gjort som ledd i utvalgvalidering som blir drøftet i 4.4.3. Videre sammenholdt jeg oppgaveutvalget med eksaminander som hadde fått karakteren 1 på sin eksamen. Dette gjorde jeg også for å se om resultater herfra kunne kaste lys over, og muligens validere, tredje og viktigste operasjon: Å se etter hvilke eksaminander jeg kjente oppgavevalget til, som i første omgang hadde fått sprikende vurdering, slik at jeg kunne regne ut hvordan oppgavevalgene fordelte seg når det kom til spriktekstene. Dette igjen for å beregne om det var signifikante forskjeller mellom enkeltoppgaver og bruttoutvalget. For å finne graden av vurderingssamsvar og som reliabilitetsmåling, beregnet jeg korrelasjoner mellom de to sensorenes karakterforslag for hele utvalget. Videre delte jeg oppgaveutvalget etter de fire oppgavetyperne, slik at jeg også kunne beregne konsistens og konsensus fra oppgave til oppgave.

4.2.2.5 Utvalg 3: Eksaminander som har bedt om klagesensur

Fylkesmannen i Østfold forvalter klagesensuren. Foreliggende data fra undersøkte eksamenskrets, en liste over eksaminander som hadde klaget, ble oversendt fra Utdanningsdirektoratet 11. august 2015 i passordbelagt excelformat. Jeg foretok deretter

manuelle søk i sprik- og oppgaveutvalget for å se etter sammenheng mellom eksaminandene som hadde klaget og henholdsvis sprikvurderinger og oppgavevalg.

4.3 SVAR PÅ SPØRSMÅL 2: TEKSTANALYSE AV EKSAMENSBE SVARELSER

4.3.1 Ekstensivutvalg av eksamensbesvarelser

Da undersøkelsen startet, var planen å gjøre et søk i et tilfeldig utvalg av sprikvurderte eksamensbesvarelser for å identifisere risikotrekk. Etter analysen av kvantitative data, er det én av oppgavene som har høyere forekomst av sprikvurderinger og lavere korrelasjon enn de andre. Dette er den kreative oppgaven, og den vil jeg derfor undersøke nærmere. Dette metodiske valget blir drøftet i kapittel 4.4.3.

I oppgaveutvalget er det 116 eksaminander som har valgt denne oppgaven, og 23 av deres tekster er sprikvurdert. Disse tekstene har jeg analysert for å finne brukbare kategorier for valg av intensivtekster. 17 av de foreliggende tekstene fikk jeg fra Fylkesmannen i Buskerud, i tillegg til at jeg fikk seks direkte fra en av de involverte skolene. En av disse var svar på en annen oppgave, og ble derfor tatt ut. Det resterende utvalget på 22 tekster ble først gjenstand for en grovsortering, før jeg bestemte hvilke som skulle inngå i intensivutvalget.

Ettersom ingen av de 22 spriktekstene hadde gått til klagesensur, men samme oppgavetype også hadde vært overrepresentert her, supplerte jeg ekstensivutvalget med klagetekster. Flere tekster kunne styrke antakelser jeg kunne være usikker på. Min oversikt viser 12 slike tekster, og 11 av disse fikk jeg fra Fylkesmannen. En av tekstene er tatt ut, for eleven har ført feil oppgavenummer. Jeg gjorde samme grovsortering av de resterende 10 tekstene, men fordi ingen av dem er sprikvurdert, er de ikke valgt til intensivutvalget.

4.3.1.1 Tekstanalyse av ekstensivutvalg

Den kvantitative undersøkelsen ga en del resultater, og den ga retning for tekstanalysen. Ekstensivutvalget på til sammen 32 tekster ble analysert, også med tanke på mulige kategorier som kunne bidra til ytterligere intensivering av utvalget. Det er brukt ulike innfallsvinkler i analysen. Først undersøkte jeg tekstenes makronivå med tanke på skrivehandlinger, sjangertrekk, hovedmotiv eller tema i tekstene. Det er gjort for å undersøke hvordan formålet med oppgaven er løst, og fordi skrivehandlingen tangerer den omdiskuterte endringen i eksamensoppgavene. Det neste var å lese tekstene ut fra vurderingskriteriene i vurderingsskjemaet som oppsummerer «Hovedinntrykk», for å se etter sammenheng mellom sprikvurdering og hvordan oppgaven var besvart. Dette er gjort fordi tidligere oppsummert forskning har vist at dette er det hyppigst diskuterte trekket ved sensur, og det er også satt i

sammenheng med sprikvurdering (Hertzberg og Andersen 2005). Videre forsøkte jeg å generalisere såkalte *risikotrekk*, tidligere gjennomgått i 2.2.3, som kan ha sammenheng med sprikende vurdering, og som skal undersøkes nærmere i intensivanalysen.

4.3.2 Intensivutvalg av sprikvurderte eksamensbesvarelser

Intensivutvalget er bestemt ut fra skrivehandlingene identifisert i ekstensivanalysen. I de tre kategoriene jeg fant her, plukket jeg ut to tekster per kategori. Dette er tekster som også kan tjene som eksempler på andre trekk og tendenser fra ekstensivanalysen. Det gir et utvalg på seks tekster. Valg knyttet til kategorier og antall blir drøftet i 4.4.3.

4.3.2.1 Tekstanalyse av intensivutvalg

Tekstanalyse kan gjøres på mange måter. I denne sammenheng har det vært viktig å identifisere trekk med innvirkning på sprikvurderingen. Kunnskap om ytringens adressivitet, som beskrevet i Bakhtins translingvistikk (Bachtin, 2005), er nyttig for å kunne si noe om teksten som svar og resepsjonen av dette svaret. Selv om jeg ikke kan si noe om sensorenes tolkning av eksamensoppgavene og -tekstene jeg analyserer, kan analyse av hvordan tekstene svarer på oppgaven og adresserer mottakeren, og hvordan de plasserer seg i ytringskjeden, vise mulig eller sannsynlig resepsjon. Imidlertid er ikke translingvistikken så systematisk at den egner seg som mer enn et teoretisk bakteppe. Derfor vil jeg bruke en del av rammeverket fra systemisk-funksjonell lingvistikk (SFL) i analysene, i tillegg til at jeg vil trekke på andre perspektiver fra kapittel 2 og 3, eksempelvis knyttet til literacy og skrivehandlinger.

I teorikapittelet redegjorde jeg for hovedlinjer innenfor SFL. Jeg tar utgangspunkt i Hallidays lingvistikk i analysene (Halliday, 2014), og her vil også Eva Maagerøs innføring (Maagerø, 2005) bli brukt. Det er knapt grenser for hvor omfattende en metafunksjonell analyse kan være. Jeg har måttet begrense detaljnivået og trekker inn rammeverket der det kan bidra til å belyse sprikvurderingen. Det vil også variere på hvilken måte jeg har tatt hensyn til Hallidays foreskriving av at metafunksjonene realiseres samtidig, men der funksjonene omtales hver for seg, betyr ikke det at de opptrer isolert. I tillegg til Hallidays metafunksjoner og todelte kontekstbegrep, bruker jeg kategoriene felt, relasjon og mediering der jeg undersøker tekstens forhold til situasjonskonteksten. Dette er begreper som ble mer utførlig omtalt i kapittel 2. Når jeg undersøker hvordan metafunksjonene er realisert, vil jeg bruke tilhørende begreper. For å ta et eksempel, kan en ved å fokusere på prosesser se hvordan virkeligheten eller ideasjonell mening er representert. Det finnes seks typer, hvorav materielle og relasjonelle opptrer med høyest frekvens (Halliday, 2014, s. 215). Det er forskjell på indre og ytre erfaring, skriver Halliday, og mens prototypisk ytre erfaring er handlinger og hendelser, er den indre

vanskeligere å sortere, «(...) but it is partly a kind of replay of the outer, recording it, reacting to it, reflecting on it, and partly a separate awareness of our states of being» (op.cit, s. 214). Grammatisk realiseres skillet mellom materielle og mentale prosesser, og følgelig forventes flere mentale prosesser i en reflekterende enn i en fortellende tekst. I tillegg vil jeg bruke begrepene deltakere og omstendigheter i ideasjonelle beskrivelser, modus og modalitet på mellompersonlig og ulike kohesjonsmekanismer på tekstuel nivå.

Kategorier for analyse

Det er vanskelig å ta utgangspunkt i en analysemodell som presiserer skrivehandlingens viktighet (Evensen, 2010), noe som blir drøftet i kapittel 4.4.2. Oppgavens vurderingskriterier er en viktig del av situasjonskonteksten, og blir derfor brukt som analyseverktøy. Det er 20 kriterier i vurderingsskjemaet, der syv er spesifikke for oppgaven (Utdanningsdirektoratet, 2015a). For at analysene ikke skulle bli for oppstykket, ser jeg noen kriterier i sammenheng, i tillegg til at to ikke er relevante til oppgaveformuleringen. Resultatet er første utgave av analysemodellen (tall i parentes er antall kjennetegn i opprinnelig vurderingsskjema):

Hovedinntrykk	Relevante svar på oppgavene (2)
	Forståelse og bruk av vedlagte tekster
	Formuleringsevne og ordforråd
Språk og formelle ferdigheter	Språk og formelle ferdigheter (4)
Svar på kortsvaroppgaven	Struktur og sammenheng
	Faglig kunnskap, fagspråk (2)
	Redegjørelse for karakteristiske trekk
Svar på langsvaroppgaven	Struktur og sammenheng
	Relevante momenter/eksempler, med god bredde
	Kreativitet i innfallsvinkler, språklige og litterære virkemidler (2)
	Reflekterer, prøver ut tanker
	Måten vedlagte tekster brukes på

Dette gir en konsentrert utgave av vurderingsskjemaet (siste del av vedlegg 4), men det er fortsatt overlappende analysekategorier, i tillegg til at jeg ønsker en modell der lese- og skrivekompetansen i kriteriene under *Hovedinntrykk* kan fordeles på de to oppgavene og språk-kategorien. Dette gir en revidert og endelig modell, som også angir analysenes struktur:

Svar på kortsvaroppgaven	Struktur og sammenheng
	Faglig kunnskap, redegjørelse for karakteristiske trekk, fagspråk (3)
	Relevant svar på oppgaven, forståelse og bruk av vedlagt tekst (3)
Svar på langsvaroppgaven	Struktur og sammenheng
	Bredde og relevans i momenter/eksempler
	Kreativitet i innfallsvinkler, språklige og litterære virkemidler (2)
	Reflekterer, prøver ut tanker
	Relevant svar på oppgaven, forståelse og bruk av vedlagte tekster (4)
Språk og formelle ferdigheter	Formuleringsevne, ordforråd, formelle ferdigheter (5)

4.4 DRØFTING AV METODE OG MATERIALE

4.4.1 Materialets nøyaktighet og pålitelighet

Sensorenes vurderinger er foreliggende data, men det er likevel potensielle feilkilder i oversikten jeg kaller *sprikutvalget*, fordi det var en manuell jobb å isolere sprikvurderingene. I gjennomgangen av de 4556 tabellradene kan jeg ha oversett enkeltrader. Dette er forsøkt motvirket ved å gjenta operasjonen. Eventuell overskriving av enkeltrader er også motvirket ved at operasjonen er gjentatt. På denne bakgrunn er det sannsynlig at dataene er korrekte og repliserbare. Beregningene er også repliserbare, fordi de er gjort i tilgjengelige programmer. Beregninger av Pearsons r og gjennomsnittsberegninger er gjort i *Excel*, signifikanstesting i *Zighe* (Aardal & Berglund, 2014) og beregning av Scotts pi, Cohens kappa og prosentvis sammenfall i *ReCal2* (Freelon, 2011). De to sistnevnte programmene er fritt tilgjengelige på Internett.

Oppgaveutvalget baserer seg på manuelt utfylte skjema fra sensorene. Fordi jeg ikke har tilgang til besvarelsene, kan det ikke utelukkes at sensorene i noen tilfeller kan ha fylt ut feil eksaminand- eller oppgavenummer. Eksaminandnumrene har jeg imidlertid sjekket opp mot excelfilen fra Utdanningsdirektoratet. Oppgavenumrene er en feilkilde som også i ett tilfelle ble bekreftet, fordi det ved innhenting av ekstensivutvalget viste seg at to av 24 tekster ikke var oppgave 2. I det ene tilfellet skyldtes det feilføring fra sensor, mens i det andre var det eksaminanden som hadde skrevet feil. Disse to ble følgelig tatt ut av utvalget.

Begrepet reliabilitet (og validitet) er opprinnelig knyttet til kvantitativ forskning, men selv om det er dem som mener det er hensiktsmessig med andre begreper når en skal drøfte kvalitativt materiale, er ikke det nedfelt i praksis (Thagaard, 2013). For å avgrense prosjektet har jeg ikke hentet inn kommentarer fra sensorer, og det kan være en metodisk svakhet ikke å vite hva sensorene selv mener er grunnen til sitt karakterforslag. Dette gir usikkerhet til grad av nøyaktighet og pålitelighet. Men formålet mitt har ikke vært å vurdere vurderingen. Det har snarere vært å se på hvordan tekstenes måter å svare på oppgavene på synes formålstjenlige eller ikke, og hvordan dette kan vise sannsynlige eller mulige årsaker til sprikvurderingen.

4.4.2 Studentens pålitelighet: Deltakelse, roller og fortolkning

De fleste dataene er foreliggende og jeg har i prinsippet kunnet beskrive dem utenfra. Tove Thagaard skriver at «Distansen mellom forskere og deltakere kan knyttes til at det ikke er noen direkte kontakt mellom partene» (Thagaard, 2013, s. 19). Dette gjelder hovedtyngden av både kvantitative (karakterer) og kvalitative (tekster) data. Det er likevel viktige unntak der det er nødvendig å reflektere over mulig innvirkning jeg kan ha hatt på forskningsprosessen

og derved resultatene (ibid.). Dette gjelder det kvantitative *oppgaveutvalget*, som blir drøftet for seg i 4.4.3, og de sammenhengene jeg selv har deltatt i og refererer fra. Av sistnevnte sammenhenger er sensorskoleringen 3. juni 2015 en del av datamaterialet, men i tillegg kommer arbeid med sensur, fellessensur, egen praksis som norsklærer, samtaler med *Fagnettverket for norsklærere i Buskerud* 26. november 2015 og møte med *Fagnemnda i norsk* 18. februar 2016, som jeg tidvis refererer til i analysen og drøftingen.

I disse sammenhengene har jeg hatt følgende roller: Masterstudent, sensor i norsk hovedmål og lektor i skolen. Masterstudenten er en rolle jeg har hatt i alle sammenhengene, sensorrollen har jeg hatt på sensorskolering, under sensur og på fellessensur, mens rollen som norsklærer har jeg der jeg innledningsvis refererer fra klasserommet. Det er fordeler og ulemper ved de ulike rollene. Masterstudenten er presumptivt en rolle der jeg skal beherske et objektivt utenfraperspektiv til dataene (Atkins, Wallace, & Spenceley, 2012), mens sensoren og norsklæreren er mer preget av innenfraperspektivet. Hvor pålitelige er for eksempel notatene mine fra sensorskoleringen når jeg selv har deltatt for å bli skolert? Thagaard skriver at det i slike tilfeller er viktig å redegjøre for roller og relasjoner, i tillegg til å skille mellom informasjon og egne vurderinger (Thagaard, 2013, s. 194). Det meste av referatet fra sensorskoleringen baserer seg på foreliggende data i form av Utdanningsdirektoratets presentasjon, men det var også momenter som fremkom i møtet, der mitt referat kan være influert av tolkning. Jeg kunne valgt å la være og delta som sensor, slik at jeg i større grad kunne referert objektivt fra møtet, men det ville fortsatt ha vært tolkninger. Potensiell usikkerhet er forsøkt motvirket ved at jeg gjør oppmerksom på når jeg benytter egne notater, slik at mottaker kan gjøre en pålitelighetsvurdering. Andre sammenhenger jeg har deltatt i, som jeg har valgt å referere fra, er ikke del av empirien. Møtene jeg som masterstudent har hatt refererer jeg til for å utdype drøftingen, fordi jeg har vurdert det som fruktbart å vise frem andre stemmer fra sammenhenger der jeg har hatt mulighet til å drøfte undersøkelsen.

Spørsmålet om fortolknings pålitelighet og validitet gjelder samtlige kvalitative data, også eksamensoppgaver og styringsdokumenter, fordi tekstanalyse og -vurdering innebærer skjønnsmessig betraktning og tolkning. Fordi kvalitative studier kan gjøres på så ulike måter, er transparens et viktig krav (Johannessen, Christoffersen, & Tufte, 2010, s. 82). Her har jeg forsøkt å redegjøre for hvordan analysen gir grunnlag for konklusjonene, samt gi kvalifiserte begrunnelser. I drøftingen vurderer jeg om alternative perspektiver kan gi relevant forståelse. Thagaard skriver her at det skal finnes begrunnelser for enhver tolkning av datamaterialet (Thagaard, 2013). Her kan den største usikkerheten knyttes til tolkningene av

ekstensivutvalget, fordi disse tekstene ikke er gjengitt. Dette er en usikkerhet som kan motvirkes av kvaliteten på de begrunnelsene jeg klarer å gi i tolkningene av intensivutvalget, men også her vil det være avveininger mellom omfanget av begrunnelser sett opp mot antallet gjentakelser, overforklaringer og hyppigheten av eksempler når tekstene også er tilgjengelige.

Nevnte usikkerhet kan også reduseres om en kan finne sammenfallende funn i liknende studier, om funn ser ut til å støtte teoretiske perspektiver eller ved å bruke relevante analysekategorier. Jeg har redegjort for en del tilsvarende funn i KAL-prosjektet, noe som kan gi tolkninger ekstern validitet og relevans i andre sammenhenger (Thagaard, 2013, s. 208). Her skal det sies at det i KAL ble funnet enkelte forekomster av flere språklige og innholdsmessige risikotrekk enn jeg har gjort, selv om det ikke alltid var knyttet til sprikvurdering. Dette kan skyldes at KAL-materialet er vesentlig større, men kanskje også at det er forskjeller på skrivere som er 15 og 18 år gamle. I analyse og drøfting av denne vil jeg søke begrunnelser ut fra tidligere omtalte teoretiske perspektiv. En kan argumentere for overførbarhet når en kan knytte til gjenkjennelse fra personer med erfaring fra fenomenene som studeres (op.cit, s. 213), som i de refererte sammenhengene over. Dette gir grunnlag for intersubjektiv mening. Når det kommer til relevante analysekategorier, fremhever sentrale forskningsprosjekter, som utvalgsprøvene og Normprosjektet, primærtrekkmetode ved vurdering, fordi både intuitiv/holistisk og analytisk vurderingsstrategi vil kunne undervurdere skrivehandlingenes betydning. Jeg valgte likevel å prioritere eksamenskonteksten da jeg bestemte analysemodell for tekstutvalget, som derfor ligger tett opp mot vurderingsstrategi og -kriterier fastsatt av Utdanningsdirektoratet. Fordi flere kriterier overlapper, ønsket jeg å redusere antall kategorier, men ikke så mye at det ikke var synlig at utgangspunkt var vurderingsskjemaet, og derfor bærer analysene preg av en viss overlapping. Når det kommer til eventuell repliserbarhet, er det spørsmål til om dette er relevant i kvalitative undersøkelser (Thagaard, 2013, s. 202). Uansett håper jeg at jeg har vært så tydelig om valg og fremgangsmåter at det er mulighet for å teste dette.

4.4.3 Utvalgsusikkerhet og validitet

Innhentet materiale har relevans for problemstillingen. Noe av materialet, som utvalget av eksamenstekster som er klagebehandlet men ikke sprikvurdert, kan sies å være noe på siden av problemstillingen. De er likevel tatt med for å støtte eller belyse funn i spriktekstene, noe som også kan gi intern validitet ved at årsakssammenhenger støttes innenfor samme undersøkelse (op.cit, s. 204). I tillegg har jeg sett på karakterfordeling og stryk i kvantitativ

undersøkelse. Dette angår for så vidt ikke sprikvurdering, men viser variasjon i vurdering, som også er en del av problemstillingen, i tillegg til at det bidrar til utvalgvalidering.

På spørsmålet om undersøkelsen er gyldig til å måle det den er tiltenkt å måle, kan en svare at i streng vitenskapelig forstand kan innsamlede data bare si noe om akkurat de som er med i undersøkelsen. Begrepet overførbarhet knyttes til «(...) vurderingen av spørsmålet om tolkninger som er basert på en enkelt undersøkelse, også kan gjelde i andre sammenhenger» (Thagaard, 2013, s. 23). Dette henger sammen med intern og ekstern validitet, som drøftet i 4.4.2, men det er også viktig å bringe på det rene om undersøkelsens kvantitative og kvalitative utvalg kan være representativt for populasjonen (Hellevik, 2002).

Bruttoutvalget utgjør 14,8% av populasjonen og samtlige vurderte eksamenstekster fra Telemark, Vestfold og Buskerud. Fordi data er hentet fra Utdanningsdirektoratets database, er det ikke problemer med bortfall. Det er sannsynlig at utvalg av sprikvurderinger og beregning av vurderingssamsvar er representativt. En innvendig kan være at normfellesskap er bedre utviklet i andre landsdeler, ved at eksempelvis sensorene på sørvestlandet er mer enhetlig innstilt til litterære elevsvar enn sensorene på østlandet. Jeg har ikke funnet kilder som tilsier slik variasjon, men snarere generaliserte innvendinger til normfellesskapet som helhet (Berge, 1996). Korrelasjon i bruttoutvalget tilsvarer også funn i tidligere undersøkelser (Berge, 1994).

Oppgaveutvalget er hentet fra samme bruttoutvalg som sprikutvalget, alle 4556 eksaminander i valgte eksamenskrets. I 4.2.2.4 beskrev jeg hvordan informasjonen om oppgavevalgene til 18,6% av eksaminandene i kretsen ble innhentet fra sensorene. Det er ganske avgjørende for undersøkelsens validitet at dette utvalget er representativt, all den tid flere sentrale funn er gjort her. Det er vanlig å si at det avgjørende prinsippet ved utvelgelse av representative utvalg som kan benyttes til å gjøre statistiske generaliseringer, er tilfeldighet eller randomisering (Johannessen et al., 2010, s. 241).

Metodelitteraturen er ikke absolutt med tanke på hvor stort et utvalg skal være. Johannessen skriver at størrelsen på eventuelle undergrupper må beregnes i hvert tilfelle, men antyder et sted mellom 30 og 100 (op.cit, s. 249). Fordi jeg undersøker variabelen sprikvurdering opp mot eksaminandenes valgte langvaroppgave, er undergruppene i mitt utvalg de fire ulike oppgavene. De er representert med mellom 94 og 468 besvarelser, noe som ifølge Johannessen bør være tilstrekkelig. Det kan argumenteres for at karakterskalaen også gir ulike grupperinger, og derfor har jeg også sett på dette. I utvalget er gruppen med karakteren 1 akkurat innenfor de 30 som Johannessen sier er et minimum, mens gruppen med karakteren 6

er under (19 stykker). Her skal en huske at svært få eksaminander får karakteren 6 på eksamen i norsk hovedmål, og i 2015 var det 91 i den utvalgte kretsen på 4556. Karakterene 2, 3, 4 og 5 utgjør tilstrekkelig store undergrupper på mellom 100 og 293. Ved signifikanstesting av forskjellene mellom karakterene i brutto- og oppgaveutvalget, gir forskjellen på 2,3 utslag på signifikans på karakteren 2. Forskjellene på de resterende karakterene er ikke signifikant.

Ottar Hellevik har undersøkt betydningen av lav svarprosent for representativitet i intervjuundersøkelser (Hellevik, 2015). Det kan knyttes en tilsvarende usikkerhet til representativitet til min innsamling av data om oppgavevalg. Hellevik konkluderer med at «(...) utvalgsundersøkelser er mer robuste enn mange nok har forestilt seg», men dette forutsetter at det ligger tilfeldig utvelging i bunn, «(...) og det ikke foreligger forhold som gir grunn til å tro at bortfallet har en klart systematisk karakter» (op.cit, s. 228). Dette er funn som stemmer overens med tilsvarende internasjonale undersøkelser (op.cit, s. 213).

Systematisk bortfall har en når det kan forventes sammenheng mellom svartilbøyelighet og interesse for eller meninger om fenomenet som undersøkes. Jeg kan ikke utelukke at dette har vært bakgrunnen for noen av dem som ikke fylte ut skjemaet. Imidlertid har det sannsynligvis også vært andre årsaker. Noen av sensorene var ikke på sensorskoleringen. Det er også å forvente at ikke alle som var på skoleringen fikk med seg informasjon om undersøkelsen, da den ikke var direkte relevant for årsaken til at de var der. På fellessensuren plasseres sensorene i ulike grupper og rom, og det ble derfor ikke systematikk i innsamlingen. Det viste seg også at enkelte hadde levert skjema direkte til Fylkesmannen i Buskerud, som ettersendte disse. Til sist må en også anta at flere sensorer ikke fylte ut skjemaet fordi det tok ekstra tid. I tråd med kriterier for tilfeldig bortfall beskrevet av Hellevik (op.cit, s. 214), taler altså fire av fem årsaker for at bortfallet, og derved sammensetningen av utvalget, har vært tilfeldig.

Den ene mulige systematiske årsaken må likevel vurderes nærmere. Det er usikkert om interesse og meninger om emnet har hatt sammenheng med om sensorene har fylt ut skjema eller ikke, og slik fått konsekvenser for resultatet av undersøkelsen. En innvending kan være at det har vært en viss type, uvisst hva slags, sensorer som har levert inn skjema. Samtidig skal det godt gjøres å fastslå at det skulle være mer eller mindre sammenheng mellom akkurat de som leverte og sprikvurdering eller ikke, også fordi sensorene gjerne hadde fylt ut skjemaet før de kom til fellessensuren, og dermed ikke visste i hvilken grad de ville ha samsvar med medsensorer. Det er også slik at noen av sensorene som har levert skjema har høy, mens andre har lav forekomst av sprikvurdering, noe som også kan tale mot at deres holdninger har hatt konsekvenser for resultatene. Det er heller ikke så sannsynlig at sensorer som ikke har villet

fylle ut og levere skjemaet, har hatt eksaminander som i større grad har valgt én oppgave fremfor en annen. Det ser snarere ut til å være ulikheter mellom de ulike skolene, at enkelte skoler har få eksaminander som for eksempel skriver kreativt eller argumenterende, men fordi sensorene vurderer flere skoler, ser denne tendensen ut til å være spredt på sensorene.

«Jo mer tilfeldig bortfallet er, desto mindre grunn til å bekymre seg for representativiteten», skriver Hellevik (op.cit, s. 215). På bakgrunn av drøftingen over er det nærliggende å slutte at oppgaveutvalget er så tilfeldig satt sammen at det tilsvarende sammensetningen i populasjonen. Samtidig er dette med respondentbortfall omdiskutert, og Hellevik sammenlikner også flere utvalgsundersøkelser med henholdsvis lav og høy svarprosent, og finner stort sett at de får de samme resultatene (op.cit, s. 227). Jeg har sammenliknet undersøkelsens funn knyttet til Pearsons r med tilsvarende beregninger gjort i en større undersøkelse av vurderingssvar av eksamen i 1993 (Berge, 1994), og som vist i kapittel 5.2.1 er det sammenfall. Akkurat her trenger en ikke overveie om metodeforskjeller spiller inn på resultatene, men det bør likevel overveies om forskjeller i tid, eksamensform og -oppgaver har gjort det.

En annen grunn til å tro at utvalget er representativt, er at jeg, som redegjort for, måtte redusere oppgaveutvalget fra 1230 til 847. Jeg oppdaget ikke behovet for den største reduksjonen, av 341 besvarelser med annen eksamenskode, før jeg hadde gjort en del beregninger, og derfor vil jeg kunne vise at det er marginale forskjeller i hvordan oppgavevalgene fordeler seg i utvalget på 1230 versus de 847. Utvalget i kvantitativ undersøkelse gir funn med betydning for eventuell sammenheng mellom sprikvurdering og oppgavevalg. Jeg håper at jeg har vist tilstrekkelige argumenter for at dette er representativt og gyldig, fordi det både er stort nok og fordi det er sannsynlig at det er tilfeldig sammensatt.

Den kvalitative delen av undersøkelsen skal belyse om det er trekk ved henholdsvis oppgaver eller elevtekster som kan ha sammenheng med sprikvurderingen. Fordelen ved slike data er at de kan utdype og nyansere tallmaterialet. Et avgjørende utvalgsprinsipp i kvalitative studier er at utvalget er egnet til å utforske problemstillingen (Thagaard, 2013, s. 65). Her er retningslinjer både at antallet ikke skal være større enn at det er mulig å gjøre omfattende analyser, men en kan også vurdere størrelsen ut fra et såkalt *metningspunkt*: «Når studier av flere enheter ikke synes å gi ytterligere forståelse av de fenomener vi studerer, kan utvalget betraktes som tilstrekkelig stort» (op.cit, s. 66). En kan si at det kan knyttes usikkerhet til at både ekstensiv- (32 tekster) og intensivutvalget (6) er relativt lite, men samtidig er både tekstanalyse og nærlesning tid- og plasskrevende, og utvalgene er ikke mindre enn vanlig i undersøkelser på dette nivået. Når det gjelder ekstensivutvalget, er det kun trekk som går

igjen i flere tekster som er trukket frem og undersøkt nærmere i intensivtekstene, og intensivtekstene representerer eksempler på trekkene som er funnet. Slik kan en hevde at et metningspunkt er nådd. Samtidig er det slik at ekstensivutvalget ikke er så stort at det rommer eksempler på lyriske oppgavesvar, noe som kan indikere at det også kan være andre sjeldne måter å svare på som ikke er representert. De aller svakeste tekstene er heller ikke representert, fordi det har vært vanskelig å finne annen forklaring på sprikvurdering enn ulik vektning. Eneste unntak er kanskje en av to tekster som utmerket seg ved at de kunne fremstå som provoserende. Den klareste tendensen er i S8, som etter sprik mellom 1 og 4 fikk 2, og her kan det tenkes at provokasjonen er hovedårsak til sprik. Dette er likevel for få tekster til at jeg valgte å undersøke dette nærmere. Enkelte tekster refererer også til populær- eller ungdomskulturelle referanser, som jeg forventet som mulig risikotrekk, men fordi det er såpass få tekster som viser tendensen, er bloggvarianten S10 den eneste som får representere noe av dette. I tillegg er det trekk i noen tekster som kan tolkes å bryte med forventning om seriøsitet, men dette er kun fremtredende i én tekst. Det er likevel til en viss grad belyst i analysen av metakommunikasjon i S13. Tekstutvalgene kan derfor ikke sies å kunne svare uttømmende på problemstillingen, men fordi jeg har valgt å se nærmere på trekk som er funnet i flere tekster, styrker det sannsynlig sammenheng med sprikvurderingen. Derfor kan utvalgene trolig gi relevante svar på forskningsspørsmål 2.

Jeg valgte å kategorisere intensivtekstene ut fra skrivehandlingen *reflektere*. Dette gjorde jeg fordi *reflektere* er sentralt i oppgaven, sentralt i diskusjonen av endring og formulering av eksamensoppgavene, og fordi ekstensivanalysen viste ulike måter å realisere refleksjon på. Fordelen ved kategoriene *reflekterende*, *fortellende* og *blandingstekster* er at de illustrerer de tilgjengelige valgene som er gjort, men ulempen er at kategoriene kan bidra til å kamuflere andre viktige trekk ved tekstene. I tillegg ble det vanskelig å skille reflekterende og blandingstekster fordi begge kategorier inneholder reflekterende skrivning. De fortellende tekstene kunne i sin tur deles inn i ytterligere kategorier; fiksjons- og jeg-fortellinger eller knyttet til om de realiserte refleksjon direkte eller indirekte. Jeg har også forsøkt å synliggjøre andre trekk knyttet til de ulike kategoriene, slik at analysen er blitt så uttømmende som mulig.

4.4.3.1 Utvalg av kvalitative data: Tekster innenfor én oppgavetype

Analysen av oppgaver og veiledninger, og resultatet fra kvantitativ analyse, har pekt ut en retning for analysen av eksamenstekster, men det er likevel flere mulige veier å gå. Her tenker jeg på de to oppgavene som er over- og underrepresentert i sprikutvalget, nemlig oppgave 2 (kreativ tekst) og oppgave 4 (retorisk analyse). Jeg tenker også på at oppgave 1 (informativ og

argumenterende tekst) er overrepresentert når det kommer til stryk-karakter og klager, samtidig som den utmerker seg med høyt samsvar og flere gode resultater. Disse funnene åpner også for muligheter til å velge et tekstutvalg for å belyse årsaker til hvorfor vurderingssamsvaret er såpass lavt som det er i de tradisjonelle analyseoppgavene, eller årsaker til at det ble såpass mange klager på oppgave 1 på tross av at vurderingssamsvaret på denne oppgaven var det høyeste. Jeg valgte å studere oppgave 2 nærmere og vil i det følgende drøfte bakgrunnen for dette valget.

Først og fremst er valget tatt fordi oppgaven utmerker seg negativt i alle de kvantitative beregningene, og fordi dette kan ses i sammenheng med at oppgaven også utmerker seg med sin åpne og til dels uklare formulering. Dette gjør at jeg kan være trygg på å finne svar på både problemstilling og forskningsspørsmål.

Oppgaveutvalgets størrelse ble drøftet tidligere i dette kapittelet, og eventuell usikkerhet kan forsterke behovet for å belyse den kreative oppgaven. Her tenker jeg på at funnene knyttet til oppgave 4 er mer sikre enn dem knyttet til oppgave 2, all den tid oppgave 4 utgjør over halvparten av oppgaveutvalget, mens oppgave 2 bare utgjør 13,7% (henholdsvis 468 og 116 enkelttekster). Dette betyr at eventuelle funn i et intensivutvalg som utelukkende består av kreative tekster, også kan bidra til å styrke det kvantitative materialets gyldighet og relevans. Det kan også være mer hensiktsmessig å gå i retning av det som tilsynelatende spriker mest, og derfor har det største utviklingspotensialet. Utdanningsdirektoratet ved Tone Vindegg bekrefter at de også i 2014 erfarte at den kreative oppgaven var overrepresentert ved klagesensur: «Etter denne eksamenen fekk vi relativt mange tilbakemeldingar om at kreativ tekst-oppgåva hadde slått uheldig ut» (Vindegg 2015). Ett siste moment som talte for å studere de kreative tekstene, er statusen de har i fagmiljøene. Oppgaven er som tidligere vist omdiskutert, og sprikvurderingen kan ha avdekket ambivalens knyttet til trekk ved slike tekster. I tillegg viste analysen av både veiledning, sensorskolering og -rapport at mange sensorer mente at flere av vurderingskriteriene var vanskelige å tolke og benytte.

4.4.4 **Forskningsetikk**

Etiske dilemmaer har en fremtredende plass når mennesker er involvert. Da jeg søkte Utdanningsdirektoratet om tilgang til karakterlistene, var dette gjennom en søknad om «dispensasjon fra taushetsplikt i forbindelse med taushetsbelagte opplysninger utlevert til forskningsprosjekt». Dette kunne angå personopplysninger. De nasjonale forskningsetiske komiteer (NESH) definerer personopplysninger som: «Opplysninger som direkte eller indirekte kan knyttes til en enkeltperson» (Kalleberg, 2006, s. 14). Indirekte

personopplysninger er «Opplysninger som registreres med et referansenummer, og viser til en atskilt liste med for eksempel navn eller personnummer» (ibid.). Personopplysningsloven (2001) påbyr meldeplikt for forskningsprosjekter, og hovedregelen ifølge NESH (op.cit, s. 13) er innhenting av informert samtykke før igangsetting. For undersøkelsen vil det si at avsendere av utvalgte tekster er å betrakte som deltakere. Samtidig er eksaminandnummer også indirekte personopplysninger, og i utgangpunktet gjør det alle eksaminandene til deltakere. Tidlig i prosessen antydet Utdanningsdirektoratet nettopp nødvendigheten av å innhente samtykke fra alle eksaminander i kretsen før eksamen, uavhengig om jeg skulle lese tekstene deres. Dette var ikke helt klart, fordi jeg skulle omtale utvalgte tekster uten å referere eksaminandnummer, og der det skulle koples til andre data, ville det kun bli referert i generaliseringer.

Johannessen et al. (2010, s. 95) klargjør ved å si at prosjekter ikke er melde- eller konsesjonspliktige hvis «Enkelt personer ikke kan identifiseres av dem som har tilgang til opplysningene» eller enkelt personer kan identifiseres, men «(...) opplysningene er ikke sensitive, og opplysningene behandles ikke elektronisk» (ibid.). Det første gir at jeg ikke trengte å melde prosjektet inn til NSD. Det siste betyr at intensivutvalget heller ikke måtte meldes, om jeg oppbevarte personopplysningene på papir. Juristene i Utdanningsdirektoratet åpnet også etter hvert for en utleveringsavtale av karakterlister, rapporter om klagebehandling og et mindre utvalg av besvarelser. For at opplysningene behandles i tråd med Personopplysningslovens bestemmelser, ble det stilt krav om taushetserklæring, oppbevaring og destruering av dataene. Særlig oppbevaringen innebar egne rutiner, fordi disse kun skulle finnes på frittstående pc, noe som også gjaldt sikkerhetskopier.

Fordi intensivtekstene skal gjengis, sier Personopplysningsloven at informert samtykke må innhentes: «Hvis enkelt personer kan identifiseres, skal de samtykke i å delta (...) Samtykket skal være en frivillig, uttrykkelig og informert erklæring fra den opplysningene gjelder, om at vedkommende godtar behandling av opplysninger om seg selv (...). At samtykket er informert, betyr at de som skal delta skal ha nødvendige opplysninger om undersøkelsen» (op.cit, s. 96). For innhenting av samtykke, kontaktet jeg skolene til de respektive eksaminandene og deretter eksaminandene selv. Samtykkeskjema er vedlagt (vedlegg 2). Selve samtykket er gitt personlig og viser personopplysninger, og kan vises på oppfordring.

Til intensivanalysene er det også knyttet etiske dilemma til deltakernes fremstilling. Thagaard spør om hvordan forskeren kan ivareta selvforståelsen som preger personers fremstilling

(Thagaard, 2013, s. 153). Fordi tekstene innebærer selvrefleksjon og selvframstilling, aktualiserer dette potensiell spenning mellom beskrivelser som oppleves som meningsfylt for deltakeren og fortolkningens faglige forståelse. Ulik fortolkning kan ikke unngås, men ved å hindre gjenkjennelse, vil deltakerens integritet være beskyttet. Det vil likevel være teoretisk mulighet for at deltakerne kjenner seg igjen når lange utdrag er gjengitt, men konsekvensene for disse deltakerne er vurdert som små. Det kan være snakk om eventuelt ubehag knyttet til at fortolkninger ikke stemmer helt overens med deres selvforståelse, noe som altså er forsøkt motvirket gjennom anonymisering å respektere deltakernes privatliv og gjennom innhenting av informert samtykke å ivareta deres rett til autonomi (Johannessen et al., 2010, s. 91).

5 ANALYSE OG RESULTAT

5.1 EKSAMEN I NORSK HOVEDMÅL 2015: SKRIVEOPPGAVER I KONTEKST

I kapittel 2 refererte jeg at det er forsket relativt lite på skriveoppgaver. I kapittel 3 var jeg inne på endringene av eksamensoppgavene i norsk i 2014 og reaksjoner som kom av dette. Det generelle behovet for kunnskap og engasjementet rundt eksamensoppgavene spesielt, er også grunner til at de er tatt med i analysen. Datagrunnlaget som blir analysert i dette delkapittelet er oppgavene våren 2015, vurderingsveiledninger og sensorskolering, samt sensorenes tilbakemeldinger på dette. Detaljer rundt datagrunnlaget ble redegjort for i 4.2.1.

5.1.1 Eksamensoppgavene i norsk hovedmål 2015

Felles for de enkelte eksamensoppgavene er at det overordnede formålet er at elevene «(...) skal vise den norskfaglige kompetansen sin. De skal lese ulike tekster og kunne sette dem inn i en faglig sammenheng, og de skal skrive en kort og en lengre tekst» (Utdanningsdirektoratet, 2015a, s. 2). Dette er knyttet til kompetansemålene fra hovedområdene *Skriftlig kommunikasjon* og *Språk, litteratur og kultur* på VG2 og VG3. Det utdypes også hva kompetanse er: «Kompetanse er i denne sammenhengen definert som evnen til å møte en kompleks utfordring eller utføre en kompleks aktivitet eller oppgave (op.cit, s. 5).

Det vil si at alle oppgavene skal gi mulighet til å utføre en kompleks oppgave som viser lese- og skrivekompetanse i tillegg til annen norskfaglig kompetanse. Våren 2015 besto eksamen i norsk hovedmål (vedlegg 4) av en felles kortsvaroppgave og fire ulike alternativer for langsvareoppgaven. Kortsvaroppgaven gikk ut på å skrive en informativ tekst som forklarer, med konkrete teksteksempler, hvorfor et vedlagt utdrag fra sluttscenen i *Soga om Gisle Sursson* er typisk for sagalitteraturen. I analysen av de fire langsvareoppgavene gjengir jeg

først hver enkelt oppgave, før jeg kort analyserer hvilken kompetanse den etterspør. Jeg vurderer også oppgaven med utgangspunkt i tidligere refererte erfaringer fra *Normprosjektet* knyttet til sammenheng mellom elevenes prestasjoner og oppgavens utforming, herunder krav til innhold og oppbygging samt felles referanser fra undervisning (Smemo og Solem). Analysens struktur følger eksamens overordnede formål og kompetansen som skal vises og vurderes på eksamen: *Lesekompetanse, Skrivekompetanse og annen norskfaglig kompetanse*.

Oppgave 1

Vedlegg:

- Arnulf Øverland: «Fri sprogutvikling»

Arnulf Øverland hadde en sentral posisjon i språkdebatten på midten av 1900-tallet. Den vedlagte teksten er et innlegg i debatten om samnorsk. Gjør greie for noen av synspunktene i teksten, og forklar kort den språkhistoriske sammenhengen. Drøft deretter denne påstanden fra teksten: «Sprogutviklingen er en naturlig prosess. Den lar seg hverken hindre eller påskynde.»

Kommentar:

Oppgaven er todelt. Du skal vise at du forstår Øverlands tekst og har kunnskaper om språkdebatten på hans tid. I den siste delen skal du vise at du har kunnskaper om språkutvikling. Du kan trekke inn eksempler du kjenner fra både fortid og nåtid.

Skrivekompetanse: Formålet er å skrive en informativ og argumenterende tekst.

Oppgavekommentarens «Oppgaven er todelt ...» utdyper dette gjennom en ramme for hvordan skriveren skal strukturere teksten sin. Eksaminandens «kunnskap om språkutvikling» skal danne underlag for drøftingsdelen. Drøftende tekster er gjerne de vanskeligste tekstene for elever å skrive (Igland, 2007), noe som kan bidra til å gjøre oppgaven mer kompleks.

Lesekompetanse: Eksaminanden skal trekke ut synspunkter fra tekst av Arnulf Øverland, forklare språkhistorisk sammenheng og drøfte en av påstandene i teksten. Den faglige tematikken krever relativt høy lesekompetanse. Det gjør også Øverlands vanskelige stil, som kan ses i konstruksjoner som: «partikularistisk sneversyn» og: «I det politiske klokkebytte står kulturverdiene høyt, og på Stortinget er vårt sprog til salgs for en spottstyver».

Annen norskfaglig kompetanse: Eksaminanden skal vise språkhistorisk kunnskap fra midten av 1900-tallet og nyere tid. En kan innvende at samnorskens vekst og fall tilhører en vanskelig del av pensum, fordi det er koplet til både fortidig politikk og en rekke rettskrivingsreformer med tilhørende fagbegreper. I tillegg skal kunnskap om et annet språklig

emne, språkutvikling, brukes i argumentasjon. Dette innebærer risiko for at eleven kan velge og støtte seg for mye til egne notater, noe flere tilbakemeldinger i sensorrapporten bekrefter.

Oppgave 2

Vedlegg:

- Tarjei Vesaas: «Kvart menneske er ei øy»
- John Donne: «Intet menneske er en øy»

Ta utgangspunkt i de to tekstvedleggene, og skriv en kreativ tekst der du reflekterer over tematikken i tekstene. Lag overskrift selv.

Kommentar:

Denne oppgaven åpner for ulike løsninger, men det skal komme tydelig fram i svaret ditt at du tar utgangspunkt i vedleggene. Du kan vise kreativitet gjennom de innfallsvinklene og perspektivene du velger, og gjennom bruken av språklige virkemidler.

Skrivekompetanse: Formålet er todelt. Eksaminanden skal reflektere gjennom en kreativ tekst. Om en ser på presiseringen av at det skal komme tydelig fram at utgangspunktet er tatt i de vedlagte tekstene, kan en si at hovedformålet med oppgaven er å reflektere, en skrivehandling som tidligere er problematisert (Dagsland, 2015). Oppgavens kommentar «(...) åpner for ulike løsninger» understreker åpenhet. Kommentaren forklarer hva som *kan* menes med kreativitet, men ikke hva som ligger i skrivehandlingen reflektere, noe som kan bidra til at eksaminander nedprioriterer dette formålet. Det er ikke eksplisitte krav til struktur i oppgaven, så eksaminanden er overlatt til sin sjangerkunnskap (Hertzberg, 2001). Mens flere elever kan ha problemer med å dekode hva reflekterende og kreative tekster er (Blikstad-Balas, 2015b, s. 30), kan en høykompetent kandidat få rom til å briljere (Kvistad & Smemo, 2015, s. 239).

Lesekompetanse: Refleksjonen skal ta utgangspunkt i tematikken i to korte tekstutdrag, som eksaminanden selv må definere. Dette momentet gjør oppgaveformuleringen mer åpen enn tilsvarende oppgave i eksamenssettet fra 2014, der tematikken ble verbalisert. Ovennevnte uklarhet rundt oppgavens formål, som kan komme av oppgavekommentarens utdyping av hva kreativ, men ikke reflekterende skriving er, vil kunne kompenseres av eksaminandens evne til å lese oppgaveteksten nøye og oppfatte den.

Annen norskfaglig kompetanse: Utover årstall og forfatternavn er tekstvedleggene dekontekstualisert og eksaminanden skal selv tolke hva som er tema i dem. Dette gir få felles referanser til hva elevene skal skrive om og overlater dem i stor grad til egne erfaringer (Tønnessen & Maagerø, 2015). Dette kan redusere en del av eksaminandenes mulighet til å nå

overordnet formål med eksamen, som blant annet er å «(...) lese ulike tekster og kunne sette dem inn i en faglig sammenheng» (Utdanningsdirektoratet, 2015a, s. 2).

Oppgave 3

Vedlegg:

- Erna Osland: «Mikke Mus»

Tolk novella.

Kommentar:

Du skal skrive om både form og innhold. Du skal gjøre greie for virkemidlene og den funksjonen de har, og begrunne dine egne tanker om teksten. Bruk relevant fagspråk.

Skrivekompetanse: Formålet er å skrive en novelletolkning. Instruksjonen gir rammer for hvordan skriveren skal strukturere teksten sin ved å presisere at hun skal skrive om både form og innhold. Videre skal eksaminanden redegjøre for virkemidler og deres funksjon, og skal begrunne tolkningene.

Lesekompetanse: Novellen som eksaminanden skal tolke er kort, skrevet på nynorsk og med et lett tilgjengelig språk uten vanskelige ord, som innledningen: «Midt på natta er det han hører den. Gråten. Er det nokon som græt?». Likevel kan den tolkes på flere måter. Kravet om begrunnede tolkninger vektlegger at eksaminanden skal vise frem sin lesekompetanse.

Annen norskfaglig kompetanse: Oppgavekommentarens instruks om at eksaminanden skal bruke relevant fagspråk, peker mot kunnskap om analyse av episke tekster. Instruksjonen om virkemidler og deres funksjon peker i samme retning. Dette gir felles referanser.

Oppgave 4

Vedlegg:

- «Barnebrud»

Denne sammensatte teksten er de tre første sidene av et kampanjemagasin fra bistandsorganisasjonen Plan Norge. Magasinet ble sendt i posten til medlemmene av organisasjonen og var vedlegg i Dagbladet 20.09.2014. Analyser utdraget. Bruk begreper fra retorikken. Vurder til slutt hvor godt du mener kampanjen fungerer etter formålet.

Kommentar:

Oppgaven er todelt. I den første delen skal du analysere både tekster og bilder og samspillet mellom dem. I den andre delen skal du begrunne vurderingene dine og bruke konkrete eksempler fra teksten.

Skrivekompetanse: Formålet er å skrive en retorisk analyse av en sammensatt tekst. Instruksjonen gir rammer for hvordan skriveren skal strukturere teksten sin ved å oppgi at den skal være todelt, der eleven først skal analysere tekst og bilde og deretter begrunne egen vurdering.

Lesekompetanse: Eksaminanden skal analysere tre sider fra en brosjyre. Tekstvedlegget er kontekstualisert og forklart. De tre sidene består av verbaltekst og bilder, i tillegg til typografi og fargevalg som skal forstås multimodalt og retorisk. Kravet om begrunnede vurderinger og teksteksempler understreker lesekompetansens viktighet.

Annen norskfaglig kompetanse: Oppgaveformuleringen ber om begreper fra retorikken, mens i kommentaren til oppgaven kommer det også frem at både tekst og bilder skal analyseres alene og i samspill, som ikonotekst. Dette gir felles referanser fordi det peker til flere ulike deler av læreplanen: retorikk, bildeanalyse og sammensatte tekster.

Analysen viser at alle oppgavene krever spesifikk lese- og skrivekompetanse. Når det kommer til annen norskfaglig kompetanse og kompleksitet, kan det se ut til at oppgave 2 ikke eksplisitt krever annen norskfaglig kompetanse, uten at det gjør den mindre kompleks. En kan si at 1 og 4 er mer faglig krevende enn 2 og 3, fordi de ber om norskfaglig kunnskap fra ulike deler av læreplanen, men samtidig er de relativt eksplisitte på innholdet, noe som igjen kan gjøre dem mer overkommelige. I teorikapitlet 2.1.4 om skriveoppgaver refererte jeg nylige funn knyttet til sammenheng mellom skriveoppgaver og gode besvarelser, der et funn nettopp var at de viste til felles referanser og inneholdt tydelige krav til innhold og struktur (Kvistad & Smemo, 2015, s. 239). Kategoriene i tabellen under tar utgangspunkt i dette, og jeg har forsøkt å vise hvordan eksamensoppgavene møter slike kjennetegn på gode skriveoppgaver:

	Felles referanser	Krav til innhold	Krav til oppbygging
Oppgave 1	Språkdebatten på midten av 1900-tallet. Debatten om samnorsk. Språkutvikling i fortid og nåtid.	Gjør greie for noen synspunkt i teksten Forklar kort språkhistorisk sammenheng Drøft påstand	Oppgaven er todelt: Først vise forståelse av Øverland og tidens språkdebatt, så kunnskaper om språkutvikling
Oppgave 2		Reflekter over tematikken i to vedlegg	Skriv en kreativ tekst der du reflekterer Bruk av språklige og litterære virkemidler Det skal komme tydelig frem at utgangspunkt er tatt i vedleggene
Oppgave 3	Episke tekster Fagspråk	Tolk novelle. Gjør greie for virkemidlene og funksjonen de har. Bruk relevant fagspråk.	Skriv om både form og innhold.
Oppgave 4	Retorikk Sammensatte tekster Bildeanalyse Fagspråk	Analyser utdraget. Bruk begreper fra retorikken.	Oppgaven er todelt. I første del analysere tekst, bilder og samspill.

		Vurder hvor godt du mener kampanjen fungerer etter formålet.	I andre del begrunne vurderinger og bruke konkrete teksteksempler.
--	--	--	--

Oppsummert har oppgave 1, 3 og 4 det felles at de viser til felles referanser ved å etterspørre annen norskfaglig kompetanse enn lese- og skrivekompetanse. Disse tre har også konkrete krav til innhold og struktur. Oppgave 2 viser til felles referanser som angår skrivekompetanse, men ingen med betydning for innholdet i teksten som skal skrives. Det gjør den mer åpen. Oppgaven inneholder krav, men de er uklare fordi oppgavens forklaring av kreativ tekst kun sier noe om hva det *kan* være og fordi det ikke forklares hva som ligger i å *reflektere*. En forklaring her kunne klargjort krav til innhold og oppbygging, i tillegg til å understreke dette formålet med oppgaven. Eneste eksplisitte krav som formidles, er at det skal være tydelig at utgangspunkt er tatt i tekstvedlegg. Derfor kan en forvente at sensorene har lagt vekt på det.

5.1.2 Kontekst: Grunnlag for vurdering av eksamensoppgavene

Sentrale myndigheter evaluerer gjennom eksamen i hvilken grad målene i læreplanen er nådd. Cecilie Bach skriver om eksamenskonteksten at «(...) fordi oppgavene når så mange, gir de signaler om hva som er viktig i faget» (Bach, 2015, s. 30). Eksamensoppgavene kan altså oppleves som styrende fortolkning av læreplanen. Videre kan eksamensveiledning, vurderingskriterier og andre rammebetingelser ses som styrt vurdering, noe som også er noe av formålet ved sentralgitt eksamen. I denne sammenheng er det imidlertid interessant å lese dette som retningslinjer og grunnlag for sensorenes vurderingsarbeid.

Utdanningsdirektoratet utarbeider styringsdokumentene som skal danne grunnlaget for normfellesskapet blant de offentlig oppnevnte sensorene. For konkretisering og avklaring av spørsmål rundt oppgaver og vurderingsveiledning, arrangeres en sensorskolering noen dager etter at eksamen er gjennomført. Møtet ledes av sensorer fra *Fagnemnda i norsk*, som representerer Utdanningsdirektoratets normer. Jeg vil gå gjennom retningslinjene som er gitt for sensorenes arbeid, i både dokumenter og på sensorskoleringen. Der det er relevant vil jeg trekke inn sensorenes tilbakemeldinger på rammene for vurderingsarbeidet, slik disse fremkommer i *Sensorrapporten* fra 2015. I kapittel 2.3.1 ble det vist til diskusjoner i skrivefeltet i etterkant av endringene som ble gjort i eksamensoppgavene i 2014, og jeg vil her starte med å se på hvordan de samme endringene er beskrevet av Utdanningsdirektoratet.

5.1.2.1 Veiledning til og kriterier for vurdering av eksamen

Jannike Ohrem Bakke og Karianne Skovholt definerer samtalen i klasserommet som «(...) viktigste redskap for å introdusere, forstå og diskutere skriveoppgaver» (Bakke & Skovholt,

2015 s. 139). Det er ikke bare oppgaveformuleringen som gir premisser for utforming av teksten, men også lærerens introduksjon (ibid). Til eksamen er eksaminanden prisgitt den skriftlige oppgaveformuleringen og hvordan den utdypes i veiledning og vurderingskriterier. Kriteriene skal også kunne øke vurderingens reliabilitet (Ecclestone, 2001).

Eksamensveiledning med vurderingskriterier 2014

Utdanningsdirektoratet oppdaterer eksamensveiledningen i norsk hvert år, og veiledningen fra 2014 ble også revidert i tråd med oppgaveendringene. Derfor ser jeg først på denne.

Eksamensveiledningen gjøres offentlig på vårparten og skal være kjent for elever, lærere og foresatte. Sensorene skal også bruke veiledningen «(...) som en felles referanseramme i arbeidet sitt» (Utdanningsdirektoratet, 2014 s. 2). Dette er med andre ord en viktig brikke i normfellesskapet som skal sikre reliabel vurdering.

I første av fem deler redegjør Utdanningsdirektoratet for hvordan skriftlig eksamen i norsk er organisert. Her antydes det også noe om hvordan oppgavene kan bli formulert. Det opplyses også om at hver oppgave er utstyrt med en *oppgavekommentar* «(...) med felles informasjon til eksaminand og sensor om hva som blir forventet av svaret» (op.cit, s. 3). Dette begrunnes i at formålet er å redusere misforståelser. Både kort- og langsvareoppgavene tar utgangspunkt i vedlagte tekster, fordi det ikke bare er skrivekompetansen som skal vises og vurderes til eksamen, men også elevenes kompetanse til å «(...) lese og forstå tekster og vurdere hva som er relevant for å svare på oppgaven» (op.cit, s. 4). Andre del beskriver revidert læreplan og endringer i oppgaveformuleringene. I del tre beskrives hjelpemidler elevene kan bruke. Bruk av kilder kreves ikke, men direktoratet presiserer at «Ukritisk bruk av fagstoff som ikke er relevant for oppgaven vil trekke ned» (op.cit, s. 4). Fordi elevene har mulighet til å bruke de fleste hjelpemidler, kan det gå inflasjon i kildegjengivelser hos de som er usikre.

Del fire og fem omhandler vurdering, som forankres i læreplaner og relevante forskrifter til opplæringsloven. Det er to sensorer som vurderer hvert eksamenssvar, og det gjentas at det er samlet kompetanse som eleven viser som ligger til grunn for vurderingen. I del fem gjengis opplæringslovens generelle karakterbeskrivelser, som blir nærmere forklart gjennom *kjennetegn på måloppnåelse*. Nevnte kjennetegn er organisert i tre deler og skiller lav fra middels og høy måloppnåelse. Kjennetegnene skal brukes av sensorene «(...) til å grunngi karakterforslaget sitt» (op.cit, s. 5), og nytt av året 2014 – etter innspill fra sensorer året før – er at kjennetegnene er tilpasset ulike oppgavetyper (ibid). Kjennetegnene er så listet opp i et tosidert skjema, før de gjentas i fire ulike vurderingsskjema som sensorene skal bruke, der det

første skjemaet er felles for alle, mens de tre neste er for ulike typer av langsvaroppgaver (vedlegg 4). Det at de generelle karakterbeskrivelsene her er nærmere forklart, gjør at *kjennetegn på måloppnåelse* i eksamensveiledningen ikke bare er konkrete utdypinger til sensorene, men at det i skrivende stund også er den mest konkrete instruksjonen til skriveledere om vurdering av skriftlige tekster. Dette siste er ikke et formål med veiledningen, men en praktisk konsekvens fordi læreplanen i norsk ikke inneholder andre kjennetegn på måloppnåelse enn forskriftens generelle karakterbeskrivelser.

Eksamensveiledning med vurderingskriterier 2015

Eksamensveiledningen for 2015 er relativt lik den fra året før. Langt på vei beskrives oppgavene på samme måte som i 2014, men begrepet *teksttyper* er mer eller mindre erstattet med begrepet *oppgavetyper*. Imidlertid får den «gamle» oppgavetypen *Tolkninger og analyser* en avgrensning: «Til eksamen bruker vi *tolkning* i forbindelse med skjønnlitterære tekster og *analyse* i forbindelse med sakprosa» (Utdanningsdirektoratet, 2015a, s. 3). I tillegg kommer en presisering av at kreativitet skal honoreres i alle oppgaver.

Informerende, argumenterende og resonnerende tekster beskrives som en oppgave som gis når elevene skal skrive saklige tekster, mens det i 2014 ble beskrevet som en teksttype som viste til formålet med teksten (Utdanningsdirektoratet, 2014). Som i 2014 utdypes de tre skrivehandlingene på denne måten: *Informere* suppleres med verbene presentere, forklare eller gjøre greie for, *argumentere* med grunngi og forsvare, og *resonnerere* med drøfte eller reflektere (Utdanningsdirektoratet, 2015a).

Den *kreative* teksten omtales noe annerledes i 2015. Den er fortsatt et samlebegrep, men nå dekker den skrivehandlingene *fortelle, reflektere og underholde*, mens det i 2014 var snakk om å «(...) fortelle og dikte, eller å reflektere og prøve ut tanker» (Utdanningsdirektoratet, 2014). At oppgaven er mer åpen for individuelle løsninger enn de tidligere sjangeroppgavene ble også presisert i 2014, men i 2015 legges det til at sensor skal honorere selvstendige valg når svaret er relevant, samtidig som det er «(...) viktig å bruke de vedlagte tekstene og svare på det oppgaven ber om» (Utdanningsdirektoratet, 2015a: 4). I 2015 blir det også gitt noe mer informasjon om hva slags oppdrag elevene kan få i de kreative oppgavene. Deretter heter det at «Uavhengig av hvilken oppgavetype de velger å svare på, kan elevene vise selvstendighet og kreativitet i tekstene sine. Det kommer til uttrykk i hvordan de bruker språket, og i ideer, perspektiver og innfallsvinkler. Dette skal sensor honorere når svaret er relevant» (ibid.).

Dette sto ikke i veiledningen fra 2014, og kan indikere et svar på tidligere omtalte kritikk av løsrivelse av kreativ skrivekompetanse (Smidt, 2015).

Når det kommer til vurderingskriterier og kjennetegn på måloppnåelse (vedlegg 4), er 12 kjennetegn felles for langsvaroppgavene. Disse gjelder de tre kategoriene *Hovedinntrykk*, *Språk og formelle ferdigheter* og *Kortsvar*. I tillegg er det ni kjennetegn for *Informative*, *argumenterende* og *resonnerende* tekster, ti for *Analyser*, *tolkninger* og *sammenlikninger* og syv for *Kreative tekster*. Kjennetegnene er fortsatt gradert i tre nivåer. Middels måloppnåelse er i tre av kriteriene i 2015 skilt gjennom en skråstrek (eksempel: «viser nokså god/god faglig kunnskap»), noe som kan antyde en forskjell mellom karakterene 3 og 4. Fordi vurderingskriteriene ikke inngår i eksamenssettet, kan det variere om elevene har lest dem. Analysen av oppgavene viste at det var færre klare krav til innhold og oppbygging i oppgave 2, og derfor kan flere her ha hatt behov for vurderingskriterienes utdyping. Jeg gjengir her kjennetegn for høy måloppnåelse, og da kun de som er unike for de respektive oppgavetyperne (to er satt i parentes da de ikke er relevante til årets eksamensoppgaver):

Vurderingsveiledningens kjennetegn for høy måloppnåelse i de ulike oppgavene:

Informative, argumenterende og resonnerende tekster	Analyser, tolkninger og sammenlikninger	Kreative tekster
Oppgave 1	Oppgave 3 og 4	Oppgave 2
gjør greie for faglige emner på en presis måte	bruker presist og relevant fagspråk	viser kreativitet i valg av innfalls- vinkler og språklige virkemidler
formulerer hovedsyn presist	gjør greie for virkemidler og viser hvordan de fungerer i teksten	reflekterer, prøver ut tanker på en selvstendig måte
argumenterer saklig, på en overbevisende måte	gjør presist greie for argumentasjon/bruk av appell- former og funksjonen de har	bruker litterære virkemidler på en gjennomført måte
har klare resonnementer, viser evne til å drøfte et emne saklig og nyansert	gjør greie for tema/budskap/hovedsynspunkt på en selvstendig måte	(viser til og reflekterer over språklig/kulturhistorisk kunnskap dersom det er relevant)
har poengterte oppsummeringer eller konklusjoner	reflekterer over tekster på en selvstendig måte, har vurderinger og grunngitte synspunkter	bruker vedlagte tekster på en kreativ og selvstendig måte
viser meget god evne til refleksjon		
bruker eventuelle kilder relevant, etterprøvbart og selvstendig	(sammenlikner tekster på en oversiktlig og god måte)	
	viser at eleven kan plassere tekster kulturhistorisk, med forståelse for sammenhengen	
	bruker eventuelle kilder relevant, etterprøvbart og selvstendig	

Alle oppgavetyperne har kriterier som åpner for skjønn, for eksempel «meget god evne til refleksjon» til oppgave 1. Likevel viser referansene til «faglige emner» i oppgave 1 og «fagspråk» i oppgave 3 og 4 at det er felles referanser til annet norskfaglig innhold i disse oppgavene, noe som kun er en opsjon i oppgave 2. Oppstillingen viser også at det er færrest kriterier for den kreative oppgaven, noe som gir færre holdepunkter til vurderingsarbeidet av denne. De beste svarene skal bruke litterære virkemidler på en gjennomført måte, reflektere selvstendig og bruke tekstvedlegg på kreativ og selvstendig måte. Lest sammen med oppgaveformuleringen, ser en at det her konkretiseres hva som menes med skrivehandlingen *reflektere*. Det er å prøve ut tanker, noe som er et sentralt sjangertrekk ved essayet (Aarnes & Aarnes, 1977, s. 61-63) (Federl, 2009). Det er også interessant at oppgavens opsjonelle *kan* i vurderingsveiledningen er tydeligere her: Kreativitet vises gjennom valg av innfallsvinkler, språklige virkemidler, måten vedlagte tekster blir brukt på, og det er et eget kriterium knyttet til bruk av litterære virkemidler. Selv om ett av kriteriene antyder sjanger, er ikke sjanger spesifisert. Dette kan gjøre vurderingen krevende, også fordi noen velger utradisjonelle sjangrer som krever andre lesemåter (Skar & Aasen, 2015, s. 13).

5.1.2.2 Sensorskolering 2015

På sensorskoleringen er Fylkesmannen representert, men avsender er Utdanningsdirektoratet og møtet ledes av representant fra *Fagnemnda i norsk*, som representerer direktoratets normer. Innspill i møtet og gjennom sensorrapporten viste enighet om at sensorskoleringen bør være obligatorisk. Samtidig er det grunn til å tro at de fleste sensorene deltar, fordi 83% av de som har bidratt til sensorrapporten oppgir å ha vært her (Utdanningsdirektoratet, 2015e). Direktoratet presiserte også på fellessensuren at ved uenighet er det sensorene som har vært på skolering som har mest å si (Utdanningsdirektoratet, 2015b).

Møtet ble innledet med retningslinjer for vurderingen og presentasjon av oppgavene, før sensorene skulle lese og vurdere fem eksempelsvar i grupper. Vurderingene ble oppsummert i plenum. Sensorene ble oppfordret til å oppsøke sine medsensorer. Det var en endring som ifølge sensorrapporten fikk gode tilbakemeldinger (Utdanningsdirektoratet, 2015e). Formålet med skoleringen er at nasjonal sensur skal være rettferdig og pålitelig. Det krever dialog og tolkningsfellesskap (Utdanningsdirektoratet, 2015d, s. 3). Vurderingen skal være valid og reliabel, der validitet ble definert til «(...) at eksamen måler elevens fagkompetanse på en troverdig måte» (op.cit, s. 6) og at oppgavene er basert på læreplanen. Reliabilitet er «(...) at sensorene vurderer svarene på samme måte», at det forutsetter likeverdige forhold, og er avhengig av sensorenes kompetanse og holdninger (ibid.). Her ble det sagt at åpne oppgaver

kan utfordre sensuren, men at det er en utfordring som følger av læreplan og fagets egenart. Det ble også presisert at om sensor mener oppgaven er uklar eller vanskelig, skal det meldes direktoratet og ikke gå ut over eleven. Sensor skal bruke vurderingsskjemaet, være lojal mot tolkningsfellesskapet og se etter kompetansen som eleven faktisk viser (op.cit, s. 10).

På sensorskoleringen stilte Utdanningsdirektoratet spørsmålstegn ved om sensorene er enige om hva norskfaglig kompetanse er (op.cit, s. 5). Det ble minnet om at formålet med eksamen er at elevene skal vise *lesekompetanse*, *skrivekompetanse* og *anvendt kunnskap* (ibid).

Sistnevnte kalte jeg, i tråd med eksamensveiledningen, *annen norskfaglig kompetanse* i analysen. Det kan være interessant å se dette innspillet opp mot fagnemndmedlem Bachs gjennomgang av sensorrapporter fra 2012 og 2014, der en av to gjengangere var nettopp at mange sensorer kritiserte mangelen på «norskfaglig kunnskap» i noen av oppgavene (Bach, 2015, s. 29). Under sensorskoleringen svarte altså Utdanningsdirektoratet på dette ved å si at mange sidestiller begrepet «norskfaglig» med fagstoffet i hovedområdet *Språk, litteratur og kultur*, mens sensorene må se på hele læreplanen (Utdanningsdirektoratet, 2015d, s. 5). Det er naturlig å tenke seg at tolkningsfellesskapet som sensorene i norsk skal utgjøre både består av dem som mener det er for mye og for lite dannelse gjennom språk- og litteraturhistorie i både læreplan og eksamensoppgaver, og at dette kan utfordre det samme fellesskapet, kanskje spesielt i møte med uklare eller åpne oppgaveformuleringer.

Da årets oppgaver ble presentert, ble spørsmålet om vekting av kort- og langvaroppgave tatt opp fra salen. Her ble sensorene minnet på at alt skal helhetsvurderes (Utdanningsdirektoratet, 2015c). I sensorrapporten fra 2015 er det forholdsvis mange som gir uttrykk for at det er uheldig og frustrerende at det ikke gis klarere føringer her (Utdanningsdirektoratet, 2015e).

Til **oppgave 1** ble det sagt at «(...) dette er kanskje den vanskeligste oppgaven i settet» (Utdanningsdirektoratet, 2015c). Det kom frem fra salen at svarene ofte led under unødig mye parafisering, og her ble det vist til at sensorene fikk vurdere kandidatens kompetanse til å bruke relevant kildetilfang (ibid.). I sensorrapporten er det enighet om at denne oppgaven var vanskelig. Noen sensorer mente at dette gjorde det enkelt å skille mellom kandidatene, mens andre pekte på at de svakeste kandidatene var dømt til å mislykkes fordi Øverlands tekst motsier seg selv (Utdanningsdirektoratet, 2015e). **Oppgave 2** kunne besvares med andre tekster enn essay, og refleksjon gjøres ved «(...) å vise fram ulike sider ved tematikken». (Utdanningsdirektoratet, 2015d, s. 23). Om det reflekteres uten kreativitet, skulle sensor vurdere i hvilken grad teksten ellers svarer på oppgaven, i tillegg til å se på kortsvaret (ibid.). På spørsmål fra salen om hvor synlig de to tekstene måtte være i besvarelsen, ble det svart

muntlig at «(...) tankegods fra tekstene må være tydelig i svaret, så tydelig at du kan se det er skrevet samme dag» (Utdanningsdirektoratet, 2015c). Tilbakemeldinger i sensorrapporten uttrykker at sensorene hadde behov for flere presiseringer av denne oppgaven.

Oppgave 3, tolkningen, ble ikke omtalt som noe annet enn at det er en kjent oppgavetype (Utdanningsdirektoratet, 2015d, s. 27). Den var heller ikke representert blant tekstene som skulle vurderes, noe som kanskje reflekterer at Utdanningsdirektoratet ikke anser at den utfordrer normfellesskapet. I sensorrapporten mente noen at dette var uheldig, men det ble ikke meldt at den var vanskelig å vurdere (Utdanningsdirektoratet, 2015e). **Oppgave 4** hadde mange valgt, og det ble antydning at det nå kunne være på tide å gi denne oppgavetypen en pause (Utdanningsdirektoratet, 2015c). Kommentaren ble gitt muntlig, og jeg er usikker på om dette var et signal fra direktoratet eller en subjektiv betraktning, og om det var retoriske analyse eller sammensatt tekst som det nå var behov for pause fra. Mange sensorer viser gjennom sensorrapporten at de delte dette synet.

Eksempeltekstene som sensorene skulle vurdere og diskutere var ett svar på oppgave 1, to på oppgave 2 og to på oppgave 4. Under en oversikt over sensorenes vurderinger i plenum, basert på mine notater, der karakteren med størst oppslutning er uthevet:

Eksempel- oppgave	Karakterfordeling blant sensorene						Udirs karakter	Min karakter	Gjennom- snitt
	1	2	3	4	5	6			
1				13	25	7	6	5	4,87
2a	4	28	11				2	2	2,16
2b		2	19	21	1		4	4	3,49
4a			3	12	17	3	5	5	4,57
4b		1	14	17	4		3	4	3,67

Totalt antall vurderinger varierer mellom 35 og 45, siden ikke alle sensorer deltok på hele møtet. Jeg har tatt med egne vurderinger, for å vise hvor jeg plasserer meg, fordi jeg skal analysere tekstutvalget senere. Sensorene fremstår i større grad som et fellesskap i vurdering av eksempeloppgave 1 og 2a enn de tre andre, men Utdanningsdirektoratets vurdering av 1 avviker fra majoriteten. I sensorrapporten uttrykte flere sensorer frustrasjon rundt det at et ganske klart tolkningsfellesskap fikk direktiv om en annen vurdering, en praksis som ikke støttet konsensus-oppfordringen tidligere i møtet (Utdanningsdirektoratet, 2015e).

Sensorkorpset sprer seg noe mer i vurderingene av eksempeloppgavene 2b, 4a og 4b. Et knapt flertall deler Utdanningsdirektoratets vurdering av 2b og 4a, mens et knapt flertall er en karakter over direktoratets vurdering av 4b. Det kan ses som bekymringsfullt med tanke på

vurderingens reliabilitet at sensorene kun samler seg om vurderingen av to (1 og 2a) av de fem besvarelsene. Det er av samme grunn bekymringsfullt at Utdanningsdirektoratet og flertallet i sensorkorpset er i utakt i vurderingen av to av fem besvarelser (1 og 4b).

Eksempeloppgavene er nok valgt fordi Utdanningsdirektoratets representanter har ment at de kan utfordre normfellesskapet. Det er derfor rimelig å anta at de til dels sprikende vurderingene ikke er helt representative for vurderingen av populasjonens eksamenstekster. Om en igjen ser på vurderingene av de tre besvarelsene der det er mindre samsvar (2b, 4a, 4b), kan en med velvilje se at sensorene fremstår som samlet, selv om de altså fordeler seg på to karakterer. Dette fordi det er svært få av sensorene som gir andre karakterer enn disse to, og når for eksempel 21 og 19 sensorer (altså 40 av 45) gir henholdsvis 4 og 3 på samme besvarelse, trenger ikke det være mer dramatisk enn at denne besvarelsen *har* deler som står til både 3 og 4, og at det *er* vanskelig å gi en helhetsvurdering når en kandidat mestrer ulike delkompetanser i ulik grad. I praksis vil ikke to medsensorer ha store problemer med å la eventuell tvil komme en slik elev til gode ved fellessensuren. Likevel er det mer sprikvurdering av disse tekstene, og en kan spørre om elevene får reliabel vurdering om to avvikere skulle inngå i samme sensorpar, noe en kan regne med skjer, særlig hvis den kvantitative undersøkelsen skulle finne lavt vurderingssamsvar.

5.1.2.3 Sensorrapporten 2015

Jeg har sett på resultatene og kommentarene til spørsmålene knyttet til veiledninger (3, 4, 5 og 6) og oppgaver (8 og 9). Rapporten er presentert og utvalget er redegjort for i kapittel 3.2.1.

Spørsmål	Helt enig/ Delvis enig	Delvis uenig/ Helt uenig	Vet ikke/ Ikke besvart
3 Sensorskoleringen var nyttig for sensuren	81,48	1,06	17,46
4 Eksamensveiledningen var nyttig for sensuren	85,71	7,41	6,88
5 Kjennetegn på måloppnåelse var nyttig for sensuren	87,3	6,35	6,35
6 Vurderingsskjemaet var nyttig for sensuren	91,53	6,88	1,59
8 Det er godt samsvar mellom kompetansemålene og oppgavene	88,36	6,88	4,76
9 Oppgavene gir mulighet til å vise kompetanse på ulike nivå	88,89	7,94	3,17

Resultatet av den kvantitative undersøkelsen (Utdanningsdirektoratet, 2015e, s. 1-2) er at de fleste sensorene her er fornøyde med både oppgaver og vurderingskriterier. Kvalitative tilbakemeldinger til hvert av spørsmålene nyanserer dette inntrykket. Flere gir uttrykk for forbedringsbehov. I metodekapittelet redegjorde jeg for hvordan rapporten er lest, og jeg forsøker å oppsummere de kvalitative tilbakemeldingene.

5.1.2.4 Sensorene om sensorskolering, vurderingsveiledning og -kriterier

Kritikken fra sensorskoleringen, om at det er uheldig at det ikke gis klarere føringer om vekting av kortsvaret, gjentas i rapporten. Flere skriver at dette gir et inntrykk av at denne oppgaven er uviktig. Tilbakemeldingen som de fleste kommer med er at vurderingskriteriene er for generelle eller overflatiske. Dette er stort sett en generell tilbakemelding, men i noen kommentarer knyttes dette spesielt til skjemaet for kreative tekster. En del melder at det var bra at noe ble avklart på sensorskoleringen, men ikke alt det var behov for. Fordi skoleringen ikke er obligatorisk, knyttes det usikkerhet til normfellesskapet. En usikkerhet som ble forsterket av at Utdanningsdirektoratet fremsto som normative da de på skoleringen var i utakt med sensorcorpsets flertall i vurdering av to av eksempeloppgavene. Det ønskes også mer differensiering av vurderingsmatrisen, spesielt mellom karakterene 3 og 4, men flere etterspør også et skille mellom 2 og 1. Noen etterspør i den anledning konkretiseringer og eksempler på hva som forventes. Det er også en del sensorer som klager på at vurderingsskjemaene er for omfattende eller mange, og her tenkes det nok på at det nå er ett skjema per oppgavetype.

5.1.2.5 Sensorene om oppgavene

Spørsmålet om oppgavene får flest tilbakemeldinger. Oppgavene får generelt bedre tilbakemeldinger i 2015 enn i 2014 (Bach, 2015), men det er flere negative enn positive kommentarer. Det nevnes blant annet at det er uheldig at det er ulike krav til faglighet i langsvareoppgavene. Den retoriske analysen, kortsvareoppgaven og den kreative oppgaven får flest negative tilbakemeldinger.

Til **kortsvareoppgaven** er det anslagsvis fire ganger så mange negative som positive innspill. Sensorene mener det var vanskelig å vurdere eller differensiere, fordi så mange parafraserte lærebøker om sjangertrekk, mens oppgaven derimot sa at de ved hjelp av teksteksempler skulle påvise trekk i utdraget. Andre mener det motsatte; oppgaven er god fordi den gjør det enklere å skille ut de gode svarene til de som støttet seg på egen lesekompetanse.

Oppgave 1 var vanskelig, spesielt på grunn av kompleksiteten i tekstvedlegget. Det var uenighet om dette var bra eller dårlig, men det er anslagsvis fire ganger så mange negative som positive tilbakemeldinger. Med unntak av noen få som uttrykker seg positivt om **oppgave 2**, mener et flertall (fem til seks ganger så mange) at oppgaven har svakheter: Den er vanskelig å besvare godt, åpner for det subjektive og derfor ulik bedømming, og den fører til analyse eller tolkning. Videre måler den i liten grad konkrete kompetansemål, konkretiserer ikke sjanger, i tillegg til at begrepet «kreativ» er vagt og derfor vanskelig å vurdere. Det sies at det er problematisk at oppgaven ber om en mulig indre motsetning ved bruk av begrepene

kreativ og reflektere, noe som kan forvirre både kandidat og sensor (Utdanningsdirektoratet, 2015e, s. 28). Enkelte mener også at den på grunn av tidspress ikke egner seg på eksamen. Selv om enkelte uttrykker seg positivt, er det ingen som mener at den har vært grei å vurdere. Det uttrykkes blant annet usikkerhet rundt hva det vil si «å ta utgangspunkt i vedleggene».

De forholdsvis få som kommenterer **oppgave 3**, mener noe om selve analyseteksten. De som er negative, om lag dobbelt så mange som de positive, begrunner det med at novellen er for enkel for en eksamensoppgave, mens andre mente den er for vanskelig. **Oppgave 4** har fått flest tilbakemeldinger. De er i overveiende grad negative, blant annet fordi det er så mange som valgte oppgaven. En del spør om retoriske reklameanalyser nå er gitt så hyppig at mange av elevene har spesialisert seg: «En slik forutsigbarhet ikke er i samsvar med den bredden læreplanen representerer» (op.cit, s. 12). Dette er kritikk som ikke går på oppgavens utforming, men mulige konsekvenser av den. Vedlegget som skulle analyseres er dessuten for omfattende. I tillegg til etterlyser noen verbaltekst som analysetekst.

5.1.3 Oppsummering av dokumentanalyse av eksamensoppgaver i kontekst

Analysen av oppgavene viser at oppgave 1, 3 og 4 også etterspør annen norskfaglig kompetanse. Alle oppgavene har krav til innhold og struktur, men oppgave 2 er åpnere enn de andre i sin mangel på felles referanser med betydning for innholdet i teksten. Krav til struktur i oppgave 2 er også uklar fordi den kun sier hva en kreativ tekst *kan* være og fordi det andre formålet; *reflektere*, ikke er forklart. Tidligere referert forskning om skriveoppgaver tilsier på denne bakgrunn en mulig sammenheng mellom oppgave 2 og sprikvurdering. Analysen av vurderingskriteriene viser at oppgave 1, 3 og 4 også refererer til annen norskfaglig kompetanse i tillegg til lese- og skrivekompetanse. Disse har også noen flere vurderingskriterier enn oppgave 2. Refleksjonskriteriet til oppgave 2 antyder et essay, og kreativitetskriteriet er tydeligere på det som i selve oppgaven uttrykkes som «kan». Alle oppgavene har kriterier som overlater til skjønnsmessig vurdering, noe som på bakgrunn av referert forskning gir uklarhet som kan indikere lavere konsensus, og igjen utmerker oppgave 2 seg som den dette eventuelt vil ramme i størst grad.

På sensorskoleringen ble det formidlet at refleksjon i oppgave 2 kunne vises indirekte ved litterær skriving. Sensorene etterlyste likevel flere presiseringer, og her er det spesielt oppgave 2 og vektning av kortsvaret som reiste flest spørsmål. De fem besvarelsene som ble vurdert på sensorskoleringen fikk frem svakheter ved normfellesskapet som garantist for reliabilitet. Vurderingen av to av oppgavene spredte seg på tre ulike karakterer og

vurderingene av tre av dem spredte seg over fire. I tillegg avvek direktoratets vurderinger av to av tekstene fra majoriteten av sensorene.

Analysen av sensorrapporten viser at av langsvaroppgavene er det oppgave 2 og 4 som får mest kritikk. Forholdsvis mye av kritikken til oppgave 4 går på at så mange valgte den, mens mye av kritikken av oppgave 2 er knyttet til den kreative oppgavetypen. Kritikk til vurderingskriterier er mest generell, for eksempel knyttet til vekting mellom de to svarene og behov for differensiering mellom flere av karakterene, men der uklarhet er knyttet til konkrete langsvaroppgaver, er det kriterier til oppgave 2 som får mest kritikk. I Bachs analyse av sensorrapporten fra 2014, skrev hun at det er en mulig forventning at den kreative oppgavens «backwash-effekt vil gi mindre usikkerhet ved eksamen våren 2015» (Bach, 2015, s. 43). Kommentarene i sensorrapporten i 2015 viser noe annet. Med få hederlige unntak inneholder denne også en rekke tilbakemeldinger om at dette ble en vanskelig oppgave å sensurere, både fordi oppgaven er for åpen og vurderingskriteriene for generelle. Dette i kombinasjon med at sensorene ikke fikk avklaringer de etterlyste på skoleringen, er nok mye av grunnen til at mange også i 2015 foreslo å droppe oppgavetypen.

5.2 KVANTITATIV ANALYSE

5.2.1 Samlet sprikvurdering, korrelasjon og vurderingssamsvar i bruttoutvalget

Bruttoutvalget av 4556 eksaminanders resultater inneholder 633 sprikvurderinger.

Eksaminander som før sensormøtet har fått vurderinger som sprikte med to eller flere karakterer utgjør 13,9% av eksamenstekstene fra de tre fylkene. Herunder utgjør besvarelser som har fått sprikvurdering på to karakterer 12,1%, mens de som har fått sprikvurdering på tre og fire karakterer utgjør henholdsvis 1,6% og 0,2% av bruttoutvalget.

For å vise hvordan de to sensorenes karakterer er relatert til hverandre, er konsistens og korrelasjon beregnet ved å bruke Pearsons r . Her er resultatet 0.58 (avrundet fra 0.576). Det er relativt likt resultatet på 0.56 fra Berges undersøkelse av eksamen etter videregående (Berge, 1994, s. 254). Det er noe ulikt hva som anses som tilstrekkelig for reliabel konsistens. Barrett skriver at verdier over 0.7 kan være det (Barrett, 2001, s. 24). 0.58 er et stykke unna. Som jeg var inne på i metodekapittelet, blir det anbefalt å kombinere beskrivelser av konsistens og konsensus når en vurderer vurderingssamsvarets pålitelighet (Stemler, 2004, s. 1).

For å få et mål for i hvilken grad det er nettopp de identiske verdiene på de to variablene som kombineres, brukes en koeffisient for konsensus eller samsvar. «Målet for samsvar varierer mellom 1 for fullstendig samsvar og 0 når samsvaret ikke er større enn det som kan forklares

av tilfeldigheter» (Hellevik, 2002, s. 261). Ved utregning av prosentvis sammenfall, Scotts pi og Cohens kappa i programmet *Recal2* (Freelon, 2011), er resultatet for bruttoutvalget:

Prosentandel enighet	Scotts pi	Cohens kappa
39.4%	0.2 (avrundet fra 0.204)	0.2 (avrundet fra 0.205)

Prosentvis sammenfall er her en av flere måter å beskrive konsensus på. Stemler skriver at om skalaen går fra 1-7, kan det betraktes som konsensus selv om det er et skille på én verdi i de to vurderingene av samme tekst (Stemler, 2004, s. 4). Ved å legge dette perspektivet til grunn vil det i bruttoutvalget være sammenfall på 86%. Samtidig er det mer nedslående at målene Scotts pi og Cohens kappa er på 0.2. Det betyr at identisk resultat eller overensstemmelse i karakter kun inntreffer i en femtedel av tilfellene ut over det som kan forklares av tilfeldigheter. Det trenger altså ikke være mangel på enighet, men den er ikke reliabel: 0.2 ligger på grensen mellom det som beskrives som «poor» og «fair» (Landis & Koch, 1977). At Pearsons r er høyere, viser at tendensen er lik: høye karakterer hos en sensor går sammen med høye hos den andre og så videre, men altså ikke nødvendigvis med lik karakter. Det er mulig å få et lavt konsistensestimert samtidig som konsensusestimert er høyt, eller omvendt, som i dette tilfellet (Stemler, 2004, s. 5). Det understreker nødvendigheten av å beregne begge.

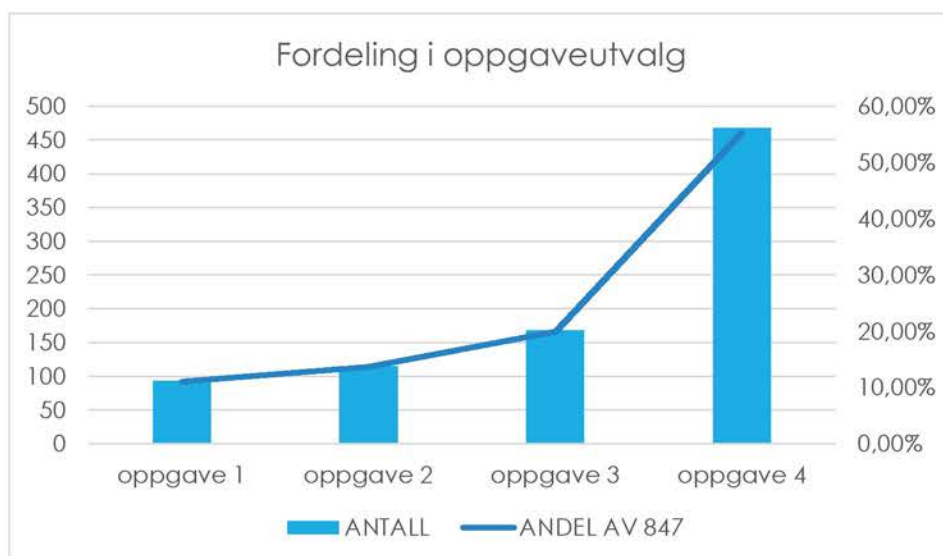
Beskrivelser av konsensus gjennom kappa er vanskelig å tolke (Stemler, 2004, s. 4). Det er også et konservativt mål, i tillegg til at resultatet kan bli lavt uavhengig av om årsakene til det er alvorlige eller ikke (op.cit, s. 5). I tillegg er det tungvint fordi hver sensorgruppe må regnes hver for seg (ibid). Fordi intervurderingsreliabiliten er på siden av problemstillingen, er ikke dette gjort for alle sensorgruppene. Beregninger gjort i et tilfeldig utvalg av fire sensorgrupper på fire sensorer hver, viser at variasjonen under er representativ:

Sensorgruppe 2	Prosent enighet	Scott's Pi	Cohen's Kappa
Førstesensor 20, andresensor 37	33,9%	0.128	0.146
Førstesensor 98, andresensor 20	34,4%	0.149	0.152
Førstesensor 37, andresensor 71	55,4%	0.401	0.402
Førstesensor 71, andresensor 98	27,7%	0.041	0.051

Her ser en at konsensus varierer mellom 0 og 0.4. Det kan indikere krevende samarbeid i flere sensorgrupper, noe jeg kommer tilbake til i tekstanalysen.

5.2.2 Eksaminanders oppgavevalg

I metodekapittelet redegjorde jeg for reduksjonen jeg måtte foreta i oversikten over eksaminanders oppgavevalg. Dette utvalget endte på 847 valg, og fordelingen mellom langsvaralternativene er operasjonalisert i tabellen under:



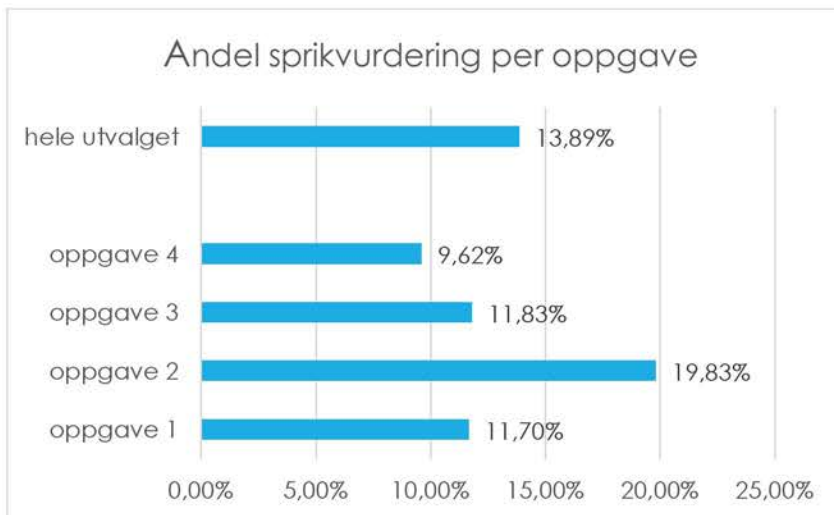
Oppgave 4 er valgt av 55%, noe som sammenfaller med refererte tilbakemeldinger i sensorrapporten (Utdanningsdirektoratet, 2015e). Rundt 20% valgte oppgave 3, novelletolkningen, 14% valgte den kreative oppgave 2 og tilslutt er den informative og argumenterende oppgave 1 valgt av 11% av eksaminandene. I KAL-studiene fant forskerne at relativt få skrivere i grunnskolen velger saksorienterte oppgaver, og at fortellende og ekspressiv skriving dominerer (Vagle & Evensen, 2005, s. 198). I oppgavesettet til eksamen etter videregående skole er det langt færre anledninger til dette, men at kun 14% her benytter eneste mulighet til fortellende og ekspressiv skriving, kan vitne om at det i mellomtiden har vært en betydelig endring i preferanser. Endringen kan skyldes endret skriftkultur, alder eller læreplaner, men angår ikke undersøkelsens problemstilling, og blir ikke drøftet nærmere.

5.2.3 Sprikvurdering og oppgaver

Ved å sammenholde oppgaveutvalget med sprikutvalget, har jeg beregnet andel sprikvurdering innenfor hver av de fire oppgavene. Resultatet er som vist i tabellen under:

	Andel sprikvurdert
Oppgave 1	11,7%
Oppgave 2	19,8%
Oppgave 3	11,8%
Oppgave 4	9,6%

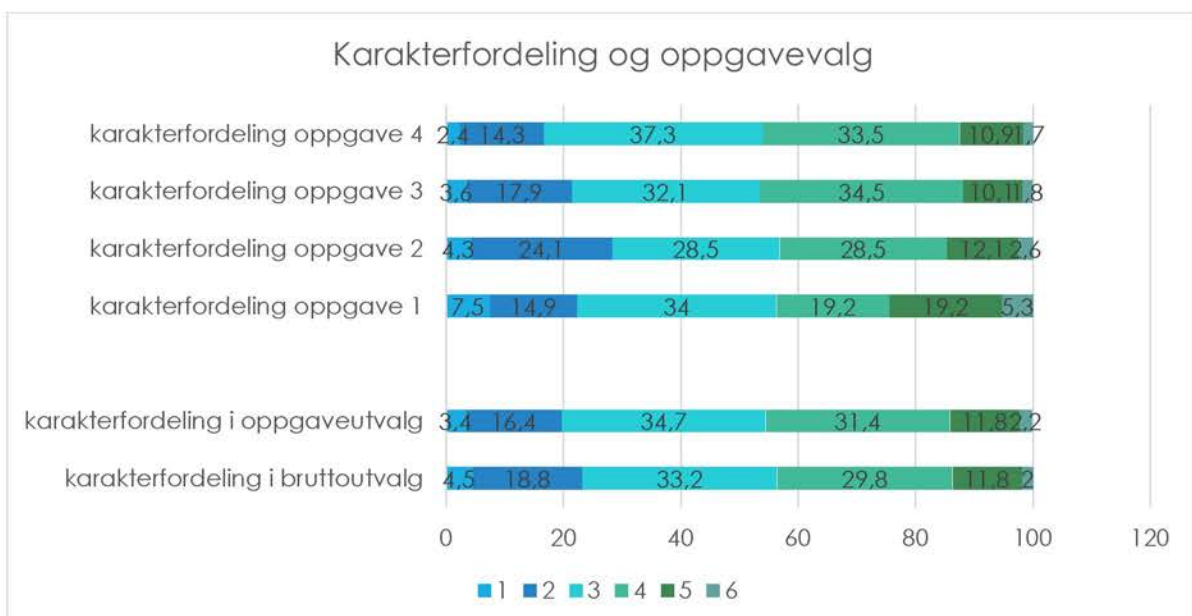
Dette kan også operasjonaliseres ved å bli vist opp mot sprikvurderingers andel av bruttoutvalget (633 av 4556 = 13,9%):



Det er bare oppgave 2 som har høyere frekvens av sprikvurdering enn det som gjelder for bruttoutvalget. Prosentvis andel av sprikvurderinger av oppgave 1 og 3 ligger noe under, mens andelen til oppgave 4 ligger vesentlig under andel i bruttoutvalget. Sprikvurderingene av oppgave 2 gjelder en av fem vurderte oppgaver, noe som har bidratt til å trekke samlet andel sprikvurderinger opp. Beregning i Zigne (Aardal og Berglund 2014) viser signifikante forskjeller mellom sprikvurderinger for hele utvalget og hver av de enkelte oppgavene.

5.2.4 Karakterfordeling i brutto- og oppgaveutvalg

Beregninger opp mot endelige karakterer til besvarelser i oppgaveutvalget er gjort for å se etter andre variasjoner i vurderingen enn sprikvurdering, men også som et ledd i utvalgsvalideringen, som ble drøftet i 4.4.3, av oppgaveutvalget. I den forbindelse dukket det opp noen sammenhenger mellom oppgavevalg og karakterer. Disse kan ses i diagrammet som operasjonaliserer sammenheng mellom karakterer og oppgavevalg:



Det er flere tendenser her, men jeg omtaler i hovedsak forskjellene som er signifikante.

Resultat etter signifikanstesting av forskjeller i karakterer i oppgaveutvalg i Zigne:

Oppgave 1	Flere har fått karakterene 1, 5 og 6 enn i resten av utvalget. Færre har fått karakteren 4
Oppgave 2	Flere har fått karakteren 2 enn resten av utvalget. Færre har fått karakteren 3.
Oppgave 3	Ingen signifikant forskjell med karakterfordeling i hele oppgaveutvalget
Oppgave 4	Ingen signifikant forskjell med karakterfordeling i hele oppgaveutvalget

Oppgave 1 og 2 skiller seg ut når det kommer til karakterfordeling. Oppgave 1 er den med størst spredning mellom karakterene. Størst andel av både stryk (7,5% av de som valgte oppgaven) og høy måloppnåelse (24,5%) finnes her. Den har også høyest gjennomsnittskarakter (3,43). Oppgave 2 har lavest gjennomsnittskarakter (3,27), noe som kan forklares av at 28,4% av dem som valgte oppgave 2 fikk lav eller ingen måloppnåelse. Det er også større andel som får ingen og høy måloppnåelse her enn blant oppgave 3 og 4, men disse forskjellene er ikke signifikante.

5.2.5 Korrelasjon og samsvar i oppgaveutvalget

Pearsons r er 0.58 og Scotts pi 0.2 for bruttoutvalget på 4556 besvarelser. Ved beregning av konsistens og korrelasjon i oppgaveutvalget på 847 besvarelser, viser dette seg å være omtrent det samme: Pearsons r 0.58 (avrundet fra 0.5804). Jeg ønsker også å finne hvordan de to sensorenes karakterer er relatert til hverandre innenfor de ulike oppgavene. Beregning ved Pearson korrelasjonskoeffisient gir dette resultatet:

Oppgave	Pearsons r
Oppgave 1	0.69 (avrundet fra 0.687)
Oppgave 2	0.46 (avrundet fra 0.462)
Oppgave 3	0.59 (avrundet fra 0.588)
Oppgave 4	0.56 (avrundet fra 0.558)

De analytiske oppgavene 3 og 4, som 75% valgte, utgjør hoveddelen av materialet og avviker i liten grad fra konsistens i bruttoutvalget. Den saksorienterte oppgave 1 har høyest korrelasjon. I følge Barrett (2001) er dette reliabel konsistens, noe som kun gjelder oppgave 1. Dette funnet (0.69) samsvarer med korrelasjonsfunn for hele KAL-utvalget (Berge, 2005a, s. 104), noe som ble omtalt som relativt høyt sammenliknet med internasjonal forskning (Evensen, 2005, s. 206). Det er imidlertid kun 11% av eksaminandene som har valgt oppgave 1. Det er også et funn som samsvarer med refererte funn i KAL, der høyest korrelasjon også ble funnet i oppgavene som var innrettet mot sak, og som få valgte (Vagle & Evensen, 2005, s. 173). Det samsvarer også med Berges undersøkelse fra videregående på 1990-tallet. Videre har den kreative oppgave 2 lavest korrelasjon. Resultatet 0.46 er nesten like lavt som det som

i forbindelse med de nasjonale skriveprøvene ble omtalt som svak pålitelighet (Berge, 2014, s. 8). Det avdekker mindre utviklet normfelleskap rundt denne oppgavetypen, alternativt at fellesskapet har fått større problemer med å møte det åpne og til dels uklare skriveoppdraget.

Ved beregning av konsensus gjennom prosentvis enighet og samsvar gjennom Scotts pi og Cohens kappa, sammenfaller resultatet for de ulike oppgavene med resultatet for bruttoutvalget, altså enighet i rundt 40% av vurderingene og samsvarskoeffisienter rundt 0.2.

Oppgave	Prosentandel enighet	Scotts pi	Cohens kappa
Oppgave 1	37.2	0.2 (avrundet fra 0.201)	0.2 (avrundet fra 0.208)
Oppgave 2	37.9	0.2 (avrundet fra 0.189)	0.2 (avrundet fra 0.201)
Oppgave 3	42.0	0.2 (avrundet fra 0.219)	0.2 (avrundet fra 0.222)
Oppgave 4	39.3	0.2 (avrundet fra 0.174)	0.2 (avrundet fra 0.175)

Det kan her se ut til å være tendenser som delvis peker i andre retninger enn korrelasjonsfunn til ulikheter mellom oppgavene. Det er imidlertid såpass små forskjeller mellom resultatene at de kan skyldes tilfeldigheter. Dermed ser det ikke ut til at oppgavetypen har betydning for konsensus mellom sensorene, noe som kanskje også skyldes at resultatet er så lavt at samsvaret er tilnærmet tilfeldig.

Når det kommer til prosentandel sensorer som har gitt sammenfallende karakterer, viser det seg etter utregning av prosentdifferansene i Zigne (Aardal & Berglund, 2014) at det ikke er signifikante forskjeller mellom de ulike oppgavene eller mellom enkeltoppgaver og bruttoutvalget. Det er uansett tankevekkende at et flertall av eksamenstekstene også her har fått ulike karakterer fra de to sensorene og at samsvaret er såpass tilfeldig.

5.2.6 Klageutvalget og oppgaver

Det er klaget på 283 av 4556 sensurerte besvarelser, og det utgjør 6,2%. Dette ligger noe høyere enn totalt antall klager etter eksamen i norsk hovedmål våren 2014, som er 5,4% (Utdanningsdirektoratet, 2015d), men tallene er ikke helt sammenliknbare. Tallene fra 2015 gjelder de tre undersøkte fylkene, mens tallene fra 2014 gjelder hele landet.

Ved å kople eksaminandene som har klaget opp mot utvalget av 633 sprikvurderinger fra sensuren, finnes ikke sammenheng mellom sprikvurdering og klagefrekvens. 44 av klagerne viser seg å ha fått sprikende vurderinger i forkant av fellessensuren. Det utgjør 15,6% av dette utvalget. Sprikutvalget i seg selv utgjør 13,9% av bruttoutvalget, og etter utregning i Zigne, viser det seg at denne forskjellen ikke er signifikant.

De 283 eksaminandene som har klaget er også koplet mot oppgaveutvalget. Tabellen under viser hvordan 49 treff fordeler seg på de fire langsvaroppgavene, sammenliknet med fordeling i bruttoutvalget:

	Klager fordelt på oppgavevalg	Fordeling av oppgavevalg i bruttoutvalget	Signifikans
Oppgave 1	18,4 %	11%	Signifikant
Oppgave 2	24,5 %	14%	Signifikant
Oppgave 3	12,2 %	20%	Signifikant
Oppgave 4	44,9 %	55%	Signifikant

Fordi dette er såpass få besvarelser, er det usikkert hvor stor vekt jeg kan legge på oversikten. Samtidig er det signifikante forskjeller mellom antall klager og valg av oppgaver. Oppgave 1 og 2 er valgt av 25% av eksaminandene, men er representert i 42,9% av klagen. Oppgave 3 og 4 ble valgt av 75% og utgjør 57,1 av klagen. Det kan indikere at eksaminandene og/eller sensorene har vært mer usikre på oppgave 1 og 2 enn 3 og 4. Når det kommer til oppgaven som er overrepresentert i sprikutvalget, oppgave 2, er det ingen av disse sprikoppgavene som er å finne på klagestatistikken, men altså en del andre besvarelser til oppgave 2. Det er for få treff mellom oppgavevalg, sprik og klager til at det gir mening å regne på det. Funn av manglende sammenheng mellom sprikvurdering og klager reduserer også behovet for dette.

5.2.7 Oppsummering av kvantitativ analyse

Andel sprikvurderinger er 13,9% av bruttoutvalget, og korrelasjonskoeffisienten mellom de to sensorene på 0.58 (Pearsons r). Utregning av koeffisienter for samsvar (Scotts pi og Cohens kappa) viser 0.2, og det gir tilnærmet tilfeldig konsensus. Sett opp mot oppgavevalg, er det én oppgave som gjennomgående utpeker seg. Den kreative oppgave 2 er mer sprikvurdert enn de andre i utvalget. Av de vi kjenner til som har valgt oppgave 2, ble 19,8% sprikvurdert, noe som er signifikant forskjellig fra de 13,9% som ble sprikvurdert i bruttoutvalget. De tre andre oppgavene ligger signifikant under dette snittet. Oppgave 2 har også lavest korrelasjon (0.46) og gjør seg bemerket på både strykstatistikk og klagesensur (utgjør 24,5% av klagen). Dette er funn som validerer dokumentanalysens forslag om at denne oppgavens åpenhet kan gi usikker vurdering. Oppgave 1 utmerker seg også når det kommer til stryk og klagesensur (18,4%), mens denne oppgaven har høyest korrelasjon (0.69). Sistnevnte gir reliabel konsistens (Barrett, 2001) og reduserer usikkerheten til denne oppgaven. Alle fire oppgaver har like lave koeffisienter for samsvar. Disse er igjen like lave som i bruttoutvalget.

5.3 TEKSTANALYSE AV EKSAMENSBE SVARELSER

Ekstensivutvalget vil først bli presentert overordnet. Deretter presenterer jeg det ut fra skrivehandlinger, svar på oppgaven og vurderingen tekstene har fått. Intensivanalysene er organisert etter vurderingskriteriene for oppgaven, og hver tekst oppsummeres med mulige årsaker til sprikvurderingen. Utvalgene, teoretisk bakgrunn og fremgangsmåter for tekstanalysene er beskrevet nærmere i kapittel 4.3.1 og 4.3.2. Metodiske valg og forskningsetiske problemstillinger ble drøftet i kapittel 4.4.

Oversikten over ekstensivutvalget fremstilles i to tabeller, en for spriktekstene og en for klagetekstene. Klagetekstene er lest for å kunne supplere funn i spriktekstene, og blir ikke analysert intensivt. I oversiktene har jeg av personvern hensyn erstattet eksaminandnumrene med et navn som består av en kombinasjon av bokstav og tall: S eller K for henholdsvis sprik- eller klagetekst og 1, 2, 3 og så videre for hvilket nummer teksten er i utvalget. Videre har jeg valgt å ta med overskriften, fordi det gir et bilde av teksten. For å fylle ut dette bildet, har jeg tatt med tolkninger av tekstens makronivå, herunder skrivehandlinger, sjangertrekk eller andre tekstmønstre, i tillegg til en antydning av hovedinnholdet. Fordi sprikvurderingen er sentral, har jeg valgt å ta med de to sensorenes karakterer samt endelig eksamens karakter.

Spriktekstene

Navn	Overskrift	Skrivehandling	Sjanger/tekstmønster og hovedinnhold	Kar. 1	Kar. 2	Slutt kar.
S1	«Er hvert menneske en øy?»	Reflektere Fortelle	Dagboksekvenser i nåtid/eksamenssituasjon illustrerer essayistiske refleksjoner rundt tematikk	4	6	6
S2	«En del av noe større»	Fortelle Reflektere Argumentere	Jeg-fortelling som rammer inn refleksjon rundt fellesskap. Kan leses som et selvmord.	6	3	5
S3	«Kjærlighet binder enhver øy sammen»	Reflektere	Essayistisk refleksjon rundt tematikken, men bryter med formen avslutningsvis ved å bli analytisk.	4	2	3
S4	«Hjertesmeden»	Fortelle	Novelle/fortelling om en som oppsøker en hjertesmed for å reparere sitt ødelagte hjerte.	6	3	4
S5	«En samtale»	Fortelle	Fiksjonell dialog/polemikk mellom Donne og Vesaas. Refleksjoner gjennom fortellemåten replikker	5	3	4
S6	«Hadde han bare bygd broer»	Fortelle	Fortelling/novelle om to brødre, der lillebroren er egosentrisk, destruktiv og tilslutt dør.	3	5	4
S7	«Halv eller hel øy?»	Reflektere Fortelle	Ekspressiv og fiktiv samtale med sneglevenn på øde øy, om metaforene i vedleggene.	2	5	3
S8	«Mennesket – kun en jækla øy?»	Reflektere Argumentere	Ekspressivt og lukket essay om hvor dumme forfatterne av vedleggene er.	4	1	2
S9	«Et smil mot mobbing»	Reflektere Fortelle Argumentere	Dagbokliknende refleksjon om det å stå alene, men tar etter hvert form som leserbrev mot mobbing	4	2	3

S10	«Mitt eget øylandskap»	Fortelle Bekjenne Reflektere	Blogg/dagbok/jeg-fortelling om hverdag med eneomsorg for to barn.	5	2	3
S11	«Ett folk, en verden»	Informere Analysere Reflektere	Artikkel/analyse med mye kontekstualisering av vedleggene, men også noe essayistisk.	4	2	3
S12	«Tropiske turisthalvøyer»	Reflektere Informere Analysere	Lukket refleksjon med trekk fra både essay og kåseri. Blir så analytisk med umotivert kunnskap om kontekst.	4	2	2
S13	«Et spørsmål om perspektiv»	Reflektere	Essay som prøver ut tanker knyttet til de geografiske metaforene	6	3	5
S14	«Mennesker som øyer, sa du?»	Reflektere	Essay som starter utprøvende, blir litt flåsete underveis, men tar seg opp	6	4	5
S15	«Brobyggeren»	Reflektere	Essayistisk refleksjon rundt tematikk. Løsriver til å handle om nedvurdering av mennesket.	3	5	4
S16	«Jeg bygger bare broer»	Fortelle Reflektere	Rammefortelling (om jeg-personen som skal skrive artikkel om dikt) rundt essay som handler om identitet.	3	5	4
S17	«Å bygge en bro»	Fortelle	Fortelling om to jenters møte, samtale og flørt i et bryllup. Over halvveis koples det til tematikk/vedlegg.	6	4	6
S18	«Musikkens magi»	Fortelle	Fortelling om tre menn som reiser til hest, tar inn på et vertshus og hengir seg til musikken.	4	2	3
S19	«Det grønne havs øyer»	Fortelle	Ekspressiv jeg-fortelling om familieband/kjernefamilie. Refleksjoner gjennom tankereferat.	4	6	5
S20	«Om meg og mennesket»	Reflektere Argumentere	Ekspressivt, bloggete, essayistisk (vandrer i tema, flerstemt). Kaller tekstenes tematikk for påstander, og leter etter logikken i dem.	4	6	6
S21	«Jeg er ild»	Reflektere	«Fiksjonelt brev» fra og til seg selv. Egen tematikk, men litt uklart hva, kanskje hva som er viktig i livet.	1	3	2
S22	«Hvorfor skulle mennesket vært en øy?»	Reflektere	Essayistisk. Kort tekst som bruker ord fra vedleggene og prøver å svare på hvordan mennesket kan være en øy.	1	4	2

Klagetekstene

Navn	Overskrift	Skrivehandling	Sjanger/tekstmønster	Kar. 1	Kar. 2	Slutt kar.
K1	«Et fastland med uendelig mange øyer med broer imellom»	Reflektere Fortelle	Essayistisk	4	4	4
K2	«Er mennesket en øy?»	Fortelle, dikte Argumentere	Kåseri innledningvis, så fortelling rammet inn av drøfting?	2	2	2
K3	«Er vi mennesker bundet sammen?»	Reflektere	Ekspressivt, personlig essay/dagbok?	2	1	1
K4	«Innesluttet i en bro av betong»	Fortelle Reflektere	Personlig fortelling	4	4	4
K5	«Øy, hva skjer med verden?»	Reflektere Informere	Lite helhet, kort essayistisk artikkel	2	1	1
K6	«Har vi alle plass i fellesskapet?»	Fortelle Reflektere	Personlig essay rammet inn av fortelling/beskrivelse av 1.skoledag/ny start/tematikk	3	3	3
K7	«Å være eller ikke være – en øy?»	Reflektere	Essay. Kunnskapsbasert ut fra vedleggenes kontekst	3	3	3

K8	«Selvstendig øy, eller et stykke fastland»	Reflektere Argumentere	Essayistisk. Gjengir vedleggene, noe refleksjon, så litt «egne meninger»	1	2	1
K9	«Veien tilbake»	Fortelle	Metatekster innlednings- og avslutningsvis for å forklare. Personlig fortelling om brudd.	3	2	2
K10	«Kontakt søkes»	Fortelle	Personlig fortelling om vanskelig relasjon	4	3	3

5.3.1 Skrivehandlinger i ekstensivutvalget

Kategoriseringen etter skrivehandlinger, som er noe av det som vil bli drøftet i kapittel 5.2, gir følgende tre hovedkategorier for spriktekstene:

1. Tekster som blander skrivehandlinger: S1, S2, S7, S8, S9, S10, S11, S12, S16
2. Tekster som i hovedsak reflekterer: S3, S13, S14, S15, S20, S21, S22
3. Tekster som i hovedsak forteller: S4, S5, S6, S17, S18, S19

Eksaminandene skulle reflektere i oppgaven, noe som kjennetegnet for høy måloppnåelse utdyper med «prøver ut tanker på en selvstendig måte» (Utdanningsdirektoratet, 2015a, s. 8). Derfor velger jeg å betrakte bruk av andre skrivehandlinger som en risikostrategi. Oppgavens formulering legger ikke direkte opp til det, men likevel er *fortelle* en av eller eneste skrivehandling i halvparten av tekstene i både sprik- og klageutvalg. I tillegg bruker noen skrivehandlingene *beskrive*, *informere* eller *argumentere*. Andre skrivehandlinger er et trekk i 2/3 av sprikutvalget, noe som sannsynliggjør en sammenheng med sprikvurderingen. Det vil si at normfellesskapet kan være lite omforent om hvordan indirekte, eller mangel på, refleksjon skal vurderes. At trekket også finnes i seks av ti klagetekster bidrar til å gi inntrykk av at dette er en risikabel strategi, og at sensorer i et lite omforent normfellesskap i noen tilfeller har truffet en som prioriterer annerledes (spriktekstene), og i andre én de deler syn med (klagetekstene).

5.3.2 Ekstensivutvalget og svar på oppgaven

Vurderingsskjemaet for *kreative tekster* består av 15 vurderingskriterier. Fire av dem tilhører kategorien *Hovedinntrykk*. Analysen organiseres med utgangspunkt i disse.

5.3.2.1 Svar på kortvaroppgaven

Kjennetegnet til høy måloppnåelse er at svaret skal være presist og relevant. Presisjonsnivået i besvarelsene varierer, men det er ingen som ikke skriver relevant, selv om flere delvis unngår analyse av tekstvedlegget. Foreløpig ser jeg derfor ikke sammenheng mellom sprikvurderingen og hvordan kortsvarene er besvart. Samtidig er det en mulighet for at de to sensorenes vektning av kortsvaret har bidratt til sprikvurderingen. Det at sensorene etterlyste avklaring av vektning av kortsvarene (Utdanningsdirektoratet, 2015e), kan også tale for dette.

5.3.2.2 Svar på langvaroppgaven

Kjennetegnet til høy måloppnåelse er at svaret skal være utfyllende og relevant, og begge deler varierer. Relevans handler blant annet om å svare på det oppgaven ber om, noe flere av tekstene ikke gjør. Oppgaven var refleksjon med utgangspunkt i tematikken i to tekstvedlegg, i tillegg til at det kunne vises kreativitet gjennom valg av innfallsvinkler og språk. Det skulle dessuten synes at refleksjonen tok utgangspunkt i de to vedleggene.

Når det gjelder refleksjon, viser analysen at syv av 22 spriktekster inneholder *reflektere* som eneste skrivehandling. Selv om enkelte viser refleksjon gjennom andre skrivehandlinger, antar jeg at både blanding og fortelling kan ha sammenheng med sprikvurderingen. Der oppgaven ber om kreativitet gjennom valg av innfallsvinkler og språk, er det noen av tekstene som ikke svarer relevant. Dette er et utpreget trekk ved fire av 22 spriktekster (S3, S9, S11 og S22). Etersom oppgavetyper heter kreativ tekst, er det sannsynlig at sensorene har tillagt denne mangelen vekt. I klagetekstene kan en også finne tendensen til at det er flere som ikke svarer på alle deler av oppgaven. Seks av ti tekster gjør det, men i ulik grad.

At refleksjonen skal ta synlig utgangspunkt i vedleggene, er tolket ulikt av eksaminandene. Av de som synes ikke å følge instruksjonen, er det to tendenser: De som i hovedsak har valgt annen tematikk, og de som skriver seg over i annen tematikk. På sensorskoleringen ble det sagt at tankegods fra vedleggene skulle være så tydelig i tekstene at en kunne se at de var skrevet eksamensdagen (Utdanningsdirektoratet, 2015c). Et fellestrekk mellom tekstene som i hovedsak er narrative, og da helst fiksjonsfortellingene fremfor jeg-fortellingene, er at de langt på vei ikke innfrir dette kravet. En kan derfor tvile på om disse svarer på oppgaven. Det er likevel vanskelig å avvise at de reflekterer tematikk i vedleggene, også fordi en leserorientert lese måte (Claudi, 2013, s. 115) åpner for subjektive tolkninger. I tillegg er tema som menneskets plass i verden, relasjoner og kommunikasjon av såpass generell karakter at veldig mange fortellinger allegorisk kan leses til å handle om dette. Ett av KAL-funnene var at fortelling som strategi er risikabelt når oppgaven ikke legger opp til det (Hertzberg & Andersen, 2005, s. 289). Imidlertid viser de fortellende tekstene ofte kreativitet, og dette kan i noen fall se ut til å ha blitt mer honorert enn refleksjon, også ut fra at disse spriktekstene har et karaktersnitt på 4,2 versus 3,5 i de reflekterende og 3,6 i blandingstekstene.

5.3.2.3 Forståelse og bruk av vedlagte tekster

Eksaminanden skulle reflektere med utgangspunkt i tekster, så dette er også et kriterium som angår svar på oppgaven. I tillegg angår det spesifikk lesekompetanse eksaminanden har vist. Kjennetegnet til høy måloppnåelse ber om *god forståelse og bruk* av de vedlagte tekstene. Det

er vanskelig å avgjøre i hvilken grad eksaminanden viser lesekompetanse når tekstvedleggene i liten grad er brukt, og det er sannsynlig at dette har sammenheng med sprikvurderingen av flere besvarelser. Dette gjelder spesielt i flere fortellinger (S4, S6, S17, S18), som må tolkes for å avgjøre om de tar utgangspunkt i vedleggene, noe det da viser seg at de i varierende grad gjør. Det gjelder også i de tekstene som benytter innhold fra dagliglivets domene eller skriver seg over i annen tematikk etter å ha startet i vedleggene (S9, S10, S15, S21, S22). Når for eksempel en alenemor skriver om sin slitsomme hverdag, er det vanskelig å avgjøre om dette er avsenderens egen tematikk eller om hun hekter seg på Donne og Vesaas. Sannsynligvis er det en kombinasjon, men det er lettere å vurdere tekster som mer tydelig skriver frem og reflekterer over egen tekstforståelse av de to vedleggene. Det er altså forholdsvis mange spriktekster som kan karakteriseres av at tematikk i tekstvedlegg ikke eller i liten grad er synlig i teksten. Disse får følgelig ikke vist særlig lesekompetanse.

Tendensen til å skrive seg over i annen tematikk enn den en finner i tekstvedleggene sammenfaller ofte med et annet relativt utbredt trekk, nemlig ekspressiv skriving og personlig innhold. Her kan det være snakk om en tydelig jeg-stemme eller jeg-perspektiv, og et personlig innhold som i flere tekster er hentet fra dagliglivets domene. Et annet brudd med den såkalt skolske diskursen kan ses i tekster med referanser til ungdomskulturelle domener, for eksempel blogger, dataspill og tv-serier. Dette står også i motsetning til eksaminander som i stedet har brukt tekstvedleggene til å assosiere og trekke inn terminologi fra andre skolefag eller spesialiserte domener. Når noen av tekstene skal leses intensivt, vil jeg blant annet utforske mulige relasjoner som kan ha oppstått mellom avsender og adressat som følge av både ekspressiv skriving og bruk av innhold eller referanser fra dagliglivets domene.

En del av tekstene behandler tematikken i vedleggene på en måte som kan tolkes som useriøs eller lite faglig. Dette kan sies å være et trekk i S7, S8, S12, S13, S14 og S20. Dette gjelder forholdsvis mange tekster, men det varierer hva det er som kan tolkes som useriøst. Enkelte kan fremstå useriøse fordi tekstkompetansen er lav (S12), enkelte kan se ut til å bomme på humor som litterært virkemiddel ved å flyte over i det flåsete (S7 og S14), mens andre kan ha overdrevet sin mottakerorientering (S13), noe jeg skriver mer om under neste punkt. Det kan imidlertid bare unntaksvis (S8) sies å være et fremstående trekk ved tekstene.

5.3.2.4 Formuleringssevne og ordforråd

I dette kriteriet blir kjennetegn som *gjennomgående godt formulert og hensiktsmessig og nyansert ordforråd* brukt. Selv om sprikvurdering kan ha sammenheng med ulik vektning av formell skrivekompetanse, har jeg i ekstensivutvalget ikke vurdert denne, men snarere sett

etter språklige trekk som sannsynligvis har utløst uenighet rundt språkets hensiktsmessighet. Her har jeg tidligere vært inne på ekspressiv språkfunksjon, som kan avdekkes språklig, eksempelvis gjennom mye bruk av personlig pronomen. Manglende bruk av språklige virkemidler vil også kunne spille inn under dette punktet, noe jeg var inne på over.

Noen av de fortellende tekstene har et språk som kan likne trivillitteraturens, som her i S17: «Lyset og latteren fra festsalen strålte ut gjennom de mektige vinduene og fylte den hvite terrassen. Hotellet var fullbooket denne weekenden». Det kan her være snakk om bruk av mange og litt overdrevne eller polariserte adjektiv (S6), eller litt høystemte eller antikverte vendinger, som her i S4: «Byen er ikke stor, men hver gate er fylt med bragd, eventyrlystne mennesker og butikker som utgjør mirakler». Det siste er blant annet et kjennetegn i fortellingene som ser ut til å være satt i en fortidig kontekst (S4 og S18), og slike språklige valg kan bidra til å kommunisere akkurat det. I KAL-studien fremheves det høystemte som risikotrekk: «Det er få ting en leser reagerer mer forskjellig på enn tekster som gir mistanke om et for høyt ambisjonsnivå (...) eller et språk som oppfattes som svulstig eller pompøst» (Hertzberg & Andersen, 2005, s. 283). I KAL ble også innslag av både trivielt og høystemt språk funnet hos gode skrivere, noe som også gjelder her. Tekstene S4, S6, S17, S18 og S19 kan representere denne tendensen, men det er også til en viss grad funnet i andre tekster.

En annen språklig tendens er å finne i de reflekterende tekstene. Det er relevant å operasjonalisere refleksjon gjennom bruk av spørsmål og metakommunikasjon (Evensen, 2010). Samtidig tas denne måten å kommunisere på såpass langt i noen tekster at de kan leses som forsøk på å gå i dialog med mottaker, noe som tidligere er funnet risikabelt (Hertzberg & Andersen, 2005, s. 299). Tekstene S1, S7, S8, S10, S13, S15, S20 og S21 representerer denne tendensen, men det er et tolkningsspørsmål og vanskelig å avgjøre om dialogen egentlig er med sensor, og dette kommer jeg tilbake til i intensivanalysen. Ovennevnte språklige trekk er funnet i over halvparten av tekstene, men hvert enkelt er ikke funnet i så mange. Det at trekkene også tidligere er funnet i sammenheng med sprikvurdering, gjør at det likevel bør studeres nærmere, og jeg kommer tilbake til dette i intensivanalysen.

5.3.3 Ekstensivutvalget og vurdering

Sprikteknene har stort sett fått gode karakterer. Syv av 22 har nådd høy måloppnåelse, fem har fått karakteren 4, seks har fått 3, mens fire fikk lav måloppnåelse. Dette gir et snitt på 3,8, som er godt over snittet i populasjonen, som i 2015 var 3,4 (Utdanningsdirektoratet).

Karaktersnittet på klagetekstene er til sammenlikning 2,5.

Ni av 22 spriktekster er klare svar på oppgaven. Disse er vurdert til gode karakterer (4, 5 og 6). Av de 13 som ikke har svart på alt (både vært kreative og reflektert om det de er bedt om), har de fleste oppnådd karakteren 2 eller 3, mens to av dem (S4, S6) har fått 4 og én 6 (S17). Med andre ord er det tre besvarelser som mangler svar på hele oppgaven, men som likevel har fått gode vurderinger. Dette kan skyldes usikkerhet, blant annet som følge av sprikvurdering, eller en oppvurdering av andre kvaliteter. Det bør bemerkes at alle disse tre har det felles at de er fiksjonsfortellinger. Vurderinger i sprikutvalget kan tyde på at fortelling premieres høyere enn refleksjon, med et karaktersnitt på 4,2 versus 3,5. Denne tendensen er ikke like tilstede i klageutvalget. Her er de to beste karakterene (4) gitt tekster som blander inn narrative elementer, men de to som kun forteller har fått 2 og 3. Akkurat disse to kjennetegnes også av å være jeg-fortellinger med svært personlig eller privat innhold.

Ingen av klagetekstene er sprikvurdert. Det at de er sendt klagesensur, indikerer at de er vurdert strengere enn dem i sprikutvalget. Fem av ti har fått samme karakter av begge sensorer. På de andre fem skiller det én karakter, og her er sensorene blitt enige om laveste karakterforslag, noe som også førte til at tre endte med strykkarakter. Mens det i spriktekstene er en karakttermessig forskjell mellom tekstene som har svart på hele oppgaven og de andre, ved at førstnevnte har fått gode karakterer og sistnevnte med noen unntak ikke så gode, har bortimot alle klagetekstene fått mindre gode karakterer.

De 22 spriktekstene er sensurert av til sammen ni sensorpar. Ett par har sensurert hele ni av disse. De to sensorene bytter på å være den strengeste, og det samme paret er også representert med to av ti klagetekster. De øvrige åtte sensorparene deler resten av tekstene i utvalget mellom seg. Dette er i og for seg ikke gjenstand for analyse, men det kan være interessant å registrere at enkelte sensorpar kan ha mer krevende samarbeid enn andre. Akkurat dette kan også ses av sensorgruppen som ble brukt som eksempel på konsensus i 5.2.1.

5.3.4 Ekstensivutvalget oppsummer

Oppgavekommentarens «Denne oppgaven åpner for ulike løsninger» ser ut til å være tatt på alvor, for utvalget av sprikvurderte tekster er mangfoldig. Som leser må en være forberedt på både sakpregete og litterære tekster med respektive kontraktspremisses (Tønnesson, 2012). Langsvarene anvender en rekke skrivehandlinger, sjangertrekk, fortellemåter og perspektiv. Det er likevel mulig å identifisere trekk som kan ha sammenheng med sprikvurderingen, og som skal utforskes nærmere i intensivutvalget (hvilke intensivtekster i parentes):

- Blanding av skrivehandlinger eller bruk av andre skrivehandlinger enn *reflektere* (S1, S10, S6, S19)
- Litterær tekst som innfallsvinkel, herunder tilsynelatende fiksjonsfortellinger (S6) og mer virkelighetsnære jeg-fortellinger (S19)
- Ikke eller lite bruk av kreative vinklinger, perspektiver eller språkbruk (S3)
- Tematikk i tekstvedlegg er ikke eller lite synlig i teksten (S6), eller teksten preges i hovedsak av sin egen tematikk og da gjerne hentet fra dagliglivets domene (S6, S10)
- Språklige trekk, som trivielle eller høystemte ordvalg (S6, S10, S13, 19) og metakommunikasjon (S1, S10, S13)

5.3.5 Intensivutvalget: Situasjonsteksten

Vurdering krever analyse (Aasen et al., 2011, s. 161), og analysemodellen som blir brukt på intensivutvalget ble beskrevet i kapittel 3.3.2 og drøftes i 6.2. Kategoriene er, som tidligere skrevet, hentet fra overordnet formål med eksamen og vurderingsskjemaet for oppgavetypen (Utdanningsdirektoratet, 2015a). Fordi oppgave 2 med kommentar er gjengitt før, gjentar jeg her kun selve oppgaveformuleringen. Fordi vedlagte tekster er en sentral del av både oppgaven og situasjonsteksten, er de gjengitt og kommentert under:

Oppgaveformulering: Ta utgangspunkt i de to tekstvedleggene, og skriv en kreativ tekst der du reflekterer over tematikken i tekstene. Lag overskrift selv.

Kvart menneske er ei øy

Kvart menneske er ei øy, som kjent. Så det må bruer til. Uendeleg mange slags bruer. Naturleg nok har vi lært den bygningskunsten grundig.

Det finst finare bruer enn ord over til andre, men berre i den rette, sjeldne stund. Den lydause stund når alt står stille for ikkje å gå i stykke, medan den minste rørsle er ei gåve på same tid – då er ord ubrukelege. Eller når ein fingertupp langs ein arm er som springande bruer av gull.

Likevel: tusen gonger oftare er ordet bru.

Eit under av ei bru.

Tarjei Vesaas 1964

Intet menneske er en øy

Intet menneske er en øy, hel og ubeskåren i seg selv. Hvert menneske er et stykke av Fastlandet, en del av det Hele. Om en Jordklump skylles bort av Havet, blir Europa mindre, på samme vis som en Landtunge blir det, eller som et Jordegods, tilhørende Dine Venner eller Deg selv må bli det. Hvert menneskes Død forminsker Meg, ti jeg er innesluttet i Menneskeheten. Gå derfor aldri ut for å spørre: Hvem ringer Klokkene for? Klokkene ringer for deg.

John Donne 1624

Oppgaven ber om spor av andres ytringer, noe som gjør den dialogisk og intertekstuell, for å bruke begreper fra kapittel 2.1.1. Den ønsker et svar med utgangspunkt i to refererte tekster eller stemmer, og eleven må gå i dialog med disse. Dette tydeliggjør ytringskjeden som eksamensteksten inngår i. Begge vedleggstekstene bruker øymetaforen for å beskrive mennesket. De respektive titlene, «Kvart menneske er ei øy» og «Intet menneske er en øy», kan ideasjonelt gi inntrykk av to polariserte menneskesyn. Vesaas fremhever mennesket som enestående og derfor avhengig av relasjoner til eller kommunikasjon med andre, uttrykt gjennom brumetaforen. Donne fremhever mennesket som del av helheten av andre mennesker, uttrykt gjennom fastlandsmetaforen. Vesaas skriver at mennesket behersker dette med bruer, at ord som oftest tjener formålet, men i andre avsnitt er det kommunikasjonen gjennom nærheten mellom to mennesker som oppvurderes. Selv om tittelen antyder en slags isolasjon av mennesket, preges teksten av det viktige forholdet til andre. Fellesskapet mellom mennesker er et likhetstrekk med Donnes tekst. Donne fremhever helheten mennesket er en del av, og sammenlikner dette med forholdet mellom Europas fastland og det av det som havet tar. Dette konkretiseres videre i at mennesket forminskes av andres død, og at alles menneskers død på den måten angår alle mennesker. Mens Vesaas er opptatt av de nære relasjonene, er Donne også opptatt av menneskeheten.

Når eksaminandene i sin oppgave skal ta utgangspunkt i en tolkning av hva som er vedleggenes tematikk, er kanskje det mest åpenbare mennesker og forholdet mellom dem, uttrykt gjennom metaforene øy, bru og fastland. Samtidig bør en også kunne forvente tekster som tar utgangspunkt i tematikk som kommunikasjon, kjærlighet, død og solidaritet. På sensorskoleringen anførte Utdanningsdirektoratet denne tematikken: «Ensomhet/fellesskap, kontakt/isolasjon, å nå fram til hverandre, individ/kollektiv osv. Hos John Donne handler det om at menneskene i den katolske kirka er ett, og det som angår en, angår alle» (Utdanningsdirektoratet, 2015d, s. 23).

Alle intensivtekstene er gjengitt i sin helhet i vedlegg 3.

5.3.6 Analyser av tekster i intensivutvalget som blander teksthandlinger

Blanding av teksthandlinger forekommer i ni av 22 spriktekster. Det går an å dele disse inn i to underkategorier; på den ene siden tekstene som bærer preg av gjennomtenkt struktur og bevisst forhold til skrivehandlinger og sjangere (S1, S2, S7, S16), og på den andre de som «bare skriver» (S8, S9, S10, S11, S12). De to valgte tekstene representerer henholdsvis det bevisste formbruddet som kreativ innfallsvinkel (S1) versus manglende bevissthet rundt dette (S10). De representerer også et annet trekk som er godt representert i ekstensivutvalget, og

som ble kritisert i sensorrapporten, nemlig innfallsvinkelen personlig skriving. Begge tekstene har også metakommunikative trekk som kan ses i sammenheng med sprikvurderingen.

S1 «Er hvert menneske en øy?»: Sjangerbruddet som fungerer

Langsvaret kan leses som et essay der det reflekteres over menneskers plass i fellesskapet.

Disse refleksjonene rammes inn og brytes av underveis av jeg-fortellende og dagbokliknende sekvenser som følger et forløp fra jegets siste eksamensforberedelser, via nattevåk før eksamen, til eksamensdagen. Sensor 1 innstilte besvarelsen til karakteren 4, men ble på fellessensuren enig med forslaget til andresensor; karakteren 6.

Svar på kortsvaroppgaven:

Struktur og sammenheng: Kortsvaret går med sine 253 ord så vidt utover oppgavens ramme på 250 ord. Det første av fire avsnitt presenterer teksten, overordnet tematikk og hva kortsvaret skal handle om. De neste to avsnittene handler primært om form, synsvinkel og andre typiske trekk ved sagastilen, og det siste avsnittet går tilbake til innholdet. Dette gir hensiktsmessig struktur og god sammenheng.

Faglig kunnskap, fagspråk, redegjørelse for karakteristiske trekk: Det gjøres godt og presist rede for sentrale og karakteristiske sagatrekk, som her i tredje avsnitt: «Dette trekket finner vi blant annet i avsnitt tre og fire i teksten, hvor avsnitt tre er en scene med replikker og dramatik som beskriver Gisles siste øyeblikk i live. Det hele er preget av en underdrivelse, også typisk for sagastilen». Eksaminanden henviser til hvilke avsnitt i teksten det kommenteres eller siteres fra. Sitater er markert og kildene er ryddig og nøyaktig oppgitt. Andre fagbegreper, som aural refererende synsvinkel, scener og referater er også brukt presist og formålstjenlig. Det er også tematiske referanser til ættesamfunn, ære og blodhevn.

Relevant svar på oppgaven, forståelse og bruk av vedlagt tekst: Kortsvaret er relevant. Det er også vist høy lesekompetanse, med ett unntak: Eksaminanden feilleser hvem som går ut av *Gislesoga* med æren i behold: «Vi ser også i siste avsnitt en beskrivelse av den hederen og æren Øyolv Grå får i ettertid av drapet».

Svar på langsvareoppgaven:

Struktur og sammenheng: Teksten er todelt og blander skrivehandlingene fortelle og reflektere: Det kan leses som et bevisst formbrudd og en mulig årsak til sprikvurderingen. Todelingen kan også leses som hensiktsmessig fordi den rammes inn en tradisjonell refleksjon med mer ekspressive eller metakommunikative beskrivelser, noe som også realiserer mening

om hvordan teksten plasserer seg i situasjonskonteksten. Tekstuelte er todelingen markert typografisk gjennom skiftende bruk av kursiv og normal stil. Den ene delen, som består av fem avsnitt i kursiv, forteller om et forløp forbundet med situasjonskonteksten, nemlig eksamensperioden og spesielt det siste halve døgnet vei mot eksamen. Disse fortellende avsnittene har et jeg-perspektiv, blant annet realisert gjennom pronomenbruk, for eksempel syv «jeg» i første avsnitt. De personlige avsnittene rammer inn resten av teksten, tilsammen ni avsnitt med refleksjoner som tar direkte utgangspunkt i tekstvedleggene. Denne delen står følgelig i et mer direkte forhold til situasjonskontekstens oppgaveinstruks. Jeget er også tilstede i denne delen, men her brukes de mer generelle og inviterende flertallsformene i større grad, i tillegg til mer upersonlige tredjepersonformer, som «en», «man» og «seg» og determinativer som «mange» og «noen», en leksikalsk kjede som kopler det personlige med det allmenne. Dette gir en dialogisk modellsender (Grüters, 2011, s. 169).

Relevante momenter og/eller eksempler, med god bredde: I hovedsak kan relevante momenter sies å være innfridd gjennom tekstnær og bred refleksjon, noe jeg kommer inn på i refleksjonskriteriet. Nevnte todeling som innledningsvis fremhever og tematiserer jegets opplevelse av eksamen, er kanskje ikke vurdert som relevant av sensor 1. I tråd med funn i KAL kan det også tenkes at dette er lest som uønsket metakommunikasjon (Hertzberg & Andersen, 2005, s. 298). Analysen av ekstensivutvalget avdekket mulig sammenheng mellom ekspressiv skriving og sprikvurdering, på tross av at en forventer selvbetraktning og - fremstilling i reflekterende tekster (Grüters, 2011, s. 162). Derfor kan det være interessant å se på hvordan det ekspressive her korresponderer med feltet i situasjonskonteksten. Ideasjonell mening finnes blant annet i beskrivelsene av feltet som i hovedsak er den stressende eksamenstiden og ensomheten som nedprioriteringen av det sosiale har ført til, vist gjennom leksikalsk kohesjon mellom innholdsord som «eksamensperiode», «lite sosialt samvær», «mye lesing», «hardt arbeid», «frustrasjon», «stress», «nerver», «kontrollen» og «prøven». Mottaker inviteres også inn i avsenders soverom gjennom prosessen «prøver å sove», noe som fort kan bli privat og av noen lesere avvises som manglende relevans. Dette er dagliglivets domene (Tønnessen & Maagerø, 2015, s. 26), men samtidig som subjektet «føler stresset og nervene tar godt tak i alt jeg er», og konstaterer at det er ingen å snakke med, generaliseres erfaringen i «Vi forsøker alle å være selvstendige, men vi føler oss nok alle litt isolerte. Litt som en øy». Dette gir et subjekt som evner å se utover seg og sitt, slik Aase foreskriver som avgjørende for at selvrefleksjonen skal unngå «ekstrem subjektivitet» (Aase, 2012, s. 52). Samtidig nyanseres egne erfaringer, blant annet gjennom bruk av «nok», slik at det uttrykkes

usikkerhet til om dette er en allmenn erfaring. Både nyansering og generalisering av erfaringer kan forbindes med det spesialiserte domenet (Tønnessen & Maagerø, 2015, s. 27). Dette kan åpne for at adressaten godtar virkelighetsbeskrivelsen, fremfor mer polariserte eller entydig jeg-orienterte beskrivelser. Ikke minst koples ytringen gjennom øymetaforen til vedleggene. Slik gjøres det subjektive relevant, i tillegg til at avsender viser åpenhet og forhandlingsvilje.

Veien fra den personlige fortellingen om hvordan det er for jeget å grue seg til eksamen, til det gjøres til et uttrykk for andres erfaringer, finnes også i de andre jeg-orienterte kursivavsnittene. I fjerde avsnitt er det individets nattevåkne ensomhetsfølelse som speiles i streng prat til seg selv om at dette også gjelder andre («*Alle andre skal gjennom det samme som du skal*», *sier jeg til meg selv*). I sjette avsnitt lindres stresset gjennom bekreftende sms-kontakt med en venninne, i åttende «innser» avsender i møte med flere venner på bussen til eksamen at de har det på samme måte, før det oppsummeres i en samlet erkjennelse i aller siste avsnitt, at det er nettopp fellesskapet som vaksinerer mot isolasjonen man kan føle på. Så mens setninger som «Jeg ligger her og tenker på hvor stort behov jeg har for å snakke med noen» (avsnitt fire), eller «Men jeg våkner, urolig og stresset. (...) Jeg føler fortsatt behov for å snakke med noen» (avsnitt seks), er ekspressive og av den grunn kanskje risikable, går likevel jeg-avsnittenes funksjon forbi selvframstillingens, fordi de belyser fenomen eller spenningsforhold mellom det individuelle og kollektive (Aase, 2012). Dette dialogiske, i følge med ovennevnte åpenhet, kan ha bidratt til å dempe det risikable førsteinntrykket og derfor ha sammenheng med at de to sensorene ble enige om beste karakter.

Kreativitet i valg av innfallsvinkler, språklige og litterære virkemidler: Valget av eksamenskonteksten for å belyse tematikken er en kreativ innfallsvinkel. Språklig vises det også kreativitet, for eksempel gjennom omskrivingen av klisjeen «nerver av stål» i første avsnitt. Stresset og nervene som tar tak i samme avsnitt viser billedbruk via personifisering. I samme avsnitt brukes også kontrast (ingen og alle), paradoks (selvstendige og isolert) og sammenlikning (som en øy). At regnet «pisker mot ruta» er et eksempel på mer opplagt og mindre kreativ billedbruk, men ikke representativ for teksten som helhet. Nevnte strukturelle todeling kan kalles et litterært virkemiddel eller en kreativ innfallsvinkel. Vurderingskriteriet sier at dette i beste fall (høy måloppnåelse) kan være utført på en «gjennomført måte». Jeg har over vist hvordan de fem jeg-avsnittene både skiller seg typografisk og grammatisk fra de andre, hvilke metafunksjonelle konsekvenser dette kan ha og hvordan det er gjennomført. Også her er det denne innfallsvinkelen som trolig har sammenheng men sprikvurderingen.

Reflekterer, prøver ut tanker: Eksaminanden viser refleksjon blant annet gjennom verbbruk. I første avsnitt viser «føler» og «trodde» til avsenders indre, og det går en kjede av det Halliday kaller mentale prosesser (Halliday, 2014, s. 215) gjennom teksten, via eksemplene «tror» og «ser» i andre og tredje avsnitt, «vite» i femte, «liker» i tiende, til «følte», «vet» og «tror» i siste. Denne tekstuelle kohesjonskjeden skaper sammenheng, i tillegg til at den representerer den indre erfaringen, eller refleksjonen som speiler ytre hendelser. Slik bidrar tekstuell sammenheng til forståelse av det ideasjonelle. Dette vurderingskriteriet antyder som tidligere nevnt refleksjon innenfor essay-formatet. For den høye måloppnåelsen skal refleksjonen være «på en selvstendig måte», og mye av det jeg allerede har identifisert av tanker rundt individets plass i det kollektive kan sies å være det. I tillegg åpnes det for adressatens reaksjoner og mellompersonlig mening gjennom flere hv-spørsmål, men også et par polariserte spørsmål, for eksempel i tittelen. Spørsmålene speiler avsenders refleksjon, men er også dialogiske fordi de foregriper svaret (Bachtin, 2005, s. 40). Modalitet har også «sterk mellompersonlig kraft fordi vi her signaliserer holdningen vår til det vi uttrykker i setningen. Vi gjør en form for evaluering eller vurdering av utsagnet, og så kan lytteren (eller leseren) ta stilling til denne» (Maagerø, 2005, s. 150). Slike nyanser i kommunikasjonen ses eksempelvis når de to tekstvedleggene i andre avsnitt introduseres ved at avsender foreslår tematikken i dem. Dette er et avgjørende punkt i teksten, fordi tolkningen av tematikken legger premiss for senere refleksjon, og avsender kan ha behov for å forsikre seg om at adressaten ikke blir avvist av denne tolkningen: Etter å ha skrevet at «Tekstene har motsigende titler, men jeg tror hovedpoenget er det samme», sier avsender: «Hovedpoenget ser for meg ut til å være at vi kan forsøke å isolere oss, vi kan føle oss så alene som bare det, men likevel er vi alle en del av noe». Her er behovet for å gardere seg realisert gjennom modulisasjonen *sannsynlighet* (Maagerø, 2005) i flere verbalgrupper, noe som åpner for andre tolkninger, også mottakers. Det er usikkert om dette kriteriet kan forklare sprikvurdering.

Relevant svar på oppgaven, forståelse og bruk av de vedlagte tekstene: Fordi tidligere beskrevne tematiske todeling av langsvaret er så synlig, og fordi fortellingssekvensene kun indirekte berører vedleggstematikk, er det trolig her dissensen mellom sensorene lå. Avsnittene som tar direkte utgangspunkt i vedleggene følger en rød tråd, der eksaminanden først (avsnitt to) etablerer hovedpoenget i tekstene, før de leses mer inngående og reflekteres over hver for seg (avsnitt tre og fem). Når dette er gjort, tas refleksjonen noe videre (avsnitt syv). Her reflekteres det videre rundt fellesskapet ved å introdusere vannet som metafor for det, noe som også viser selvstendig billedbruk. I avsnitt ni til tretten reflekteres det nærmere

rundt ulike sider ved tematikken. På denne måten blir tekstene brukt gjennomgående, og det blir vist god forståelse av innholdet i dem. Nevnte avsnitt rammes inn og avbrytes av tidligere nevnte jeg-avsnitt som forteller om denne tematikken i jegets liv, og på den måten bidrar til enda et perspektiv, noe som også viser selvstendighet og kreativitet. I jeg-sekvensene er det situasjonskonteksten som tematiseres, både gjennom fortellingen fra eksamensforberedelser og ved å gå i relasjon med adressaten ved å stille spørsmål underveis. Relasjonen mellom teksten og situasjonskonteksten kan også ses gjennom bruk av modalitet, i tillegg til ovennevnte dialog mellom subjektet og samfunnsfenomenet det erfarer.

Formuleringsevne, ordforråd og formelle ferdigheter:

En tekst som behersker formverk, ortografi og tegnsetting vil også sikre god kommunikasjon med adressaten. I kortsvaret er det et par unøyaktigheter i syntaks og tegnsetting, noe som kan ha bidratt til å dempe førsteinntrykket, men hovedinntrykket er at dette mestres. Det samme gjelder sammenbinding mellom setninger og avsnitt. Avsender bruker ressurser som kausale koplinger når det i langsvaret skrives eksplisitt om vedleggteksene, for eksempel «så» og «for» i avsnitt fem, som uttrykker forbindelser mellom trekk i tekstene som omtales og mening avsender tillegger det. I samme avsnitt brukes også additive og implisitte koplinger for å uttrykke dette forholdet mellom lest tekst og egne tolkninger. På denne måten fremstår setningskoplingen variert og språklig nyanserende, og den tjener formålet som er å tydeliggjøre hvilke tanker som kommer fra vedleggene og hva som er egne refleksjoner. Med noen unntak fremstår avsnittsinndelingen gjennomtenkt og hensiktsmessig, selv om avsnitt ni til tretten har en noe løsere struktur. Med unntak av et mulig dempende førsteinntrykk i deler av kortsvaret, er det lite sannsynlig at skrivekompetansen har forårsaket sprikvurderingen.

Oppsummering av mulige årsaker til sprikvurdering (4 og 6) av S1:

- Ulik vektning av en feil i tolkningen av teksten i et ellers faglig solid kort svar
- Todeling og formbrudd i langsvaret, også typografisk
- Personlig innhold og ekspressiv og metakommunikativ stil i delen som jeg-forteller

S10 «Mitt eget øylandskap»: Mammabloggeren

Jeget i langsvaret både forteller om og reflekterer rundt sin hverdag som eneforsørger for to barn. Teksten trekker på både jeg-fortellingen og dagboksjangeren, en hybrid en i dag gjerne ser i blogginnlegg. Besvarelsen ble innstilt til henholdsvis 5 og 2 av de to sensorene, og de ble enige om karakteren 3.

Svar på kortsvaroppgaven:

Struktur og sammenheng: Svaret er innenfor rammen av 250 ord og fordelt på fire avsnitt. Første avsnitt innleder og plasserer utdraget, andre og tredje avsnitt beskriver typisk innhold og form, mens fjerde avsnitt avslutter med en gjentakelse av at innhold og form er typisk for sagastilen. Dette gir en overordnet struktur, men på spørsmålet om strukturen er hensiktsmessig (middels og høy måloppnåelse), er svaret todelt. Avsnittet om typisk innhold handler mest om hevndrapet i sagalitteratur generelt, mens oppgaven ba om utgangspunkt i vedlegget. I tillegg blir det en tendens til gjentakelse når det blir mer om hevndrapet i andre avsnitt. På denne måten blir det vist få trekk med mange ord.

Faglig kunnskap, fagspråk, redegjørelse for karakteristiske trekk: I tredje avsnitt er det bruk av en del fagbegreper. Imidlertid knyttes det ikke alltid til vedlegget, for eksempel *kronologisk*. Samtidig vises formålstjenlig fagspråk når «kampsener med et detaljert hendelse forløp» nevnes i forbindelse med et relevant teksteksempel. Å kalle oppramsing av navn for personskildringer og starten på et utdrag for «in medias res» er galt. Det er et pre at kvadet trekkes frem, men når det påstås at det er uten fast rytme og rim, er det ikke mer enn «noe faglig kunnskap» (lav måloppnåelse) som kommer frem.

Relevant svar på oppgaven, forståelse og bruk av vedlagte tekst: Dvelingen ved fagstoff om sagalitteraturen generelt fremfor teksten som skulle analyseres, kan vurderes som «i noen grad relevant». Det at eksaminanden tar i bruk utdragets «hemnedrap» og ikke egen målform, viser usikkerhet rundt innholdet som beskrives. Språklig uklarhet kommer også til uttrykk der eksaminanden tolker «frender» til å handle om viktigheten av slektskap, noe som også angår lesekompetansen. Kortsvaret kan potensielt gi uenighet fordi det varierer mellom lav og middels måloppnåelse. Etersom sensorenes uenighet var mellom lav og høy, og sensor 1 trolig ikke har vektlagt det, finnes det nok flere forklaringer til sprikvurderingen i langsvaret.

Svar på langsvareoppgaven:

Struktur og sammenheng: De fire første avsnittene reflekterer rundt fellesskap med utgangspunkt i egne erfaringer, mens avsnitt fem til ni er en jeg-fortelling som løper fra middags- til leggetid, før siste avsnitt reflekterer noe mer. I de fortellende avsnittene kan en se en tendens til at situasjonskontekst og vedleggstematikk blir perifer, fordi avsenders egen historie blir hovedtema. Den røde tråden er stort sett der, i tillegg til at siste avsnitt sikrer avslutning og helhet, men det er sannsynlig at noe av sprikvurderingen skyldes diskusjon rundt strukturens hensiktsmessighet. Blandingen av skrivehandlinger kommer uten forvarsel

eller overgang, og det kunne kanskje vært avhjulpet med høyere bevissthet rundt sjanger og skrivehandlinger. I eksplisitt sjangerundervisning brukes blant annet modelltekster for å bidra til å viske ut literacyforskjeller (Hertzberg, 2001), noe eksaminanden kunne profitert på.

Bredde og relevans i momenter/eksempler: I teorikapitlet beskrev jeg blant annet literacy gjennom skillet mellom dagliglivets og det spesialiserte eller refleksjonsdomenet (Tønnessen & Maagerø, 2015). De som ikke behersker skolens sekundærdiskurs legger for eksempel mest vekt på konstruksjon av narrativ meningsstruktur (Penne, 2010, s. 30). Temaet i «Mitt eget øylandskap» er virkelighetsnært og stort sett narrativt konstruert, men samtidig er dette en avsender som langt på vei behersker sekundærdiskursen, noe en blant annet ser gjennom den krigsmetaforiske kohesjonskjeden i femte avsnitt og fordi hun klarer å kople bildene fra Vesaas' og Donnes tekster til fremstillingen av egne erfaringer. Det siste gjøres flere steder, for eksempel når fellesskapet til egne barn beskrives i fjerde avsnitt:

Som ei utvaska alenemor er det nok av konflikter og utfordringer i det lille samfunnet hjemmet som krever broer av fine ord. Broer av gull. For det er nemlig slik at fine ord om hverandre og til hverandre lager broer. Den lille fingertuppen til min sønn, langs min arm, lager en bro av gull. Smilet til min datter bekrefter et under av en bru. Klisjé? Det får så være.

I sitatet er det også mellompersonlig meningspotensiale i det metakommunikative og polariserte spørsmålet avslutningsvis, som foregriper mottakers tolkning av forutgående setning. Det metakommunikative ble kategorisert som et risikotrekk i KAL-utvalget (Hertzberg & Andersen, 2005, s. 298), og i denne teksten er det enda tydeligere innledningsvis i tredje avsnitt: «Dere tenker sikkert at min historie ikke er unik. Det er flere med et anspent forhold til foreldrene sine (...), men det er alt annet enn en tårevåt velkomst jeg vi ha». Dette er en mottakerhenvendelse som både gjør teksten dialogisk og bidrar til at jeget generaliserer erfaringene sine. Imidlertid er det ikke sikkert at det er sensorene i situasjonskonteksten det henvender seg til. Slike henvendelser er vanlig i en del blogger, og fungerer da gjerne som speiling (Grüters, 2011, s. 166). En av de mest populære norske bloggene, *Mamma til Michelle*, hadde 45.000 faste lesere 1. kvartal 2015 (Lunde, 2015). Et tilfeldig oppslag på bloggen 2. desember 2015 var til et innlegg om forholdet mellom mamma Rasmussens egne barndomserfaringer og ønskene hun har til egne barns oppvekst (Rasmussen, 2015). Refleksjonene er til forveksling like dem en finner her. Tilbake til henvendelsesform: Det er også en rekke du-henvendelser i siste avsnitt, nå mer appellativt: «Velg ditt eget fellesskap, din egen vei. Still deg inn på...». Fordi teksten tar utgangspunkt i hverdagsfeltets tematikk, kan det være nærliggende å tolke leserhenvendelsene til å relatere seg mer til selvet enn situasjonskonteksten (Grüters, 2011). Dette gir særlig avslutningvis et

monologisk og selvhjelpende preg, noe som kan ha utløst uenighet rundt hensiktsmessighet. Blant annet sensorrapporten forteller om forventning i situasjonskonteksten om det spesialiserte eller refleksjonsdomenet, noe som i mindre grad innfris når jeget tematiseres. Dette er innholdsvalg som også kan gi uenighet om relevans.

Kreativitet i innfallsvinkler, språklige og litterære virkemidler: Nevnte krigsmetaforer og selvstendig bruk av vedleggsmetaforene er eksempler på språklig kreativitet. Mens forrige analysetekst åpnet for det ekspressive og personlige, domineres «Mitt eget øylandskap» av det samme. Det er en innfallsvinkel som kan likne flere samtidstekster, kanskje mest mammabloggen (Rasmussen, 2015), mens det ideasjonelle også slekter på mer litterære jegfortellinger, eksempelvis noen av Vigdis Hjorth sine (Hjorth, 2001). Det gjør innfallsvinkelen kreativ, men kan også forklare sprikvurderingen, fordi det er delte meninger om hvor personlig det er hensiktsmessig å være og om tekstmodeller fra populærkulturen er tjenlige.

Reflekterer, prøver ut tanker: Halvparten av teksten er tydelig reflekterende, men det varierer i hvilken grad utgangspunkt tas i vedleggene. Avsender skriver med troverdighet om barndom og egen hverdag, og fremstår som noen år eldre enn nittenåringen, som ikke har erfaringer som har ført til brudd i skolegangen. Avsender reflekterer også rundt dette og erkjennelse av manglende dannelse i andre avsnitt: «Samfunnsengasjementet kom ikke før jeg var voksen nok til å forstå at samfunnet forventer noe av meg» (...) «Og mens andre bygde gode relasjoner og broer allerede i barnehagen, var jeg en sosial outsider, en øde øy, som brukte lang tid på å finne min plass på fastlandet». Flere av disse selvrefleksjonene gir sår innsikt i vanskelige relasjoner, noe sensor 1 sannsynligvis har vurdert som gjort på «selvstendig» måte, mens andresensors forslag tyder på at denne mener dette kun er «i noen grad».

Relevant svar på oppgaven og forståelse og måten vedlagte tekster er brukt på: Langsvarets relevans er diskutert over, og tvil rundt dette gir følgelig tvil rundt lesekompetansen. Sensor 2 kan ha vektlagt at eksaminanden begrenser bruken av vedleggene til å skrive om fellesskap, og kan ha tillagt referansene til billedbruk i vedleggene mindre vekt. Dette kan henge sammen med at avsender ikke reflekterer rundt eksplisitt innhold i vedleggene, men refererer relativt hyppig til dem ved å kople dem til fortellingen, som vist over. Innledningsvis reflekterer jeget over en vanskelig oppvekst der hun ikke fikk den sosialiseringen hun senere har sett er viktig. Ansvaret legges på en alkoholisert mor og av den grunn en tidlig: «svikt i bro konstruksjonen mellom mor og datter». Etter at avsender «våget å være alt annet enn min egen mor», ble hun «En mamma på egne ben og et brennende ønske om å være en del av noe mer», og deretter overtar fortellingen om hverdagen med to barn, en av dem med ADHD-diagnose. Også denne

avsenderen evner å generalisere egne erfaringer til noe som kan gjelde andre, for eksempel i tredje avsnitt: «Det er flere med et anspent forhold til foreldrene sine». Når det gjøres, er det blant annet gjennom skifte av pronomenform, fra første til andre eller tredje person, eller fra «jeg» til «vi». Det analysen av forrige tekst fant å være noe av grunnen til at det subjektive ble akseptert av adressaten, er altså også gjort her. Likevel vaksinerer det kanskje ikke like godt, og andresensors avslag kan skyldes at veksling mellom skrivehandlinger og innfallsvinkler ikke fremstår like tydelig. Dette kan igjen skyldes språklige eller tekstuelle begrensninger og det ideasjonelle utgangspunktet i jegets hverdagsliv fremfor vedleggene. At det personlige går over i følsom tematikk kan også være en faktor. I KAL-studien valgte forfatterne å droppe «personlig utleverende tekster» som risikotrekk. Det var i høyeste grad tilstede i tekstene, men vanskelig å avgjøre om det var fiksjon eller ikke (Hertzberg & Andersen, 2005, s. 276). Dette er en faktor her også, men det forhindrer ikke at alkoholiserede foreldre, bitterhet etter skilsmisse og barn med diagnose er følsomme tema som adressater kan reagere ulikt på, en reaksjon som kan forsterkes når jeg-formen gjør det personlig. Sensor 1 har nok akseptert dette innholdet som en «kreativ og selvstendig måte» å bruke vedlagte tekster på, mens andresensors forslag uttrykker at dette er gjort på en «enkel måte».

Formuleringsevne, ordforråd og formelle ferdigheter:

Ovennevnte eksempler kan illustrere et langsvar som i hovedsak kommuniserer godt. Likevel kan mangler i presisjon i ordvalg gi tendens til triviell stil, som i avsnitt ni: «vel inne i sengen min», «jeg har ikke hjertet til å flytte dem» og «før drømmeland henter meg». Dette er også eksempler på enkelt og dagligdags språk, som forbindes med dagliglivets domene (Tønnessen & Maagerø, 2015). Det er også en del formelle feil, blant annet feil gjengivelse av personnavn i kortsvaret og særskrivingsfeil i begge svar, som «hendelses forløpet» i kortsvarets tredje og fjerde avsnitt. Det er derfor rimelig å vurdere at dette er «stort sett», som kjennetegnet til middels måloppnåelse sier. Samtidig har sensorene foreslått henholdsvis høy og lav måloppnåelse, og i tråd med lesningen over er det mest sannsynlig at det som språklig er mest kontroversielt, er det ekspressive og hverdagslige.

Oppsummering av mulige årsaker til sprikvurdering (5 og 2) av S10:

- Ideasjonelt innhold er hentet fra dagliglivets domene, og utover i langsvaret prioriteres dette fremfor tekstvedleggenes tematikk
- Noe av innholdet kan fremstå følsomt og personlig utleverende (bittert skilsmisseoppgjør, alkoholisert mamma, diagnostisert barn)

- Bruk av en populærkulturell form som blander skrivehandlinger og bruker en ekspressiv og tidvis triviell stil
- Metakommunikasjon som kan tolkes som utidige leserhenvendelser, men også som myntet på andre lesere og slik trekke teksten ytterligere bort fra situasjonskonteksten

5.3.7 Analyser av tekster i intensivutvalget som i hovedsak forteller

Det er seks fortellende tekster i sprikutvalget. Fire av dem (S4, S6, S17, S18) har det felles at de må tolkes før en kan vurdere om de reflekterer med utgangspunkt i tekstvedleggenes tematikk. S5 svarer tydelig på denne delen av oppgaven ved at eksaminanden, gjennom fortellemåten replikker, har konstruert teksten som en samtale mellom Donne og Vesaas. S19 er en personlig fortelling som realiserer refleksjon gjennom fortellemåten tankereferat. En av tekstene er plukket ut for å vise tendensen særlig fiksjonsfortellingene har med å utfordre oppgavekontraktens refleksjonsinstruks, samtidig som sluttvurderingen blir god (S6), mens den andre (S19) gjennom valg av fortellemåte tydeligere følger kontrakten. Tekstene er også valgt for å belyse språklige eller stilistiske trekk som kan utløse uenighet.

S6 «Hadde han bare bygd broen»: Fiksjonen

Dette er en novelle om en tilsynelatende vellykket, men usympatisk lillebror, som tilslutt dør av egen hybris. Fortelleren er storebroren, som også beskriver sin bror gjennom retrospeksjon til oppveksten. De to sensorene foreslo karakterene 3 og 5, før de ble enige om 4.

Svar på kortsvaroppgaven:

Struktur og sammenheng: Kortsvarets 294 ord er betydelig lenger enn oppgavens ramme på 250. En setning tjener som innledning, men det er ingen tilsvarende avslutning. De fem avsnittene forklarer hvert sitt typiske sagatrekk ut fra tekstvedlegget. Dette gir «stort sett» hensiktsmessig struktur.

Faglig kunnskap, fagspråk, redegjørelse for karakteristiske trekk: Eksaminanden påviser og forklarer aural refererende synsvinkel, objektiv fortellemåte, referat, scenisk fremstilling og persontyper/idealtyper. Fagspråk kan presiseres der underdrivelsen kalles «ekstreme språklige bilder», men ellers vises meget god faglig kunnskap og presist fagspråk, som her: «Synsvinkelen er aural og refererende» og «Komposisjonen veksler mellom referat og scener».

Relevant svar på oppgaven, forståelse og bruk av vedlagte tekst: Både oppgaven og innholdet i *Gislesoga* er godt forstått. Svaret tar tydelig utgangspunkt i utdraget og bruker flere relevante teksteksempler, som i første avsnitt, der eksaminanden først påviser synsvinkel,

forklarer synsvinkelen: «Fortelleren står utenfor handlingen og skildrer personene utenfra, uten å informere leserne direkte hva personene føler og tenker», før dette vises gjennom et relevant teksteksempel. Det er argumenter for både høy og middels måloppnåelse i svaret som helhet, så førstesensors karakterforslag 3 er det her vanskelig å finne årsaken til.

Svar på langsvaroppgaven:

Struktur og sammenheng: Langsvaret består av elleve avsnitt som varierer i lengde.

Førstesensor kan ha ment at svaret er for kort (873 ord), men det er ikke åpenbar mangel på innholdsmomenter, noe som kan gi støtte til at andresensor ikke har vektlagt dette. Enkelte avsnitt, for eksempel seks og syv, kan sies logisk å tilhøre hverandre, men oppdelingen kan også sies å fungere fordi det fremhever innholdet i disse avsnittene. Det er noe mangel på sammenheng avslutningsvis fordi to replikker her er markert med listestil, i tillegg til at det kan være uenighet om siste replikk fungerer som avslutning. Flere trekk i langsvaret bidrar til å klassifisere det som en litterær fortelling eller novelle. Kronologi er et viktig organiseringsprinsipp i fortellingen (Skjelbred, 2014, s. 47), noe bruken av temporale koplinger kan tyde på, som i første avsnitt; «da», «etter at», «så» og «før» og «i dag». Teksten inneholder også flere tempusskifter, for eksempel der retrospeksjon i fjerde avsnitt innledes ved skifte fra preteritum til pluskvamperfektum, fulgt av tidsadverb som «allerede», «aldri» og «fortsatt». Bruk av kompositoriske grep som *in medias res* og vendepunkt kan bidra til å bestemme at teksten er en novelle. I tillegg er kataforisk referanse mer vanlig i litterære tekster (Maagerø, 2005, s. 177). Ett eksempel er fra fjerde avsnitt: «I kabinen på et fly satt en smilende mann iført sin pilotuniform. Denne mannen het Rolf, og han hadde....».

Bredde og relevans i momenter/eksempler: Ideasjonelt er dette en fortelling om to brødre, Petter og Rolf, og som nevnt må den tolkes for å vurdere relevans. Fortelleren har innsyn i begge tanker, men det meste av forløpet blir fortalt gjennom storebroren Rolf. Petter er hovedpersonen, og han introduseres i første avsnitt sammen med flere statiske attributter for overflatisk vellykkethet: «Han hadde pent lagt kofferten inn i bagasjerommet på den speilblanke, sorte Mercedesen». I femte setning kan prosessen; «kunne skjule» og omstendighetene «fra kvelden før», fungere som det fortellingsteorien kaller frampek (Aarseth, 1976). Petters fasade skjuler noe som kan bli bestemmende for videre forløp. Siste setning i avsnittet gir innsyn i hvordan Petter ser ned på andre når han øker farten gjennom sølepyttene og «sikkert» sprutet «en masse på ordinære menn og damer». I andre avsnitt er det gått fem år. Kohesjonskjeden «som vanlig», «nok en», «sin daglige» og «pleide», forteller at en morgen i bakrus er rutine for Petter. I tredje avsnitt beskrives et sølvfat «med et hvitt

pulver med kapasitet til å gjøre han til et rovdyr uten like», noe mange vil forbinde med kokain, og fortelleren har etablert Petters rolle som en materielt vellykket, men likevel usympatisk misbruker, sannsynligvis på vei nedover.

Fra fjerde avsnitt er det Rolfs tanker som gir fortellingen fremdrift. Deltakerne «kone og tre barn», «godt betalt jobb i SAS» og «fredelig enebolig med stor hage», innlemmes for å fortelle om Rolfs A4-tilværelse, den også materielt vellykket. Prosesser som «smilte (til alle)» representerer at Rolf ikke er som broren sin, før en ny kontrast beskriver Rolf om kvelden, for da «smilte han ikke». Han bekymrer seg for lillebroren, vist gjennom en lengre retrospeksjon til oppveksten. Rolf beskriver Petter direkte som «notorisk mobber» og indirekte ved å fortelle om «en alvorlig overfallsepisode». Her introduseres familiens tredje deltaker, faren, «en alkoholisert forfatter med godt hjerte». I femte avsnitt blir hele Donne-teksten lagt i munnen på faren, som leste dette for brødrene i barndommen. I sjette og syvende avsnitt tegnes det videre på det kontrastfylte personportrettet av Petter. Deltakere som «de som leste Dagens Næringsliv» og «de yngre jentene som jobbet i utelivsbransjen», innlemmes for å tro det beste om ham, mens det bare var Rolf og faren som «visste» om den egentlige Petter bak fasaden; «kokainavhengig og hadde drept mennesker på bakgrunn av lyst og frustrasjon».

Etter retrospeksjonen får Rolf vite at broren har forulykket etter ruskjøring. Igjen understrekes ytre vellykkethet via nyheten om at «viktige forretningsmenn fra hele verden» skal delta i begravelsen. Neste setning er ved graven, der Rolf leser deler av Vesaas-vedlegget ved Petters grav. En mer selvstendig tilnærming til vedlegget kommer i siste avsnitt: «Hadde han bare bygd broer i stedet for sin egen skyskraper ville han ikke falt så voldsomt». Dette refererer både til Vesaas og tittelen «Hadde han bare bygd broer». Det kan fungere som en sirkelkomposisjon som rammer inn det som ideasjonelt kan leses til Petter som øy, her uttrykt gjennom en annen kontrasterende metafor til bruer, nemlig skyskraperen, som kan representere Petters hybris. En slik tolkning kan konkludere med at momentene som er med er relevante, mens uenigheten mellom sensorene kan ha sammenheng med om disse momentene har «bredde», som er kjennetegnet til høy måloppnåelse.

Kreativitet i innfallsvinkler og virkemidler: Teksten har mange litterære virkemidler ved bruk av in medias res, kronologi, frampek, retrospeksjon, vendepunkt og sirkelkomposisjon.

Synsvinkelen veksler, og det er usikkert om dette er bevisst, men uansett vises kompetanse om episke tekster i måten teksten er konstruert på. Bruk av språklige virkemidler kan det være uenighet om, fordi dette kommer an på om en er villig til å tolke teksten som en allegori på tekstvedleggene. Et eksempel er attributtene på brødrenes vellykkethet, som kan aksepteres

som et allegorisk svar på vedleggene, fordi det viser mennesket som setter det materielle fremfor mennesker. Det kan også avslås som karikerte uttrykk for avsenders verdier, nettopp fordi det ikke blir nyansert eller reflektert over. Sistnevnte oppfatning kan forsterkes av at kontrasten er hyppig brukt, mens det også kan tolkes som «gjennomført» bruk av virkemidler.

Reflekterer, prøver ut tanker: Det er Rolfs tanker som i hovedsak fremstilles i teksten. De kretser hovedsakelig rundt den vanskelige relasjonen til Petter. Her kommer det frem at det er paradoksalt at Petter hadde alle muligheter, men likevel valgte et egoistisk og destruktivt liv, men dette uttrykkes gjennom beskrivelser fremfor refleksjon. Det blir ikke gitt eksplisitt forklaring på hvorfor Petter er blitt slik han fremstilles, men det er noe upålitelig ved en forteller som hevder at Petter hadde alle muligheter samtidig som faren beskrives såpass ynkelig. En overtolkende leser kan ane en oppvekst i rus, samlivsbrudd, en fraværende mamma og brødre som har knivet om knappe ressurser. Vi husker den alkoholiserede forelderen i S10, og det kan være interessant å registrere at det blir større avstand til en slik gjennom fiksjonen, noe som også kan ha innvirkning på hvordan en så personlig tematikk kan bli mottatt. Uansett er det lite i teksten som viser refleksjon rundt temaene i vedleggene, noe en blant annet ser gjennom fremhevet og polarisert fremstilling av materiell vellykkethet.

Relevant svar på oppgaven, forståelse og bruk av vedlagte tekster: «Det skal komme tydelig fram i svaret ditt at du tar utgangspunkt i vedleggene», står det i eksamensoppgaven. Det må spørres om ovennevnte forløp viser tilstrekkelig utgangspunkt i eller forståelse av de to vedleggene. Frem til Donne-teksten siteres midtveis i teksten, er det ingen eksplisitte referanser som viser dette utgangspunktet, noe som kan være grunnen til at sensor 1 har foreslått middels måloppnåelse. En motvillig leser kan si at det viser «enkel bruk av vedlagte tekster» (lav måloppnåelse) at hele vedlegget gjengis i en replikk. Når momentet så forlates med en gang, kan leseren tekstuelte oppfatte det som mangel på sammenheng (Maagerø, 2005, s. 97), og avslå denne måten å løse oppgaven på. Samtidig er det ikke uvanlig at adressater av fiksjonstekster vil akseptere et slikt tilbud om selv å skape sammenheng (Maagerø, 2005). Vedleggene handler om relasjoner, og fortellingen kan leses som en allegori over dette. Da blir øymetaforen som begge forfattere bruker representert av Petter. En adressat som blir med på denne tolkningen, som sensor 2, vil kanskje også lese gjengivelsen av Donnes tekst, og da referansen til døden, som et nytt frampek til hvordan det skal gå med Petter. Det kan likevel innvendes, etter å studere kjeder mellom ideasjonelle innholdsord for å få kunnskap om feltet i situasjonskonteksten, at tekstens ideasjonelle mening er på siden av vedleggene. Forskjellige felt har forskjellige vokabular og aktiviteter (op.cit, s. 181), og gjennom leksikalske

kohesjonskjeder kan minst fem felt avdekkes; bilkjøring, vellykket fasade, forretningsliv, rusmisbruk og familie/oppvekst. Med unntak av familiefeltet, som angår relasjoner og fellesskap, reflekterer ikke de andre feltene vedleggstematikk. I hovedsak er de også hentet fra dagliglivets domene (Tønnessen & Maagerø, 2015).

Tekstuell metafunksjon kan også svare på om teksten er effektiv ut fra hensikt og omgivelser (Maagerø, 2005, s. 159). Eksamenstekstens verbale omgivelser er oppgavetekstens krav om å ta utgangspunkt i vedleggene. Jeg har vist hvordan dette eksplisitt er realisert og foreslått tolkninger av hvordan det kan være implisitt. En kort analyse av hvordan eksaminanden behersker å referere gjennom både *endofori* og *homofori* (op.cit, s. 177), til referenter både innen- og utenfor teksten, som bruken av *homofori* i syvende avsnitt; «de som leste Dagens Næringsliv», tilsier at dette er en skribent med forutsetninger til å referere mer til vedleggene også. Det finnes hold for at fortellingen om Petter, som ødelegger for både seg selv og fellesskapet, viser frem sider ved tematikken i vedleggene, og det på en selvstendig og kreativ måte, som er kjennetegnet til høy måloppnåelse. Samtidig blir det vist lite lesekompetanse, og det kan fremstå paradoksalt at det refereres både internt i og til felt utenfor teksten, men i liten grad til vedleggene. I den forbindelse kan bruk av referanser fra rusfeltet og fremheving av materielle statusmarkører være en mulig risiko, fordi heller ikke de befester tekstens forhold til situasjonskonteksten.

Formuleringsevne, ordforråd og formelle ferdigheter:

Nevnte bruk av referanser og vekslingen mellom måten de fremtrer i syntaksen, skaper sammenheng og vitner om evne til å variere ordforråd og setninger. Samtidig er det en del kommafeil i teksten, i tillegg til flere anglisismer, som «var rapportert å komme» i avsnitt åtte og «i fronten for sitt eget firma» i avsnitt fire. Det er mulig risiko i bruk av tabuord, men det fremstår ikke umotivert. Jeg har vært inne på hvordan livet til Petter fortelles gjennom positiv og negativ polaritet, gjerne uttrykt i trivielle ordvalg, noe flere av eksemplene over kan illustrere. Det relative fraværet av modalitet (Maagerø, 2005, s. 150) uttrykker mangel på nyanser til kommunikasjonen, noe som igjen kan føre til både aksept og avslag fra adressaten. Fordi teksten skal reflektere, vil dette kunne leses som en mangel. En slik negativ vurdering kan forsterkes av momenter som fremstår overdrevet, en tendens som toppes i påstanden om at Petter skal ha drept flere mennesker uten andre konsekvenser enn at familien har vendt ham ryggen. Det er likevel eksempler på at ordforrådet er variert. I første avsnitt benyttes *merotymi* mellom «Mercedes» og «bagasjerommet», «kabinen», «gasspedalen». «Stockholms travle morgentrafikk» kan leses som et nivå over «bilen», noe som gjør både den, «sølepyttene» og

«lysene ble grønne» til *kohyponomer*. Språklig klassifisering i over- og undergrupper forbindes med det spesialiserte domenet (Tønnessen & Maagerø, 2015), men det spørres om eksaminandens bilinteresse overbeviser mottaker om at dette gjør teksten mer faglig. Uansett gir slike leksikalske kohesjonskjeder variasjon og sammenheng. Dette er skrivekompetanse som kan virke formildende i forhold til tidligere omtalte polarisering, men det kan ha vært uenighet om dette «viser variasjon» (middels) eller om det gjør ordforrådet «nyansert» (høy).

Oppsummering av mulige årsaker til sprikvurdering (5 og 3) av S6:

- Novellen som innfallsvinkel viser kompetanse om sjangeren, men det kan ha vært uenighet om en litterær tekst løser formålet med oppgaven, som var å reflektere
- Innhold hentet fra andre ideasjonelle felt enn vedleggene gir uklart forhold til situasjonskonteksten og ytterligere tvil rundt hvorvidt oppgavekontrakten er innfridd
- Polarisert og tidvis overdrevet fremstilling av hovedpersonen Petter
- Språket kjennetegnes også av polarisering, mangel på modalitet, til dels triviell språklig stil og et par forekomster av tabuord

Akkurat denne oppgaven fikk jeg anledning til å diskutere med erfarne norsklærere i forbindelse med et foredrag jeg holdt om denne undersøkelsen for *Fagnettverket for norsklærere i Buskerud* 26.11.2015. Her møtte jeg 20 kolleger som jeg ba vurdere S6. Konteksten var sprikvurdering, lærerne visste at dette var en spriktekst, og en kan forvente at dette var styrende for vurderingen som ble gitt. I tillegg ble det noe samsnacking deltakerne imellom, på den måten at flere gjerne sjekket ut egne med andres vurderingsnormer før de bestemte seg for eget karakterforslag. Dette var med andre ord et stykke unna reell sensur, men kan likevel belyse analysen. 19 av lærerne delte sitt karakterforslag, og de aller fleste foreslo 4, karakteren teksten også hadde fått etter fellessensur. Fire av lærerne foreslo 3, og tre foreslo 5, noe som reflekterer uenigheten mellom sensor 1 og 2. Det var enighet om at det var vanskeligst å vurdere om langsvaret svarte på oppgavens «reflektere om tematikken i vedleggene», og at dette berørte flere av vurderingskriteriene. Enkelte, som hadde satt 3, mente at løsrevet omgang med tekstvedleggene ikke var tilstrekkelig svar på oppgaven, og at den eneste grunnen til at karakteren likevel ble 3, var andre kvaliteter og ikke minst et godt kort svar som trakk opp. Andre begrunnet karakteren 5 med at nettopp denne måten å vise frem tematikken på, viste et selvstendig svar på oppgaven.

S19 «Det grønne havs øyer»: Jeg-fortellingen som reflekterer

Dette er en relativt kort og kronologisk fortelling som tematiserer menneskers fellesskap gjennom både familie- og universelle bånd. Den er en av to sprikfortellinger i utvalget som kombinerer et forløp med refleksjon rundt vedleggene. Refleksjon realiseres direkte gjennom jegets tankereferat og indirekte gjennom handlingen. I et vendepunkt mot slutten får mottaker vite hva «Det grønne havs øyer» er. Sensorene vurderte besvarelsen til karakterene 4 og 6, og ble enige om 5.

Svar på kortsvaroppgaven:

Struktur og sammenheng: Svaret er med sine 275 ord over rammen på 250 ord. Det er strukturert i fire avsnitt, hvorav første og siste er svært korte fordi de skal tjene som innledning og avslutning. Avsnitt tre og fire utgjør hoveddelen og svaret på oppgaven. Både lengden og de to korte avsnittene kan en være uenig i vurderingen av. I et kortsvar forventes en slags innledning og avslutning, men kanskje heller gjennom hvordan første og siste setning formuleres. Her blir avslutningsavsnittet nærmest en gjentakelse av det som tidligere er sagt, og det kan ses som unødvendig, særlig når svaret er for langt. Andre vil kanskje ikke legge vekt på hverken lengde eller avsnitt, fordi svaret ikke er så mye lenger enn det skal være, og fordi eksamensveiledningen ikke utdyper strukturen utover at det skal skrives «en strukturert og sammenhengende tekst», noe som blant annet innebærer inndeling i avsnitt.

Faglig kunnskap, fagspråk, redegjørelse for karakteristiske trekk: I andre avsnitt påpeker eksaminanden typiske sagatrekk knyttet til synsvinkel og komposisjon, som her etter første teksteksempel: «Fortelleren står utenfor situasjonen og forteller handlingen sett utenfra uten indre tanke- eller følelsesbeskrivelser». I tredje avsnitt gjøres det rede for personschildringer, tematikk og underdrivelser. Et eksempel på presist og formålstjenlig fagspråk ses i tredje avsnitt: «Sagakomposisjonen veksler mellom handlingsreferater og scener».

Relevant svar på oppgaven, forståelse og bruk av vedlagte tekst: Når eksaminanden først nevner et trekk ved å bruke fagbegrepet, viser det gjennom et teksteksempel, før det forklares, viser det en klar og relevant måte å svare på. Flere direkte og indirekte referanser til utdraget viser aktiv bruk av vedlagt tekst, og det er god lesekompetanse i måten utdraget beskrives på, som her: «Dette utdraget er i hovedsak en scene, som vil si at situasjonen er spenningsfylt og avgjørende for historien». Uenigheten mellom sensorene sto mellom middels og høy måloppnåelse, og det er grunn til å tro at kortsvaret har kunnet bidra til enigheten om høy.

Svar på langsvareoppgaven:

Struktur og sammenheng: 745 ord over åtte avsnitt gir et relativt kort langsvar, noe som kan svekke vurderingen. I tillegg er enkelte avsnitt korte, noe som kan svekke inntrykket av sammenheng. Imidlertid kan kreative innfallsvinkler forsvare både kortere svar og avsnitt enn i andre oppgavetyper, og det sensor skal vurdere er om strukturen er «hensiktsmessig». Når det gjelder avsnittene, følger de en logisk inndeling, selv om det ikke hadde vært noe i veien for å slå et par av dem sammen, for eksempel fjerde og femte. Fortellingen som helhet er strukturert rundt et forløp der vi følger en kjernefamilie på spasertur, og er oppbygd av naturskildringer, handlingsreferat, enkelte replikker og hovedpersonens tanker. Fordi det er et tydelig vendepunkt i nest siste avsnitt, som også gir et tolkningsrom som utvider teksten, trenger ikke fortellingen flere momenter, men fordi også handlingen er enkel, kan det forklare en mulig uenighet om strukturen er «stort sett» eller helt hensiktsmessig.

Bredde og relevans i momenter/eksempler: Fortellingen følger en familie en vårdag, og ved første øyekast kan det virke som handlingsoppbyggingens enkelhet (de går i vårværet, jeget reflekterer og studerer lillesøsteren sin) kun tjener som et spinkelt rammeverk for refleksjoner rundt menneskers forhold til hverandre, som kanskje er den mest opplagte tolkningen av tekstvedleggene. Eksempler på dette kommer noe senere i analysen. Imidlertid introduseres et nytt moment mot slutten av sjette avsnitt, som følges ut teksten, og det er referansen til dødstematikken i Donnes tekst. Dette kommer når jeget skal svare på egne retoriske spørsmål om hva som skjer når bånd og broer brytes og faller:

Vi kan segle uten vind og vi kan ro åreløse, og gråte, det kan vi og. Og le og synge og leve og dø, men før den tid skal vi (...) På den måten mister vi ikke oss selv. Vi gjør oss bare større og lar noen andre beholde det som er blitt ekstra når vi selv ikke kan være hos dem mer.

I syvende avsnitt antydes målet for tittelen og familiens tur: «Jeg ser over havet av steiner som titter opp som øyer i det store, grønne gressvannet». Tidligere i teksten, på slutten av andre avsnitt, er det fremhevet at gresset er grønt. Der denne opplysningen kommer, kan den leses som noe umotivert, men i syvende avsnitt kan det hjelpe forståelsen av det refererte bildet når det utdypes av: «Blomster nesten overalt, men ikke alle steder. Det er trist». Det sies ikke direkte, men denne leseren tolket det til at familien nå var kommet til en kirkegård. De står der en stund, og når de i siste avsnitt går derfra er de «glade, men litt tyngre enn da vi kom. Og alle har en dråpe i øyekroken og er litt varmere på hendene enn vi pleier å være», noe som kan bekrefte at de har besøkt en avdød. Denne tolkningen styrkes av at jeget, etter en ny referanse til den tause vissheten i Vesaas' tekst og båndene som gjør dem til en helhet, avslutter med en direkte Donne-referanse: «Og vi går tilbake den samme veien som vi kom og hører kirkeklokkene som ringer for oss alle». De to sensorenes forslag kan fortelle om

uenighet i vurderingen av dette kriteriet. De har nok vært enige om at fortellingen er relevant, men sensor 2, som foreslo 6, må i tillegg ha ment at den viser «bredde».

Kreativitet i innfallsvinkler og virkemidler: Det narrative og subjektive er primærdiskursive trekk som det kan ha oppstått uenighet rundt (Penne, 2010, s. 30). Det omtalte vendepunktet gir et ekstra tolkningsnivå og dybde til teksten, noe som kan betraktes som et litterært kvalitetstegn (Hagen, 2004, s. 39). Det er ikke urimelig å anta at andresensors innstilling på karakter 6 kan uttrykke et slikt tekstsyn, og at vedkommende har brukt kriterier for høy måloppnåelse, som «viser kreativitet» og at bruken av litterære virkemidler er «gjennomført», i vurderingen. Videre kan det antas at førstesensors 4-forslag uttrykte at teksten viser «noe kreativitet» og «bruker litterære virkemidler». Imidlertid kan det være vanskelig å avgjøre hva i en tekst som skiller virkemiddelbruk fra gjennomført sådan. Sensorene uttrykte også tvil til om vurderingskriteriene til denne oppgaven var presise nok til å avgjøre blant annet dette (Utdanningsdirektoratet, 2015e). Virkemiddelbruken er allerede illustrert i flere av teksteksemplene. Her kan nevnes hvordan eksaminanden både har jobbet videre med metaforene fra vedleggene og utviklet selvstendige bilder. Det andre er variasjon i fortellermåter, oppbyggingen til vendepunktet, bruk av anaforiske gjentakelser i flere setninger (ses for eksempel i første sjetten avsnitt), billedbruk og allusjonen til den svenske folkesangen «Hvem kan segla förutan vind» i sjetten avsnitt. I tillegg kommer den enkle, naivistiske stilen som brukes noen steder, som i observasjonen av lillesøsteren på kirkegården i syvende avsnitt: «Og lillesøster vil telle øyene, men klarer ikke, for hun kan ikke telle lenger enn til tretti». Under språk-kriteriet beskriver jeg at denne enkle stilen kanskje ikke er like vellykket alle steder, og at den kombinert med trivielle ordvalg kan være et trekk ved teksten det kan ha vært uenighet om. Tidligere nevnte tolkningsnivåer kan også være et risikabelt grep i seg selv, fordi en eksamenstekst skal leses og vurderes i løpet av 26 minutter, noe som kan være til hinder for en estetisk lesning (Rosenblatt, 1994) og føre til at det tolkningsrommet (Eco, 1979) som teksten åpner for (Eco & Rognlien, 1994) blir stående tomt. Dette samsvarer også med tidligere funn av sprikvurdering knyttet til sjangerløse oppgaver (Skar og Aasen, 2015). Samtidig er det vanskelig å se for seg at det har oppstått uenighet når det kommer til bredden eller variasjonen i virkemiddelbruken, da denne er relativt åpenbar.

Reflekterer, prøver ut tanker: I langsvaret vises refleksjon gjennom tanker og beskrivelser, og eksempler på dette ble vist over og flere kommer under neste kriterium. Det reflekteres direkte gjennom tankereferatene, og de to karakterforslagene uttrykker enighet om at dette er

utprøvende, mens det bare er forslaget fra sensor 2 som uttrykker at dette er selvstendig. Denne har kanskje lest refleksjonen som fremkommer indirekte gjennom de andre fortellemåtene på en annen måte enn sensor 1. Her tenker jeg på det som blir vist frem av bekreftelser på fellesskap i handlingen, som når familien venter på søsteren som har sakkett akterut i andre avsnitt, i tredje holder hverandre i hendene eller tilbyr hjelp i sjette. Om en er villig til å lese dette slik, kan en si at indirekte og direkte refleksjon også bidrar til dybde og selvstendighet. Eksempler viser også evne til abstraksjon, evaluering og generalisering av enkelterfaringer, noe som viser kompetanse i det spesialiserte og refleksjonsdomenet (Tønnessen & Maagerø, 2015). Dette kan også tale for at teksten, på tross av sin subjektive stil, er mer dialogisk enn selvreflekterende og evner å balansere forholdet mellom det individuelle og kollektive (Aase, 2012).

Relevant svar på oppgaven, forståelse og bruk av vedlagte tekster: Svaret er både kreativt og reflekterer med utgangspunkt i vedlegg, så oppgaveformuleringen er forstått, men bruken av vedleggene kan ha ført til uenighet. Det er ikke referanser til forfatterne av vedleggene, men enkelte eksplisitte og en rekke implisitte referanser til tekstene. Første direkte referanse kommer i tittelen, der øymetaforen fra begge vedleggstekster er brukt. Indirekte møter en allerede i første avsnitt et «vi», som refererer til fellesskapet i familien. Fellesskapet beskrives også gjennom naturen, som denne referansen til Donnes tekst i første avsnitt: «Kun ett av mange, men likevel en del av hele det store treet». I tredje avsnitt reflekterer jeget rundt båndene som oppstår mellom mennesker, og at de som har bånd mellom seg vet om det, uten ord, noe som tjener som en referanse til Vesaas' tekst: «Man trenger ikke si noe». I fjerde avsnitt tas denne tanken videre til å gjelde alle mennesker, beskrevet som «en felles forståelse» også mellom de ulike ukjente menneskene familien passerer på turen. I femte avsnitt blir lillesøsteren sliten, og får i sjette tilbud om å bli båret av faren, noe hun avslår. Dette åpner igjen for refleksjon rundt fellesskapets motsats: «Alltid skal mennesket bevise at det klarer ting selv. Ene og alene», før denne indirekte referansen til Vesaas blir mer direkte når brumetaforen utdypes: «Båndene og broene mellom menneskene som gjennom livet oppstår, enten vi vil eller ikke. Og hva skjer når båndet brytes? Hva skjer når den bygde broa faller? En mister en del av seg selv». Her ser refereres det også til Donnes tekst, noe som kan tolkes å formidle «god forståelse» (høy måloppnåelse) av begge de vedlagte tekstene ved at det fremhever felles tematikk, noe det er rimelig å tro er førstesensors vurdering. Imidlertid er kjennetegnet til middels måloppnåelse (andresensors forslag) at «eleven forstår og bruker de vedlagte tekstene», og det kan være vanskelig å avgjøre hva i en tekst som kan tjene som

bevis for at forståelsen er «god» eller ikke. Uansett brukes vedleggene gjennomgående, også i flere eksempler enn de nevnte. Som tidligere nevnt skiller det denne sprikteksten fra de fleste andre fortellinger i utvalget. Det kan tyde på at dette har vært vanskelig å få til for andre.

Formuleringsevne, ordforråd og formelle ferdigheter:

Med unntak av kommafeil, er det få formelle feil i teksten, men det kan stilles spørsmål til ordforrådet. Her kan det pekes på tre tendenser. Den ene er bruk av enkle ord som gjerne gjentas, kombinert med korte setningsemner, som her innledningsvis: «Det er sol ute. Endelig sol. En fin dag å ha sol på». Det andre er trivielle eller upresise ordvalg, som i femte avsnitt (min utheving): «Lillesøster *bumser* ned på en stor stein (...) i denne *underforståtte* voksenverdenen». Det tredje er en tendens til høystemthet, som her i første avsnitt: «Trærne (...) tårner høyt over oss, og langt der oppe kan vinden rykke tak i tretoppene så de svaier langsomt frem og tilbake og bladene risler stillferdig». I 2.2.3 var jeg inne på at det trivielle og høystemte tidligere er satt i sammenheng med sprikvurdering, og at trekket også ble funnet i gode elevers tekster (Hertzberg & Andersen, 2005, s. 283). Enkelheten kan tolkes til å kamuflere et lite nyansert ordforråd. Det enkle, konkrete og dagligdagse assosieres med dagliglivets språk (Tønnessen & Maagerø, 2015), men kan også leses som en naivistisk innfallsvinkel til en fortelling der et barn spiller en av hovedrollene. Det siste kan fungere i skildringene av lillesøsteren, men bør kanskje ikke være så gjennomgående at en også kan finne det i jegets egne refleksjoner, som her i andre avsnitt: «Det er så rart med mennesker». En mottaker som allerede har reagert på enkelheten, vil kanskje få bekreftet enkel formuleringsevne (kriterium for lav måloppnåelse) i trivielle ordvalg som i eksempelet over, og det kan være bakgrunnen for førstesensors vektlegging. Samtidig har mange VG3-elever lest Erlend Loe, en forfatter som i sin tid gjorde karriere av sin enkle stil, og når enkelheten til dels åpner tolkningsrom, bør den kanskje leses som kreativitet i innfallsvinkler og språklige virkemidler, noe det er sannsynlig at andresensor har gjort.

Oppsummering av mulige årsaker til sprikvurdering (4 og 6) av S19:

- Tendens til enkelt, uklart og trivielt språk kan gi uenighet om formuleringsevne og ordforråd, som kan videreføres til uenighet om refleksjonen er selvstendig
- Bruk av gjentakelser og korte setningsemner, som kan befeste et inntrykk av enkelthet eller leses som kreativ språkbruk
- Enkel handlingsoppbygging av en relativt kort tekst, som kan føre til uenighet rundt strukturens hensiktsmessighet og «bredde» i valg av innholdsmomenter

- Et vendepunkt som, fordi det antydes gjennom billedbruk, kan gå tapt for en lite litterært eller estetisk innstilt leser

5.3.8 Analyser av tekster i intensivutvalget som i hovedsak reflekterer

Fire av de syv tekstene i denne kategorien, S13, S14, S15, S20, er relativt tydelige essay. Disse er alle vurdert til gode karakterer (4, 5 og 6), mens S3, S21, S22 er vurdert til 3 og 2. Jo flere ganger jeg leste tekstene i ekstensivutvalget, jo flere har forsvunnet fra denne til blandingskategorien. Det kan fortsatt knyttes tvil til kategoriseringen av enkelte tekster, og da kanskje særlig de svakeste. Jeg har valgt å se på en av disse, S3, som er et tvilstilfelle fordi den avslutningsvis ser ut til å gå over i det analytiske. Den representerer et trekk som finnes i flere tekster, nemlig refleksjon uten kreativitet, noe enkelte sensorer fremhevet som en risiko når oppgaven ber om begge deler. Jeg vil også se på en av de mer vellykkete, S13, som løser oppgaven ved blant annet å være tverrfaglig; finne støtte til refleksjonen i geografifaget (samme innfallsvinkel som S14). Den er også valgt fordi det også her kan ha vært uenighet rundt hvordan det kreative er realisert. Her er det aspekter ved både billedbruk og metakommunikasjon som kan ha overskygget vurderingen av refleksjonen.

S3 «Kjærlighet binder enhver øy sammen»: Ikke-kreativ refleksjon

Langsvaret representerer en tendens i de reflekterende tekstene: Å reflektere rundt tematikken, men uten å vise kreativitet. Usikkerhet rundt oppgavetyper kommer også til uttrykk gjennom at eksaminanden avslutningsvis går over i det analytiske. De to sensorene foreslo på hver sin kant karakterene 4 og 2, og ble enige om 3.

Svar på kortsvaroppgaven:

Struktur og sammenheng: Kortsvaret er noe lenger enn de 250 ordene som oppgaven ber om. I første avsnitt presenteres en del generell informasjon om form og innhold i sagalitteraturen, og fordi det ikke knyttes til tekstutdraget, er dette på siden av oppgaven. I tillegg gir det gjentakelser når underdrivelse og ære igjen omtales i andre avsnitt. I andre avsnitt vises det både indirekte og direkte til et sted i teksten, og det brukes til å påvise underdrivelse. Dette er relevant, og fagbegrepet er korrekt. Tredje avsnitt inneholder også et sjangertrekk som eksemplifiseres, før det avsluttes med: «På denne måten kommer det frem i teksten», noe som gjør avsnittet og avslutningen ufullstendig.

Faglig kunnskap, fagspråk, redegjørelse for karakteristiske trekk: Det er noen karakteristiske trekk som påvises ved hjelp av noe fagspråk: underdrivelse og «kort og enkelt språk». Dette

viser noe eller nokså god faglig kunnskap, men usikkert om dette kan forklare sprikvurdering mellom karakterene 2 og 4.

Relevant svar på oppgaven, forståelse og bruk av vedlagte tekst: Henvisning til generelle sjangertrekk uten å vise til teksten gjør deler av svaret mindre relevant, noe som også svekker lesekompetansen. Utsagn som «Setningsbygningen i sagastilen er veldig korte og enkle», viser språklig usikkerhet, i tillegg til at det ikke forankres til teksten. Om dette kortsvaret skal tillegges vekt ved sprikvurdering, vil det sannsynligvis bidra til å trekke hovedinntrykket ned.

Svar på langvaroppgaven:

Struktur og sammenheng: Teksten består av seks avsnitt, og de er stort sett markert på enhetlig måte. Første avsnitt er ingen åpenbar innledning, og siste avsnitt startes som en oppsummering, men går helt avslutningsvis over til å analysere vedleggenes oppbygging. Dette er en fin observasjon, og den passer til tekstlesningen som er presentert, men det er et nytt moment som ikke passer som avslutning. Det kan også bidra til usikkerhet knyttet til om avsender har forstått at oppgaven har vært å reflektere, ikke analysere. Usikkerheten til skrivehandlingen viser manglende sjangerkompetanse, noe som åpner for at eksaminanden kunne fått mer hjelp av sjangerinstruks i oppgaven (Blikstad-Balas & Hertzberg, 2015). Det svekker oppgavesvaret, noe arbeid med skrivestrategier og modelltekster kunne motvirket (Hertzberg & Dysthe, 2012). Det kan se ut til at en av sensorene har trukket mye for dette, mens den andre ikke har tillagt det så stor vekt.

Bredde og relevans i momenter/eksempler: Langsvaret innledes ved å gå rett på et sitat fra Vesaas' tekst. Avsender viser deretter refleksjon rundt sitatet, og de tre første avsnittene handler for det meste om refleksjoner eller forklaringer av dette vedlegget. Refleksjonen er realisert ved at en utprøvende tankeprosess speiles, som her i første avsnitt: «...men for å bygge en bru til et annet menneske, må du åpne deg. Du må slippe inn flere folk på øyen din, for det er sjelden du står alene. Du står egentlig aldri alene». I andre avsnitt vises det samme gjennom spørsmål: «Er vi en egen øy eller er vi en del av fastlandet? Er vi alle sammensveiset?». Det at langsvaret tar så tydelig utgangspunkt i vedleggene er både en styrke og svakhet. Ettersom oppgaven har bedt om refleksjon rundt vedleggene, er det en styrke å være tekstnær, blant annet gjennom utstrakt bruk av sitater. I de tre første avsnittene ser en også både vilje og evne til å utdype eller forklare metaforene til Vesaas. I fjerde og femte avsnitt ses det samme med utgangspunkt i Donnes tekst. Innfallsvinkelen er også en svakhet ved teksten, fordi den gir lite rom for kreativitet. Fordi mye av teksten blir sitater og egne

forklaringer som ligger tett opp til sitatene, blir det heller ikke så mye plass igjen til de selvstendige refleksjonene. Noe er det, som for eksempel der avsender reflekterer rundt relasjoner i andre avsnitt. Likevel bidrar altså innfallsvinkelen til at det i større grad er analytisk tekstforståelse enn egen refleksjon ut fra vedleggene som preger teksten. Samtidig vises god lesekompetanse flere steder, som i et eksempel fra siste avsnitt:

John Donne som skrev på 1600-tallet, ser mer på den helhetlige verden, og hvordan vi er avhengige av hverandre, mens Tarjei Vesaas ser mer på det som at hver person er for seg selv, og bygger sin egen sosiale krets. Her er det mindre fokus på fellesskapet, men kanskje mer fokus på seg selv og sine.

Her ses enkelte ting som ikke fungerer så godt språklig, men det er likevel en relevant sammenlikning. Det viser som nevnt lesekompetanse, men representerer også en analytisk innfallsvinkel som har fortrenget det kreative og derfor kan vurderes som mindre relevant. Imidlertid er momentene som er tatt fra tekstene relevante (middels), og for Vesaas tekst er det også god bredde (høy).

Kreativitet i innfallsvinkler og virkemidler: Det er et hovedinntrykk at teksten er lite kreativ. Klisjéen «like sikkert som at jorda er rund» i første avsnitt er en sammenlikning og hører hjemme under kjennetegnet til lav måloppnåelse. Tendensen til å gjenta ord og formuleringer, som i andre avsnitt, kan leses som et virkemiddel som realiserer utprøvende refleksjon, men utover i teksten bærer det mer bud om at avsender ikke har mer å skrive om. Refleksjonen rundt forfatterens billedbruk i for eksempel tredje og fjerde avsnitt, tilfører ikke noe nytt til bildene, og kan derfor ikke sies å bidra til kreativ språkføring. Innfallsvinkelen som er valgt er heller ikke kreativ, som beskrevet over. Imidlertid er kjennetegnet til lav måloppnåelse her at teksten "viser at eleven har valgt noen innfallsvinkler», så innfallsvinkler må ikke være kreative før karakteren 3 skal foreslås. Andresensors forslag kan uttrykke at denne mangelen er tillagt stor vekt, mens førstesensor ser ut til å ha lagt mindre vekt på denne kompetansen.

Reflekterer, prøver ut tanker: Det varierer i hvilken grad avsender reflekterer. Det kan argumenteres for at dette er «i noen grad» (lav), men jeg har også vist at avsender skriver utprøvende (middels), særlig i første del av svaret. Jeg har tidligere uttrykt usikkerhet til om denne teksten hører helt hjemme i kategorien, fordi det siste analytiske trekket kan plassere den blant blandingstekstene. Samtidig er det først i siste del av avslutningsavsnittet at formen brytes, og selv om refleksjonen flere steder preges av å være mer forklarende enn utprøvende, mener jeg fortsatt at dette er en tekst som i hovedsak reflekterer. Det kan heller være ujevn lesekompetanse som gjør at avsender viser tendens til å «fylle ut» teksten med sitater og

forklaringer. Likevel har eksaminanden fått frem såpass god lesekompetanse at en kan anta at klarere rammer for struktur og innhold i oppgaven kunne hjulpet til å vise høyere kompetanse.

Relevante svar på oppgavene, forståelse og bruk av de vedlagte tekstene: Eksaminanden reflekterer som nevnt med utgangspunkt i vedleggene, men har ikke skrevet en kreativ tekst, så svaret viser ikke at oppgaven i sin helhet er lest og forstått, noe jeg har vært inne på kan forklare sprikvurderingen. De vedlagte tekstene vises det lesekompetanse til, og da noe mer til Vesaas' tekst enn Donnes. I langsvaret både «forstår og bruker avsender de vedlagte tekstene», som er kjennetegnet til middels måloppnåelse, og når de to tekstene leses opp mot hverandre vises til dels «god forståelse», som er kjennetegnet til høy. Når det kommer til måten de er brukt på, må en kanskje velge å vurdere at de er brukt på «en enkel måte» (lav måloppnåelse), fordi ovennevnte strategi med å vise egen forståelse av utvalgte teksteksempler viser at de ikke er brukt «selvstendig», som er kjennetegn til middels og høy. Dette gir en såpass variert lesekompetanse at det kan forklare begge sensorers forslag.

Formuleringsevne, ordforråd og formelle ferdigheter:

Det er en del formelle feil. Det angår blant annet bruk av funksjonsord, som «noen tegn for nederlag» og «Dette viser store tegn for underdrivelse» i henholdsvis første og andre avsnitt i kortsvaret. Enkelte steder vises ortografisk usikkerhet, som når det i kortsvarets første avsnitt skrives «Setningsbygningen i sagastilen». Det er også en del kommafeil og ikke minst feil tegnsetting flere steder der det er brukt sitater, som dette eksempelet fra femte avsnitt: «Og som John Donne sier «hvert menneskes død forminsker meg.» så mener han at ...». Her mangler kolon og komma, mens punktum er feilplassert. Fordi sitatbruken er såpass utstrakt, svekkes også sammenhengen noe når det blir såpass uklart tegnsetting.

Formuleringsevnen og ordforrådet bærer preg av gjentakelser av ord og formuleringer, som «lagt mer fokus på» og «legger altså fokus på» i slutten av andre avsnitt i kortsvaret. I tillegg er det en del anglisisme og upresise formuleringer, for eksempel «...der du rekker ut til en du er glad i» i langsvarets tredje avsnitt, eller «Det stemmer fint som de sier det her» i første avsnitt. En kan også spørre seg hvor mye innhold som kommuniseres i denne formuleringen fra fjerde avsnitt: «Han tar for seg livets gang, og hvordan livet henger sammen, hvordan det går for seg». Som helhet er teksten likevel ikke gjennomgående «uklart formulert», som er kjennetegnet til lav måloppnåelse, men den er kanskje heller ikke «stort sett godt formulert», som er kjennetegnet til middels. Igjen kan det se ut til at andresensors forslag uttrykker en streng vurdering, mens førstesensor har lagt mindre vekt på dette.

Oppsummering av mulige årsaker til sprikvurdering (2 og 4) av S3:

Analysen har vist at det er trekk som kan vurderes til både lav, middels og enkelte steder høy måloppnåelse. Fordi sensorer som vektlegger ulikt vil kunne komme frem til ulike karakterforslag, gir ujevn kompetanse generell risiko for sprikvurdering:

- Ulik vektlegging av usikkerhet i språk og struktur som understrekes av formbruddet avslutningsvis
- Ulik vektlegging av mangelen på kreativitet i innfallsvinkler og språkbruk
- Ulik vektlegging av lesekompetansen knyttet til langsvarets vedleggstekster
- Ulik vektlegging av tendens til å lese vedleggstekstene mer analytisk enn utprøvende

S13 «Et spørsmål om perspektiv»: Den selvbevisste avsenderen

Langsvaret prøver ut tanker rundt hva mennesket er med utgangspunkt i de geografiske metaforene øy og fastland, før eksaminanden reflekterer videre gjennom egen billedbruk. Metakommunikasjon og ordbruk fra andre oppgavetyper kan leses som både lekenhet og uønskete brudd. Til forhåndssensuren fikk besvarelsen karakterene 6 og 3. De to sensorene ble enige om karakteren 5.

Svar på kortsvaroppgaven:

Struktur og sammenheng: Eksaminanden holder seg innenfor rammen av 250 ord. Gjennom tre avsnitt gjøres det greie for noen karakteristiske trekk ved ættesagaene, primært fortellemåte og oppbygging. Første avsnitt har en innledende setning, men svaret avsluttes uten avrunding, men ellers er strukturen hensiktsmessig.

Faglig kunnskap, fagspråk, redegjørelse for karakteristiske trekk: I andre avsnitt påvises «at mye handling beskrives med få ord», noe som begrunnes med et teksteksempel som viser underdrivelse. Dette avslører en usikkerhet rundt hva som er hva, noe en kan se flere steder: For eksempel i første avsnitt der fortelleren defineres som helt upartisk, mens en kanskje forventer at de sterkeste svarene får frem at det var slik denne fortelleren kunne ønske å fremstå. Mot slutten av første avsnitt leser eleven eksamenssettets kontekstualisering av utdraget som innledningen til selve sagaen, noe som viser manglende forståelse, også fordi utdraget er en sluttscene. Tematisk blir hevndrapet tillagt vekt, men svaret tar ikke frem andre tematiske trekk, noe som gjør at avsender ikke får vist så mye lesekompetanse om utdragets innhold. Dette gir middels måloppnåelse på faglig kunnskap og lav på fagspråk.

Relevant svar på oppgaven, forståelse og bruk av vedlagt tekst: Første halvdel av svaret, til slutten av første avsnitt, er en generell beskrivelse av sjangeren, som her: «fortelleren står utenfor teksten og kun beskriver handlingen». Dette er på siden av oppgaven, fordi den har bedt om at slike trekk koples til utdraget, og derfor viser dette også lavere lesekompetanse. I siste halvdel viser eksaminanden direkte og indirekte til tekstvedlegget. Det er lite trolig at kortsvaret kan forklare sprikvurderingen, men manglende vekting av det kan forklare hvorfor en av sensorene har foreslått karakteren 6.

Svar på langsvaroppgaven:

Struktur og sammenheng: Avsnittsinndelingen er ikke helt enhetlig. Det er åtte avsnitt som er markert med luft, men tre av disse er igjen delt opp ved hjelp av linjeskift. Det er ikke uvanlig at elever på den måten markerer en slags pause, men det er ikke hensiktsmessig. I en reflekterende tekst er det ikke de samme kravene til logisk struktur som i andre oppgavetyper (Lothe et al., 1997), slik at en får følge tankeprosessen. Tekstanalysen viser grep som er gjort for sammenheng i teksten, men samtidig kan en også ønske mer ordening av tankene før de presenteres i tekst (Hoel, 2008). Her tenker jeg spesielt på at vedleggenes tematikk, som det skal tas utgangspunkt i, ikke presenteres før i fjerde og femte avsnitt. Dette kan vurderes som lite hensiktsmessig når det er disse lesningene refleksjonen tar utgangspunkt i.

Bredde og relevans i momenter/eksempler: Innledningsvis tar teksten utgangspunkt i at essensen i tekstvedleggene er to ulike påstander om mennesket. Ordvalg som nettopp «påstand», fulgt av «påstå», «spørsmålet» og «kritisk spørsmål», utgjør en kohesjonskjede som kan minne om en argumenterende oppgavetype. Det antyder at avsender ønsker å drøfte tekstvedleggene som påstander som eventuelt kan verifiseres eller ikke, oppsummert i avsnittets avslutning: «Kan et menneske virkelig beskrives som noe så simpelt som en bit med land?». Innfallsvinkelen likner drøftingens og kan av andresensor ha blitt lest som et brudd med det mange forbinder med utprøvende refleksjon. Litteraturvitenskapelig leksikon kaller essayet en «drøftende sakprosa» (Lothe et al., 1997, s. 67), og uansett åpner innfallsvinkelen for en tekst som skal besvare et spørsmål avsender har fått av å lese tekstene til eksamensoppgaven. Dette anslaget følges opp i neste avsnitt, og med bakgrunn i geografifaget tilbakeviser avsender at mennesket kan sammenliknes med land. Noe er imidlertid likt: «Landstykket er der for en grunn, og er formet av historien sin, mye slik som menneskene». I hovedsak er det likevel landstykkets komplekse geografiske egenart som fremheves. Dette understrekes mot slutten av avsnittet når landstykket besjeles: «Som land ville jeg blitt fornærmet hvis jeg hadde blitt sammenlignet med noe så mykt og skjørt som et

menneske», før avsender vurderer dette gjennom en egen sammenlikning: «Blir litt som å sammenligne en gammel kampherdet løve med en døgnflue». Avsender bruker altså kunnskap fra andre skolefag til å reflektere over billedbruk i vedleggene, noe som kan tolkes som en kreativ innfallsvinkel, i tillegg til at egen billedbruk viser språklig kreativitet.

I tredje avsnitt endres spørsmålet til «Er land og broer virkelig den eneste måten å beskrive mennesker på?». Denne referansen til forrige avsnitt: «Mennesket er jo ikke et fast objekt som står stille mens tiden går», kan tyde på at fastlandmetaforen nå er forkastet. Dette baner veien for avsenders egen sammenlikning, nemlig at mennesker «er båter som flyter i et enormt hav». Dette utdypes gjennom flere kohyponomer til båt, når eksempelvis barn blir «joller», voksne «cabin cruisere eller seilbåter» og viktige personer «cruiseskip». Avsender utvider også båtsimilen til en allegorisk tankerekke (op.cit) ved å bruke referanser fra feltet, som «skjær», «går på grunn» og behovet for «fendere mellom seg slik at man ikke krasjer og skader noe». På denne måten reflekterer avsender videre rundt det å være menneske i verden.

I avsnitt fire og fem går avsender over til en mer direkte lesning av vedleggene. Først fellesskapet fra Donnes tekst, så Vesaas' bruer mellom øyene. I begge avsnittene finnes eksempler på god lesekompetanse og refleksjonsevne, som i starten av femte avsnitt: «Broer som da er laget av ord og følelser. Man går da ut ifra at alle mennesker er alene, og at man selv må stå for innsatsen når det gjelder kontakt med andre». Vedleggene leses her som motsetninger, men i sjette avsnitt åpnes det for en «gylden middelvei» mellom dem. Skjærgården foreslås som et bilde som kan forene fastland og øyer, og avsender forklarer dette ved å vise hvordan denne skjærgården kan arte seg i praksis. Dette viser vilje til refleksjon for å svare bredt på oppgaven. I åttende avsnitt forbindes skjærgårds- med båtallégorien fra tredje avsnitt. Dette sikrer en viss sammenheng, i tillegg til at noe har skjedd siden sist, fordi noe må forkastes: «Nå innser jeg at det ikke er plass til cruiseskip i en skjærgård, så la oss droppe den teorien nå». Igjen brukes termer fra andre oppgavetyper, her kanskje mer det vitenskapelige forsøket enn tidligere foreslåtte drøfting. Det kan både leses som formbrudd og et humoristisk virkemiddel. Nå forenes også bildene fra avsnitt tre og syv til en ny middelvei: «Skjærene er de båtene som har lagt seg til ro» og «yngre mennesker er hver sin skjærgårdsbåt» som «farer rundt». For å nyansere denne polariserte fremstillingen av generasjonene, er avsender bevisst ved å si at dette ikke betyr at de eldre er «blitt til stein, fossiler av det de var», noe som også peker tilbake til geografireferansene tidlig i teksten. Det er naturlig å se for seg at noe av uenigheten mellom sensorene angikk om nevnte momenter er relevante, noe som vil kunne ha forplantet seg i vurderingen av både kreativitet og refleksjon.

Kreativitet i innfallsvinkler og virkemidler: Avsender velger tidlig i teksten to innfallsvinkler som trenger aksept fra mottaker. Det ene kan leses av valget av den verbale prosessen «påstå» og andre ordvalg, som peker mot en tekst som skal drøfte en bokstavelig lesning av metaforer i vedleggene. Det andre er referansen til geografifaget og andre avsnitts lesning av nevnte metaforer på dette fagets premisser. Innfallsvinklene kan leses som kreative og originale måter å løse oppgaven på, men som nevnt kan de også leses til å rote til svaret. Bruken av egne similer, metaforer eller allegorier viser også kreativitet. Førstesensor har trolig ment at dette også er gjort på en gjennomført måte, som er kjennetegnet til høy måloppnåelse. Det å introdusere egne bilder som bidrar til andre perspektiv til tematikken i vedleggene, viser en selvstendig tilnærming til tekstene. Det samme kan sies om innfallsvinklene jeg refererte innledningsvis i tekstanalysen, i tillegg til bruken av besjeling og humor. Samtidig uttrykker andresensors forslag at for eksempel billedbruken er vurdert annerledes. Det kan skyldes at de noe omfattende båt- og skjærgårdsallegoriene ikke er tatt på alvor som tankerekker.

Reflekterer, prøver ut tanker: Under relevans var jeg inne på flere momenter som kunne leses som utprøvende refleksjon. Der det i starten kan se ut til at vedleggene skal drøftes ut fra geografifaget, kan av andresensor ha blitt vurdert som å reflektere på siden av tematikken i vedleggene, men førstesensor har nok vurdert det som selvstendig refleksjon. På tross av sin selvbevissthet, er det lite selvrefleksjon i teksten, men kanskje snarere det Grütters kalte en «dialogisk modellsender på jakt etter en faglig dialogpartner» (Grütters, 2011, s. 169). Der avsender lanserer egne bilder, er det kanskje ikke kunnskapsutviklende, men det bidrar til å prøve ut ulike tanker rundt hvordan de kan fungere som utdypinger av tematikken. Avsenders lesekompetanse vises ikke direkte før i fjerde avsnitt, og dette ankepunktet til strukturen kan også ha fått følger for vurderingen av refleksjonen, som jo skulle ta utgangspunkt i vedlegg. En motvillig leser vil kunne ønske seg at avsender tidlig i sin tekst klargjør sin tolkning, mens andre vil si at dette er pirk, all den tid dette til en viss grad berøres innledningsvis før det går i dybden på det senere. Den reflekterende teksten skal også kunne vandre i emnet sitt (Lothe et al., 1997, s. 69). Oppsummeringen i siste avsnitt viser modalitet og åpenhet ved å fremheve at det ene perspektivet ikke er mer riktig enn det andre. Den viser også lesekompetanse ved å få frem at vedleggene er forbundet til hverandre og avsenders refleksjon: «Enten man mener man er båt, øy, skjær eller fastland vil man forbli det mennesket man er, så det forandrer lite. Hvordan du enn ser på det, lever vi i et samfunn der vi alle er bundet til hverandre til en viss grad». På denne måten avsluttes teksten ved å si at vedleggene, som tidligere er lest både polarisert og ved hjelp av middeelveien, også sier det samme, men er ulike fordi de er skrevet

av ulike mennesker. Denne evnen til språklig nyansering og evaluering plasserer teksten både i det faglige og kritiske domenet (Tønnessen & Maagerø, 2015).

Avsender er ikke redd for å vise seg frem i teksten. Allerede i avsnitt to fortelles det om faglig fordypning i geografifaget. Denne måten å bruke sin egen bakgrunn aktivt i refleksjonen, gjør teksten metakommunikativ. Dette ses også tydelig avslutningsvis, i denne referansen til eksamenskonteksten: «Kun fantasien setter begrensninger på hvor mange sammenlikninger man kan finne, og tiden man har på å finne dem på naturligvis, slik som i mitt tilfelle». Metakommunikasjon er tidligere funnet risikabelt i eksamenstekster (Hertzberg & Andersen, 2005, s. 298), men dette kan også leses som en mottakerbevissthet en gjerne forventer i tekster som denne (Lothe et al., 1997, s. 70). Mottakerbevissthet finnes flere steder, for eksempel i inngangene til de fleste avsnittene, som her til avsnitt to og tre: «Misforstå meg ikke» og «Men det var faktisk et poeng jeg ville fram til». I KAL utdypes uenigheten i spriktekster som har dette trekket med at noen sensorer mener at det viser utprøving, mens andre vurderer det negativt (op.cit, s. 299).

Relevante svar på oppgavene, forståelse og bruk av de vedlagte tekstene: Tekstanalysen viser at langsvaret gjør det oppgaven har bedt om og viser god forståelse og bruk av de vedlagte tekstene. Det er lite sannsynlig at lesekompetansen har forårsaket sprikvurderingen.

Formuleringsevne, ordforråd og formelle ferdigheter:

I all hovedsak en korrekt og velformulert tekst med et nyansert ordforråd. Enkelte steder kan ordvalg varieres, for eksempel i langsvarets første avsnitt der ordet «spørsmål» kommer tre ganger i samme setning. En muntlig vending som «den dag i dag» passer ikke så godt der det er brukt (siste setning i femte avsnitt), men er ikke representativ for teksten som helhet. Det er også enkelte tunge konstruksjoner eller overforklaringer, valg av feil funksjonsord og svakheter i tegnsetting, men ikke nok til å si at formverket ikke mestres. Tendensen til teknisk, vitenskapelig stil, som bidrar til assosiasjoner til andre oppgavetyper, er kanskje vurdert av den ene sensoren som malplassert eller at det gir et stivt preg, mens den andre kanskje har foretrukket dette fremfor ordvalg som «tenker», «føler» og litt for mange retoriske spørsmål, som er en mer åpenbar måte å realisere refleksjon i tekst på.

Oppsummering av mulige årsaker til sprikvurdering (6 og 3) av S13:

- Analyse av kortsvaret viser kompetanse tilsvarende andresensors forslag. Det er derfor mulig at sensor 1 ikke har tillagt kortsvaret vekt.

- Språklig stil som gir assosiasjoner til drøftende eller vitenskapelig skriving. Uenigheten kan gå på om det er uhensiktsmessig eller selvstendig.
- Båt- og skjærgårdsallegorien kan vurderes som selvstendig og kreativ språkbruk, eller som useriøst eller mindre relevant. Fordi «Båtavsnittet» kommer før avsender skriver konkret om tematikken i vedleggene, kan det bli vurdert som mindre relevant.
- Metakommunikativ stil kan leses som dialogisk mottakerbevissthet eller som uønskete leserhenvendelser som kan gi et useriøst preg

5.3.9 Intensivtekstene oppsummert

Svar på kortsvaroppgaven: Det er usikkert om det er trekk ved kortsvarene som kan forklare sprikvurderingen, men ulik vektning av dem kan i noen tilfeller ha sammenheng. Kortsvarene til S10 og S13 viser såpass lite norskfaglig kompetanse at det er vanskelig å se for seg at sensorene med de høyeste karakterforslagene (5 og 6) kan ha tillagt dem vekt. Sluttkarakteren de har fått reflekterer at kortsvaret her til slutt er vektlagt, men det gjelder kanskje i større grad for til S10 (fikk karakteren 3) enn S13 (fikk 5), med mindre sensorene har vurdert langsvaret i S13 til karakteren 6. Motsatt kan en registrere at sensorene med de laveste karakterforslagene (3 og 4) til S6 og S19 kanskje heller ikke har vektlagt kortsvarene, for i disse besvarelsene viser analysen at kortsvarene holder høyt nivå. I begge disse tilfellene kan imidlertid sluttkarakteren tyde på at kortsvarene har fått betydning. I drøftingen kommer jeg tilbake til vektning av kortsvaret og dets betydning ved tvil rundt vurdering av langsvaret.

Svar på langsvaroppgave: De fleste analysene viser eksaminander som har lest, forstått og tatt utgangspunkt i vedlagte tekster, selv om noen av dem i større grad viser forståelsen indirekte gjennom refleksjon og fortelling (S6, S10, S19) og enkelte ikke får vist så mye lesekompetanse (S3, S6, S10). Fire av de seks tekstene har svart på alle delene av oppgaven: Alle unntatt S6 reflekterer i hele eller deler av teksten sin, og alle unntatt S3 viser kreativitet i innfallsvinkler eller språkbruk. Det som varierer mest når det kommer til svar på oppgaven, er altså i hvilken grad eksaminandene får vist lesekompetanse av vedleggene, og usikkerhet rundt dette har trolig sammenheng med sprikvurderingen av S3, S6, S10 og S19. Denne antakelsen styrkes av at oppgaven fremhever at vedleggene skal komme tydelig frem, noe som også er vektlagt i eksamensveiledningen og ble gjentatt på sensorskoleringen.

Det er også sannsynlig at strukturens formålstjenlighet er diskutert i flere av oppgavene. Oppgaven krever ikke stringent struktur, men det at det forekommer mer eller mindre bevisste formbrudd i fire av de seks tekstene, sannsynliggjør at dette likevel er blitt diskutert av

sensorene. S1 er teksten med det klareste formbruddet, blant annet fordi det er vist typografisk og er brukt så gjennomført at det fremstår som et valg av innfallsvinkel. Byttet av form eller skrivehandlinger fremstår ikke som så gjennomtenkt i S3 og S10, og jeg har foreslått at begge disse eksaminandene kunne profitert på klarere krav til struktur i oppgaven. S13 har en springende form som holder seg innenfor essayformatet, men som likevel kan ha blitt diskutert, uten at dette er sannsynlig årsak til sprik. De to som har valgt fortellingen som innfallsvinkel (S6 og S19) har trolig unngått ulik vurdering på grunn av strukturvalg, om en ser bort fra at innfallsvinkelen i seg selv kan ha blitt diskutert.

Språk og formelle ferdigheter: Analysen viser høy eller middels skrivekompetanse i de seks besvarelsene. Eksaminandene som har skrevet S3 og S10 viser noe mer usikkerhet til formverket enn de andre, for eksempel gjennom feil i tegnsetting eller særskriving.

Sprikvurderingen kan skyldes ulik vektning av formell skrivekompetanse, men det er trolig andre sider ved det språklige som har ført til større uenighet, og jeg har kun foreslått det som en mulig årsak når det gjelder S3. Det som er mer gjennomgående i tekstene, og som det kan ha vært uenighet rundt, er måten språket er brukt på eller evne til å skrive med nyansert ordforråd. Her tenker jeg på metakommunikasjon og ordvalg som gir en subjektiv, enkel, hverdagslig, triviell, polarisert eller høystemt språklig stil. Uenigheten kan være knyttet til språkbruks relevans, kreativitet eller selvstendighet, og det er sannsynlig at måten refleksjonen eller kreativiteten er vist språklig også har sammenheng med sprikvurderingene.

Dette gir at sprikvurderingen trolig henger sammen med ulik vektning av kortsvaret, tvil rundt i hvilken grad langoppgaven er besvart – eller uenighet rundt måten denne er besvart.

6 DRØFTING AV RESULTATER

6.1 SVAR PÅ SPØRSMÅL OM SAMMENHENG MELLOM OPPGAVER OG VURDERING

Det første forskningsspørsmålet er: Er det sammenheng mellom oppgavetyper og variasjon i vurdering?

6.1.1 Sammenfatning av dokumentanalyse og kvantitativ undersøkelse

Analysen av eksamensoppgaver og -kriterier viser at oppgave 2 er mer åpen enn de andre, noe forskning har vist kan gi usikkerhet rundt vurderingen (Solheim & Matre, 2014). Mens de andre oppgavene på ulike måter etterspør innhold som angår annen norskfaglig kunnskap enn lese- og skrivekompetanse, gjør ikke denne det. Fordi kreativ, men ikke reflekterende skriving, ble konkretisert i oppgavekommentaren, gir det uklarhet rundt formålet. Det er også

grunn til å tro at vurderingskriteriene ikke har fanget opp mangfoldet av egenskaper ved eksamenstekstene som svarer på denne oppgaven (Skar & Aasen, 2015). Sensorskoleringen avdekket stor variasjon i vurderingen av tre av fem eksempeltekster, i tillegg til at fellesskapet av sensorer var i utakt med direktoratets vurderinger av to av eksempeltekstene. I sensorrapporten kom det frem at sensorene stort sett er fornøyde med oppgaver og vurderingsveiledning. Flere uttrykker likevel at åpenheten i oppgave 2 gjorde den vanskelig å vurdere, i tillegg til at vurderingskriteriene ikke fanget opp mangfoldet, særskilt i de litterære tekstene. Sensorene uttrykte både på skolering og i rapporten at vurderingsarbeidet vil bli mer forutsigbart ved mer differensierte kriterier og vektning av kortsvaret.

Kvantitativ analyse viser at 14% av eksamensbesvarelsene i de tre undersøkte fylkene er sprikvurdert. Korrelasjon beregnet etter Pearsons r er 0.58, mens samsvarskoeffisienter ble 0.2. Undersøkelsen av oppgavevalg viser at den kreative oppgave 2 førte til signifikant flere sprikvurderinger enn de andre: 20% av eksaminandene som skrev denne er sprikvurdert. Her er også korrelasjonen lavere enn i de andre oppgavene: 0.46. Den informative og argumenterende oppgave 1 skiller seg motsatt ut ved å ha høyeste korrelasjon: 0.69. Oppgave 1 er den de færreste eksaminandene har valgt (11%), og den skiller seg også ut ved at en større andel besvarelser fikk karakterene 1, 5 og 6. Den er også overrepresentert på klagestatistikken. Det er også oppgave 2. Den utmerker seg også ved å ha flere besvarelser med lav måloppnåelse etter fellessensuren. Det er ikke funnet sammenheng mellom sprikvurdering og klager.

Foreløpig svar på forskningsspørsmål 1 er at det er en mulig sammenheng. Oppgaven som viser dette klart, er den kreative oppgave 2, og årsaken kan knyttes til oppgaveformulering og vurderingskriterier, noe som også kan støttes av tidligere referert forskning. Det er også funnet at oppgave 1 utpeker seg. Denne er klarere formulert og har blant annet faglig orienterte vurderingskriterier. Den viser også forholdsvis høy korrelasjon og lavere sprikvurdering. I tillegg er det funnet at oppgave 2 har signifikant flere besvarelser som fikk lav måloppnåelse etter fellessensuren, mens oppgave 1 har flere med høy måloppnåelse.

6.1.2 Drøfting av kvalitative funn i eksamensoppgaver og rammer for vurdering

Den kreative oppgaven skiller seg ut ved et åpnere skriveoppdrag enn de andre oppgavene. Det gjelder både innhold og struktur. Tidligere empiri har vist at det kan være risikabelt å overlate eleven i så stor grad til egne erfaringer og bakgrunn (Tønnessen & Maagerø, 2015) fremfor felles referanser fra arbeidet i klasserommet (Kvistad & Smemo, 2015), noe som forsterkes av oppgavens mangel på krav til struktur (Blikstad-Balas & Hertzberg, 2015).

Oppgavens åpenhet *i seg selv og for* det subjektive inviterer også eleven til skriving som tidligere forskning har vist at det er problematisk å vurdere (Nilsen, 2005).

Dette betyr ikke at eleven ikke *kunne* vise annen norskfaglig kompetanse enn lese- og skrivekompetanse i svaret på oppgave 2, men tekstanalysen viser at det stort sett ikke er gjort. Det er tidligere funnet at mange sensorer misliker oppgaver uten eksplisitt forventning til annen norskfaglig kunnskap (Bach, 2015). En kan spørre om det er litt enøyd å insistere på dette, for norskfaglig kunnskap og lese- og skrivekompetanse er ikke motsetninger, noe direktoratet også tok opp under sensorskoleringen 2015 (Utdanningsdirektoratet, 2015d). Sensorers skepsis til den kreative oppgaven kan tolkes på minst to måter. Noen vil si at den uttrykker et syn som vektlegger dannelsesstradisjonen i faget, en slags snerpete tradisjonalisme uten åpenhet for redskapsfaget. Men det kan også være uttrykk for at sensorer ser behovet for at alle har noe relevant å skrive om, fordi det bedrer elevenes muligheter til å skrive en god oppgave, noe forskning også har vist (Smemo og Solem, 2014). Det meste av kritikken som angår oppgave 2 i sensorrapporten fra 2015, går på at den er vanskelig å skrive og sensurere.

Analysen av oppgave 2 viser at denne også inneholder en uklarhet rundt prioritering av formålene reflekterende og kreativ tekst. I kapittel 3 viste jeg at endringen fra sjanger til skrivehandlinger eller oppgavetyper har vært kontroversiell, blant annet fordi det kan være utfordrende for mange elever å skulle velge sjanger selv (Blikstad-Balas & Hertzberg, 2015). Det kan fremstå paradoksalt at revidert læreplan fremhever eksplisitt sjangerundervisning, men toner ned sjanger i eksamensoppgavene ved avsluttet opplæring. Dette kan skyldes at læreplanen vektlegger sjangerarbeid i undervisningen, men når læringsutbyttet skal måles, står eksaminandene friere med tanke på hvor i sjangerrepertoaret de vil hente sin tekststruktur. På LNU-seminaret som tematiserte dette (Botheim, 2015), mente innleiderne at sjangerundervisning fortsatt er viktig. Samtidig ble det beskrevet en begrepsforvirring som sannsynligvis går ut over elevene. Når eleven skal nå formål gjennom skrivehandlinger, er de helt avhengige av å kjenne sjangrer å velge blant. I undervisningen har skriveleeren mulighet til å peke på ulike måter å realisere samme skrivehandling *og* sjanger på, og vise effekten av sjangerbrudd, men på eksamen er eleven prisgitt oppgaveformuleringen. En del av eksamenstekstene viser usikkerhet som kunne vært redusert ved å etablere en mer forutsigbar kontrakt mellom eksaminand og sensorer, gjennom angivelse av for eksempel sjanger.

Sensorrapporten i 2015 bekrefter at det fortsatt er knyttet særskilt usikkerhet til vurderingen av kreative tekster. Det samme viste Bachs analyse av samme rapport året før (Bach, 2015). Usikkerheten er ikke av ny dato, som vist i kapittel 3. Det historiske tilbakeblikket på

novellens kortvarige skjebne på starten av 1990-tallet, kan aktualisere spørsmålet om det lar seg gjøre å gi så klare kriterier at en sikrer begrepsvaliditet, som igjen kan gi delte og forutsigbare vurderingsnormer til oppgaver som i sin natur er åpne, fordi de skal føre til nettopp kreativ og selvstendig skriving. Det er på den ene siden viktig at *Fagnemnda i norsk* lager eksamensoppgaver som reflekterer læreplanen. På den andre er det påpekt at det kan være «noen potensielle utfordringer» ved skriving og oppgavelaging ut fra hjulets å *reflektere* (Dagsland:218). I tillegg er kreativ skriving en kanskje fremmed fugl i skrivekolen som prioriterer skrivekompetansen samfunnet trenger og skriving som grunnleggende ferdighet i alle fag. Den kreative oppgaven er kritisert fra flere hold fordi den løsriver det kreative fra annen skriving (Smidt 2015 og Blikstad-Balas, 2015b). Dette synes forsøkt løst ved at eksamensveiledningen i 2015 ble oppdatert ved å trekke frem kreativitet som noe sensor skal honorere i alle oppgavetyper (Utdanningsdirektoratet, 2015a). Her kan det være snakk om et rom for elevens stemme i teksten som er skrevet for å innfri samfunnets behov, eller koplingen mellom det private og offentlige (Miller, 2001). Dette er for eksempel mulig når en ber eksaminander reflektere rundt norskfaglig tematikk, som til eksamen i sidemål høsten 2015, der den kreative oppgaven ba om refleksjon rundt «forestillinger om det norske».

Analysen av eksamensveiledningen viser at alle oppgavens vurderingskriterier gir relativt stort rom for skjønnsmessig vurdering. Kjennetegn på måloppnåelse er organisert i tre nivå; lav, middels og høy. De skiller ikke karakterene 3 og 4 eller 5 og 6. Her kan en tenke seg at det er i hvilken *grad* besvarelsen møter kjennetegnene som kan tjene differensierende, men da overlates det til den enkelte sensor å skille, noe som kan svekke mulighet for enhetlig vurdering. Dette kan forsterkes når kriteriene oppleves uklare eller ikke rommer de uvanlige måtene å løse oppgaven på (Skar & Aasen, 2015). Behov for presisering og differensiering er også en tilbakemelding fra flere sensorer (Utdanningsdirektoratet, 2015e). Behovet for å skille mellom karakterene 3 og 4 kan synes ekstra viktig når hele 66,5% av kullet oppnår disse på eksamen i norsk hovedmål: 34,3% fikk 3 og 32,2% fikk 4 i 2015 (Utdanningsdirektoratet). Det gir veldig mange besvarelser, også for hver enkelt sensor, der en skal skille mellom 3 og 4 uten støtte i differensierte kriterier.

Utdanningsdirektoratet sier at reliabiliteten sikres når sensorene bruker vurderingsskjema og -veiledning (Utdanningsdirektoratet, 2015d, s. 9), og på ett punkt fremstår sensorene spesielt innstilt på å gjøre dette. Dette gjelder vektning av den obligatoriske kortsvaroppgaven. Den skal sikre at alle eksaminander, uansett valgt langsvaer, får vist fagspråk og -kunnskap, men det fordrer at alle sensorene faktisk vektlegger den, og helst på samme måte. Sensorene

etterlyser en formalisering av dette. I tillegg viser undersøkelsens tekstanalyser at kortsvarets innflytelse på helhetsvurderingen varierer. Behovet for vektning kan synes ekstra viktig i den kreative oppgaven, også fordi kvantitativ undersøkelse sannsynliggjør at denne svekker gyldighet ved å åpne for ulike bedømmermønstre.

Mer presise vurderingskriterier og eksplisitt vektning av kortsvaret kan altså gjøre det enklere for sensorene å differensiere, men her kan også oppgaveformuleringer bidra. Det viser funn av sammenheng mellom åpen eller uklar oppgaveformulering og sprikvurdering. Forskning har funnet at en kan utforme oppgaver innenfor rammer som gjør vurdering enklere, men det kan utfordre elevens «(...) mulighet til å vise en mer kompleks og mangfoldig skrivekompetanse» (Bach, 2015, s. 31), noe som er ett av formålene ved eksamen etter videregående. En konkret erfaring fra Normprosjektet var sammenheng mellom oppgaver og prestasjoner (Solheim & Matre, 2014). En studie av skriveoppgaver gitt på 3. og 4. trinn er ikke nødvendigvis overførbart til eksamen etter videregående opplæring. Eksamensoppgavene er her såpass omfattende at selv med visse skriverammer vil de åpne for individuelle valg. Innholdet i oppgaven vil, selv om det er pensumrelatert, heller ikke være av så spesifikk art i en eksamensoppgave. Dette fordi den baserer seg på ett eller flere tekstvedlegg, og et tekstsyn som åpner for elevens individuelle lesning og forståelse. Samtidig kan en se noen fellestrekk. Oppgaver om sentrale deler av pensum og med eksplisitte krav til struktur, vil gi mer forutsigbare svar og vurderinger. Oppgave 1 kan ha vært en slik oppgave. Den viser til fagstoff fra språkhistorien og inneholder en overordnet skriveramme. Undersøkelsen viser at oppgave 1 har høyest vurderingssamsvar og at hele karakterskalaen er godt representert. Normprosjektets undersøkelse viste at eksplisitte skriveoppdrag hever kvaliteten på de fleste elevsvarene, men at egenarten og muligheten til å briljere forsvinner sammen med muligheten for enhetlig vurdering (op.cit). Dette gjelder ikke for oppgave 1, som nettopp utmerker seg ved å ha størst andel eksamenstekster med høy måloppnåelse, samtidig som den mest åpne oppgaven i settet, oppgave 2, utmerker seg ved å ha størst andel med lav måloppnåelse. Dette taler for at kompleksiteten i eksamensoppgavene motvirker tendensen til at styrte oppgaver kan være til hinder for at de beste skriverne får vist kompetanse.

Når det kommer til ulik bedømming, viser undersøkelsen at tolkningsfellesskapet ikke var utpreget på sensorskoleringen våren 2015. Dette kan bekrefte et behov for videre jobbing med normfellesskapet, og ut fra hvordan en arbeider systematisk med dette i skriveprøvene, henger normfellesskap tett sammen med kvalitetssikret dokumentasjon av kompetanse innenfor delferdigheter (Berge, 2005b) eller at eksplisitte forventningsnormer blir operasjonalisert

(Solheim & Matre, 2014). Bildet fra sensorskoleringen ble komplisert av at direktoratets karakterer til to av tekstene lå henholdsvis en karakter høyere og lavere enn gjennomsnittet i sensorkorpset. Dette kan være et uttrykk for at de fleste sensorene i noen tilfeller tolker oppgaver og kriterier annerledes enn myndighetene, men også at forskjellene er tilfeldige og forårsaket av generell usikkerhet til anvendelse av vurderingskriteriene. I sensorrapporten tok enkelte til orde for at de ønsket eksempler på andre måter å løse oppgavene på. Det er grenser for hvor mange tekster sensorene kan vurdere på én dags skolering, men det er også mulig å gjøre eksempeltekster tilgjengelig som del av vurderingsveiledningen (Skar & Aasen, 2015).

6.1.3 Drøfting av kvantitative funn

Utdanningsdirektoratet definerer eksamens reliabilitet som tilstede når sensorene vurderer svarene på samme måte (Utdanningsdirektoratet, 2015d). Intervurderingsresultatene i undersøkelsen kan gi usikkerhet til at alle besvarelser vurderes under likeverdige forhold og om resultatene er til å stole på. Så store variasjoner kan svekke eksamens reliabilitet, men kan også nyanseres ved å se på erfaringer fra tidligere forskning (kap.2.3) som viser at forhandlingene under fellessensuren bedrer reliabiliteten.

På spørsmål om hvor mange vurderere en besvarelse må ha for å oppnå *true score*, sier teori rundt normfellesskap at det er vanskelig å generalisere et svar, fordi noen besvarelser trenger flere vurdere enn andre (Borgström & Ledin, 2014, s. 8). Den norske modellen med oppmann på sensormøtet er i teorien sensur med tre vurderere, som er det som trengs for de fleste besvarelser (ibid.), men vaksinerer kanskje ikke tilstrekkelig når sensorene selv avgjør behov for å trekke vedkommende inn. En økt styring kan potensielt dempe tvilen. Samtidig vil ikke tre vurderere sikre gyldig vurdering av det mindretallet av tekster som er ekstra utfordrende å vurdere (ibid.). Dette kan gjøre det nødvendig å styrke vurderingsarbeidet på andre måter, for eksempel ved å utvikle veiledninger (Skar & Aasen, 2015). I tillegg viser undersøkelsen at spriketekstene ikke er overrepresentert på klagesensuren. Det indikerer at de ikke i større grad har fått dårligere karakter enn kandidaten har ment rimelig, men funn av forholdsvis lav konsistens og lav konsensus åpner mulighet for at vurderingen fortsatt kan ha vært vilkårlig.

En kan også avdramatisere manglende samsvar og sprikvurdering, all den tid medsensorene blir enige i sensormøtet, og forutsatt at karakteren de blir enige om uttrykker *true score*.

Borgström og Ledin skriver at det er umenneskelig å kreve fullstendig samsvar, samtidig som de kaller det et sort hull når man via Berges avhandling fra 1996 vil ta seg fra doxa til det faktiske vurderingssamsvaret som var på 0.56 (Borgström & Ledin, 2014, s. 146). Her kan det være nødvendig å understreke behovet for systematisk arbeid med normfellesskapet, men

også å nyansere Berges undersøkelse. Om en ser nærmere på sensorenes forutsigbarhet, og resultatene til gruppene på både tre og fem sensorer som Berge fikk til å vurdere et intensivutvalg, og selve sensuren etter forhandling, var korrelasjonen for 5-sensorsnittet 0.89 for Pearsons r. For 3-sensorsnittet var det 0.78 (Berge, 1994, s. 255). Det er noe ulikt hva som betraktes som forutsigbar vurdering. Borgström og Ledin, og Barrett, skriver 0.7. Berge skriver 0.8, og kommenterer også at funnene over «innebærer at de to sensorenes forhandlingsarbeide fører til at sensurens pålitelighet blir bedre enn resultatene i kapittel 3 kunne tyde på» (ibid.). Dette underslår ikke at Berge fortsatt finner store uenigheter rundt en del besvarelser, men kombinert med at snittet også «heves betraktelig» (op.cit, s. 256) på fellessensuren, kan det være grunn til å foreslå at manglende samsvar i seg selv ikke behøver å svekke eksaminandens mulighet for å få en karakter som uttrykker besvarelsens *true score*.

Kvantitativ analyse av prosentvis enighet kan også tolkes på to ulike måter. Her tenker jeg på funnet av 39,4% enighet i bruttoutvalget, men at Stemler sier at om skalaen går fra 1-7, kan det betraktes som konsensus selv om det er et skille på én verdi i de to vurderingene av samme tekst (Stemler, 2004, s. 4). Det gir sammenfall på 86% i bruttoutvalget. Imidlertid skriver Stemler også at en ved vurdering av pålitelighet bør kombinere beskrivelser av konsistens og konsensus (op.cit, s. 1), og når det er gjort, må denne tolkningen tillegges mindre vekt. I tillegg får funn av sprikvurdering, kombinert med relativt lav konsistens og lav konsensus, støtte av tidligere svake pålitelighetsfunn i forskning rundt skrivevurdering. I forbindelse med arbeidet med de nasjonale skriveprøvene ble det forklart med mangelfullt tolkningsfellesskap (Evensen 2014, s. 250). Dette kan også gjelde i videregående skole, der mange ikke har lærerutdanning utover PPU, og ikke nødvendigvis fagdidaktikk i alle undervisningsfagene sine. Dette kan også tale for at det er nødvendig å utvikle veiledninger slik at de i større grad tar høyde for og viser frem eksempler på trekk eller hele besvarelser som det er vanskelig å vurdere. Det er en mulighet for at dette behovet kan ha blitt større etter overgangen fra sjanger til oppgavetyper. Funn i sjangerløse og sprikvurderte elevtekster fra utvalgsprøvene viser at de fleste velger samme sjanger, men at noen ikke gjør det, og at vurderer da mangler redskap til å vurdere akkurat de tekstene (Skar & Aasen, 2015, s. 5).

Til nå har jeg drøftet lav korrelasjon i bruttoutvalget, men det er én oppgave som utfordrer normfellesskapet mer enn de andre. Analysen viser at oppgave 2 trakk andelen sprikvurderte besvarelser opp, mens de tre andre oppgavene ligger signifikant under snittet på 14%. Utrengningen av korrelasjon og oppgavetyper viser at oppgave 1 og 2 skiller seg ut med merkbart høyere (0.69) og lavere (0.46) samsvar enn 3 og 4. Det er vist at dette kan knyttes til

at de to oppgavene har ulik utforming. Tidligere empiri fra norskeksamen etter videregående (Berge, 1994) har også funnet pålitelighetsproblemer ved ulikt bedømmermønster fra oppgave til oppgave, og Berge kalte de ulike oppgavene kvalitativt ulike prøver. Endringen av eksamensoppgavene i 2014 gjorde kortsvaret obligatorisk. Det kan bidra til å svekke denne konklusjonen, men når samme tendens ses i undersøkelsen, kan det skyldes at kortsvaret ikke vektet på samme måte. Dette siste kan bety at oppgavens potensiale ikke er fullt ut realisert. Jeg tenker her på formålet den ved å dekke flere skrivemåter har til å gi: «(...) eleven mulighet til å vise mer av fagkompetansen sin til eksamen» (Utdanningsdirektoratet, 2013).

De to omtalte oppgavene er også overrepresentert på klagestatistikken. Oppgave 1 og 2 er valgt av 24,8% av eksaminandene, men utgjør 42,9% av klagen. Dette kan ha sammenheng med utforming, men analysen viser her forskjeller mellom dem. En supplerende forklaring finnes i tilbakemeldingene fra sensorene. Her er det pekt på at oppgave 2 er vanskelig å vurdere på grunn av sin åpenhet. Oppgave 1 er vanskelig for mange elever fordi Øverlands tekst krevde høy lesekompetanse. Emnet åpner også for ukritisk kildeparafrasering, noe vurderingsveiledningen sier «vil trekke ned» vurderingen. Dette forklarer sannsynligvis hvorfor forholdsvis mange (7,5%) eksaminander som valgte oppgave 1 strøk. Fordi oppgave 3 og 4 er uendrete oppgavetyper, mens 1 og 2 er reviderte oppgavetyper, kan det ha vært større usikkerhet i normfellesskapet her. Denne forklaringen bør ikke tillegges vekt, for oppgave 1 viser høyest korrelasjon. Sannsynlig forklaring av overrepresentasjonen er derfor ulik for de to oppgavene. Oppgave 1 fordi den på grunn av sin vanskegrad også er overrepresentert på strykstatistikken, oppgave 2 på grunn av usikkerheten den har ført til i sensorkorpset, som allerede rommer en del ambivalens til kreativ skriving.

Funn av manglende sammenheng mellom sprik og klager, gir mulighet til å foreslå at sprikvurderinger ikke øker risiko for klager. Om en antar at klager er et uttrykk for at eksaminanden mener at det ikke er gitt riktig karakter, kan en si at sprikvurdering ikke nødvendigvis gjør at eksaminandens mulighet for *true score* avtar. Dette er i så fall et positivt funn. Det åpner for å foreslå at sensorene som har sprikvurdert stort sett blir enige om *true score* på fellessensuren. Denne forhandlingsviljen er bekreftet av referert empiri fra KAL og Berge (1994). Imidlertid ville dette vært et sikrere funn om konsistens og konsensus var høyere, som er reliabilitetsmål på hvordan normfellesskapet fungerer. Det er også lite trolig at eksaminanden klager på for høy karakter, så selv om hun er fornøyd, kan karaktersetningen ha vært vilkårlig. I tillegg er det få eksaminander som klager (6,21% i undersøkte krets), men det betyr ikke nødvendigvis at det er få som mener karakteren de har fått uttrykker usann

vurdering. I 2014 ble resultatet uendret for 78,7% av klagerne, samtidig som enkelte ble satt ned (Utdanningsdirektoratet, 2015d). Det skal mye til for å få medhold, og det kan derfor være grunn til å tro at mange antar at det ikke er verdt bryet. Ikke dermed sagt at tallenes tale ikke skal tillegges vekt, for basert på beregningene er det altså ikke grunn til å tro at de sprikvurderte oppgavene vurderes mer urimelig enn de andre.

6.2 SVAR PÅ SPØRSMÅL OM SAMMENHENG MELLOM TEKSTER OG VURDERING

Det andre forskningsspørsmålet er: Er det trekk som er utpreget i eksamenstekster som har utfordret normfellesskapet spesielt?

6.2.1 Sammenfatning av tekstanalysene

Intensivanalysen avdekker at kortsvaret er avgjørende. Kortsvar kan trekke helhetsvurderingen opp eller ned, men noen ignoreres. Det finnes kortsvar med så lite faglighet at de lite trolig er tillagt vekt av sensorene som har foreslått de beste karakterene. Det finnes også faglig gode kortsvar som trolig er oversett av sensorene med de laveste karakterforslagene. Ulik vektning har derfor trolig sammenheng med sprikvurderingen.

Analysen av ekstensivutvalget viser at oppgave 2 gir et mangfoldig utvalg av elevtekster. Skrivehandlingen *reflektere* er realisert gjennom både sakpreget og litterær skriving. I flere tekster er refleksjon vist indirekte gjennom andre skrivehandlinger, eller det er heller vanskelig å avdekke refleksjon. Indirekte refleksjon er mest påtakelig i fiksjonsfortellingene, mens de mer virkelighetsnære jeg-fortellingene også kan ha problemer med å holde seg til tematikken. Det er funnet tre måter å utfordre oppgavekontrakten på: Ved ikke å reflektere, ikke ta utgangspunkt i vedlagte tekster eller ikke vise kreativitet. Etter å ha tolket skrivehandlingen i vid forstand og tatt høyde for usikkerhet rundt flere tekster, er det fortsatt ti av 22 spriktekster som bryter med deler av oppgavekontrakten. Tre av disse har, på tross av det uklare forholdet til oppgaven, fått gode karakterer til slutt.

Analysen av intensivutvalget viser at sprikvurdering trolig kan knyttes til ovennevnte uklare svar på langvaroppgaven. Det er også funnet trekk som kan knyttes til at teksten skulle reflektere, og at uenighet dermed kan handle om *måten* denne delen av oppgaven er løst på. Dette angår metakommunikasjon, subjektiv skriving eller selvrefleksjon, og fordi dette er trekk ved skrivehandlingen (Aase 2012, Solheim & Matre, 2014, Skjelbred 2014), kan uenighet uttrykke usikkerhet til vurdering av denne. Det er også funnet språklige eller stilistiske trekk i tekstene, som muntlig, enkelt, høystemt eller trivielt språk, i tillegg til strukturvalg som avviker fra mer logisk oppbygging. Dette er også følger av skrivehandlingen

(Lothe et al., 1997, Tønnesson 2012), eller realisering av kreative innfallsvinkler eller språkbruk, så også her kan sprikvurdering skyldes motstridende normer til måten det er gjort på, eller ambivalens til oppgavetypen. Noen eksaminander viser også usikkerhet i møte med den åpne oppgaven. Det kan ses gjennom mangel på hensiktsmessig struktur, i tillegg til språklige og innholdsmessige valg, som har gjort tekstene mer dagligdagse enn spesialiserte. Disse eksaminandene kunne profitert på mer konkrete skriverammer.

6.2.2 Drøfting av kvalitative funn i eksamenstekster

I drøftingen av oppgaver og kriterier ble enhetlig vektning av kortsvaret foreslått som svar på reliabilitetsutfordringene knyttet til samsvar og sprikvurderinger. Det er tidligere referert til at eksamen skal måle norskfaglig kompetanse, og at sensorene utfordrer (Bach, 2015) ulike oppgavers kvalitet når det gjelder å invitere eksaminander til å vise slik kompetanse. Kreative tekster er her ekstra hyppig utfordret. Om en ser for seg at de to sensorene i sensormøtet fortsatt er uenige om hvordan språk, struktur og innhold i et kreativt langsvare uttrykker «annen norskfaglig kompetanse», kan de ty til kortsvaret for å vurdere dette. Et annet argument for vektning er årsaken til at kortsvaret ble gjort obligatorisk i 2014. Da ble det blant annet begrunnet med at eksamen skal dekke en større «(...) bredde av teksttyper og skrivemåter» (Utdanningsdirektoratet, 2013, s. 2). Dette kan også ses i sammenheng med tidligere forskning som har konkludert med at skrivekompetanse ikke kan generaliseres på bakgrunn av én type skriving (Berge 2005b, s. 16).

Trekk i langsvarene som er funnet som mulige årsaker til sprikvurdering, faller sammen med trekk i utvalgets klagetekster. Dette betyr mulighet for at trekk som i sensorforhandlingene langt på vei aksepteres av sensorene som har vært uenige, kan avslås av vel så mange sensorer som aksepterer dem. Det kan bidra til vilkårlig vurdering. Dette kan bekreftes av oppgavens overrepresentasjon ved klagesensuren, og jeg gir noen eksempler i drøftingen av enkelttrekk.

Det er funnet tre ulike risikotrekk knyttet til svar på oppgaven. I KAL-prosjektet ble dette kalt «gambling med kontrakten» (Hertzberg & Andersen, 2005). Er det slik at dette har sammenheng med sprikvurderingen av tekstene, betyr det at det i normfellesskapet er mye uenighet rundt hvordan slik gambling skal vurderes. Det er overraskende at det ikke er mer enighet om vurdering av eksamenstekster som ikke følger kontrakten, men dette kan kanskje forklares av kontraktens relative åpenhet. Analysen av vurderingsveiledningen viser at sensorene skulle honorere selvstendige valg, men kun «når svaret er relevant», og at det samtidig var «viktig å bruke de vedlagte tekstene og svare på det oppgaven ber om» (Utdanningsdirektoratet, 2015a, s. 4). Oppgaven ber blant annet om skrivehandlingen

reflektere. Endrete oppgaver skulle som kjent bidra til en dreining mot skrivingens formål. Da er det utfordrende for vurdererne når beskjeden som ble gitt på sensorskoleringen var at det også er greit å gjøre noe annet: «reflektere ved å vise fram ulike sider ved tematikken» (Utdanningsdirektoratet, 2015d, s. 23). Når skrivehandlingen kan realiseres indirekte, er det vanskelig å avgjøre at noe ikke er resultat av refleksjon. Det relativiserer oppgavekontrakten og kan forklare sensorenes usikkerhet. Det har også gjort analysearbeidet vanskeligere.

Karakterene som sprikte tekstene fikk til slutt, kan tyde på at sensorers usikkerhet er blitt dempet etter fellessensuren. Dette fordi det stort sett er sammenheng mellom endelige karakterer og det at oppgaven til dels ikke er besvart. Dette kan imidlertid diskuteres i møte med tre godt vurderte besvarelser, alle fiksjonsfortellinger. En av dem var intensivteksten S6, og analysen viste at det trolig var kreativitet og selvstendighet som ga karakter 4. Det gjelder trolig de to andre også; «Hjertesmeden» (S4) og «Å bygge en bro» (S17). Fordi alle disse tre tekstene er fiksjonsfortellinger, kan det være grunn til å tro at det er noe med denne måten å skrive på som gjør at noen sensorer kan glemme dette med relevans.

Dette går over i forholdet mellom det reflekterende og kreative. Det er sensorer som har kommentert at det er noe motstridende ved en oppgave som ber om refleksjon og kreativ skriving. Det er kanskje et paradoksalt funn at det er tekstene som realiserer skrivehandlingen direkte som sliter med å være kreative. Analysen av oppgaveformuleringen viser at denne prioriterer refleksjon fremfor det kreative ved å presisere at vedleggstematikken refleksjonen skulle ta utgangspunkt i skulle være tydelig i svaret. At sensorene likevel ser ut til å verdsette det kreative fremfor det reflekterende, viser både ambivalens og usikkerhet til vurderingen av disse tekstene. Det kan også henge sammen med at oppgavekommentaren utdyper hvordan det kreative kan realiseres, men ikke hvordan skrivehandlingen kan det, noe som av både eksaminand og sensor kan leses som en oppveining av nettopp den kreative delen av oppgaven. Dette kan også supplere forklaringen av de godt vurderte fiksjonstekstene over.

I de to andre kategoriene, fortellingene og blandingstekstene, fremstår ikke det kreative som like vanskelig å få til, kanskje nettopp på grunn av utradisjonelle innfallsvinkler som til dels fører til indirekte refleksjon eller andre utgangspunkt enn oppgavens. Dette gir spørsmål til både sensur og oppgaveformulering. Det kan være svakheter ved en oppgaveformulering som kan se ut til å tvinge eleven til risikable valg for å realisere oppgaven fullt ut, og det kan være svakheter ved vurderingen når aksepten av slike valg varierer såpass mye. Fordi vurderingen av de kreative tekstene i mindre grad korrelerer, er det vanskelig å se for seg at eksaminander er på trygg grunn når de skal velge sine innfallsvinkler. For hva er en kreativ innfallsvinkel?

Kan den være gjennomført kreativ uten å bryte forventning til form eller innhold? Her kan intensivteksten S1 tjene som eksempel. Her ble et essay rammet inn av dagboksfortellende jeg-sekvenser. Det kan forvirre en leser med tanke på hvilken kontrakt som ligger til grunn for vurderingen, men er både gjennomført kreativt og gir dybde til refleksjonen. Den ene sensoren vurderte teksten til 6, og på fellessensuren ble også den andre sensoren med på det. I klageutvalget er en tekst (K6) med tilsvarende innfallsvinkel gitt karakteren 3 av begge sensorer. Det er mye som tilsier at K6 og S1 skal vurderes ulikt, men *om* grunnen til at K6 har fått såpass streng vurdering at den måtte klagebehandles, er at de to sensorene var enige om at innfallsvinkelen er lite formålstjenlig, kan det illustrere nevnte risiko oppgaveformuleringen kan legge opp til. Det er blitt foreslått at en fordel ved å fjerne sjangerinstruksene var å motvirke formalisme, men formalisme kan knyttes til mer enn sjanger.

Det er viktig å unngå formalistisk vurdering av skrivehandlinger, noe modellen som langt på vei ligger til grunn for de nye oppgaveformuleringene også kan åpne for. Dagsland skriver at Skrivehjulet er en idealtypisk modell, og i det ligger at «(...) skrivehandlingene ikke skal kunne finnes i «ren form» i tekstene man vurderer eller forsker på» (Dagsland, 2015, s. 219). Dette fører til at «(...) beskrivelser av eller kjennetegn på ulike skrivehandlinger ikke bør behandles formalistisk» (ibid.). På samme måte som sjangrenes kjennetegn er relativt stabile, men også endrer seg fordi konvensjoner i samfunnet gjør det (Skjelbred, 2014:39), er det grunn til å tro at tilsvarende dynamikk finnes i forventninger til skrivehandlinger, samtidig som de ikke vaksinerer mot formalisme. Det er likevel viktig at oppgavene inneholder visse rammer for teksten som skal skrives og vurderes. Mens et vellykket sjangerbrudd kan vise nettopp god skrivekompetanse, og derfor ikke skal trekke ned fordi en sensor ikke kan krysse av alle punktene i en gitt sjekklister, kan det være uheldig at samme sensor blir usikker på om det skal trekke ned at skrivehandlingen som erstatter sjangeren ikke vises i teksten. Igjen kan det tenkes at resepten mot slik mulig relativisering kan være nettopp sjanger, noe refererte respons på endringene kom inn på (Blikstad-Balas & Hertzberg, 2015). Skrivehandlinger er også et nivå over sjanger, og de står derfor ikke i motsetning til hverandre (Aasen, 2015).

Analysen viser bruk av *fortelle* i halvparten av tekstene, noe som trolig gir sammenheng med sprikvurdering. Sprikende vurdering kan knyttes til hva slags tekst de to sensorene har forventet, sakpreget eller litterær, der sistnevntes kontrakt med leseren innebærer at denne er med på det fiktive universet og da heller ikke forventer en eksplisitt tekst (Skar & Aasen, 2015, s. 3). Her kan de to ha hatt ulike innstillinger i utgangspunktet, eller de kan ha ment ulikt om hvordan dette fungerer etter formålet. I analysen av den fortellende S6 refererte jeg

fra et møte med *Fagnettverket for norsklærere i Buskerud* 26.11.2015, der 20 kolleger bidro med forklaringer av sprikvurderingen av denne. Dette førte til diskusjon rundt den kreative oppgaven, og det kom forslag til hvorfor det kan se ut til at denne blir hyppigere sprikvurdert. Flere pekte på at kreative tekster kan appellere svært ulikt til ulike mottakere. «Bare se på litteraturen», som en sa. I dette ligger en risiko for at den som vurderer tillegger enkeltstående trekk i teksten for mye vekt, noe som kan slå både positivt og negativt ut. Noen stilte også spørsmål til om det kan være slik at «vi ikke vil at elevene skal være forfattere», fordi vi mener at det er på siden av samfunnsoppdraget. Det var enighet om at det i vurdering av slike tekster er viktig å vektlegge at oppgavetyper ikke skal stå i veien for teksten og eleven, at egne preferanser må legges til side når helhetsvurderinger skal gis, og ikke minst det «alle» vet, også fordi det presiseres årlig på sensorskoleringene, at tvil skal komme eleven til gode. Noen pekte også på at når langsvaret på grunn av nevnte tvil er så vanskelig å vurdere, bør kanskje kortsvaret vektles tyngre. Kortsvaret skal nettopp sikre at eksaminanden får vist fagspråk og -kompetanse, og kan på den måten bidra til å redusere tvil.

Subjektiv skriving og metakommunikasjon er utpreget i en rekke sprikteksster. I *Skrivehjulet* uttrykkes formålet med skrivehandlingen *reflektere* slik: «Identitetsdanning og selvrefleksjon, metakommunikasjon og egenvurdering» (Skrivesenteret, 2013). Noe av det som særpreger reflekterende skriving, er at den er subjektiv og selv fremstillende (Aase, 2012). I 3.3.1 var jeg inne på at det kan realiseres på monologisk og dialogisk måte, og at det dialogiske er å foretrekke (Grüters, 2011). Dette ble også konkretisert i analysene av de to blandingstekstene. Metakommunikasjon er en vanlig måte å realisere det dialogiske på, også det tett forbundet med skrivehandlingen (Solheim & Matre, 2014), og enda et valg som fremstår som risikabelt.

Ekspressivt språk og personlig form og innhold er altså naturlig i svar på en oppgave som ber om refleksjon. Litt polarisert kan en spørre seg om en såkalt objektiv refleksjon tilhører andre skrivehandlinger eller oppgavetyper? Samtidig har noen besvarelser en ekspressiv form fulgt av mer privat enn personlig innhold, noe som kan ta refleksjonen over i bekjennelser fra dagliglivets domene. I kapittel 2 refererte jeg problematisering av at skolen legger opp til skriving som prioriterer elevens prosjekt og nedprioriterer skolsk sekundærdiskurs. Tidligere forskning har også funnet subjektiv skriving risikabelt (Hertzberg & Andersen, 2005). I intensivutvalget er det S10 som viser dette tydeligst, men det er flere slike personlige og såre refleksjoner rundt vanskelige relasjoner i ekstensivutvalget, for eksempel i de lavt vurderte klagetekstene K9 og K10. Det er flere andre kvaliteter i disse tre tekstene som tilsier at sensorer i videregående i mindre grad oppvurderer slik skriving.

I omtalte møte med *Fagnettverket for norsklærere i Buskerud* benyttet jeg også anledningen til å drøfte subjektiv skriving. Deltakerne bekreftet at det er problematisk å vurdere subjektive og ekspressive tekster, og det på tross av at kreative og reflekterende oppgaver inviterer til det. Flere pekte på at erkjennelsen av at «vi ikke vil ha det», samtidig som det er et så fremstående trekk ved kulturkontekstens samtidstekster. Det gjør at «vi må ta inn over oss at dette må være greit», som en sa. Dette kan uttrykke motstridende meninger som kan forklare sprikvurdering. Uansett er det nok flere elever i videregående skole som leser for eksempel blogger fra dagliglivets domene, kanskje nettopp fordi de er personlige, enn aviser eller kanoniske bøker, og en såpass åpen oppgave som denne, overlater på mange måter eksaminandene til deres litterære forbilder eller summen av det de har lest. Vår tid er av litteraturprofessor Eirik Vassenden beskrevet som en der grensene mellom offentlige og private sfærer kollapser, og at tekstene er preget av senmoderne identitetsproblematikk, pseudodokumentarisme og selviakttakelse (Vassenden, 2007, s. 358). Derfor kan det tenkes at en i normfellesskapet bør ha et bevisst forhold til samtidstekster, også de som blir lest av mange. Sensuren av tekster til oppgaver som åpner for det subjektive, bør også vise åpenhet til ethvert forsøk på svar, uansett om en skulle ønske seg at elevene fikk oppgaver med klarere rammer og føringer til form og innhold.

Det er funnet en del stilistiske eller språklige trekk som er foreslått som forklaring av sprikvurdering. I de litterære tekstene er det funnet der språket brukes for å gi farge til en fortelling. Eksempler er antikvert stil når settingen er fortidig, fremheving og gjentakelse av statusmarkører når hovedpersonene prioriterer materiell vellykkethet eller naivistisk skrivemåte der hovedpersonen er et barn. I de reflekterende tekstene kan det også se ut til at språklige valg vurderes ulikt, for eksempel fremtredende billedbruk eller ordvalg forbundet med andre oppgavetyper. Dette er funn som sannsynligvis er påvirket av tidligere drøftet fortolkning. Usikkerheten kan reduseres av tilsvarende og refererte funn i tidligere forskning (Hertzberg & Andersen, 2005, s. 276). Smak og behag, eller skjønnsmessig uenighet, kan bli ekstra stor i tekster som inviterer til kreativ språkbruk og innfallsvinkler, samtidig som det er lite trolig at det er mulig å presisere kriteriene så mye at smak og behag forsvinner. I tillegg kommer at en oppgave som åpner for det litterære både utfordrer hvilken lese- og skrivemåte tekstene skal møtes med (Tønnesson, 2012, s. 18). Den utfordrer også forholdet mellom vurdering og kritikk, eller det mellom praktisk og kunstnerisk aktivitet, slik at forholdet til mottaker kan gruskes til av praktiske hensyn (Mukařovský, 2003, s. 26). Manglende samsvar tyder på at mottakerne her er langt fra omforente. I kritikk er det gjennom tidene operert med ulike

kvalitetskriterier, der ett er graden av enighet mellom kritikere (Hagen, 2004, s. 28). Dette er trolig lite anvendbart i vurdering av eksamenstekster med skrivestil som ser ut til å polarisere vurderingen: Enkelte nedvurderer, mens andre oppvurderer den, og igjen kan en mangel på omforent holdning i normfellesskapet, og lavt vurderingssamsvar, gi risiko for at to hauker eller duer (Borgström & Ledin, 2014) møtes i samme sensorpar.

Oppgavens oppfordring til kreative innfallsvinkler kan altså tvinge eksaminanden til å ta risikable valg. Det kreative kan også realiseres gjennom språklige valg, og resultatet av tekstanalysene viser at også dette innebærer risiko. I tillegg kan det altså se ut til at oppgaven ber om en risikabel skrivehandling. Det kan derfor bety at eksaminander som holder seg til kontrakten også må gamble, fordi det ser ut til å være så ulike vurderinger av trekk det kan være nødvendig å ha med. Hva ville for eksempel S1 vært uten sin todelte struktur og S10 uten sine såre barndomsminner? Oppsummert gir dette et normfellesskap som fremstår lite samlet, både når eksaminanden lar være å gjøre det oppgaven sier, og når han faktisk gjør det.

6.3 AVSLUTTENDE REFLEKSJONER

Hovedfunnene er oppsummert i kapitlene 5.1.3, 5.2.7, 5.3.4, 5.3.9, 6.1.1 og 6.2.1. Jeg vil anføre noen avsluttende refleksjoner rundt implikasjoner undersøkelsen kan ha for norskfaglig undervisning og forskning.

Det er lav konsistens og konsensus i vurderingene som er gitt i bruttoutvalget. Med unntak av den informative og resonnerende oppgaven, gjelder dette alle oppgavene. Undersøkelsen har i hovedsak utforsket sammenhenger knyttet til den kreative oppgavetypen, fordi den hadde lavest konsistens og høyest sprikvurdering. Det skyldes at motstridende vurderingsnormer her kan knyttes til både åpen oppgaveformulering og normfellesskapets ambivalens til kreativ og reflekterende skriving. Derfor kan det være behov for mer kunnskap om felles normer og forventninger til tekstkvalitet i slike tekster. Det fins kanskje andre forklaringer også, og derfor kan det være nyttig å studere et større tekstutvalg, alternativt undersøke forklaringer til at også tolkninger og analyser har lav korrelasjon. Her kan en innfallsvinkel, eller en studie i seg selv, være oppgaveformuleringer og vurderingsveiledninger, for eksempel ved å undersøke om vurderingen blir mer forutsigbar av veiledninger med modelltekster eller mer differensierte kriterier. Videre vil det være interessant å utforske den eneste oppgavetypen som ser ut til å gi pålitelig vurdering, i tillegg til flere godt vurderte tekster enn de andre

oppgavene. Det er eksempler på at kortsvaroppgaven ikke er tillagt vekt. Til dette kan det foreslås et prøveprosjekt der denne oppgaven skal vektes på en bestemt måte.

Didaktiske implikasjoner vil være skriveundervisning som tydeliggjør formålet med skrivingen, og på bakgrunn av dette jobber med hvordan ulike og konkrete sjangrer kan realisere skrivehandlinger. I tillegg kan det være viktig å jobbe spesielt med hvordan en kan løse åpne oppgaveformuleringer. Hva er hensiktsmessig form og innhold i oppgaver der elevene overlates til egne erfaringer? Det kan være fruktbart å utfordre elevene på hvor personlig en bør være i skriftlige skole- og eksamenstekster, og her kan en bruke modelltekster fra både eksamen og populærkulturen. Når det kommer til form, har det vist seg at gjennomført kreativ språk- og virkemiddelbruk kan ha sammenheng med sprikvurdering, men det synes problematisk å skulle oppfordre elevene til å la dette være, for hvordan skal de ellers løse en slik oppgave? Her vil det å lese og diskutere modelltekster kunne være viktige verktøy i arbeidet med hvordan språk og virkemiddelbruk kan fungere.

7 LITTERATUR

- Aardal, B., & Berglund, F. (2014). ZIGNE. Signifikans 5.9. Hentet fra <http://www.aardal.info/dataprogram.html>
- Aarnes, A., & Aarnes, A. (1977). *Litterært leksikon : begreper og betegnelser : teori - kritikk - historie* (3. utv. og rev. utg. utg.). Oslo: Tanum-Norli.
- Aarseth, A. (1976). *Episke strukturer : innføring i anvendt fortellingsteori*. Oslo: Universitetsforlaget.
- Aase, L. (2005). Norskfagets dannelsespotensial i fortid og samtid. I S. Nome & A. J. Aasen (Red.), *Det Nye norskfaget* (B. nr. 161). Bergen: Fagbokforl.
- Aase, L. (2012). *Skriveprosesser som dannelse*. Oslo: Universitetsforl., cop. 2012.
- Aasen, A. J. (2015). *Å handle med/ gjennom/ i ord: Begreperne teksttype og sjanger i norskfaget – perspektiver fra forskning*. Paper presentert på Teksttyper versus sjangere: fra sjangerformalisme til sjangeranarki?, Oslo. Powerpointpresentasjon hentet fra <http://www.norskundervisning.no/ukategorisert/2015/09/teksttyper-versus-sjangere-fra-sjangerformalisme-til-sjangeranarki/>
- Aasen, A. J., Solheim, R., & Smidt, J. (2011). *På sporet av god skriveopplæring : ei bok for lærere i alle fag*. Trondheim: Tapir akademisk forl.
- Atkins, L., Wallace, S., & Spenceley, L. (2012). *Qualitative research in education*. Los Angeles: SAGE.
- Bach, C. (2015). Eksamensoppgaver i norsk etter 13 år : sensorstemmer fra våren 2012 og våren 2014. Bergen[Oslo]: Fagbokforl. : Landslaget for norskundervisning, cop. 2015.
- Bachtin, M. (2005). *Spørsmålet om talegenrane*. Oslo: Pensumtjeneste.
- Bakke, J. O. (2014). *Å "late tanker" med penn og tastatur : om Skrivehjulet og skriving som grunnleggende ferdighet*. Bergen: Fagbokforl. : Landslaget for norskundervisning, 2014.
- Bakke, J. O., & Skovholt, K. (2015). "Å forestille seg betyr å bruke fantasien til å tenke på noe som ikke fins" : introduksjon av skriveoppgaver på 7. trinn. Bergen[Oslo]: Fagbokforl. : Landslaget for norskundervisning, cop. 2015.
- Barrett, P. (2001). Assessing the reliability of rating data. Hentet fra <http://www.liv.ac.uk/~pbarrett/rater.pdf>
- Berg, C. (2015). *Teksttyper vs. sjangre*. Paper presentert på Teksttyper versus sjangere: fra sjangerformalisme til sjangeranarki? , Oslo. Powerpointpresentasjon hentet fra <http://www.norskundervisning.no/ukategorisert/2015/09/teksttyper-versus-sjangere-fra-sjangerformalisme-til-sjangeranarki/>
- Berge, K. L. (1994). *Norsksensorenes vurderinger av eksamensbesvarelser i norsk hovedmål, allmennfaglig studieretning: en undersøkelse av norsksensorenes bedømmersamsvar og tekstnormer, sensurens pålitelighet og gyldighet, med vekt på karaktersettingen* (B. nr. 14). Trondheim: Senter for samfunnsforskning, Universitetet i Trondheim.
- Berge, K. L. (1996). *Norsksensorenes tekstnormer og doxa: en kultursemiotisk og sosiotekstologisk analyse*. Trondheim: Institutt for anvendt språkvitenskap, Det historisk-filosofiske fakultet, Norges teknisk-naturvitenskapelige universitet.
- Berge, K. L. (2005a). Skriveprøvenes pålitelighet. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse : B. 1 : Norsksensuren som kvalitetsvurdering*. Oslo: Universitetsforl.
- Berge, K. L. (2005b). *Skriving som grunnleggende ferdighet og som nasjonal prøve – ideologi og strategier*. folk.uio.no/kjellbe/Skriving.pdf
- Berge, K. L. (2005c). *Ungdommers skrivekompetanse* (B. [3]). Trondheim.
- Berge, K. L. (2009). Er tolkningsfellesskap mulig å oppnå i skriveprøver? Oslo: Universitetsforl., cop. 2009.

- Berge, K. L. (2014). *NORMPROSJEKTET vurderingskvalitet & tekstkvalitet: noen kvantitative observasjoner*.
<http://webcache.googleusercontent.com/search?q=cache:fbNILFsgHnMJ:norm.skrivesentert.no/wp-content/uploads/2014/02/BERGEGardermoen-januar2014.pdf+&cd=1&hl=en&ct=clnk&gl=no>
- Berge, K. L., Maagerø, E., Coppock, P. J., Halliday, M. A. K., Martin, J. R., & Hasan, R. (1998). *Å skape mening med språk : en samling artikler* (B. nr 112). Oslo: Landslaget for norskundervisning Cappelen akademisk forl.
- Bjølseth, G., Havnes, A., & Lauvås, P. (2011). Lavt sensorsamsvar – kan det bedres? *Uniped*, 4(04), 4-17.
- Blikstad-Balas, M. (2015a). *Fra sjangerformalisme til sjangeranarki?* Paper presentert på Teksttyper versus sjangere: fra sjangerformalisme til sjangeranarki? , Oslo. Powerpointpresentasjon hentet fra <http://www.norskundervisning.no/ukategorisert/2015/09/teksttyper-versus-sjanger-fra-sjangerformalisme-til-sjangeranarki/>
- Blikstad-Balas, M. (2015b). Gåten kreativ tekst: Hvordan kan vi fremme elevenes kreative skriving? *Norsklæraren*(2), 30-34.
- Blikstad-Balas, M., & Hertzberg, F. (2015). Fra sjangerformalisme til sjangeranarki? *Norsklæraren*, 1.
- Borgström, E., & Ledin, P. (2014). Bedömarvariation: Balansen mellan teknisk och hermeneutisk rationalitet vid bedömning av skrivprov. *Språk & Stil*, 24, 133-165.
- Botheim, S. S. (2015). Frå sjangerformalisme til sjangeranarki? *Norsklæraren*(4).
- Buffon, A. M. R., & Marti, K. T. (2014). Skrivning i skjæringspunktet mellom det faglige og det private: Om reflekterende tekster. *Norsklæreren*(1), 70-76.
- Claudi, M. B. (2013). *Litteraturteori* (B. 192). Bergen: Fagbokforl.
- Dagsland, S. (2015). Om relasjonen skrivehandling-skriveoppgave-elevtekst : utfordringer ved antatt reflekterende elevtekster. Bergen[Oslo]: Fagbokforl. : Landslaget for norskundervisning, cop. 2015.
- Dysthe, O., & Hertzberg, F. (2014). Skriveopplæring med vekt på prosess og produkt. Bergen: Fagbokforl. : Landslaget for norskundervisning, 2014.
- Dysthe, O., Hoel, T. L., & Hertzberg, F. (2010). *Skrive for å lære : skriving i høyere utdanning* (2. utg. utg.). Oslo: Abstrakt.
- Ecclestone, K. (2001). 'I know a 2:1 when I see it': Understanding criteria for degree classifications in franchised university programmes. *Journal of Further and Higher Education*, 25(3), 301-313. doi: 10.1080/03098770126527
- Eco, U. (1979). *The role of the reader : explorations in the semiotics of texts*. Bloomington, Ind: Indiana University Press.
- Eco, U., & Rognlien, J. (1994). *Seks turer i fortellingenes skoger*. Oslo: Tiden.
- Evensen, L. S. (2005). Pålitelighet og betydning. Etterord. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse : B. 1 : Norsksensuren som kvalitetsvurdering*. Oslo: Universitetsforl.
- Evensen, L. S. (2010). En gyldig vurdering av elevers skrivekompetanse? Trondheim: Tapir akademisk forl., cop. 2010.
- Evensen, L. S. (2014). Men kan det komme noe godt fra lærerværelset? om validitet i læreres vurdering. Oslo: Novus, 2014.
- Federl, M. (2009, 09.07.2015). *Essay*. Hentet 6. februar 2016, fra <http://ndla.no/nb/node/13316?fag=27>
- Freelon, D. (2011). ReCal2. Hentet fra <http://dfreelon.org/utills/recalfront/recal2/>
- Gorman, T. P., Purves, A. C., & Degenhart, R. E. (1988). *The IEA study of written composition*. Oxford: Published for the International Association for the evaluation of Educational Achievement by Pergamon Press.
- Grüters, R. (2011). En retorisk inspirert studie av refleksjonstekster i bloggformat. Trondheim: Tapir akademisk forl., cop. 2011.

- Gustafsson, J.-E., Cliffordson, C., & Erickson, G. (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*.
- Hagen, E. B. (2004). *Litteraturkritikk : en introduksjon*. Oslo: Universitetsforl.
- Halliday, M. A. K. (2014). *Halliday's introduction to functional grammar*. Abingdon: Routledge.
- Haugaløkken, O. K., Evensen, L. S., Hertzberg, F., & Otnes, H. (2009). *Tekstvurdering som didaktisk utfordring*. Oslo: Universitetsforl.
- Heian, B. (2015). *LNU: Teksttyper versus sjangere*. Paper presentert på Teksttyper versus sjangere: fra sjangerformalisme til sjangeranarki?, Oslo. Powerpointpresentasjon hentet fra <http://www.norskundervisning.no/ukategorisert/2015/09/teksttyper-versus-sjangerer-fra-sjangerformalisme-til-sjangeranarki/>
- Hellevik, O. (2002). *Forskningsmetode i sosiologi og statsvitenskap* (7. utg. utg.). Oslo: Universitetsforl.
- Hellevik, O. (2015). Hva betyr respondentbortfallet i intervjuundersøkelser? *Tidsskrift for samfunnsforskning*, 56(2), 211-231.
- Hertzberg, F. (2001). Tusenbenets vakre dans: Forholdet mellom formkunnskap og sjangerbeherskelse. *Rhetorica Scandinavica*(18), 92-105.
- Hertzberg, F. (2006). Skrivekompetanse på tvers av fag. Oslo: Universitetsforl., cop. 2006.
- Hertzberg, F., & Andersen, S. (2005). Å ta en risk - "Utfordrerne" blant eksamensskriverne *Ungdommers skrivekompetanse. Bind II. Norskeksamen som tekst* (B. B. 2, s. 411 s. : ill.). Oslo: Universitetsforl.
- Hertzberg, F., & Dysthe, O. (2012). *Prosesskriving - hvor står vi i dag?* Oslo: Universitetsforl., cop. 2012.
- Hjorth, V. (2001). *Hva er det med mor : roman*. Oslo: Cappelen.
- Hoel, T. L. (2008). *Utprøvande skriving i læringsprosessen*. Oslo: Universitetsforl., cop. 2008.
- Holter, H., & Kalleberg, R. (1996). *Kvalitative metoder i samfunnsforskning* (2. utg. utg.). Oslo: Universitetsforl.
- Igland, M.-A. (2007). "Svinaktig vanskelig"? : skriftleg argumentasjon på ungdomssteget. *Skriving i arbeidsliv og skole*.
- Johannessen, A., Christoffersen, L., & Tufta, P. A. (2010). *Introduksjon til samfunnsvitenskapelig metode* (4. utg. utg.). Oslo: Abstrakt.
- Johnsen, E. B. (1994). *Oppgavetekst og dannelse : artiumsstilens emner, formuleringer og forvaltning 1880-1991*. E.B. Johnsen, Oslo.
- Kalleberg, R. (2006). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*.
- Kittang, A. (1975). *Litteraturkritiske problem : teori og analyse*. Bergen: Universitetsforlaget.
- Kringstad, T., & Kvithyld, T. (2014). *Vurdering av skriving*. Bergen: Fagbokforl.
- Kristeva, J. (1969). *Sēmeiōtikē : recherches pour une sémanalyse*. Paris: Seul.
- Kvistad, A. H., & Smemo, J. (2015). Den gode skriveoppgaven? : en studie av fellestrekk ved vellykkede skriveoppgaver fra Normprosjektet. Bergen[Oslo]: Fagbokforl. : Landslaget for norskundervisning, cop. 2015.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 16.
- Lothe, J., Refsum, C., & Solberg, U. (1997). *Litteraturvitenskapelig leksikon*. Oslo: Kunnskapsforl.
- Lunde, M. G. (2015, 17.04.2015). Færre lesere, dobler lønnen, *Dagens Næringsliv*.
- Maagerø, E. (2005). *Språket som mening : innføring i funksjonell lingvistikk for studenter og lærere*. Oslo: Universitetsforl.
- Metliaas, I., & Solbu, K. R. (2015). *Eksamenstekster som modelltekster i opplæringa*. Hentet 02.01 2016, fra <http://skrivesenteret.no/prosjekter/eksamensbesvarelser-som-modelltekster-i-opplaeringa/>
- Miller, C. (2001). Genre som sosial handling. *Rhetorica Scandinavica :tidsskrift for skandinavisk retorikforskning*(18).

- Mukařovský, J. (2003). Intenjonaltet og ikke-intenjonaltet i kunsten *Záměrnost a nezáměrnost v umění*. Oslo: Universitetsforl., cop. 2003.
- Nilsen, H. (2005). Hva sier sensor? Sensorkommentarer om tekstkvalitet. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse : B. 1 : Norsksensuren som kvalitetsvurdering*. Oslo: Universitetsforl.
- OECD. (2005). *THE DEFINITION AND SELECTION OF KEY COMPETENCIES*. Hentet 03.01 2016, fra <http://www.oecd.org/education/skills-beyond-school/definitionandselectionofcompetenciesdeseco.htm>
- Otnes, H. (2015). *Å invitere elever til skrivning : ulike perspektiver på skriveoppgaver*. Bergen: Fagbokforl. Landslaget for norskundervisning.
- Penne, S. (2010). *Litteratur og film i klasserommet : didaktikk for ungdomstrinn og videregående skole*. Oslo: Universitetsforl.
- Rasmussen, A. (2015). Jeg er lei av å søke etter å bli eldre... Hentet fra <http://mammatilmichelle.blogg.no/>
- Rosenblatt, L. (1994). *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*.
- Seland, I., Lødding, B., & Prøitz, T. S. (2015). *Delrapport 1 fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk. Litteraturstudie*. Oslo: NIFU.
- Skaftun, A. (2009). *Litteraturens nytteverdi* (B. nr. 178). Bergen: Fagbokforl.
- Skar, G. B., & Aasen, A. J. (2015). Risikotrekk og skjulte kvalitetar i elevtekstar. *Nordic Journal of Literacy Research*, 2(1), 20.
- Skjelbred, D. (2014). *Elevens tekst : et utgangspunkt for skriveopplæring* (4. utg. utg. B. [63]). Oslo: Cappelen Damm akademisk.
- Skjelbred, D., & Veum, A. (2013). *Literacy i læringskontekster*. Oslo: Cappelen Damm akademisk. Skrivesenteret. (2013), fra <http://www.skrivesenteret.no/ressurser/skrivehjulet/>
- Skårderund, J. R. (2015, 21.11.2015). Mattefall skyldes eksamen, *Klassekampen*.
- Smemo, J., & Solem, H. (2015). *Skriveoppgaver i skriveprøvene : oppgavedesign, utfordringer og erfaringer*. Bergen[Oslo]: Fagbokforl. : Landslaget for norskundervisning, cop. 2015.
- Smidt, J. (2009). *Norskdidaktikk : ei grunnbok* (3. utg. utg.). Oslo: Universitetsforl.
- Smidt, J. (2014). *Skriftens mange scener : skriveopplæringens mål og midler i vurderingstider*. Oslo: Novus, 2014.
- Smidt, J. (2015). Kreative tekster? I alle fall kreative norsklærere og (kan vi håpe) kreative elever. *Norsklæraren*(2), 19-27.
- Smidt, J., Tønnessen, E. S., & Aamotsbakken, B. (2011). *Tekster i bevegelse : tekster, tegn og grunnleggende ferdigheter*. Trondheim: Tapir akademisk forl., cop. 2011.
- Solheim, R., & Matre, S. (2014). Forventninger om skrivekompetanse. Perspektiver på skrivning, skriveopplæring og vurdering i "Normprosjektet". *Viden om læsning*(15).
- Solheim, R., Sjøhelle, D. K., & Matre, S. (2012). *Teorier om tekst i møte med skolens lese- og skrivepraksiser*. Oslo: Universitetsforl.
- Stemler, S. E. (2004). A Comparison of Consensus, Consistency, and Measurement Approaches to Estimating Interrater Reliability. *Practical Assessment, Research & Evaluation*, 9(4).
- Thagaard, T. (2013). *Systematikk og innlevelse : en innføring i kvalitativ metode* (4. utg. utg.). Bergen: Fagbokforl.
- Thoresen, M. (2010). *Sprikvurdering av elevtekster til grunnskoleeksamen*. Oslo: M. Thoresen.
- Tønnessen, E. S., & Maagerø, E. (2015). *Å lese i alle fag* (2. utgave. utg.). Oslo: Universitetsforlaget.
- Tønnesson, J. L. (2012). *Hva er sakprosa* (2. utg. utg. B. 25). Oslo: Universitetsforl.
- Utdanningsdirektoratet. Skoleporten. Hentet 02.01.2016 <https://statistikkportalen.udir.no/vgs/Pages/Karakterer-i-videreg%C3%A5ende-skole.aspx>
- Utdanningsdirektoratet. (2013). *Revidert læreplan og eksamen. Eksempeloppgaver*.
- Utdanningsdirektoratet. (2014). *Eksamensveiledning - om vurdering av eksamenssvar*.
- Utdanningsdirektoratet. (2015a). *Eksamensveiledning – om vurdering av eksamenssvar*.
- Utdanningsdirektoratet. (2015b, 18. juni). *Fellessensur*. Paper presentert, Drammen.
- Utdanningsdirektoratet. (2015c, 3. juni). *Sensorskolering*. Paper presentert, Drammen.

- Utdanningsdirektoratet. (2015d). *Sensorskolering våren 2015*. Paper presentert på Sensorskolering våren 2015. Norsk hovedmål og norsk sidemål. Powerpointpresentasjon/lysark hentet fra Utdanningsdirektoratet. (2015e). *Tilbakemelding fra sensorene i Kunnskapsløftet NOR1211 – NOR1218 – NOR1231*. Sammenfattet av Tone Vindegg.
- Vagle, W. (2005). Hva presterer elevene målt med gammel og ny karakterskala? Sensuren ved overgangen til nye rammer for elevvurdering. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse : B. 1 : Norsksensuren som kvalitetsvurdering*. Oslo: Universitetsforl.
- Vagle, W., & Evensen, L. S. (2005). Oppgavesettene og elevenes oppgavevalg i KAL-årene. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse : B. 1 : Norsksensuren som kvalitetsvurdering*. Oslo: Universitetsforl.
- Vassenden, E. (2007). Hva er «samtidslitteratur», og hvorfor leser vi den? – Noen begrepshistoriske og fagkritiske bemerkninger. *Edda*, 107(04), 357-371.
- Östlund-Stjärnegårdh, E. (2002). *Godkänd i svenska?: bedömning och analys av gymnasieelevers texter* (B. 57). Uppsala: Institutionen.