

ULEMGOPPG - Masteroppgave

«Jeg tenker det er en sterk firer, jeg da. Men det kan hende at det egentlig er stryk.»

Rettferdig sluttvurdering i norskfaget
i videregående skole

Synnøve Skare Baade

Erfaringsbasert master i undervisning og læring,
fordypning norsk

2022



HØGSKULEN
I VOLDA

Abstract

This study explores the factors which affect Norwegian teachers' final assessment of students' written texts in Norwegian upper secondary schools. Different teachers' assessments of the same texts have been proven to vary quite significantly, and this study therefore aims to answer the following thesis question: *Which factors influence three Norwegian teachers' assessment of a student text?* The foundation of knowledge accounts for relevant steering documents and terms of reference, scientific foundation, and previous research in the field.

The source of data has been collected through individual semi-structured interviews where three Norwegian teachers from different upper secondary schools explain their assessment of the same student's text. Through a descriptive analysis it becomes evident that the three teachers belong to professional environments with different qualities, and this appears to have led them to emphasize different factors in their assessment of the text. In the interviews, they also reflect on the term "fair assessment", and these reflections shed light on the dilemma between the universal and differential dimension in assessments (Prøitz & Borgen, 2010).

It becomes clear in the interviews that the teachers assess the same elements of the texts: *language, structure, and content*. Despite having this in common, the analysis shows that they have different perceptions of what characterizes quality in the different elements. Furthermore, the analysis shows that whether they make active use of assessment criteria or not, does not seem to affect their grading. On the other hand, programs designed to enhance assessment, like sensor schooling, seems to have significance for whether or not teachers agree on the assessment of students' texts.

Sammendrag

Denne studien undersøker hvilke faktorer som spiller inn i norsklæreres sluttvurdering av skriftlige elevtekster i norsk vg3. Temaet springer ut av, blant annet, det faktum at det i enkelte tilfeller er beviselig stor avstand mellom ulike læreres vurdering av de samme elevtekstene. Gjennom studien søkes det å besvare følgende problemstilling: *Hvilke faktorer påvirker tre norsklæreres vurdering av en elevtekst?* I kunnskapsgrunnlaget redegjøres det for relevante styringsdokumenter, vitenskapsteoretisk forankring samt tidligere forskning på feltet.

Datagrunnlaget er innhentet gjennom individuelle semistrukturerte intervjuer, der tre norsklærere fra ulike videregående skoler gjør rede for sin vurdering av den samme elevteksten. Gjennom en deskriptiv analyse blir det klart at de tre lærerne tilhører profesjonsfellesskap med ulike kvaliteter, og dette synes å ha påvirket lærerne til å vektlegge ulike faktorer i vurderingen av elevteksten. I intervjuene reflekterer de også over begrepet «rettferdig vurdering», og gjennom disse refleksjonene belyses dilemmaet mellom den universelle og den differensielle dimensjonen i vurderingsarbeidet (Prøitz & Borgen, 2010).

Gjennom intervjuene viser det seg at utvalget vurderer de samme elementene ved elevteksten: *språk, struktur og innhold*. Til tross for dette viser analysen at de har ulik oppfatning av hva som kjennetegner kvalitet i de ulike elementene. Videre viser analysen at hvorvidt de bruker vurderingskriter aktivt eller ikke, ikke ser ut til å ha betydning for samsvaret i karaktersettingen. Opplæringsaktiviteter, som for eksempel sensorskolering, ser derimot ut til å ha betydning for hvor enige lærerne er i hvilket karakternivå elevteksten befinner seg på.

Forord

Disse ordene markerer avslutningen på både hektiske og lærerike år. Da jeg for alvor bestemte meg for å endelig ta fatt på et masterløp, var det med visshet om at å sjonglere studietid med lærerjobb og familieliv ikke ville være det samme som å være student på fulltid. I prosessen har jeg likevel vært glad for at jeg har med meg erfaring fra praksisfeltet.

Når jeg nå nærmer meg målstreken er det mange som fortjener en takk. Aller først vil jeg takke de tre lærerne som stilte opp som deltakere i intervju, og som åpenhjertig delte sine tanker og refleksjoner. Uten dere hadde ikke denne oppgaven blitt den samme. Videre vil jeg rette en stor takk til min alltid positive veileder Pernille Fiskerstrand ved Høgskulen i Volda. Takk for dine gode råd og konstruktive tilbakemeldinger, og takk for at du alltid har vært tilgjengelig når jeg har hatt bruk for din kyndige veiledning. Takk også til min arbeidsgiver for å velvillig legge til rette for min fikse idé om å være fulltidsstudent ved siden av jobb, og til gode kollegaer som har gjort jobblivet til en fest. En særlig takk til Trine, min gode kollega og aller beste venn. Uten deg, din imponerende arbeidskapasitet og din vilje til å dele gode undervisningsopplegg, hadde mine elever i langt større grad lidd under mitt studieprosjekt. Til slutt vil jeg takke familien min. Takk til David og Hedvig, som gang etter gang godtar at «det må vi ta en annen gang». Nå er det straks «en annen gang», det lover jeg! Sist, men ikke minst: tusen takk Henning. Takk for at det aldri er nei i din munn, og for at du dermed har gjort det mulig for meg å kombinere familieliv med både fulltidsstudier og lærerjobb.

Samtidig som denne oppgaven leveres går videregående skole en eksamensfri avslutning i møte, for tredje år på rad. Debatten om rettferdighet i sluttvurdering har sjelden vært mer aktuell, og det føles som et privilegium å ha fått studert dette fenomenet på nært hold i akkurat dette tidsrommet. Jeg håper at mitt engasjement for fagområdet smitter over på leseren. God lesning!

Innhold

| | |
|---|-----|
| Abstract | I |
| Sammendrag | II |
| Forord..... | III |
| 1.0 Innledning..... | 1 |
| 1.1 Valg av tema | 1 |
| 1.2 Problemstilling og forskningsspørsmål..... | 1 |
| 1.3 Oppgavens oppbygning og begrepsavklaring..... | 3 |
| 1.3.2 Rettferdig vurdering | 4 |
| 1.3.3 Vurderingskriterier | 4 |
| 2.0 Kunnskapsgrunnlag | 5 |
| 2.1 Læreplan, lovverk, føringer | 5 |
| 2.1.1 Forskrift til opplæringslova | 5 |
| 2.1.2 Overordnet del av læreplanverket..... | 5 |
| 2.1.3 Læreplan i norsk | 6 |
| 2.1.4 Eksamensveiledning med vurderingskriterier..... | 7 |
| 2.2 Vitenskapsteoretisk tilnærming | 8 |
| 2.2.1 Vygotsky og sosiokulturell læringsteori | 9 |
| 2.2.2 Bakhtins dialogisme..... | 9 |
| 2.2.3 Meads speilingsteori og «de signifikante andre» | 10 |
| 2.2.4 Gadammers filosofiske hermeneutikk og Bakhtins translingvistikk | 12 |
| 2.2.5 Vurdering av elevtekster i lys av sosiokulturell læringsteori | 13 |
| 2.3 Tidligere studier | 14 |
| 2.3.1 Vurdering av skriveferdigheter | 14 |
| 2.3.2 Eksamen som skjult læreplan..... | 14 |
| 2.3.3 Bedømmersamsvar og rettferdig vurdering | 16 |
| 2.3.5 Vurdering i profesjonsfellesskapet..... | 23 |
| 2.4 Eksamensoppgave og elevtekst..... | 25 |
| 2.4.1 Eksamensoppgave | 26 |
| 2.4.2 Elevtekst | 26 |
| 3.0 Metode..... | 28 |
| 3.1 Valg av forskningsmetode | 28 |

| | | |
|-------|---|----|
| 3.1.1 | Kvalitativ forskningsmetode..... | 28 |
| 3.1.2 | Liten N- og fenomenologiske studier | 28 |
| 3.1.3 | Abduktiv tilnærming..... | 29 |
| 3.1.4 | Semistrukturert intervju..... | 30 |
| 3.2 | Forberedelser og gjennomføring..... | 30 |
| 3.2.1 | Strategisk tilgjengelighetsutvalg og snøballmetoden | 30 |
| 3.2.2 | Om deltakerne..... | 32 |
| 3.2.3 | Gjennomføring av intervju | 32 |
| 3.2.4 | Transkribering | 34 |
| 3.2.5 | Analyseprosess | 35 |
| 3.3 | Troverdighet og gyldighet..... | 36 |
| 3.3.1 | Reliabilitet | 36 |
| 3.3.2 | Validitet | 36 |
| 3.4 | Forskningsetiske betraktninger | 37 |
| 3.4.1 | Informert samtykke..... | 37 |
| 3.4.2 | Konfidensialitet | 37 |
| 3.4.3 | Konsekvenser | 38 |
| 3.4.4 | Forskerens rolle..... | 38 |
| 4.0 | Resultat..... | 39 |
| 4.1 | Faktorer som påvirker vurderingen av en elevtekst | 39 |
| 4.1.1 | Lærer1 sin vurdering av en elevtekst: vektninger og begrunnelser | 39 |
| 4.1.2 | Lærer2 sin vurdering av en elevtekst: vektninger og begrunnelser | 41 |
| 4.1.3 | Lærer3 sin vurdering av en elevtekst: vektninger og begrunnelser | 42 |
| 4.1.4 | Sammenfatting og oppsummering..... | 44 |
| 4.2 | Likheter og forskjeller mellom de ulike lærernes vektninger..... | 45 |
| 4.2.1 | Element som vektlegges av flere lærere | 45 |
| 4.2.2 | Element som vektlegges av bare en lærer | 47 |
| 4.2.3 | Sammenfatting og oppsummering..... | 48 |
| 4.3 | Rettferdig vurdering | 49 |
| 4.3.1 | Lærer1 sin forståelse av begrepet «rettferdig vurdering» og beskrivelse av egen praksis | 49 |
| 4.3.2 | Lærer2 sin forståelse av begrepet «rettferdig vurdering» og beskrivelse av egen praksis | 50 |

| | | |
|-------|---|-----|
| 4.3.3 | Lærer3 sin forståelse av begrepet «rettferdig vurdering» og beskrivelse av egen praksis | 51 |
| 4.3.4 | Sammenfatting og oppsummering..... | 52 |
| 5.0 | Drøfting | 53 |
| 5.1 | Faktorer som skal inngå i en sluttvurdering | 53 |
| 5.1.1 | Å vurdere elevenes forutsetninger og innsats | 54 |
| 5.1.2 | Å sammenligne elever | 55 |
| 5.1.3 | Forholdet mellom skriveferdighet og leseferdighet | 56 |
| 5.2 | Rettferdig vurdering på tekstenivå..... | 58 |
| 5.2.1 | Hvor går skillet mellom de ulike karakterene? | 58 |
| 5.2.2 | IRR og IRA – vurderingssamsvar og vurderingsenighet | 60 |
| 5.3 | Rettferdig vurdering på profesjonsnivå..... | 62 |
| 5.3.1 | Sensorskoleringens effekt | 63 |
| 5.3.2 | De signifikante andre i profesjonsfellesskapet | 65 |
| 6.0 | Avslutning..... | 69 |
| 6.1 | Hovedfunn og konklusjon..... | 69 |
| 6.1.1 | Hvilke faktorer peker lærerne selv på som viktige i sin vurdering?..... | 70 |
| 6.1.2 | Hvilke faktorer avgjør om en vurdering er rettferdig, ifølge lærerne?..... | 70 |
| 6.1.4 | Konklusjon | 71 |
| 6.2 | Videre forskning..... | 71 |
| 7.0 | Referanser | 73 |
| | Vedlegg I – Informasjonsskriv og samtykkeskjema for bruk av elevtekst..... | I |
| | Vedlegg II – Informasjonsskriv og samtykkeskjema for deltakere i intervju..... | III |
| | Vedlegg III – Intervjuguide..... | V |
| | Vedlegg IV – Godkjenning fra NSD (skjerm bilde) | VI |

1.0 Innledning

1.1 Valg av tema

Tema for denne studien er sluttvurdering i norskfaget i den videregående skolen, og valg av tema springer delvis ut av funnene som Bjørnson og Skar (2021) rapporterer om i *Sensorreliabilitet på skriftlig eksamen i videregående opplæring*. Rapporten ble publisert i 2021, og viser et stort sprik mellom karakterene som settes av de ulike sensorene på skriftlig eksamen i norsk, og slår fast at den samlede enigheten er så lav som 42,65 % i norsk hovedmål på studieforberedende utdanningsprogram (s. 21). Ifølge rapporten viser MFRM-analysen at det er «[...] høy sannsynlighet for at enkeltelever er prisgitt sensors vurderingsatferd» (s. 36), og det konkluderes med at det er lav reliabilitet ved vurderingen av eksamenssvar i norskfaget. Funnene er uheldige, men ikke spesielt overraskende. Sensasjonsoverskriftene om eksamenskandidater som har gått opp flere karakterer etter klagesensur er årvisse, og de viser at eksamenskandidatene, i enkelte tilfeller, virkelig er prisgitt nettopp vurderingsatferden til sensor. Valg av tema springer også delvis ut av min egen erfaringsbakgrunn som norsklærer i videregående skole gjennom åtte år. Rapporter om manglende vurderingssamsvar på eksamen går på yrkesstoltheten løs, og rettferdighetssansen trigges litt ekstra de gangene egne elever på eksamen blir vurdert til andre karakterer enn jeg selv ville vurdert dem til på samme besvarelse.

OECD etterlyser nettopp rettferdighet i vurderingspraksisen i norsk skole, men begrepet rettferdighet er ikke å finne «verken i opplæringslov eller forskrift til loven» (Dahl et al., 2015, s. 46-47). Kanskje er dette et uttrykk for at rettferdig vurdering anses som et selvsagt prinsipp, men i praksis ser det likevel ut til at ikke all sluttvurdering er rettferdig. Jeg undrer meg derfor over om enkeltelever er prisgitt læreren og sensorens vurderingsatferd i all form for sluttvurdering.

1.2 Problemstilling og forskningsspørsmål

Min tilnærming til temaet er forankret i en forståelse av at vurderingsarbeidet i videregående skole gir samfunnsmessige ringvirkninger, og en erkjennelse av at norskfaget med sine tre standpunktkarakterer og tre mulige eksamenskarakterer i så henseende står i en særstilling.

Det er altså av betydning både for den enkelte elev, og for samfunnet som sådan at vurderingene som gis er både etterrettelige og rettferdige, og det er et uttalt mål fra Utdanningsdirektoratet at lærerne i skolen skal ha en rettferdig vurderingspraksis (Utdanningsdirektoratet, 2022). Jeg ønsker derfor å finne ut om ulike lærere navigerer etter ulike prinsipper i vurderingsarbeidet, eller om de er del av et tolkningsfellesskap, og jeg søker dessuten innsikt i hvordan ulike norsklærere forholder seg til prinsippet om rettferdig vurdering. Med dette som utgangspunkt lyder problemstillingen som følger:

Hvilke faktorer påvirker tre norsklæreres vurdering av en elevtekst?

Vurdering er et komplekst fagområde, og omfanget av denne masterstudien er for begrenset til å ha ambisjoner om å favne alle tenkelige faktorer som påvirker en lærers vurdering. I tillegg til åpenbare påvirkningsfaktorer, som for eksempel skolens styringsdokumenter, vil et vurderingsarbeid dessuten også være påvirket av faktorer som ligger utenfor både vurdererens og forskerens bevissthet, og det vil derfor være umulig å ta sikte på å favne alle disse faktorene. Studien konsentreres av den grunn omkring de faktorene som lærerne selv gir uttrykk for å la seg påvirke av. Det er dessuten nødvendig å presisere at ordet «påvirker» i problemstillingen ikke er et uttrykk for at studien tar sikte på å måle effekt eller kausalitet. Problemstillingen er likevel utformet som den er, fordi studien søker å romme en spenning mellom det som lærerne vet, de skal la seg påvirke av på den ene siden, og det de føler, er riktig å la seg påvirke av på den andre siden. For å kunne si noe om hvordan de går frem når de vurderer den konkrete elevteksten som ligger til grunn for denne studien, trenger jeg også innblikk i hvordan de vanligvis tenker når de vurderer en elevtekst, og i denne sammenheng er begrepet «rettferdig vurdering» sentralt. For å konkretisere problemstillingen ytterligere er det derfor utformet to forskningsspørsmål:

1. Hvilke faktorer peker lærerne selv på som viktige i sin vurdering?
2. Hvilke faktorer avgjør om en vurdering er rettferdig, ifølge lærerne?
 - Rettferdig vurdering på tekstnivå
 - Rettferdig vurdering på profesjonsnivå

For å svare på problemstillingen har jeg benyttet kvalitativ forskningsmetode, og begge forskningsspørsmålene baserer seg på datamateriale som er innhentet gjennom semistrukturerte intervju med tre norsklærere i videregående skole. Andre forskningsspørsmål er inndelt i to nivå som en følge av at begrepet «rettferdig vurdering» rommer både tekstinterne og kontekstuelle faktorer.

1.3 Oppgavens oppbygning og begrepsavklaring

Oppgavens hoveddel består av fire kapitler. I *2.0 Kunnskapsgrunnlag* presenteres relevante vitenskapsteorier og tidligere forskning på feltet. Videre følger *3.0 Metode*, et kapittel som søker å begrunne de metodiske valgene som er tatt i arbeidet med denne studien. Datagrunnlaget er analysert og forsøkt sammenfattet i *4.0 Resultat*, og drøftes i lys av problemstilling og forskningsspørsmål i *5.0 Drøfting*.

I kunnskapsgrunnlaget redegjøres det løpende for sentrale begreper, men med hensyn til ordlyden i problemstilling og forskningsspørsmål vil jeg her gjøre rede for de gjennomgående begrepene *vurdering*, *rettferdig vurdering* og *vurderingskriterier*.

1.3.1 Vurdering

Vurdering har de siste årene utviklet seg til å bli et «trendord» i norsk skole, og kjært barn har som kjent mange navn. I litteratur om vurdering møter man begreper som *formativ vurdering*, *summativ vurdering*, *vurdering for læring*, *vurdering av læring*, *underveisvurdering* og *sluttvurdering*. I det følgende skal jeg primært se på vurdering som skjer i slutten av en opplærings situasjon, og vil således benytte meg av begrepene «vurdering av læring» og «sluttvurdering». De gangene jeg refererer til andre som bruker en variant av «summativ vurdering» kommer jeg likevel ikke til å endre ordet til for eksempel «sluttvurdering». Det er et valg som tas for å sikre at det ikke blir fjernet meningsinnhold fra andres tekster. At jeg i deler av teksten bruker ulike begreper må likevel ikke forstås som et innlegg i en eventuell debatt omkring hvilke begreper som er mest dekkende for fenomenet.

Sluttvurdering, eller summativ vurdering, kan, ifølge Engh et al. (2007), være både «et mål på hvor mye læring som har skjedd til nå enten i hver modul eller fagområde som et fag består av, eller et mål på det som totalt er oppnådd i faget» (s. 29). I forskrift til opplæringslova (2020)

slås det derimot fast at «Sluttvurderinga skal gi informasjon om kompetansen til eleven [...] ved avslutninga av opplæringa i fag» (§ 3-14). Det er altså en spenning mellom hvorvidt man kan kalle vurdering som foregår etter endt «modul» for sluttvurdering, eller ikke. I det følgende vil jeg i hovedsak forholde meg til opplæringslovens forståelse av begrepet. Det synes likevel å være vanlig praksis å legge de samme kriteriene til grunn underveis i et skoleår, som mot slutten, og jeg velger derfor å inkludere den utvidede forståelsen av begrepet i tillegg. At sluttvurderingen ikke skal ta hensyn til elevenes innsats, forutsetninger og lignende er det likevel enighet om i både forskningslitteraturen og i lovverket (jf. § 3-14; Engh et al., s. 29).

1.3.2 Rettferdig vurdering

Som en bærende samfunnsinstitusjon er skolen avhengig av tillit i befolkningen, og hvilke praksiser skolen har for nettopp sluttvurdering er, slik jeg ser det, avgjørende for graden av tillit som samfunnet har til skolen. I Utdanningsdirektoratets vurderinger og anbefalinger om fremtidens eksamen vektlegges det at standpunkt og eksamen skal gi et «rettferdig bilde av elevenes kompetanse» (Utdanningsdirektoratet, 2020, s. 53), en nødvendighet for skolen selv, så vel som for alle de enkeltelevne som til enhver tid befinner seg i skolesystemet. I forskningslitteratur (Lortie, 1975; Jencks, 1988; Resh; 2009) pekes det likevel på et dilemma: er det mest rettferdig å vurdere alle elever på samme grunnlag, eller er det mest rettferdig å tilpasse vurderingen individuelt? Spørsmålet om rettferdig karaktersetting har altså både en universell og en differensiell dimensjon (Prøitz & Borgen, 2010, s. 10). Hva angår sluttvurdering er lovverket tydelig på at elevenes forutsetninger ikke skal hensyntas, og det er derfor i hovedsak den universelle dimensjonen som ligger til grunn i det følgende. Prøitz og Borgen (2010) peker likevel på at lærere benytter en differensiell tilnærming som «kan gjenspeiles i sammenhenger med vurdering og karaktersetting» (s. 44), og det er derfor også naturlig å inkludere denne dimensjonen i denne begrepsavklaringen.

1.3.3 Vurderingskriterier

Vurderingskriterier er et mulig verktøy for å sikre universell rettferdig vurdering. I denne oppgaven brukes begrepene *vurderingskriterier* og *kjennetegn på måloppnåelse* om hverandre, og når disse begrepene brukes er det Sadlers (2010) definisjon som ligger til grunn: “By a criterion is meant a property or characteristic that is useful in the context of quality and quality determinations” (s. 544). Ifølge Matre (2017) brukes gjerne begrepet kriterier av

lærerne i skolen, «et begrep de kjenner fra vurderingsteori» (s. 2), men i eksamensveiledningen for norsk vg3, utarbeidet av Utdanningsdirektoratet, omtales de som *kjennetegn på måloppnåelse*. En veksling mellom de ulike begrepene gjennom oppgaven må forstås som et forsøk på å til enhver tid bruke det samme begrepet som brukes av kilden.

2.0 Kunnskapsgrunnlag

2.1 Læreplan, lovverk, føringer

Den norske skolens styringsdokumenter legger sentrale føringer for innholdet i et prosjekt av denne art, og i gjeldende kapittel blir disse derfor introdusert.

2.1.1 Forskrift til opplæringslova

Elevenes rettigheter og plikter i forbindelse med individuell vurdering i grunnskolen og videregående skole er nedfelt i *Forskrift til opplæringslova* (2020, §3). Elevene har rett til sluttvurdering, og denne sluttvurderingen skal gis som en hel karakter på skalaen 1-6. For elever ved studieforberedende linje i videregående er det bare standpunktkarakter og eksamenskarakter som anses som sluttvurdering. Forskriften slår fast at det er kompetansemålene som skal danne grunnlag for vurdering i fag, og innsats skal som hovedregel ikke inngå som en del av vurderingsgrunnlaget. Enkelte elever har rett til fritak fra vurdering med karakter i skriftlig sidemål, men i de tilfellene det er fattet vedtak om fritak skal eleven skal ha underveisvurdering uten karakter. Elevene har altså ikke fritak fra å verken lese eller skrive på sidemålet sitt, men sidemålsferdighetene skal ikke vurderes med karakter. Forskrift til opplæringslova regulerer også elevenes rett til en rettferdig sensur. § 3-28 (Forskrift til opplæringslova, 2020, § 3-28) slår fast at alle prøvesvar skal vurderes av to sensorer, der begge har tilfredsstillende kompetanse. Hva som regnes som tilfredsstillende kompetanse fremgår derimot ikke av forskriften.

2.1.2 Overordnet del av læreplanverket

Overordnet del av læreplanverket er også forskriftsfestet, og fastsettes ved kongelig resolusjon. Historisk har overordnet del vært preget av mer ro enn læreplanen for øvrig, og gjeldende del er fra 2017. Den gang førte den med seg en del spesifikke endringer. For eksempel defineres nå kompetansebegrepet langt mer konkret enn tidligere: «Kompetanse er å kunne tilegne seg

og anvende kunnskaper og ferdigheter til å mestre utfordringer og løse oppgaver i kjente og ukjente sammenhenger og situasjoner. Kompetanse innebærer forståelse og evne til refleksjon og kritisk tenkning.» (Overordnet del, 2017, s. 10), og denne definisjonen skal ligge til grunn når en lærer eller sensor vurderer en elevs faglige kompetanse.

Ludvigsenutvalget, nedsatt av Regjeringen Stoltenberg II i 2013, er de som først tar til orde for å utvide kompetansebegrepet i skolen, og de peker også blant annet på stofftrenghet som en av skolefagenes store utfordringer (NOU 2014: 7 og NOU 2015: 8). Meld. St. 28 bygger på rapportene fra Ludvigsenutvalget i sin redegjørelse for intensjonen med fornyelsen av innholdet i grunnopplæringen (Meld. St. 28 (2015-2016), s. 15-16), og daværende kunnskaps- og integreringsminister Jan Tore Sanner (H) vektlegger nettopp stofftrenghetsproblematikken når han presenterer Regjeringen Solbergs plan for å fornye skolefagene (Aftenposten, 2018). Overskriftene i riksdekkende medier er tabloide: «Skolefagene skal slankes!». Intensjonen med slankekur er å skape mer rom for tverrfaglighet og på den måten tilrettelegge for elevenes dybdelæring.

«Profesjonsfellesskapet» får også sin plass i den nye overordnede delen fra 2017 (Overordnet del, 2017, s. 15, 18-19). Skolen skal videreutvikles i et profesjonsfellesskap: «Kompliserte pedagogiske spørsmål har sjelden sikre svar», og vi ser konturene av et sosialkonstruksjonistisk ståsted når vi i overordnet del kan lese at spørsmål skal «besvares på nytt hver dag av alle som jobber i skolen» (Overordnet del, 2017, s. 19). I John Donnes dikt heter det seg at «No Man is an Island», og selv om ikke arkitektene bak overordnet del fra 2017 skal beskyldes for å ha brukt et spesielt poetisk språk (Engelsen, 2020. s. 212) er det likevel nettopp dette som synes å være budskapet: Det er gjennom profesjonsfellesskapet man skal utvikle sin praksis og finne svar på pedagogiske spørsmål.

2.1.3 Læreplan i norsk

Våren 2022 jobber man, som følge av innføringen av fagfornyelsen, med utgangspunkt i to ulike læreplaner i det ordinære norskfaget i videregående skole: LK20 for vg1 og vg2, og LK06 for vg3. For norskfagets vedkommende merker man dette aller tydeligst ved at den skriftlige karakteren til elevene ikke lenger er delt i hovedmål og sidemål i vg1 og vg2, i tillegg til at kravet til elevene om å gjennomføre en selvvalgt fordypningsoppgave

(Kunnskapsdepartementet, 2013, s. 12) i vg3 er borte. Til gjengjeld understrekes det i kompetansemålene for vg3 i LK20 at elevene skal kunne «analysere og tolke romaner [...]» (Kunnskapsdepartementet, 2019, s. 14), et krav man ikke finner i LK06. I tillegg er en del fagstoff flyttet fra et årstrinn til et annet, for eksempel språkhistorie, som i LK20 er samlet på vg2 i motsetning til å være delt over vg2 og vg3, som i LK06. I realiteten er det likevel få momenter som er fjernet helt fra kompetansemålene, og man kan derfor diskutere hvor effektiv den varslede slankekurten egentlig har vært for norskfaget sitt vedkommende.

Hva angår ordningene for skriftlig sluttvurdering, slik begrepet er definert i *Forskrift til opplæringslova* (2020, §3), har de noe ulik ordlyd i de to versjonene av læreplanen, men er innholdsmessig identiske: «Vg3 studieforbereende utdanningsprogram: Eleven skal opp til skriftlig eksamen i norsk hovedmål. Skriftlig eksamen blir utarbeidet og sensurert sentralt. Eleven kan også trekkes ut til skriftlig eksamen i norsk sidemål» (Kunnskapsdepartementet, 2019, s. 18). Det er varslet noen endringer for skriftlig eksamen i norsk under LK20, men endringene er forholdsvis små. Eksamensformen, skoleeksamen på et gitt antall timer, videreføres. Antall oppgaver og oppgavetyper ligger derimot an til å bli endret (Utdanningsdirektoratet, 2021a). De nye eksamensoppgavene er i skrivende stund ikke klare, og dette prosjektet baserer seg derfor på en elevbesvarelse og en sensorveiledning tilhørende LK06.

2.1.4 Eksamensveiledning med vurderingskriterier

Eksamensveiledningen er utarbeidet av Utdanningsdirektoratet, og er ment som en støtte for både elev og sensor. Eleven skal være kjent med hvordan eksamen blir vurdert, og «Sensorene må bruke veiledningen som en felles referanseramme i arbeidet sitt» (Utdanningsdirektoratet, 2018a, s. 2). Eksamensbesvarelsen som de intervjuede lærerne har blitt bedt om å vurdere er fra 2018, men den foreløpig nyeste eksamensveiledningen er fra 2021. Jeg legger derfor begge disse til grunn i det følgende. I tilfeller hvor det er betydelig avstand mellom eksamensveiledningen fra 2018 og 2021 vil det opplyses om det i den løpende teksten.

I begge eksamensveiledningene presiseres det at eksamens overordnede formål et at «elevene skal vise norskfaglig kompetanse» (s. 2), og det er verdt å merke seg at det ikke

begrenses utelukkende til et skriftlig kompetansebegrep. Elevene skal presentere svaret sitt i skriftlig form, men oppgavene krever også at de har annen norskfaglig kompetanse – for eksempel lesekompetanse, ettersom alle oppgavene har tekstvedlegg. Det er utarbeidet vurderingskriterier for de språklige og formelle ferdighetene, og til hver enkelt oppgavetype. I tillegg presiseres det at sensor, uavhengig av oppgavetype, «forventer at teksten har hensiktsmessig struktur, god sammenheng og relevante eksempler og momenter» (s. 4), og at elevens kildebruk er «ryddig og oversiktlig» (s. 5). Hva angår kjennetegn for måloppnåelse er de stort sett identiske i 2018 og 2021. Ingen av kjennetegnene for måloppnåelse er fjernet, men noen nye har kommet til:

Tabell 1: Oversikt over de kjennetegn på måloppnåelse man finner i eksamensveiledningen fra 2021, men ikke i eksamensveiledningen fra 2018.

| | Karakter 2: Eksamenssvaret | Karakter 3-4 Eksamenssvaret | Karakter 5-6 Eksamenssvaret |
|--------------------------------------|--|---|--|
| Hovedinntrykk | viser lav kompetanse | viser nokså god/god kompetanse | viser svært god kompetanse |
| | kommuniserer med leseren | kommuniserer godt med leseren | kommuniserer meget godt med leseren |
| Språk og formelle ferdigheter | har enkel setningsstruktur | har stort sett variert setningsstruktur | har formålstjenlig og variert setningsstruktur |
| Kreative tekster | viser en enkel utvikling av motiv, tematikk, personer, miljøer eller hendelser | viser noe utvikling av motiv, tematikk, personer, miljøer eller hendelser | viser en utvikling av motiv, tematikk, personer, miljøer eller hendelser |

2.2 Vitenskapsteoretisk tilnærming

Dette masterprosjektet baserer seg på en konkret elevtekst og tre ulike norsklæreres vurdering av denne. Jeg legger til grunn at både elevteksten og lærernes vurdering av den er et resultat av både en historisk og en sosial kontekst, og prosjektet er dermed forankret i den sosiokulturelle læringsteorien. Mange vil hevde at sosiokulturell læringsteori er flere teorier som trekker i samme retning heller enn en læringsteori i seg selv, men fellesnevneren mellom de ulike teoriene er uansett en grunnholdning om at læring skjer i interaksjon og samhandling mellom mennesker (Dysthe et al, 2012, s. 49). I dette kapittelet belyses aktuelle sosiokulturelle læringsteorier.

2.2.1 Vygotsky og sosiokulturell læringsteori

Vygotsky og Bakhtin er forskere innenfor sosiokulturell læringsteori som representerer et kulturhistorisk syn på kunnskap der språket og kommunikasjonen er særlig sentral. Vygotsky (2001) mener at sosial handling er en forutsetning for læring, og i *Teoretiske perspektiver på dialog og dialogbasert undervisning* forklarer Dysthe et al. (2012) at Vygotsky bryter radikalt med tidligere teorier når han hevder at «alle høyere mentale funksjoner [...] oppstår først i sosial samhandling og senere blir internalisert» (s. 49). I følge Vygotsky kan altså kunnskap aldri eksistere i et vakuum, men den vil alltid være «infiltrert i en historisk og kulturell kontekst» (Dysthe et al. 2012, s. 49), og det viktigste redskapet for kunnskapsutvikling er, ifølge Vygotsky (1978, s. 40), språket.

2.2.2 Bakhtins dialogisme

Bakhtin bygger videre på Vygotskys tanker, og det er i boken *Problems of Dostoevsky's Poetics* Bakhtin utforsker sin dialogiske tenkning mest fullstendig, ifølge Dysthe et al. (2012, s. 56). I denne boken går han så langt som å omtale hele den menneskelige eksistensen som dialog: «To be means to communicate dialogically. When dialogue ends, everything ends. Thus dialogue, by its very essence, cannot and must not come to an end» (Bakhtin, 1984, s. 252). Bakhtin er også opptatt av at alle ytringer er ledd i en kompleks, organisert ytringskjede, og beskriver forholdet mellom de ulike leddene i denne kjeden slik:

[...] han er ikkje den første som talar, [...] og han føreset ikkje berre eksistensen av språkssystemet for det språket han nyttar seg av, men også eksistensen av føregående ytringar – sine egne og framande – som hans ytring relaterer seg til på den eine eller den andre måten. (Bakhtin & Slaattelid, 1998, s. 11).

I følge Bakhtin er altså kommunikasjon grunnleggende dialogisk, og kommunikasjonen må forstås som et samarbeid mellom den talende og den lyttende. I tillegg mener han altså at en ytring alltid vil være et svar på forhenværende eller fremtidige ytringer – en holdning som er interessant for et prosjekt som dette, som handler om hvordan ulike lesere oppfatter den samme ytringen. Problemstillingen og forskningsspørsmålene i dette prosjektet retter seg i hovedsak mot lærernes sluttvurdering av en elevtekst, og det følgende vil derfor i all hovedsak handle om nettopp disse vurderingene. Siden det er elevteksten som danner grunnlaget for

samtalen med de aktuelle lærerne, vil jeg likevel her vie noe plass til å redegjøre kort for elevteksten som sjanger i lys av Bakhtins dialogismebegrep da dette danner et viktig bakteppe for lærernes vurdering.

Elevteksten kan altså, ifølge Bakhtin, ikke forstås som en isolert ytring, men man må ta høyde for at teksten er et svar på tidligere ytringer. I tillegg kan teksten leses som et svar på fremtidige ytringer ved at den er utformet med hensyn til den modell-leseren som eleven ser for seg underveis i skrivingen. I dette konkrete tilfellet har eleven svart på en eksamensoppgave, og vedkommende har med stor sannsynlighet møtt lignende oppgaver gjennom sin tid som elev i videregående skole. Det er også rimelig å anta at eleven har lest andre tekster som svarer på lignende oppgaver, og at han har vært delaktig i å vurdere styrker og svakheter ved disse tekstene, eller i det minste bevitnet andres vurdering av dem. Denne erfaringen ligger, bevisst eller ubevisst, til grunn når eleven produserer en tekst til en tenkt leser, og leseren blir dermed å regne som medprodusent av teksten, ifølge Bakhtin. Elevteksten er derfor ikke monologisk, men av natur dialogisk.

2.2.3 Meads speilingsteori og «de signifikante andre»

Dysthe et al. (2012) oppsummerer Bakhtins syn på språk, meningsskaping og kommunikasjon med at det: “ikke er jeg, men vi som skaper mening. Mening oppstår i samspillet mellom de som kommuniserer, og responsen fra den andre er det aktiverende prinsippet som skaper forståelse.” (s. 56). Her ser man tydelig Bakhtins sosiokulturelle ståsted, og man ser også likhetstrekk med Mead (1934) og hans speilingsteori, en teori som danner et interessant bakteppe for det vurderingsarbeidet som en enkelt lærer gjør av et elevarbeid. Mens russiske Vygotsky og Bakhtin er representanter for et kulturhistorisk syn på kunnskap, har amerikanske Mead, og hans forgjenger Dewey, et pragmatisk kunnskapssyn (Dysthe et al., 2012, s. 49) og en grunnleggende holdning om at interaksjon og samhandling mellom mennesker er avgjørende for utviklingen av *selvet*:

Selves can only exist in definite relationship to other selves. No hard-and-fast line can be drawn between our own selves and the selves of others, since our own selves exist and enter as such into our experience only in so far as the selves of the other members of his social group. (Mead, 1964, s. 164)

Mead har altså en holdning om at hvert enkelt *selv* utvikler seg i takt med, og i relasjon til, de *selvene* man omgir seg med, og med bakgrunn i denne holdningen regnes Mead som opphavsmannen til begrepet «signifikante andre». Berger og Luckman (1966) bruker begrepet “significant others” (s. 66) i sin klassiske tekst *The social Construction of reality*, og presiserer i en fotnote at «The term ‘significant others’ is taken from Mead. For Mead’s theory of the ontogenesis of the self, cf. his *Mind, Self and Society*” (Berger & Luckmann, 1966, s. 220)¹. I denne boken, *Mind, Self & Society*, skriver Mead følgende:

The self is something which has a development; it is not initially there, at birth, but arises in the process of social experience and activity, that is, develops in the given individual as a result of his relations to that process as a whole and to other individuals within that process. (Mead, 1934, s. 135)

Selvbevissthet er altså, for Mead, uløselig knyttet til sosiale prosesser, og utvikling skjer når vi vurderer oss selv ut fra de tilbakemeldingene vi får fra menneskene rundt oss.

Det er vanlig å skille mellom de generaliserte andre og det som altså omtales som de signifikante andre (Rye, 2013, s. 178), og det er tilbakemeldingene fra de signifikante andre som har størst betydning for vår personlige utvikling. Hvorvidt man kan snakke om signifikante andre for både barn og voksne er det derimot noe uenighet omkring:

I Meads fremstilling er «de signifikante andre» lokalisert i barndommen. Det er først når disse erstattes av «den generaliserte andre» at man har blitt et fullverdig medlem av samfunnet. De andre bidragsyterne synes å ha en oppfatning av «de signifikante andre» som vedvarende gjennom hele livet. (Rye, 2013, s. 178).

¹ At Meads verker skal være arnested for begrepet «signifikante andre», møter motbør blant enkelte. Johan Fredrik Rye, professor i sosiologi ved NTNU, hevder i artikkelen *Mead, Berger & Luckmann og de signifikante andre* at «Svært mange sosiologer krediterer den amerikanske sosialpsykologen George Herbert Mead for begrepet, med referanse til det som regnes som hans hovedverk: *Mind, Self, and Society* (Mead, 1934). Denne referansen er imidlertid gal. Mead skrev aldri om «de signifikante andre». Ikke ett eneste sted i *Mind, Self, and Society* brukes uttrykket.» (Rye, 2013, s. 169-170). Det er likevel et uomtvistelig faktum at Meads ideer danner grunnlaget for begrepets eksistens. Av hensyn til tematikken for dette prosjektet nøyer jeg meg med å gjengi denne kritikken av referansen jeg bruker, og vil ikke gå dypere inn i diskusjonen.

Jeg støtter meg her på oppfatningen om at fenomenet *signifikante andre* kan vedvare gjennom livet, og i en skolekontekst kan disse signifikante andre være for eksempel en lærer, en medelev eller en kollega. Når vi opptrer i et læringsfellesskap, det være seg blant elever eller lærere, er det derfor ikke likegyldig hvilke andre mennesker som befinner seg i den samme konteksten.

2.2.4 Gadamers filosofiske hermeneutikk og Bakhtins translingvistikk

I læringsfellesskapet er heller ikke enkeltindividenes forkunnskaper eller tidligere erfaringer likegyldige, og vi beveger oss da inn i et landskap hvor det blir uunngåelig å nevne hermeneutikken – en vitenskapelig metode som har utviklet seg til å bli en vitenskapsteoretisk posisjon. Søren Kjerup (2008) definerer hermeneutikk som «[...] fortolkning og læren om fortolkning» (s. 63), og i vår tid kretser hermeneutisk teori rundt at vi «på ein eller annan måte må tolke oss *sjølve* inn i tolkingsprosessen» (Fuglseth, 2018, s. 254). Ideen om at vi må tolke oss selv inn i tolkningsprosessen er tydeliggjort og understreket gjennom arbeidene til Gadamer, og illustreres gjennom metaforen om den hermeneutiske spiralen (Fuglseth, 2018, s. 255). Gadamer er opptatt av at vi i et fortolkningsarbeid har med oss en forforståelse som avgjør hvordan vi fortolker en tekst eller en situasjon, og sannheten er derfor, ifølge Gadamer, avhengig av fortolkerens forforståelse. Gadamer omtaler denne forforståelsen som «fordommer», men er samtidig opptatt av at en fordom ikke nødvendigvis er «en gal dom.» (Gadamer, 2001, s. 120). Snarere, skriver han, må «fordommene [...] legitimeres gjennom fornuftens erkjennelse» (Gadamer, 2001, s. 122). I *Estetika Slovesnogo Tvorchestva* fra 1979 kan vi se tilsvarende holdninger hos Bakhtin, når han skriver om translingvistikken:

Vi lærer å støype talen vår i genreformer, og når vi hører ein annans tale, gissar vi genren allereie frå dei første orda, vi anar omfanget [...], den kompositoriske oppbygginga, og dannar oss på førehand ein idé om slutten. Vi har med andre ord, heilt frå byrjinga, ein idé om taleheilskapen som seinare vert differensiert i taleprosessen. (Bakhtin, 1979, sitert i Bakhtin & Slaattelid, 1998, s. 80).

Man kan ikke sette likhetstegn mellom Bakhtins translingvistikk og Gadamers filosofiske hermeneutikk uten at man samtidig vil risikere å undergrave kompleksiteten hos dem begge.

I lys av sitatet over kan man likevel si at Bakhtin, i likhet med Gadamer, er opptatt av at menneskenes forforståelse er avgjørende for tolkningene man gjør.

2.2.5 Vurdering av elevtekster i lys av sosiokulturell læringsteori

Når man skal vurdere en elevtekst, det være seg som faglærer eller som sensor, befinner man seg midt i dette landskapet som teoriene til Bakhtin, Mead og Gadamer belyser, og det kan være krevende å fullt ut se seg selv og forstå sin egen rolle i situasjonen. Ifølge Bakhtin vil de som samtaler eller kommuniserer også samarbeide om å skape og gjenskape mening (Igland & Dysthe, 2001, s. 113), og det er også utgangspunktet for lærergjerningen. Som faglærer søker man å utvide elevenes kunnskaper og forståelse, samtidig som man selv i stor grad er den som avgjør de ytre rammene for den kunnskapen og forståelsen som elevene til enhver tid vil ha anledning til å tilegne seg. Læreren valg av metoder, utvalg av fagstoff og utarbeiding av oppgavetekster springer ut av hennes egen forkunnskap, og påvirker dermed direkte elevenes kunnskapsutvikling. Læreren har dessuten det som i forskningslitteraturen omtales som en «dobbel rolle» (Kvistad & Otnes, 2019, s. 104). I tillegg til å være veileder er man som faglærer den som teksten er skrevet for, og, ifølge Bakhtin, har man dermed selv vært med på å indirekte påvirke teksten allerede før den er skrevet.

Når man skal snakke om lærernes formelle sluttvurdering av en elevtekst møter man i tillegg et nytt nivå man må ta hensyn til – nemlig andre vurderere. På samme måten som elevene vil legge både kjente og ukjente lesere til grunn for tekstproduksjonen sin, vil læreren også kunne legge andre vurdereres antatte holdninger, både kjente og ukjente, til grunn når de foretar sin vurdering av elevteksten. Lærernes eget læringsfelleskap kalles ofte et profesjonsfelleskap, og de fleste lærere vil være en del av både et konkret fellesskap blant kollegaer på arbeidsplassen sin, og ett eller flere mer abstrakte fellesskap. De abstrakte fellesskapene kan være nåtidige og bestå av lærere som har likt pedagogisk grunnsyn, lik fagbakgrunn, som underviser i samme skoleslag eller lignende. I tillegg kan de være fortidige, og bestå av for eksempel medstudenter fra studietiden, forelesere man så opp til, tidligere kollegaer, tidligere med-sensor på eksamen og videre. Teknologien i vår tid gjør dessuten at flere har tilgang til et mangfold av slike lærings- og profesjonsfelleskap, noe som øker sannsynligheten for at en lærers signifikante andre befinner seg i nettopp disse fellesskapene. Videre virker det sannsynlig, med utgangspunkt i et sosiokulturelt perspektiv, at disse signifikante andre kan ha

indirekte betydning for vurderingen av en konkret elevtekst. Ifølge Bakhtin er altså respons fra den andre det aktiverende prinsippet for å utvikle mening (Dysthe et al., 2012, s. 56), og ifølge Mead gjør *selvet* justeringer basert på responsen til de signifikante andre. Med utgangspunkt i Meads teorier kan man derfor anta at tanken på andre læreres vurderingspraksis kan ha direkte innvirkning på hvilken karakter som settes på en elevtekst.

2.3 Tidligere studier

I dette delkapittelet vil jeg belyse tidligere forskning på vurderingsfeltet generelt, og på sluttvurdering og vurdering av skriftlige ferdigheter spesielt. I kapittelet trekkes det inn forskningsartikler fra både norsk og internasjonal kontekst, og sammen med de sosiokulturelle teoriene danner den tidligere forskningen det teoretiske grunnlaget for undersøkelsen min.

2.3.1 Vurdering av skriveferdigheter

Med sine 60 involverte land og over 100 deltakende utdanningssystemer er organisasjonen IEA, International Association for the Evaluation of Educational Achievement, essensiell i omtale av internasjonal utdanningsforskning. Organisasjonen står bak de mye omtalte TIMSS-, PIRLS- og TALIS-studiene, og er således mye omtalt i forskningsøyemed så vel som i politiske spørsmål og i praksisfeltet. På 90-tallet gjennomfører IEA en internasjonal skriveundersøkelse, og gjennom denne undersøkelsen kommer det frem at det er vanskelig å sammenligne elevers ferdigheter fra land til land, og at skriveferdighetene utmerker seg som særlig vanskelige å sammenligne (Findlay, 1993, s. 314). Til tross for dette velger jeg å i det følgende presentere utenlandske og norske studier om hverandre, og delvis sammenligne resultatene i studiene. Jeg ønsker likevel å presisere at dette er gjort med IEA-undersøkelsen i mente. De fleste utenlandske studiene det refereres til omhandler av den grunn først og fremst læreres vurderingspraksis og bruk av vurderingskriterier. Det gjøres også forsøk på å belyse disse to fenomenene i en norsk kontekst, men tematisk representerer de norske studiene et noe større mangfold. Denne vektingen oppleves som naturlig ettersom eget datagrunnlag er innsamlet i en norsk kontekst.

2.3.2 Eksamen som skjult læreplan

Dette prosjektet omhandler sluttvurdering som fenomen på generelt nivå, og ikke utelukkende eksamen som sluttvurdering. Mange av studiene det vises til er likevel nettopp

studier på eksamen, og årsaken til at jeg har valgt å inkludere disse i så stor grad er det faktum at eksamen udiskutabelt er en sluttvurdering. Når eksamensdatoen kommer er man ferdig med opplæringen i faget, og man kan derfor ikke sette likhetstegn mellom vurdering av eksamen og den vurderingen som foregår underveis i et opplæringsløp. Det er likevel en tydelig sammenheng mellom de to. Både LK06, som i skrivende stund er gjeldende for elever i vg3, og LK20, som har trådt i kraft i vg1 og vg2, er læreplaner med stor metodefrihet og lite detaljert innhold. Som lærer står man derfor i prinsippet forholdsvis fritt når man skal lede elevene på vei mot målet, «[...] men i den grad eksamen er målet, vil den for de fleste lærere oppleves som en styrende fortolkning av læreplanen» (Bach, 2015, s. 31). Eksamen har altså en washback-effekt og fungerer i praksis som en konkretisering av læreplanen. Ifølge Bach (2015) kommer dette til uttrykk ved at norsklærerne ser ut til å bruke eksamensoppgavene som styringsdokument for egen opplæring, «både innholdsmessig og når det kommer til de retningslinjene som blir gitt i løpet av vurderingsperioden.» (2015, s. 31). Det samme fenomenet understrekes også av Berge og Hertzberg når de i etterkant av KAL-prosjektet gjør seg noen tanker om framtidens utfordringer:

Det er en kjensgjerning at eksamensoppgavene i norsk fungerer som mønster for den typen oppgaver lærerne gir elevene underveis, og avgangsprøva sier dermed noe om hva slags skriveoppgaver som er sentrale på ungdomstrinnet. Likeledes må sensorvurderingen kunne tolkes som en norm for hvordan norsklærere vurderer. I sitt daglige virke er samtlige sensorer norsklærere, og de må forventes å bære med seg samme skjønn enten de sensurerer eksamensoppgaver eller veileder elever i sin egen klasse. [...] Dermed blir avgangsprøva ikke bare et mål på hvor elevene har nådd etter endt grunnskole, men også et bilde på hva slags skriveundervisning som gis.» (Berge & Hertzberg, 2005, s. 387-393)

Det er altså tydelig at lærerne har god kjennskap til eksamensoppgavene, og tydelige tanker om hva som forventes når disse skal løses. Sentralt gitt eksamen fremstår også som en institusjon som høster stor respekt, og nærmest ærefrykt, blant lærerne. Dermed oppstår det et asymmetrisk avhengighetsforhold mellom den styrende eksamenen på den ene siden, og den undervisningen og de vurderingene som gjennomføres i løpet av et opplæringsløp på den andre siden.

2.3.3 Bedømmersamsvar og rettferdig vurdering

Beskrivelse av situasjonen

Lærere bruker mye tid på vurderingsarbeid, og slagordet vurdering for læring «gjaller i skolen», ifølge Kjell Lars Berge (2014, s. 259). Kanskje kan man forstå det som en naturlig konsekvens av et endret elevsyn, og et ideal om elevaktiv læring. Når det gjelder vurdering *av* læring viser det seg derimot at vi vet forholdsvis lite om hva lærerne «ser etter når dei vurderer, og korleis dei vektlegg ulike aspekt ved tekstane» (Matre & Solheim, 2014, s. 219). At ikke alle lærere navigerer etter de samme prinsippene synes likevel å være et faktum, og er belyst av blant andre Gardner (2007). Han trekker, i artikkelen «Is teaching a `partial` profession?», frem læreryrkets forholdsvis lave status, og stiller spørsmål ved «What are teachers apparently lacking that doctors, arcitects – even lawyers – have?» (s. 19). Svaret er, ifølge Gardner, «Public trust», og en av årsakene til at læreryrket mangler tillit er nettopp utfordringen med manglende vurderingssamsvar, mener han.

Manglende vurderingssamsvar er ikke bare et problem for renomméet til lærerstanden, men store deler av utdanningssystemet hviler på nettopp dette kravet til reliabilitet. Vi må derfor kreve av en gyldig vurdering «at den er pålitelig i den forstand at den gir noenlunde samme resultat uavhengig av hvem det er som vurderer og uavhengig av når vurderingen foregår» (Evensen, 2009, s. 17). Manglende vurderingssamsvar blant de som vurderer elevarbeid er påvist i gjentatte studier, blant annet av McNamara (1996) og Weigle (1998). I norsk kontekst ble nasjonale skriveprøver stoppet allerede etter ett år på grunn av «problemer med å oppnå akseptabelt samsvar mellom lærerne som vurderte prøvene» (Skar & Aasen, 2018). Videre har Berge studert vurdering av norskeksamen i videregående skole, og konkludert med at det er ganske stor variasjon i hvilke egenskaper sensorene tillegger de ulike oppgavetyper, og at en del oppgavetyper har «forholdsvis lavt faktisk bedømmersamsvar» (Berge, 1996, s. 630). Man kan således hevde at kravet om reliabilitet ikke innfris. Dette er problematisk sett i lys av at man som bedømmer har betydelig makt over de man bedømmer, og at vurderingen som gjøres potensielt kan ha stor betydning for elevens, eller eksamenskandidatens, læringsutbytte, selvbilde og framtidige muligheter (Evensen, 2009, s. 15).

Hvordan skal karakterskalaen forstås?

Til tross for lavt bedømmersamsvar kommer man ikke unna at en betydelig del av jobben til en lærer innebærer å vurdere elevtekster, og for lærere i ungdomsskolen og på videregående skole innebærer det å overføre vurderingen til et siffer i en skala. Karakterskalaen vi bruker i skolen er en såkalt «ordinalskaala» (Berge, 2009, s. 45). Med det menes at de ulike sifrene ikke har faste referansepunkter i seg selv, men at de bare kan forstås lys av hverandre. Skalaen forteller oss bare at «6» er bedre enn «5», «5» er bedre enn «4» og så videre. Det skalaen derimot ikke forteller, er avstanden mellom de ulike tallene. Ifølge Berge (2009, s. 45) vil mange lærere intuitivt hevde at avstanden mellom «4» og «5» er kortere enn avstanden mellom «5» og «6», uten at det er et nedfelt synspunkt. Problematikken knyttet til at det er vanskelig å skille de ulike tallkarakterene fra hverandre ble belyst av Wenche Vagle allerede i forbindelse med innføringen av tallkarakterer i videregående skole, ved overgangen til *Reform 94*. Hun undersøkte den gang hvordan overgangen ble håndtert ved eksamenssensuren, og konkluderte med at innholdet i de nye karaktergradene burde spesifiseres mer presist. Hun tok til orde for å utarbeide «forholdsvis detaljerte signalementer» (Vagle, 2005, s. 97) for de forskjellige karaktergradene innenfor hvert enkelt fag – altså, det vi i dag kaller *vurderingskriterier*, eller *kjennetegn på måloppnåelse*.

Vurderingskriterier

Resultatene i studier om bruken av vurderingskriterier er derimot sprikende. På den ene siden er det gjennomført studier som tyder på at det har forkommet adekvat skolering av norsksensorene i forkant av eksamen. Anderson og Hertzberg (2005) maler for eksempel et bilde av ungdomsskolesensorer som følger opp de instruksjer de har fått, og melder om «svært få eksempler på uvettig bruk av sensorskjemaet.» (Anderson & Hertzberg, 2005, s. 301). Særlig interessant, skriver de, «er det at sensorene tar på alvor at formelle svakheter i rettskriving og tegnsetting ikke skal blåses opp og gjøres til et hovedkriterium.» (Anderson & Hertzberg, 2005, s. 301). Et annet signal om at sensorskoleringen i forkant har vært hensiktsmessig, er at sensorene virker trygge på å skulle vurdere stilistiske virkemidler: «de leter etter dem og finner dem, og vurderer om det er brukt adekvat.» (Anderson & Hertzberg, 2005, s. 301). Det kan altså virke som om det er en viss enighet om hva som kjennetegner tekstkvalitet, blant sensorene som Anderson og Hertzberg har studert. På en annen side er ikke enighet om hvilke kriterier som avgjør tekstkvalitet synonymt med tilsvarende enighet om hvor godt en tekst

faktisk svarer på oppgaven. Enighet om hvor god sjangerforståelse eleven viser eller hvor god den formelle språkkompetansen er, er derimot avhengig av en felles tolkning av vurderingskriteriene (Nilsen, 2005, s. 118-119). Manglende, eller ikke tilfredsstillende, felles tolkning av vurderingskriteriene er påvist i flere studier, og i det følgende gjøres det rede for noen av funnene i slike studier.

Som et trinn i å styrke påliteligheten til de summative resultatene trekker Paul Black et al. (2011, s. 457) frem nettopp felles tolkning av vurderingskriteriene som en forutsetning. Lærerne som deltok i studien til Black et al. mente likevel at kriteriene som var tilgjengelige i den nasjonale læreplanen ikke var tilstrekkelige, og de uttrykte behov for å bryte ned kriteriene i «sub-level criteria» (Black et al., 2011, s. 457). Det er da interessant å merke seg at forskerne på sin side uttrykker at det for dem så ut som at problemet i større grad var knyttet til at lærerne hadde manglende eierskap til læreplanen. Dette, mener de, har sammenheng med skolens krav om regelmessige tester der resultatene skal rapporteres i form av nettopp slike underkriterier som lærerne etterspurte også på læreplannivå. Man ser dermed en skolepraksis som forsterker et allerede eksisterende skille mellom læreplanforståelse og vurderingspraksis.

For å forbedre den summative vurderingspraksisen ved skolene arrangerte forskerne møter der lærerne kunne diskutere med hverandre, og intensjonen var altså at disse møtene skulle fungere som en opplæringsaktivitet. I enkelte avdelinger hadde man ikke gjort noe tilsvarende tidligere, og en del lærere ble overrasket over hvor ulikt enkeltlærere vurderte elevarbeidet. Det manglende samsvaret, skriver Black et al. (2011) kommer særlig til uttrykk der vurderingskriteriene er for lite presise, noe som altså er i tråd med det som lærerne selv rapporterte. Når man i vurderingskriteriene blir bedt om å vurdere om arbeidet er «simple» eller «quite complex» må man som lærer selv fylle disse ordene med innhold, og det er her forskjellene oppstår, ifølge Black et al. (2011, s. 458). Erfaringer fra SKRIV-prosjektet viser samme tendens i norsk skole: «[...] da utvalgte tekster ble vurdert av flere lærere [...] viste resultatene så stort sprik mellom de enkelte vurderingene at prøvene ikke ble regnet som pålitelige» (Solheim & Matre, 2014, s. 77). I mange tilfeller er et manglende tolkningsfellesskap, ifølge Solheim og Matre (2014, s. 81), et resultat av en innarbeidet tradisjon der lærerne navigerer etter erfaring og intuisjon heller enn gitte

vurderingskategorier. Det er rimelig å hevde at en slik praksis kan gå på bekostning av vurderingens validitet (Evensen, 2009, s. 16), ettersom man da ikke nødvendigvis vurderer de eksakt samme punktene man har sagt at man skal vurdere. I norskfaget prøves elevene som hovedregel i ferdigheter så vel som kunnskap om et fagstoff, og denne balansegangen kan dermed være særlig vanskelig som norsksensor.

På skriftlig eksamen i norsk i videregående skole bygger alle oppgaver på minst ett tekstvedlegg, og det er forventet at elevene skal vise kompetanse i lesing så vel som skriving. Som norsklærer og -sensor har man trolig god forståelse for det gjensidige avhengighetsforholdet mellom lese- og skrivekompetanse. Hvorvidt dette forholdet kommuniseres tydelig til elevene, i løpet av opplæringen og i forkant av en eksamenssituasjon, kan man derimot sette spørsmålsteget ved. I lys av både SKRIV-prosjektet og studien til Black et al. er det rimelig å anta at «intuisjonen» til lærerne og sensorene er ulik hva angår vekten av elevens lesekompetanse i forhold til skrivekompetanse, og at det dermed også er ulikt i hvor stor grad man som lærer gjør elever oppmerksomme på avhengighetsforholdet mellom de to ulike kompetansene.

Hvordan skal man rangere ulike teksters kvaliteter i forhold til hverandre?

Norskfaget på videregående er ikke avsluttende før etter vg3. Det betyr i praksis at elevene ikke kan komme opp til eksamen i faget på et tidligere tidspunkt, og at elevene skal prøves i kompetansemålene fra tidligere år når de skal opp til minst en norskeksamen i vg3. Inntil nylig har kompetansemålene i faget vært inndelte i tre hovedområder: 'Skriftlig kommunikasjon', 'Muntlig kommunikasjon' og 'Språk, litteratur og kultur' (LK06). LK20 bruker riktignok ikke samme inndeling, men alle de tre hovedområdene er fortsatt representerte, for eksempel gjennom disse kompetansemålene fra LK20:

Mål for opplæringen er at eleven skal kunne

- *skrive essay om utforsker og reflekterer over innhold i tekster*
- *bruke fagkunnskap og presist fagspråk i utforskende samtaler, diskusjoner og muntlige presentasjoner om norskfaglige emner*
- *utforske og reflektere over hvordan tekster fra den realistiske og den modernistiske tradisjonen framstiller menneske, natur og samfunn*

Til sammen utgjør altså kompetansemålene i læreplanen i norsk «et bredt spekter av ulike kompetanser, og elevene skal ha kunnskap om mange aspekt ved norskfaget.» (Bach, 2015, s. 30), og dermed er det en annen utfordring ved vurderingsarbeidet som melder seg: «lærere opplever det som vanskelig å bli enige om hvordan man skal rangere ulike teksters kvaliteter i forhold til hverandre.» (Berge, 1993a, sitert i Berge 2009, s. 45). For å eksemplifisere vil det altså si at lærere kan være uenige i hvorvidt det er de språklige eller innholdsmessige kvalitetene ved elevteksten som skal vektes tyngst i den samlede vurderingen. I tillegg har studier avdekt at lærere ikke alltid er enige i hva som skal forventes på ulike trinn i utdanningssystemet. I artikkelen *Vurdering av skrivekompetanse: En kompleks utfordring* belyser Evensen (2009) denne problematikken gjennom eksempler fra fjerde, sjuende og tiende årstrinn i grunnskolen:

På fjerde årstrinn varierer samsvaret for disse delaspektene mellom .28 og .36, avhengig av sjanger og oppgave; på sjuende trinn ligger den tilsvarende variasjonen mellom .30 og .46. En av grunnene til dette lave samsvaret er sannsynligvis at det ikke er etablert like klare forventningsnormer for disse trinnene. Lærere er altså ikke enige innbyrdes om hvilke forventninger en skal ha til for eksempel formelle sider ved skrivning på fjerde eller sjuende årstrinn, mens på tiende årstrinn har de en klarere formening om dette. (s. 18)

Det vil være rimelig å hevde at denne tendensen sannsynligvis også er gjeldende i videregående skole. Det er utarbeidet vurderingskriterier til bruk på eksamen etter vg3, og siden svært mange norsklærere har sensorerfaring er disse velkjente blant lærerstanden. Det er derimot ikke utarbeidet tilsvarende kriterier for elever som går i vg1 og vg2, og man må da forvente at bedømmersamsvaret kan være lavere på de årstrinnene. Hva som blir oppfattet som en god tekst er dessuten ikke entydig. Ulike sensorer trekker frem mangfoldige og uensartede kvaliteter ved elevtekstene, og Nilsen (2005, s. 158) argumenterer derfor for at tekstenes kvaliteter på mange måter realiseres i øynene som ser, og ikke i teksten alene - et syn på tekst og tekstskaping som er i tråd med den sosiokulturelle læringsteorien, og Bakhtins dialogisme spesielt.

Hva kan og bør endres?

Ambisjonene om rettferdig vurdering av eksamen i norsk er store (Bach, 2015, s. 42). Likevel bekrefter gjentatte studier at tolkningsfellesskapet blant sensorene i norsk ikke er spesielt sterkt, og i den ferske rapporten *Sensorreliabilitet på skriftlig eksamen i videregående opplæring* (Bjørnson & Skar, 2021) bekreftes altså det svake tolkningsfellesskapet blant norsklærere atter en gang.

Det faktum at to lærere setter samme karakter på en tekst betyr ikke at de nødvendigvis har lagt de samme prinsippene til grunn for vurderingen, og for å ivareta denne todelingen kan man skille mellom interrater reliability (IRR) på den ene siden og interrater agreement (IRA) på den andre:

Interrater reliability [...] provides an indication of the extent to which the variance in the ratings is attributable to differences among the objects rated. [...] In contrast, interrater agreement represents the extent to which the different judges tend to assign exactly the same rating to each object. (Tinsley & Weiss, 2000, s. 98).

I denne definisjonen ser vi et skille mellom det som av Bjørnson og Skar (2021) omtales som sensorsamsvar, IRR, og sensorenighet, IRA (s. 5). IRA forteller altså noe om i hvor stor grad ulike lærere setter den samme karakteren på den samme teksten, mens IRR forteller noe om hvorvidt de legger de samme prinsippene til grunn når de setter karakteren. De studiene jeg tidligere har vist til bekrefter at man må fortsette å jobbe for å styrke begge disse faktorene blant lærerne i skolen.

Kjell Lars Berges dom over skrivepedagogikken i norsk skole fra 1996 var knusende. Han omtalte arbeidet med avhandlingen som et «sorgarbeid» (Berge, 1996, s. 647), og avsluttet med å peke fremover:

[...] vurdering bør i enda høyere grad bli sett på som kollektivt arbeide, [...]. Det bør være en forutsetning at det utvikles slike kollektive vurderingskulturer ved alle norskseksjoner. [...] forskningsbasert kunnskap om hva skriving er, hvilke implikasjoner

denne formen for sosial praksis har og hvordan den læres, må bli en del av alle norsk læreres kunnskapsforråd. (Berge, 1996, s. 648)

Videregående skole har gått gjennom store endringer siden studien til Berge ble gjennomført i 1996, og en del av endringene kan sies å være i tråd med de ønskene han fremmet den gang. Da Berge utgav sin studie, i 1996, var den nye generelle læreplandelen fra 1993 (L93) forholdsvis fersk, og denne fikk et langt liv:

Den erstattet med en gang første kapitel i den generelle delen av M87, mens resten av den generelle delen av M87 ennå gjaldt. Den ble generell del i L94 – en ny læreplan for videregående opplæring (fra 1994). Den ble selvfølgelig generell del i L97 (fra 1997), som den egentlig var laget for [...]. Da vi, [...] fikk en ny læreplan i 2006 (LK06: Læreplan for Kunnskapsløftet av 2006), valgte man å beholde L93 som generell del. (Engelsen, 2020. s. 211-212)

Til tross for at den overordnede delen av læreplanverket bestod, fikk de grunnleggende ferdighetene sin plass i skolen med innføringen av *Læreplanverket for Kunnskapsløftet 2006* (LK06). Denne reformen omtales av Berge som en literacy-reform: «Den er en reform der skriftkyndighet åpent erklæres som grunnlaget for den type avansert læring som vi oppretter skoler for å ta seg av.» (Berge, 2005a, s. 165). Så sent som i 2017 ble L93 erstattet av den nye overordnede læreplandelen, og med dette ble også begrepet «profesjonsfelleskap» (Overordnet del, 2017 s. 15) innført – et fellesskap som «reflekterer over [...] utviklingsbehov, og bruker forskning, erfaringsbasert kunnskap [...] som grunnlag for målrettede tiltak.» (Overordnet del, 2017 s. 18-19). Sett i lys av denne utviklingen kan vi kanskje hevde at *Kunnskapsløftet* og den nye overordnede delen, mellom annet, ser ut til å være uttrykk for et ønske om å bøte på en manglende forskningsbasert kunnskap om skriving i norsk skole, så vel som manglende profesjonsfelleskap blant lærerne.

I tillegg til disse overordnede endringene i skolen, har det også vært mer konkrete og fagspesifikke endringer de siste årene. For norskfaget sin del er det, i denne sammenheng, verdt å nevne spesifikt at elevene ikke lenger skal skrive «informative og resonnerende tekster» (LK06), men nå blir bedt eksplisitt om å skrive «fagartikler» (LK20). Særlig interessant

er det, siden studien til Berge (1996) avdekte at det faktiske bedømmersamsvaret i oppgavetyperen «artikkel» var betydelig lavere enn i andre oppgavetyper, til tross for at sensorene ikke opplevde denne oppgavetyperen som vanskeligere å vurdere (Berge, 1996, s. 615-616). At korrelasjonssnittet er avhengig av skrivemåte, eller sjangertype, ble også avdekket gjennom KAL-prosjektet der de fant at «Det var betydelig høyere korrelasjoner for resonnerende skriving enn det var for fortellende og mellom-personlig skriving.» (Berge, 2005b, s. 112-113). Kanskje er gjeninnføringen av mer avgrensede sjangerkrav i norskfaget et uttrykk for at man tenker seg at lærere i dag, som en følge av skoleutviklingen, er bedre rustet til å både lære opp elevene og vurdere dem i denne sjangeren. Hvorvidt det er tilfelle eller ikke, er det derimot for tidlig å si noe om.

2.3.5 Vurdering i profesjonsfelleskapet

Opplæringsaktiviteter blant lærere

Gjennom studien til Black et al. (2011) viser det seg at enkeltlærere blir overrasket over i hvor stor grad det er variasjon i de ulike lærernes vurderinger av et elevarbeid. Hvorvidt ulike opplæringsaktiviteter er effektive for å redusere variasjonen på sikt er man derimot usikker på, ifølge Thomas Eckes (2008). Eckes viser til blant andre Lumley & McNamara (1995) og Weigle (1998) når han skriver at «[...], rater training has been shown to be much less effective at reducing rater variability than expected; that is, raters typically remained far from functioning interchangeably even after extensive training sessions» (Eckes, 2008, s. 156). Videre viser han til Elder et al. (2005) når han hevder at også individuell feedback på vurderingsarbeidet er lite effektivt for å redusere variasjonen i vurderingsarbeidet. Elder et al. (2005) er på sin side positive til resultatene i sin egen studie. For det første, skriver de, har de fleste lærerne positive holdninger til tilbakemeldingene de får, de anser det som nyttig å motta slike tilbakemeldinger, og forskerne trekker frem at tilbakemeldingene bidrar til at lærerne tenker grundigere gjennom sine egne vurderinger. Avslutningsvis trekker de likevel frem det økende samsvaret etter tilbakemeldingene som en mulig utfordring. De ulike lærerne ser nemlig ut til å innrette seg i så stor grad at det noen ganger kan gå på bekostning av vurderingens gyldighet – de blir «overcautious or fearful» og mindre tilbøyelige til å bruke «the full range of scores available to them» (Elder et al., 2005, s. 190). I kapittel 2.2.5 *Vurdering av elevtekster i lys av sosiokulturell læringsteori* ble det antydnet at en lærers forventning om

responsen fra den signifikante andre kan ha innvirkning på hvilken vurdering læreren gjør, og ifølge Elder et al. ser altså dette ut til å stemme.

Ulike vurderertyper

Med utgangspunkt i påstanden om at opplæringsaktiviteter er lite effektive søker Thomas Eckes selv, i sin studie, å teste en hypotese om at lærere med erfaring i vurderingsarbeid vil tolke vurderingskriterier ulikt, og at dette vil gi en variasjon i vurderingene. I tillegg vil han teste en annen hypotese om at de ulike lærerne: «would fall into types (classes, clusters) that were clearly distinguishable from one another with respect to the importance raters attached to scoring criteria” (Eckes, 2008, s. 156). Han har altså en tanke om at han, basert på å studere vurderingspraksisen til ulike lærere, kan plassere lærerne i «vurderingsbåser», så å si. Tanken er at ulike «raters» oppjusterer viktigheten av enkelte kriterier, og nedjusterer viktigheten av andre kriterier når de vurderer den skriftlige prestasjonen til en elev. Resultatene i studien hans bekrefter langt på vei disse hypotesene. De totalt 65 deltakerne skilte seg betydelig fra hverandre hva gjaldt vektleggingen av de ulike kriteriene, og det kom dessuten frem at de manglet en felles oppfatning av de ulike kriteriene. (Eckes, 2008, s. 177). Med bakgrunn i disse resultatene mener Eckes å kunne plassere de ulike deltakerne i seks ulike grupper, der hver gruppe representerer en markant ulik vektlegging av kriteriene enn de andre gruppene. I tillegg viste studien at ingen av disse gruppene fordelte oppmerksomheten jevnt over alle kriteriene, men det så ut til at hver gruppe forholdt seg til ulike undergrupper av de ulike kriteriene (Eckes, 2008, s. 178). Eckes mener det er et problem at ulike lærere nokså konsekvent vektlegger ulike kvaliteter ved et elevarbeid. Til tross for at han innleder artikkelen med å poengtere at opplæringsaktiviteter av «raters» har lite å bety for hvor stor variasjon det er i vurderingene, konkluderer han likevel med at «Rater training can help to resolve these problems.» (Eckes, 2008, s. 179). Da er det likevel en forutsetning, mener han, at man tar høyde for at lærere har ulik vurderingsadferd, og at man konsentrerer vurderingsopplæringen om disse vurderertypene. På denne måten, hevder han, vil vi kunne bidra til en mer balansert bruk av kriteriene som anses som relevante i vurderingen av elevarbeid.

Intuitivt vurderingsarbeid

I flere av de nevnte studiene trekkes lærernes intuisjon frem som en potensiell årsak til det lave bedømmersamsvaret. I artikkelen *'I know a 2:1 when I see it': Understanding criteria for*

degree classifications in franchised university programmes avdekker Kathryn Ecclestone (2001) at det kan være en utfordring å få erfarne og kompetente vurderere til å delta i diskusjoner om vurdering nettopp på grunn av en holdning om at vurderingsarbeidet er intuitivt - derav artikkeltittelen «'I know a 2:1 when I see it': [...]» (Ecclestone, 2001, s. 312). Ecclestone argumenterer for at vurderingskriterier ikke i seg selv gir et tolkningsfellesskap. I stedet, hevder hun, må man ha en annen og mer strategisk tilnærming når man skal sosialisere ansatte inn i et akademisk fellesskap enn å basere seg utelukkende på vurderingskriterier for å oppnå et tolkningsfellesskap. Det kan være nyttig med «assessment guidelines and criteria», men i tillegg har man i vurderingsarbeidet behov for regelmessige diskusjoner med kolleger. Vurderingskriterier vil derimot kunne spille en viktig rolle i disse diskusjonene, og i opplæringen av nye «assessors». På denne måten vil man kunne innlemme nybegynnerne i et allerede eksisterende tolkningsfellesskap.

I lys av Meads speilingsteori er det dessuten sannsynlig at de erfarne vurdererne har rollen som signifikant andre. Det er derfor en særlig utfordring dersom nettopp disse motsetter seg samarbeid og undergraver verdien av vurderingskriterier. Det vil være et hinder for utviklingen av profesjonsfellesskapet. Ecclestone er opptatt av vurderingskriterier som verktøy for å innlemme nye vurderere i et tolkningsfellesskap. Av hensyn til både profesjonsfellesskapet og de elevene eller studentene som skal vurderes vil jeg likevel argumentere for at både nye og erfarne vurderere må delta i å utvikle et stadig nytt tolkningsfellesskap. De nye vurdererne kan ikke innlemmes i en eksisterende tradisjon uten å inviteres til å utvikle den gjeldende tradisjonen. Dersom målet er utvikling kan man se både til Bakhtin og Gadamer når man argumenterer for at man som profesjonsfellesskap kontinuerlig må være i dialog med hverandre. Dersom man ikke møtes til dialog vil utviklingen stagneres, og man risikerer at de nye vurdererne innlemmes i et dysfunksjonelt profesjonsfellesskap som stadig videreføres.

2.4 Eksamensoppgave og elevtekst

Datainnsamlingen i denne studien gjøres ved intervju, og til grunn for deler av intervjuene ligger en elevbesvarelse som de intervjuede lærerne har vurdert med karakter i forkant av intervjuet. I dette delkapittelet vil jeg presentere eksamenssettet og den konkrete oppgaven som elevteksten er et svar på, samt selve elevteksten.

2.4.1 Eksamensoppgave

Elevesvarelsen ble levert inn som et svar på hovedmåleeksamen for elever i videregående skole, våren 2018. Eksamensformen er skoleeksamen med fem timers varighet og alle hjelpemidler tillat, med unntak av oversettelsesprogrammer samt «åpent internett og andre verktøy som kan brukes til kommunikasjon» (Utdanningsdirektoratet, 2018b, s. 2). Eksamenssettet består av en obligatorisk kortsvarsoppgave og fire langsvarsoppgaver, og elevene skal velge én av langsvarsoppgavene.

Kortsvarsoppgaven krever at elevene skriver ca. 250 ord. I det aktuelle eksamenssettet ber kortsvarsoppgaven elevene om å gjøre greie for naturalistiske trekk i et vedlagt tekstutdrag. Tekstutdraget er hentet fra *Hellemyrsfolket* (1887) av Amalie Skram. Videre følger de fire langsvarsoppgavene som alle krever at elevene skriver ulike teksttyper: en retorisk analyse, en kreativ tekst, en tolkningsoppgave som ber elevene sammenligne en eldre og en nyere lyrisk tekst, samt en resonnerende tekst om samfunnskritisk litteratur. Elevteksten som de intervjuede lærerne har lest i forkant av intervjuet er et svar på den sistnevnte oppgaven. I sin helhet lyder oppgaven slik:

I 1871 hevdet den danske litteraturforskeren Georg Brandes at litteraturens viktigste funksjon var «å sette problemer under debatt». Skriv en resonnerende tekst der du reflekterer over hvordan litteratur kan være samfunnskritisk. Bruk eksempler fra de vedlagte tekstene i svaret ditt. Du kan også bruke andre eksempler du kjenner.

Kommentar: Det skal komme tydelig fram av svaret ditt at du bruker eksempler fra alle de tre tekstvedleggene når du skriver.

De tre tekstvedleggene det refereres til i oppgaveteksten er et utdrag fra *Hellemyrsfolket*, det samme som kortsvaret baserer seg på, samt et utdrag fra *Mor streiker* (1981) av Ebba Haslund og et utdrag fra *Bienes historie* (2015) av Maja Lunde.

2.4.2 Elevtekst

Den aktuelle elevteksten er vurdert til middels måloppnåelse av to separate sensorpar. I den opprinnelige sensuren fikk eleven karakteren 3, men karakteren ble justert opp til 4 etter at

eleven klaget. Jeg har innhentet tillatelse fra eleven til å dele elevteksten med de intervjuede lærerne og til å opplyse om de ulike karakterene, men siden besvarelsen er elevens åndsverk vil jeg ikke legge den som vedlegg til denne studien. I det følgende gis det derfor en beskrivelse av elevtekstens oppbygging og innhold.

Elevbesvarelsen innledes av elevens kortsvar. Kortsvaret består av totalt 271 ord, og er delt i tre avsnitt: innledning, hoveddel og avslutning. I innledningen presenteres tekstvedlegget, og eleven presenterer hva han skal gjøre i hoveddelen. I hoveddelen svarer eleven på oppgaveteksten, og i avslutningen oppsummerer og konkluderer eleven med at tekstvedlegget har «et dystert preg» og at naturalismens deterministiske tankegang kommer til uttrykk i tekstvedlegget.

Langsvaret er vesentlig lenger, og teksten består av totalt 1348 ord fordelt på seks avsnitt. Første avsnitt fungerer som innledning, og her gjentas Georg Brandes-sitatet fra oppgaveteksten og alle de tre tekstvedleggene med tema blir introdusert. Eleven gjør grundigere greie for de tre tekstvedleggene i påfølgende avsnitt. Avsnitt 3-5 handler om andre tekster som eleven oppfatter «setter problemer under debatt», og her nevnes et bredt spekter av forfattere og titler, fordelt på både skjønnlitterære sjangre og sakprosa: «Dostoevsky, Tolstoj og Gorkij», *Våre små søsken* av Inger Hagerup, *Nattens sønner*, *Kunsten å være Inder*, *Das Kapital* og *Kommunistenes manifest* av Karl Marx, for å nevne noen. Det gjentas flere ganger at litteraturen er «en stemme for de stemmeløse». Siste avsnitt er ganske langt, og det fremstår som uklart hvorvidt dette er en avslutning eller et avsnitt i hoveddelen. Siste avsnitt har likevel en litt annen form enn de foregående avsnittene, og en rimelig lesning vil derfor være at eleven har tenkt at dette skal være en avslutning eller en konklusjon.

Elevteksten har to separate kildelister: en til kortsvaret, og en til langsvaret. I disse refererer eleven til tekstvedleggene, læreverket *Panorama* og til en ordbokdefinisjon av ordet *litteratur*.

3.0 Metode

3.1 Valg av forskningsmetode

3.1.1 Kvalitativ forskningsmetode

De lærde strides om hvorvidt det er de kvalitative eller kvantitative studiene som er best egnet til å beskrive sosiale fenomener og menneskelig atferd. Dette har blitt utførlig belyst i «metodedebatten» (Postholm & Jacobsen, 2018, s. 90-93), og valg om metodisk tilnærming må tas ut fra hva som egner seg i en gitt problemstilling (2018, s. 17-18, 111). Problemstillingen i dette prosjektet, *Hvilke faktorer påvirker tre norsklæreres vurdering av en elevtekst?*, søker å «forstå menneskers handlinger og meningsskaping i deres naturlige kontekst» (Postholm & Jacobsen, 2018, s. 113). Med en problemstilling som søker å finne ut hvordan lærere påvirkes i sin vurdering av en elevtekst er derfor kvalitativ metode det naturlige valget. Valget er likevel tatt med visshet om at man samtidig velger bort noe annet, og det er ikke et valg fattet med letthet. Ved kvantitative studier vil man i langt større grad få funn som kan fungere generaliserende, noe jeg i utgangspunktet hadde et ønske om. Samtidig har jeg et genuint ønske om å forstå hvilke mekanismer som ligger til grunn for det vurderingsarbeidet hver enkelt lærer gjør i skolen. Med dette formålet for øye mener jeg derfor at kvalitative studier som går i dybden og gir svar på «hvordan» og «hvorfor», er verdifulle.

3.1.2 Liten N- og fenomenologiske studier

Prosjektet kategoriseres som en *liten N-studie*, som skiller seg fra casestudier ved at kontekstens betydning tones ned (Postholm & Jacobsen, 2018, s. 74). I dette konkrete tilfellet skal jeg studere et lite antall tilfeller, i likhet med prinsippene som ligger til grunn for en casestudie, men i motsetning til i en casestudie er det fenomenet «sluttvurdering» som skal være i sentrum, ikke selve vurderingskonteksten. Postholm og Jacobsen illustrerer liten N-studier gjennom et eksempel om kollegaveiledning på ulike skoler: «Her rettes interessen mot kollegaveiledning som fenomen, noe som kan utspille seg på ulike måter, men som likevel er et fenomen det er meningsfullt å sammenligne på tvers av kontekster.» (Postholm og Jacobsen, 2018, s. 74-75). Dette eksempelet vil jeg hevde har direkte overføringsverdi til denne studien. Videre i eksempelet skriver de at kollegaveiledning som fenomen kan antas å variere mellom for eksempel store og små skoler, og mellom skoleslag. I denne studien kan

man anta at de ulike lærernes sluttvurderinger og begrunnelser kan variere ut fra for eksempel utdanningsbakgrunn, sensorerfaring og samarbeidsrutinene på arbeidsplassen.

Liten N-studier som forskningstilnærming kan deles inn i undertypene fenomenologiske og narrative studier. Mens fenomenologiske studier «legger vekt på ulike forståelser og fortolkninger av et fenomen» (Postholm & Jacobsen, 2018, s. 77) kan narrativ tilnærming defineres som «spoken or written text giving an account of an event/action or series of events/actions, chronologically connected» (Czarniawska, 2004, s. 17). Man kan således argumentere for at dette prosjektet er både en narrativ og en fenomenologisk studie, jeg vil likevel ta til orde for at denne aktuelle studien er fenomenologisk fordi også undertegnede er opptatt av å studere ulike forståelser og fortolkninger av et fenomen, nemlig sluttvurdering.

3.1.3 Abduktiv tilnærming

En abduktiv tilnærming til forskning kjennetegnes, i motsetning til induktive og deduktive tilnærminger, av en vekselvirkning mellom empiri og teori:

[...] abduction starts from an empirical basis, just like induction, but does not reject theoretical preconceptions and is in that respect closer to deduction. The analysis of the empirical facts(s) may very well be combined with, or preceded by, studies of previous theory in the literature, not as a mechanical application on single cases, but as a source of inspiration for the discovery of patterns that bring understanding. (Alvesson & Sköldberg, 2018, s. 5)

En lignende beskrivelse av abduksjon finner vi hos Postholm og Jacobsen når de skriver at de som støtter seg til denne tilnærmingen vil hevde at det er meningsløst å snakke om induksjon eller deduksjon fordi «Abduksjon er [...] en kontinuerlig veksling mellom teori og empiri, der ingen av de to kan sies å ha forrang.» (Postholm & Jacobsen, 2018, s. 103). Abduktiv tilnærming er altså mer pragmatisk enn både induktiv og deduktiv tilnærming, og det er kanskje årsaken til at de induktive og deduktive metodene sjelden vil strekke til når man skal analysere og snakke om «menneskers uforutsigbare verden» (Kvale & Brinkmann, 2015, s. 225). Nettopp derfor er det også i mitt eget tilfelle naturlig å legge til grunn en abduktiv tilnærming med en pragmatisk veksling mellom teori og empiri.

3.1.4 Semistrukturert intervju

I følge Postholm og Jacobsen (2018, s. 103) vil det i en pragmatisk tilnærming være mer fornuftig å snakke om hvor *åpen eller lukket* en datainnsamling er, enn å ta stilling til om man benytter seg av en deduktiv eller induktiv tilnærming. Spørsmålet om hvor åpen eller lukket datainnsamlingen er handler om «hvor store begrensninger forskeren bevisst legger på de data som skal samles inn, *før* han eller hun starter undersøkelsen.» (Postholm & Jacobsen, 2018, s. 103). Hvis man som forsker ikke er avhengig av å strukturere datainnsamlingen særlig sterkt på forhånd vil et ustrukturert intervju være eksempel på en åpen tilnærming fordi den er «åpen for det nye og overraskende» (Postholm & Jacobsen, 2018, s. 104). Dette prosjektet baserer seg riktignok ikke på et ustrukturert, men et semi-strukturert intervju. Påstanden er likevel gjeldende da det semi-strukturerte intervjuet «[...] har som målsetting å forstå deltakernes perspektiv» (Kvale & Brinkmann, 2015, sitert i Postholm & Jacobsen, 2018, s. 121). Gjennom intervjuene vil det foregå en kontinuerlig analyse, og en åpen datainnsamling med en veksling mellom teori og empiri vil gi muligheter for å stille ulike spørsmål for å følge opp de tankene og holdningene som bringes frem i løpet av intervjuet.

3.2 Forberedelser og gjennomføring

3.2.1 Strategisk tilgjengelighetsutvalg og snøballmetoden

En forutsetning for å kunne være deltaker i intervjuet var at man hadde undervisningserfaring fra norskfaget på videregående skole, gjerne fra alle trinn, men med undervisningserfaring fra vg3 som et minimumskrav. Av hensyn til utvalget som helhet, ble det på forhånd definert at noen av deltakerne burde mangle sensorerfaring fra norskfaget i videregående skole, mens andre burde ha slik erfaring. Siden eksamen er en vurderingssituasjon der formålet udiskutabelt er å gi en sluttvurdering, og siden det dessuten er utarbeidet oppgavespesifikke vurderingskriterium til de ulike eksamensoppgavene var det særlig interessant dersom den eller de som hadde sensorerfaring også hadde erfaring med å delta på sensorskolering. Som en forberedelse til intervjuet ble deltakerne bedt om å lese og vurdere den samme elevteksten, og for å redusere muligheten for at de ulike deltakerne samsnakket om elevteksten i forkant av intervjuet var det dessuten nødvendig at de jobbet ved ulike skoler.

Kvalitative studier har som hovedregel forholdsvis få deltakere, og med sine tre deltakere er ikke denne studien noe unntak. Når utvalget er lite er det særlig viktig at «[...]vi anvender en

utvelgingsprosess som er hensiktsmessig for problemstillingen, slik at analysen av data kan gi en forståelse av de fenomenene vi studerer.» (Thagaard, 2018, s. 54). Deltakerne i denne studien er rekrutterte gjennom et strategisk tilgjengelighetsutvalg, og for å sikre en viss bredde i utvalget ble det altså på forhånd definert et sett med kriterier for å kunne delta i studien. Derfor kan man også betegne utvalgsprosessen som «kvoteutvelging» (Grønmo, 2016, s. 114). En av deltakerne ble dessuten rekruttert gjennom «snøballmetoden» (Thagaard, 2018, s. 56-57) ettersom en av de opprinnelig kontaktede lærerne ikke passet til alle kriteriene. Vedkommende lærer satte meg derfor i kontakt med en kollega som oppfylte alle kriteriene, og denne kollegaen utgjør en av de tre deltakerne i studien.

Som en følge av at det foreligger en del kriterier for å kunne delta var det nødvendig å rekruttere gjennom et tilgjengelighetsutvalg. En slik rekruttering kan likevel være problematisk ettersom at «[...]de som er villige til å delta i undersøkelsen, kan føle at de – i større grad enn det som er vanlig – mestrer sin livssituasjon, og at de derfor ikke har noe imot innsyn fra forskeren.» (Thagaard, 2018, s. 57). I rekrutteringsarbeidet opplevde jeg likevel ikke å få negative svar, med unntak av den ene læreren som ikke oppfylte alle kriteriene og som derfor satte meg i kontakt med en kollega. De lærerne som er deltakere i studien var altså de første jeg tok kontakt med, og jeg har derfor ikke grunn til å tro verken at de føler seg spesielt trygge i en vurderingssituasjon, eller at de er å anse som «eksperter» på området. Da de takket ja til å delta i intervju begrunnet dessuten samtlige det positive svaret med at vurderingsarbeid generelt og sluttvurdering spesielt er den største utfordringen i læreryrket. De uttrykte derfor at det var interessant å kunne bidra i prosjektet. Alle de tre deltakerne har, nødvendigvis, høyere utdanning, og det er en tendens at «[...] personer med høyere utdanning stort sett er mer villige til å delta i kvalitative studier enn personer med lite utdanning.» (Thagaard, 2018, s. 57). Med dette som bakteppe mener jeg det er naturlig å forstå deres velvilje til å delta i lys av deres utdanningsbakgrunn, heller enn et uttrykk for at de mener de mestrer vurderingsarbeidet bedre enn andre lærere. Som høyt utdannede personer har de dessuten erfaring med, og er fortrolige med forskning (Thagaard, 2018, s. 57), noe som kommer til uttrykk gjennom det faktum at de opplever det som positivt å kunne delta, og på den måten bidra til forskningen. Thagaard (2018, s. 57) understreker også at personer med høyere utdanning er vant til å reflektere over eget liv, og man kan derfor ikke utelukke at de

intervjuede lærerne så deltakelsen som en anledning til å videreutvikle egen vurderingspraksis.

3.2.2 Om deltakerne

Deltakerne omtales i det følgende som Lærer1, Lærer2 og Lærer3, til fordel for å opplyse om kjønn og gi dem fiktive navn. Tallene indikerer i hvilken rekkefølge de ble intervjuet. Deltakerne er anonymiserte og når jeg i det følgende refererer til forfatteren av elevteksten som «han» og de intervjuede lærerne som «hun» eller «henne» er det for å kunne skille dem fra hverandre i teksten. Det er altså ikke et uttrykk for deltakernes faktiske kjønn. For å sikre deltakernes anonymitet nevnes heller ikke andre identitetsmarkører som for eksempel spesifikk alder eller arbeidssted. Av hensyn til analysearbeidet er det likevel nødvendig å si noe om de ulike lærernes utdannings- og erfaringsbakgrunn, men dette gjøres på en slik måte at det ikke vil være mulig å gjenkjenne enkeltpersoner i studien.

Alle de tre lærerne har høyere pedagogisk utdanning, undervisningskompetanse og undervisningserfaring i norskfaget. I tillegg har alle både undervisningskompetanse og undervisningserfaring fra andre humanistiske fag, som for eksempel religion, historie eller ulike samfunnsfaglige skolefag, og underviser i norsk ved en videregående skole inneværende skoleår. Lærer1 har jobbet som lærer siden 2003, og har vært sensor for norsk skriftlig eksamen mange ganger, både på ungdomsskolen og på videregående. Hun er den av de tre intervjuede lærerne som har mest sensorerfaring. Lærer2 har jobbet som lærer siden tidlig på 1990-tallet, og er således den deltakeren med lengst undervisningserfaring. I tillegg har hun vært sensor på skriftlig eksamen i norsk to ganger. Lærer3 har jobbet som lærer i litt over 15 år, og har ikke erfaring fra å være sensor på skriftlig eksamen i norsk. Hun er derimot den av de tre lærerne som tydeligst oppgir å ha jobbet ved en skole med kultur for samarbeid, og med faste rammer for å vurdere tekster i fellesskap.

3.2.3 Gjennomføring av intervju

Intervjuene ble gjennomført på nyåret i 2022, og den som husker tilbake til de ukene vil huske stadige pressekonferanser og hyppige nyhetsoppslag om situasjonen rundt Covid -19. Landet stengte ned, og åpnet opp igjen. Vi gikk fra å lese oss opp på deltasmitte til å lese oss opp på omikronsmitte, og i takt med økende vaksinasjonsgrad og krav om gjenåpning av samfunnet,

steg smittetallene til stadig nye høyder. Med dette som bakteppe ble det bestemt at intervjuene skulle gjennomføres digitalt, med video. Avgjørelsen ble tatt av undertegnede, i samråd med deltakerne.

Man kan argumentere for at det i et fysisk intervju vil være lettere å etablere den gode, og nødvendige, kontakten mellom intervjueren og intervjupersonen (Kvale & Brinkmann, 2015, s. 160). Det ble likevel vurdert som forsvarlig, så vel som nødvendig, å gjennomføre intervjuene digitalt. De tre intervjudeltakerne er alle erfarne lærere som skulle snakke om vurdering, et fenomen de jobber med daglig, og som de er vant til å diskutere med andre. Det hører dessuten med til historien at begge parter, på intervjutidspunktet, hadde hatt jevnlig digitale møter og samtaler med andre i nesten to år, siden mars 2020, da koronapandemien for første gang førte til nedstenging av samfunnet. Å konversere med mennesker gjennom en skjerm føles derfor ikke like fremmed som det en gang gjorde, en holdning som også ble bekreftet at de intervjuede lærerne.

Gjennomføringen av intervjuene var avhengig av at lærerne hadde forberedt seg i form av å lese og vurdere en elevtekst. For å være trygg på at de skulle få tid til å gjøre disse forberedelsene ble elevteksten, oppgaveformuleringen og tekstvedleggene sendt til deltakerne fire uker i forkant av intervjuet. Eksamensveiledningen ble med vilje utelatt fordi den inneholder kjennetegn på måloppnåelse, og jeg ville ikke legge føringer for bruken av slike. For å trygge deltakerne samt forhindre at de skulle føle seg villedet, fikk de samtidig tilsendt store deler av intervjuguiden (Kvale & Brinkmann, 2015, s. 105). Noen av spørsmålene var dessuten av en slik art at det ble vurdert som effektivt dersom deltakerne hadde fått tenkt gjennom disse i forkant, for eksempel: «Hvor lenge har du jobbet som lærer?» og «Hva legger du i 'å vurdere rettfærdig'?»». Andre spørsmål ble utelatt fordi de ville kunne legge uheldige føringer for forberedelsen som ble gjort eller svarene som ble gitt, og dermed gi mindre autentiske og spontane svar (Kvale & Brinkmann, 2015, s. 105), for eksempel: «Har du diskutert den aktuelle elevbesvarelsen med kollegaer / andre lærere?» og «Denne teksten ble opprinnelig vurdert til karakteren 3 på eksamen. Hva tenker du om det?».

3.2.4 Transkribering

Hva angår transkripsjonen av intervjuene er det, ifølge Kvale og Brinkmann (2015, s. 208) ikke utarbeidet noen universell form som skal benyttes, men intervjueren må ta en del valg i transkripsjonsarbeidet. Hvordan man velger å transkribere intervjuene må nødvendigvis sees i sammenheng med intervjuenes intensjon. I dette prosjektet er jeg først og fremst ute etter at uttalelsene til de intervjuede lærerne skal være lettlest, og at transkripsjonen skal gjengi et så korrekt bilde av lærernes holdninger, meninger og vurderinger som mulig. I transkripsjonen velger jeg derfor å utelate pauser, og detaljer omkring disse. I all hovedsak vil jeg også utelate følelsesuttrykk, men det gjøres et fåtall unntak der følelsesuttrykkene er nødvendige for å understreke at en uttalelse er ment for eksempel humoristisk eller ironisk, og i disse tilfellene er følelsesuttrykkene markerte med hakeparentes:

Og han trekker jo inn at det kan ha innvirkning på politikk, men at det var Dostojevskij, Tolstoj og Gorkin som startet den russiske revolusjon? Tja, det kan jo være at det var andre ting som spilte inn her [ler].» (Lærer1, intervju)

Ønsket om å gjengi lærernes holdninger så korrekt som mulig vurderes ikke til å være i strid med kravet om reliabilitet og validitet. Man kan derimot argumentere for at det er uetisk å gjengi intervjuene ordrett i sin helhet. Kvale og Brinkmann (2015, s. 205) minner om uttrykket «*traduttori traditore* – oversettere er forrædere», og jeg vil nødig at de intervjuede lærerne skal føle seg forrådt i transkripsjonen. Det er årsaken til at jeg velger å gjengi uttalelsene på en sammenhengende måte, til fordel for å transkribere det muntlige språket ordrett (Kvale & Brinkmann, 2015, s. 213). Dette valget innebærer at transkripsjonene er resultat av en del praktiske, overordnede valg som har blitt fattet. Jeg tar for eksempel ut en del muntlige formuleringer som «på en måte» og «liksom», der disse ikke er nødvendige for å forstå konteksten. I tilfeller hvor de intervjuede lærerne begynner på en setning, avbryter seg selv og begynner på den samme setningen en gang til, inkluderer jeg bare den fullstendige setningen i transkripsjonen. Noen steder vil jeg også tillate meg å legge til ord for at setninger skal bli fullstendige og/eller mer forståelige for leseren. Disse ordene markeres med parentes: «I tillegg (til norsk) så har jeg historie [...].» (Lærer1, intervju)

De fullstendige transkripsjonene ligger ikke vedlagt, men kan gjøres tilgjengelige på forespørsel.

3.2.5 Analyseprosess

Analyseprosessen ble innledet med det som Postholm og Jacobsen (2018) omtaler som en «deskriptiv analyse» (s. 139), en analyseform som har som hensikt å strukturere materialet og gjøre det rapportvennlig. Den deskriptive analysen ble gjort ved å lete etter mønster for å kunne sortere datamaterialet i ulike kategorier. Intervjuguiden ble utformet med hensyn til hvilke temaer som naturlig ville bygge på hverandre i samtalen, og intervjuene ble altså ikke gjennomført på en slik måte at første del ville gi svar på første forskningsspørsmål, og så videre. Gjennom den deskriptive analysen ble derfor datamaterialet sortert etter forskningsspørsmål, og gjennom denne analysen ble det tydelig at fenomenet «rettferdig vurdering» måtte få spille en større rolle enn det som i utgangspunktet var tenkt. Problemstillingen og forskningsspørsmålene ble derfor på dette tidspunktet delvis omarbeidet som en følge av at empirien tok en annen retning enn først antatt.

Videre i analysearbeidet ble det foretatt en meningsfortetting (Kvale & Brinkmann, 2015, s. 232) der intervjupersonenes uttalelser ble omgjort til kortere formuleringer, et analysearbeid som var særlig nødvendig for å kunne svare på første forskningsspørsmål. Deler av disse meningsfortettingene er gjengitt i kapittel 4.1 *Faktorer som påvirker vurderingen av en elevtekst*. Videre dannet meningsfortettingen grunnlag for temaanalytisk tilnærming, med formål om få en «utdypet forståelse at temaene i materialet» (Thagaard, 2018, s. 177), en tilnærming som var særlig nødvendig for å kunne belyse grundig første forskningsspørsmål, *Hvilke faktorer peker lærerne selv på som viktige i sin vurdering?*

En innvendig mot den temaanalytiske tilnærmingen er derimot, ifølge Thagaard (2018, s. 177), at temaene løsriver fra konteksten de er en del av. I søken etter å både finne og beskrive et helhetlig bilde foretas det derfor også en meningsfortolkning. I en meningsfortolkning går forskeren, ifølge Kvale og Brinkmann (2015), utover det som blir sagt, «og finner frem til meningsstrukturer og betydningsrelasjoner om ikke fremtrer umiddelbart i en tekst» (s. 234). Denne tilnærmingen har vært særlig viktig i analyse av datamaterialet som ligger til grunn for andre forskningsspørsmål, *Hvilke faktorer avgjør om en vurdering er rettferdig, ifølge lærerne?*

3.3 Troverdighet og gyldighet

3.3.1 *Reliabilitet*

I kvalitativ forskning kjennetegnes studier med høy reliabilitet ved at de kan etterprøves av andre, og gi samme resultat i senere studier (Holand, 2018, s. 99), men på dette området er det ikke hensiktsmessig å stille de samme kravene til kvalitative studier som man gjør til kvantitative. Det ligger dessuten i det semistrukturerte intervjuets natur at ulike deltakere både vil kunne oppfatte spørsmål ulikt, og legge vekt på ulike momenter i svarene sine. Det er derfor ikke rimelig å forvente identiske resultater i senere studier. Man kan derimot vurdere reliabiliteten i intervjuforskning ved å stille spørsmål til både gjennomføring av intervju, transkripsjon og analyse. (Kvale & Brinkmann, 2015, s. 276).

I likhet med de intervjuede deltakerne har intervjueren høyere pedagogisk utdanning, undervisningskompetanse og -erfaring fra norskfaget i videregående skole. Intervjueren har derfor befunnet seg i lignende vurderingssituasjoner som intervjudeltakerne, og i et forsøk på å skape en intervjusituasjon som i så stor grad som mulig skulle oppleves som «en samtale mellom likeverdige deltakere» (Kvale & Brinkmann, 2015, s. 22) ble alle intervjuene innledet med denne informasjonen. Man kommer likevel ikke unna at det er intervjueren som «sitter på makta» i intervjusituasjonen, og fullstendig likeverdighet er dermed vanskelig å oppnå. Å være deltaker i et forskningsprosjekt er dessuten ikke hverdagskost, og til tross for at deltakerne fikk vite om både formål med intervjuet og fikk tilsendt de aller fleste spørsmålene i intervjuguiden på forhånd, må man anta at svarene de gir i noen grad er preget av at de reagerer på at de blir studerte (Grønmo, 2016, s. 21).

3.3.2 *Validitet*

Kleven og Hjordemaal (2018, s. 133) bruker begrepet ytre validitet i omtale av resultatenes gyldighetsområde. «I de senere årene er det blitt mer vanlig å snakke om *overføring* av forskningsresultater i stedet for å bruke uttrykket *generalisering*. I kvalitativ metodologi brukes nesten bare uttrykket *overføring*, [...]» (Kleven & Hjordemaal, 2018, s. 133). Spørsmålet er altså hvorvidt de holdningene og praksisene som kommer frem gjennom dette prosjektet er representative for de deler av populasjonen, i dette tilfellet norsklærere i videregående skole, som ikke selv er deltakere i studien. Det er hevet over enhver tvil at man ikke kan generalisere, eller overføre, med utgangspunkt i en kvalitativ studie av denne størrelsen.

Intervjuguiden har likevel en del spørsmål som handler om hvordan vurdering som fenomen har blitt behandlet i studietiden, og hvordan man i praksis jobber med vurdering i kollegiet. Man kan derfor tenke seg at man via fortellingene til de aktuelle deltakerne får innblikk i en situasjon som flere kan kjenne seg igjen i, men det vil være forenklende å si at resultatene er representative for den totale populasjonen.

3.4 Forskningsetiske betraktninger

En intervjuundersøkelse er, ifølge Steinar Kvale og Svend Brinkmann (2015, s. 95, s. 110), en moralsk undersøkelse med fire usikkerhetsområder: informert samtykke, konfidensialitet, konsekvenser og forskerens rolle. I gjeldende kapittel søkes det å belyse disse fire usikkerhetsmomentene.

3.4.1 Informert samtykke

Informert samtykke baserer seg på prinsippet om individuell autonomi og prinsippet om å gjøre godt (Marzano, 2012, s. 443). Dette prosjektet behandler personopplysninger og er derfor meldepliktig til Norsk senter for forskningsdata (NSD). Jeg har benyttet meg av NSDs mal for informasjonsskriv med samtykkeskjema, og de formelle kravene til informert samtykke er derfor ivaretatt. Når det gjelder utvelging av deltakere gjennom snøballmetoden så kan også dette være problematisk med hensyn til «prinsippet om at alle som deltar i et forskningsprosjekt, skal gi sitt informerte samtykke til deltakelse.» (Thagaard, 2018, s. 57). I tilfeller hvor en forskers kontaktperson gir forskeren informasjon om andre personer uten at de har samtykket til det kan man argumentere for at det er et brudd på prinsippet om samtykke til deltakelse. I dette aktuelle tilfellet gav den foreslåtte personen samtykke til at kontaktpersonen kunne videreformidle deres navn og kontaktinformasjon til undertegnede, en håndtering som også anbefales av Vassenden og Andrews (2007, s. 154-156). Prinsippet om samtykke til deltakelse er derfor ivaretatt også når det gjelder personen som ble rekruttert og valgt ut gjennom snøballmetoden.

3.4.2 Konfidensialitet

Også de formelle kravene til *konfidensialitet* (Kaiser, 2012) ivaretas gjennom skjemaet for informert samtykke da det aktuelle samtykkeskjemaet etablerer en enighet «[...] med deltakerne om hva som kan gjøres med dataene som blir et resultat av deres deltakelse»

(Kvale & Brinkmann, 2015, s. 106). Når det kommer til konfidensialitet er det likevel viktig at det ikke bare er de formelle kravene som ivaretas, men at jeg som intervjuer trygger deltakerne og forsikrer dem om at deres private data vil bli behandlet konfidensielt og med respekt.

3.4.3 Konsekvenser

I en kvalitativ undersøkelse har man med enkeltmennesker å gjøre, og man må forholde seg til hvilke *konsekvenser* deltakelse i prosjektet kan medføre for hver enkelt – både på godt og vondt. «Det etiske prinsippet om velgjørenhet [...] betyr at risikoen for å skade en deltaker bør være lavest mulig» (Guidelines, 1992, s. 15, sitert i Kvale & Brinkmann, 2015, s. 107). I mitt prosjekt er det meningsinnholdet som er det viktigste, og for å ivareta prinsippet om velgjørenhet, unngår jeg å bruke direkte språklig transkripsjon da dette kan skape et uriktig bilde av de som har blitt intervjuet. Jeg transkriberer og gjengir derfor sitater fra intervjuene på bokmål uavhengig av dialekt – en løsning som ifølge Kvale og Brinkmann (2015, s. 214) også kan gjøre uttalelsene mer lettleste. Man kan argumentere for å bruke nynorsk dersom det språket ligger nærmere deltakernes dialekter, men av hensyn til konfidensialitet og anonymitet transkriberer og siterer jeg på den samme målformen for alle.

3.4.4 Forskerens rolle

Det siste av de fire usikkerhetsmomentene som Kvale og Brinkmann (2015, s. 108) nevner er *forskerens rolle*, og de understreker at denne rollen er avgjørende for kvaliteten på både den vitenskapelige kunnskapen og de etiske beslutningene som treffes i den kvalitative forskningen. Fordi forskeren selv er det viktigste redskapet for innhenting av kunnskap i et forskningsintervju, er betydningen av forskerens integritet sentral i gjennomføringen. Ved å være fortrolig med verdspørsmål samt etiske retningslinjer og teorier vil det være lettere for forskeren å foreta spontane valg i det som kan oppleves som en spenning mellom etiske og vitenskapelige hensyn. Når alt kommer til alt er det likevel, ifølge Kvale og Brinkmann (2015, s. 108) forskerens kunnskap, erfaring, ærlighet og rettferdighet som er den avgjørende faktor. Min egen lærerbakgrunn plasserer meg på mange måter «på innsiden» av forskningsfeltet. Av den grunn må jeg være kritisk til mine egne tolkninger for å unngå at jeg overtolker eller tillegger intervjudeltakerne meninger eller holdninger de ikke har. Jeg mener likevel at min erfaring som lærer og vurderer i den videregående skolen også kan være en styrke. Jeg har

forkunnskaper om «hvor skoen trykker», og har dermed også forutsetninger for å etablere en trygg situasjon for de som skal delta i forskningsprosjektet.

4.0 Resultat

4.1 Faktorer som påvirker vurderingen av en elevtekst

Formålet med dette kapittelet er å gi svar på spørsmålet: *Hvilke faktorer peker lærerne selv på som viktige i sin vurdering?*

4.1.1 Lærer1 sin vurdering av en elevtekst: vektinger og begrunnelser

Lærer1 har jobbet som lærer i nesten 20 år, og har lang erfaring som sensor på skriftlig eksamen i norsk, både fra ungdomsskolen og videregående. På generelt grunnlag mener Lærer1 at vurderingen av en elevtekst må være en helhetsvurdering der man tar i betraktning både tekstens språk, struktur og hvorvidt eleven svarer på det som oppgaven ber om. Det karaktermessige skillet mellom en god og en mindre god karakter går mellom karakteren 3 og karakteren 4, ifølge Lærer1. Når det gjelder hvordan hun mener man skal vekte de ulike elementene opp mot hverandre når karakteren skal settes, poengterer hun at hun selv aldri teller skrivefeil eller tenker at «hvis du har fem kommafeil så ryker sekseren», men presiserer at de ulike elementene ofte henger sammen: «Det er veldig sjelden at det er kjempedårlig innhold og knakende god rettskriving». Videre sier hun at hun har gått bort fra å bruke vurderingsmatriser aktivt i vurderingsarbeidet da hun oppfattet det som kunstig. Hun opplever at innholdet i disse matrisene, «i alle fall den eksamensmatrisen» er ganske omfattende, og hun opplyser også om at formuleringene oppleves som vanskelig tilgjengelig for elevene, for eksempel ordet «syntaks». Til tross for at hun altså ikke bruker disse matrisene aktivt i vurderingsarbeidet så har hun rutine for å alltid prøve å si noe om tekstens språk, struktur og hvorvidt eleven svarer på det oppgaven ber om. Hun understreker derfor at det er et mønster i de tilbakemeldingene hun gir, til tross for at hun ikke aktivt bruker Utdanningsdirektoratets kjennetegn på måloppnåelse, eller tilsvarende, i vurderingsarbeidet.

Når det gjelder den konkrete teksten som ligger til grunn for denne studien vurderes den av Lærer1 til å være en svak firer, men hun argumenterer også for at teksten har kvaliteter som tilsier at den kan bli vurdert til karakteren 3. Lærer1 opplever at vurderingen av den aktuelle

teksten er krevende, og hun beskriver det slik: «Jeg synes det er kjempevanskelig. Dette er nok en tekst som hadde gått på omgang med kollegaene mine hvis jeg skulle ha (rettet den på 'ekte')». Årsaken til at hun likevel har landet på karakteren 4 til fordel for karakteren 3 begrunner hun med prinsippet om å la tvilen komme elevene til gode:

Det kan være gode argument for å sette en firer, og det kan være gode argument for å sette en treer, tenker jeg. Alt etter hva du vektlegger. Så jeg kommer tilbake til dette med (at) helheten i teksten skal telle mest [...]. Det kan hende at jeg lar tvilen komme eleven til gode.

På spørsmål om hvorfor hun mener det er en vanskelig tekst å karaktersette svarer hun at det er fordi den «spriker i flere retninger». De konkrete styrkene og svakhetene som Lærer1 trekker særlig frem er fremstilt i tabellen under. Oversikten består i hovedsak av parafraaser og sammenfatninger hentet fra lydfiler og transkripsjon, og er en del av meningsfortettingen det er redegjort for i kapittel 3.2.5 *Analyseprosess*. Kategoriene *språk*, *struktur* og *innhold* har oppstått som et ledd i analyseprosessen, og de delene som er hentet ordrett fra intervjuet med læreren er markerte med anførelsestegn.

*Tabell 2: Innholdet i tabellen er hentet fra meningsfortettingen basert på intervju med Lærer1. Det redegjøres for meningsfortetting som analyseverktøy i kapittel 3.2.5 *Analyseprosess*.*

| Lærer1 | Styrker ved teksten | Svakheter ved teksten |
|-----------------|---|---|
| Språk | | <ul style="list-style-type: none"> - bruker et lite effektivt språk - har unødvendig gjentakelse av enkelte fraser - en del skrivefeil |
| Struktur | <ul style="list-style-type: none"> - den overordnede strukturen er god | |
| Innhold | <ul style="list-style-type: none"> - eleven trekker inn «en del» fagkunnskap - eleven er selvstendig - eleven har eksempler med god bredde | <ul style="list-style-type: none"> - skiller ikke tydelig mellom «form og innhold» - er «omtrentlig på en del fagkunnskap», og følger ikke opp den fagkunnskapen som trekkes inn: - trekker inn sakprosa - trekker inn retorikk i omtale av en dialog i en skjønnlitterær tekst |

| | | |
|--------------|--|--|
| | | - nevner i liten grad to av de tre vedlagte tekstene |
| Annet | - langsvaret er langt for å være en resonnerende tekst | |

Lærer1 mener altså at manglende presisjon er en svakhet ved både språket og innholdet i teksten, og vektlegger særlig elevens selvstendighet og faglige bredde når karakteren settes. Lærer1 nevner ingen svakheter ved struktur, og heller ingen styrker ved språk. Dette må likevel ikke oppfattes som at læreren mener at slike styrker og svakheter er fullstendig fraværende, men må heller forstås som at de ikke er avgjørende faktorer for karakteren som settes.

4.1.2 Lærer2 sin vurdering av en elevtekst: vektinger og begrunnelser

Lærer2 har nesten 30 års erfaring som norsklærer i videregående skole, og har vært sensor på skriftlig eksamen i norsk to ganger. På generelt grunnlag mener hun at en god besvarelse er kjennetegnet ved at den kommuniserer «greit» med leseren. Hun forteller om en opplevelse med en medsensor som «hang seg veldig opp i klumsete formuleringer», og poengterer at hun selv prøver å ikke henge seg opp i det. I tillegg vurderer hun om eleven svarer på de ulike delene i oppgavebestillingen. Videre sier hun at hun gjerne bruker vurderingsskjemaene fra Utdanningsdirektoratet som hjelpemiddel når hun skal vurdere en elevtekst: «Jeg guler ut det som gjelder for eleven [...], og så har jeg en foreløpig indre overbevisning om hva som er den rette karakteren». Skillet mellom en god og en mindre god karakter mener Lærer2 går mellom karakteren 3 og karakteren 4, og hun understreker at «man skal vise mye faglig kompetanse for å få 4».

Når det kommer til vurdering av den konkrete teksten som ligger til grunn for denne studien vurderer Lærer2 den til å være en firer. Hun understreker likevel at hun er i tvil og at «Det er nok ikke en sterk firer, men den er på firersiden». En konsentrert fremstilling av hennes begrunnelse for å sette karakteren 4 er fremstilt i tabellen under²:

² Oversikten består i hovedsak av parafraaser og sammenfatninger hentet fra lydfiler og transkripsjon, og er en del av meningsfortettingen det er redegjort for i kapittel 3.2.5 *Analyseprosess*. Kategoriene *språk*, *struktur* og *innhold* har oppstått som et ledd i analyseprosessen, og de delene som er hentet ordrett fra intervjuet med læreren er markerte med anførselstegn

Tabell 3: Innholdet i tabellen er hentet fra meningsfortettingen basert på intervju med Lærer2. Det redegjøres for meningsfortetting som analyseverktøy i kapittel 3.2.5 Analyseprosess.

| Lærer2 | Styrker ved teksten | Svakheter ved teksten |
|-----------------|---|--|
| Språk | | <ul style="list-style-type: none"> - lite variasjon i språket, uheldig gjentakelse av enkelte fraser - formelle feil - kommuniserer ikke lett |
| Struktur | <ul style="list-style-type: none"> - i kortsvaret: «OK struktur» | <ul style="list-style-type: none"> - mye oppramsing i innledningen - i langsvaret: «hoppende struktur uten god sammenheng» |
| Innhold | <ul style="list-style-type: none"> - bruker alle tekstvedleggene - trekker inn tekster som ikke nevnes i oppgaveteksten | <ul style="list-style-type: none"> - viser lite faglig kompetanse - trekker inn retorikk når han snakker om skjønnlitteratur |
| Annet | <ul style="list-style-type: none"> - teksten er lang - er «et seriøst forsøk på å skrive et godt svar» | |

Lærer2 mener altså at eleven viser manglende formell språklig kompetanse. Når Lærer2 sier at eleven også mangler «faglig kompetanse» så er det sortert under «innhold» fordi det av konteksten i intervjuet fremgår at hun mener «kunnskap om fagstoff». Når hun setter karakter vektlegger hun derimot at eleven fremstår som selvstendig når han trekker inn mange ulike tekster i tillegg til de som står nevnt i oppgaveteksten. Hun vektlegger også det faktum at teksten er lang, og at det virker som et «seriøst forsøk på å skrive et godt svar».

4.1.3 Lærer3 sin vurdering av en elevtekst: vektinger og begrunnelser

Lærer3 har jobbet som norsklærer i omtrent 15 år og har ikke erfaring som sensor på skriftlig eksamen i norsk. Hun oppgir likevel å tidligere ha jobbet i et kollegium med stort fokus på å vurdere i fellesskap, og sier at hun som nyutdannet lærer lærte mye av mer erfarne kollegaer. På generelt grunnlag mener hun at en god elevbesvarelse først og fremst må svare på oppgaven. Videre poengteres det av Lærer3 at svaret må fremstilles gjennom et klart og godt formulert språk, vise forståelse for tekstvedleggene og trekke inn fagkunnskap der eleven går i dybden. Skolen hun jobber ved nå har gjort enkle tilpasninger i Utdanningsdirektoratets kjennetegn på måloppnåelse, og Lærer3 bruker disse aktivt i vurderingsarbeidet. Skillet mellom en god og en mindre god karakter går, ifølge Lærer3, mellom karakteren 3 og 4.

Når det gjelder den konkrete elevteksten som ligger til grunn for denne studien vurderes den av Lærer3 til å være en sterk firer. Hun mener at langsvaret ligger mellom karakterene 4 og 5, og at kortsvaret ligger mellom karakterene 3 og 4. Når hun begrunner karakteren hun har satt refererer hun konsekvent til skolens kjennetegn på måloppnåelse, som i realiteten ligner svært mye på kjennetegnene utarbeidet av Utdanningsdirektoratet:

På femmer og sekser så skal du ha «formålstjenlig struktur og god sammenheng». Det har denne eleven. (Eleven) skal (også) ha «relevante moment og/eller eksempel med en god bredde». Det synes jeg eleven har. «Gjør greie for faglige emner på en presis måte». Det synes jeg ikke han får til. Det er mer «gjør greie for faglige emner på stort sett klar måte».

Her ser vi hvordan Lærer3 bruker vurderingskriteriene aktivt, vurderer elevteksten opp mot de ulike kjennetegnene og diskuterer med seg selv hvorvidt elevteksten passer inn på det ene eller andre nivået. Når hun skal sette den samlede karakteren viser hun til kollegaer med sensorerfaring: «Jeg har hørt fra disse kollegaene mine, som er sensorer, at kortsvaret er kun med på å vippe opp eller ned.». Lærer3 ender derfor opp med å gi elevteksten karakteren 4, men argumenterer altså også for at det er kvaliteter ved teksten som kan vurderes til karakteren 5. En konsentrert begrunnelse for hva hun trekker frem for å ha satt karakteren 4 er videre fremstilt i tabellen under³:

Tabell 4: Innholdet i tabellen er hentet fra meningsfortettingen basert på intervju med Lærer3. Det redegjøres for meningsfortetting som analyseverktøy i kapittel 3.2.5 Analyseprosess.

| Lærer3 | Styrker ved teksten | Svakheter ved teksten |
|-----------------|--|--|
| Språk | - stort sett gode formuleringer | - ufullstendige setninger |
| Struktur | - «har 'formålstjenlig struktur og god sammenheng' og 'relevante moment og/eller eksempel med god bredde'» - oversiktlig drøfting | |
| Innhold | - svarer på oppgaven - har et variert syn på litteratur | - bruker tekstvedleggene for lite aktivt |

³ Oversikten består i hovedsak av parafraaser og sammenfatninger hentet fra lydfiler og transkripsjon, og er en del av meningsfortettingen det er redegjort for i kapittel 3.2.5 *Analyseprosess*. Kategoriene *språk*, *struktur* og *innhold* har oppstått som et ledd i analyseprosessen, og de delene som er hentet ordrett fra intervjuet med læreren er markerte med anførselstegn

| | | |
|--------------|------------------------|---|
| | ○ viser breddekunnskap | <ul style="list-style-type: none"> - har en del påstander som ikke følges opp <ul style="list-style-type: none"> ○ manglende dybde - lite nyanse i drøftingen |
| Annet | | <ul style="list-style-type: none"> - sitatteknikk |

Lærer3 mener altså at manglende dybde er en svakhet ved innholdet i teksten, og at manglende presisjon er en svakhet ved både språk og kildeføring. Når karakteren skal settes vektlegger Lærer3 særlig at elevteksten er et tydelig svar på oppgaven, at eleven viser faglig bredde, og at teksten har en god struktur.

4.1.4 Sammenfatting og oppsummering

Samtlige av de intervjuede lærerne understreker at å sette karakter er en helhetsvurdering, og alle de tre lærerne nevner alle de tre hovedområdene *språk, struktur og innhold* når de begrunner karakteren de har satt. De har også en felles oppfatning om at skillet mellom en god og en mindre god karakter går mellom karakteren 4 og karakteren 3, altså midt på middels måloppnåelse.

Når de bes om å sette en karakter på den aktuelle elevteksten vurderer samtlige den til karakteren 4. Til tross for at de til slutt ender opp med å sette den samme karakteren, spriker likevel vurderingene fra «nesten 3» til «nesten 5», og alle tre trekker delvis frem ulike sider ved teksten når de skal peke på styrker og svakheter. Lærer1 vektlegger elevens selvstendighet og faglige bredde i vurderingen, men gir karakteren 4 til tross for manglende presisjon i både språk og innhold. Lærer2 mener at eleven viser manglende fagkunnskap i teksten, men legger vekt på tekstens lengde og det gode forsøket eleven har gjort når hun setter karakteren 4. Lærer3 trekker frem at eleven svarer på oppgaven og viser bredde i fagkunnskapen, men gir karakteren 4 i stedet for karakteren 5 på grunn av blant annet manglende språklig presisjon og lite dybdekunnskap.

Resultatene viser altså at lærerne vektlegger ulike elementer ved elevteksten, og gjør ulike vurderinger av denne, til tross for at de gir den samme karakteren.

4.2 Likheter og forskjeller mellom de ulike lærernes vektinger

Formålet med dette kapittelet er å peke på likheter og forskjeller i de ulike lærernes vurderinger. Kapittelet fungerer således som en bro mellom første og andre forskningsspørsmål, og danner et viktig grunnlag for den drøftingen som kommer i kapittel 5.0 *Drøfting*.

4.2.1 Element som vektlegges av flere lærere

Under gjennomføringen av intervjuene var jeg opptatt av at lærerne skulle få snakke fritt og få sette sitt eget preg på svarene som ble gitt, og jeg har derfor ikke lagt andre føringer enn de som gis naturlig i intervjuguiden. Som en følge av dette bruker lærerne av og til ulike begrep og forklaringer på de samme fenomenene. Funnene som presenteres i dette kapittelet er derfor et resultat av de tolkningene jeg har gjort av lærernes svar, og de gangene jeg oppfatter at de snakker om samme fenomen, men bruker ulike ord, har jeg satt likhetstegn mellom de begrepene og formuleringene som jeg oppfatter at har likt meningsinnhold.

Språk

Alle de tre lærerne oppgir *språk* som en av de kategoriene de vurderer når de skal sette en karakter. Lærer1 sier likevel at hun aldri «teller skrivefeil og tenker at hvis du har fem kommafeil så ryker sekseren», og Lærer2 opplyser på sin side om at hun er opptatt av at man ikke behøver å «henge seg opp i klumsete (formuleringer)». Lærer3 sier ikke noe om hva hun *ikke* ser etter, men er opptatt av et godt formulert språk der eleven «klarar å formidle det han vil».

Hva angår den aktuelle elevteksten mener Lærer3 at eleven har «stort sett gode formuleringer», men bortsett fra det er alle de tre lærerne enige om at språket har en del svakheter. Alle trekker frem at teksten har en del formelle feil, som skrivefeil på ordnivå, manglende tegnsetting og ufullstendige setninger. I tillegg er Lærer2 og Lærer3 opptatt av at eleven har et lite effektivt språk med mange unødvendige gjentakelser, og begge trekker frem de samme eksemplene: «dyster» og «en stemme for de stemmeløse».

Struktur

Også tekstens struktur kommenteres av alle lærerne. På dette punktet fremstår vurderingene likevel noe mer sprikende. Lærer2 mener at kortsvaret har «OK struktur», men beskriver strukturen i langsvaret som «hoppende» og «uten god sammenheng». Lærer1 og Lærer3 skiller ikke mellom kortsvar og langsvar, men begge trekker frem styrker ved strukturen når de skal begrunne karakteren de har satt. Lærer1 beskriver den overordnede strukturen som «god», mens Lærer3 viser til skolens kjennetegn på måloppnåelse og slår fast at teksten har «formålstjenlig struktur og sammenheng» samt «relevante moment og/eller eksempel med god bredde».

Innhold

I intervjuene gir alle de tre lærerne mest omfattende begrunnelser i omtale av elevtekstens *innhold*, men det er også på dette hovedområdet begrunnelsene og vektingene deres spriker i størst grad. Lærer3 mener at eleven svarer på oppgaven, men Lærer1 mener at han ikke skiller tydelig mellom «form og innhold» og legger til grunn at oppgaven krever det, uten at det står eksplisitt i oppgaveteksten. Oppgaven krever derimot at eleven skal bruke de tre tekstvedleggene aktivt. Lærer2 noterer seg at alle tre brukes, og trekker dette frem som en styrke. Lærer1 og Lærer3 mener derimot at tekstvedleggene brukes for lite aktivt. Lærer1 roser eleven for å trekke inn «en del» fagkunnskap, men alle de tre lærerne mener at eleven viser upresis eller manglende fagkunnskap på en del områder. Lærer1 beskriver fagkunnskapen som «omtrentlig», og begrunner det med at eleven trekker en del enkle slutninger. Lærer2 sier at eleven «viser lite faglig kompetanse», mens Lærer3 savner at eleven følger opp de påstandene som kommer i løpet av teksten. Denne manglende oppfølgingen mener Lærer3 er et uttrykk for manglende dybdekunnskap. Både Lærer1 og Lærer2 trekker dessuten frem at det er et uttrykk for manglende forståelse for fagstoffet at eleven bruker begrep fra retorikken i omtale av skjønnlitterære tekster.

De største forskjellene mellom de ulike lærerne kommer likevel tydeligst til uttrykk når de reflekterer over oppgavetekstens ytre rammer:

[...] Skriv en resonnerende tekst der du reflekterer over hvordan litteratur kan være samfunnskritisk. Bruk eksempler fra de vedlagte tekstene i svaret ditt. Du kan også bruke andre eksempler du kjenner.

Spriket mellom dem er i hovedsak knyttet til begrepet «litteratur». Alle de tre tekstvedleggene er skjønnlitterære, og Lærer1 mener derfor at eleven viser svak forståelse for oppgaveteksten ved å trekke inn saktekster i svaret sitt. Lærer2 og Lærer3 mener derimot at dette er en styrke ved elevbesvarelsen, og oppfatter dette som en «god» måte å løse oppgaven på. Lærer3 går så langt som å fremheve «det der med ulike typer litteratur», altså både skjønnlitteratur og sakprosa, som særlig positivt.

4.2.2 Element som vektlegges av bare en lærer

Lærer1 – selvstendighet

Allerede innledningsvis i intervjuet trekker Lærer1 inn «selvstendighet» og «selvstendig språk» som en styrke ved gode elevbesvarelser, og begrepet «selvstendighet» henter hun frem igjen når hun skal begrunne karakteren hun setter på den aktuelle elevteksten. Hun mener at eleven viser selvstendighet ved å for eksempel «hente eksempel fra egen virkelighet og eget liv», og drar frem «de selvstendige refleksjonene» som særlig positive. På spørsmål om hvorfor hun tror at eleven gikk opp fra karakteren 3 til karakteren 4 etter klage, trekker hun også frem selvstendigheten som en potensiell årsak: «Jeg tror det har talt positivt at eleven har vist selvstendighet [...]», og hun tenker seg derfor at det nye sensorparet har vektlagt dette i større grad enn det opprinnelige sensorparet.

Lærer2 – elevtekstens omfang og elevens forsøk

Både Lærer1 og Lærer2 trekker frem elevtekstens omfang i omtale av styrker og svakheter, men Lærer1 slår bare fast at «langsvaret er langt for å være en resonnerende tekst». Hun mener det er en styrke ved teksten, men bruker ikke dette som en av begrunnelsene for å sette den karakteren hun gjør. Det gjør derimot Lærer2, som den eneste av de tre:

Men grunnen til at jeg er på 4 likevel, selv om her er en del rusk og rask, det er at eleven har lagt en del arbeid i det her. [Han] skriver en mye lenger tekst en mange elever bruker å gjøre. Så her ligger et seriøst forsøk på å skrive et godt svar [...].

Her ser vi altså at lengde og innsats kommer inn som et nytt element, uten at det har blitt nevnt av verken Lærer2 eller de andre lærerne på spørsmål om hva de på generelt grunnlag vil si kjennetegner en god elevtekst

Lærer3 – forståelse og sitatteknikk

Alle de tre lærerne er opptatt av at eventuelle tekstvedlegg må brukes aktivt i elevbesvarelsen. Lærer1 sier for eksempel at eleven «nevner» tekstvedleggene i liten grad, og Lærer2 trekker frem som en styrke at eleven har «brukt eksempler fra alle tekstvedleggene som oppgaven ber om». Lærer3 har likevel en annen innfallsvinkel til tekstvedleggenes funksjon i oppgaveteksten, og i elevbesvarelsen: «De må forstå de tekstene de skal bruke. De som forstår dem svarer stort sett bedre enn de som misforstår, eller ikke forstår». Lærer3 er altså den eneste av de tre som i klartekst oppgir å vektlegge betydningen av at elevene *forstår* innholdet i tekstvedleggene, og som dermed vektlegger elevenes lesekompetanse i vurderingen sin.

I tillegg er Lærer3 den eneste som oppgir å vektlegge elevens sitatteknikk i vurderingen. Sitatteknikken beskrives som «dårlig», og oppgis som en av årsakene til at hun vurderer teksten til karakteren 4 i stedet for karakteren 5. Hun er dessuten inne på at sitatteknikken kanskje er så dårlig at eleven står i fare for å stryke på eksamen: «Jeg tenker det er en sterk firer, jeg da. Men det kan hende at det egentlig er stryk». Lærer3 har ikke erfaring som sensor, men begrunner denne avgjørelsen med å vise til andres sensorerfaring: «Jeg har hørt at de er ganske strenge på det [sitatteknikk og kildeføring] til eksamen».

4.2.3 Sammenfatting og oppsummering

Alle de tre lærerne gir tilbakemeldinger på tre hovedområder, uten at verken intervjuguide eller oppfølgingsspørsmål legger føringer for det: *språk, struktur og innhold*. Vurderingene de gjør er mest sammenfallende på hovedområdet språk, og spriker mest på det omfattende hovedområdet innhold. I vurderingen av både struktur og innhold er lærernes vurderinger delvis i strid med hverandre. For eksempel mener Lærer1 og Lærer3 at strukturen er god, mens Lærer2 beskriver strukturen som «hoppende». Når det kommer til innhold mener Lærer2 og Lærer3 det er en styrke at eleven trekker inn saktekster i refleksjonen sin, mens Lærer1 mener dette er en svakhet ved elevteksten.

I tillegg vektlegger hver av de tre lærerne minst ett element i totalvurderingen som de andre to ikke trekker inn: Lærer1 vektlegger selvstendighet, Lærer2 vektlegger elevtekstens omfang og elevens forsøk og Lærer3 vektlegger forståelse for tekstvedleggene og sitatteknikk.

4.3 Rettferdig vurdering

Formålet med dette kapitlet er å gi svar på spørsmålet: *Hvilke faktorer avgjør om en vurdering er rettferdig, ifølge lærerne?* Som en følge av at datainnsamlingen har foregått ved et semistrukturert intervju vektlegger ikke de ulike lærerne de samme aspektene ved begrepet «rettferdig vurdering». Derfor vil heller ikke forskningsspørsmålets to underkategorier, *Rettferdig vurdering på tekstnivå* og *Rettferdig vurdering på profesjonsnivå*, behandles eksplisitt i dette kapitlet. Når de ulike lærernes holdninger skal forstås i lys av teorier og tidligere forskning, i kapittel 5.0 *Drøfting*, vil de to underkategoriene til gjengjeld løftes frem.

4.3.1 Lærer1 sin forståelse av begrepet «rettferdig vurdering» og beskrivelse av egen praksis

Lærer1 har lang sensorerfaring, og dermed også lang formell erfaring med å vurdere tekster sammen med andre, og hun trekker frem sensorskoleringen som særlig viktig og nyttig. Gjennom denne opplever hun at hun får «kalibrere» seg, og på den måten blir hun tryggere i eget vurderingsarbeid. Disse erfaringene trekker hun frem på spørsmål om hva hun legger i begrepet «rettferdig vurdering»:

Da forstår jeg at min karakter ikke er noe annerledes enn det de ville fått på en annen skole. Fordi du har såpass mange kollegaer rundt omkring [...] som du vet at du ligger litt på nivå med. Så jeg synes det er veldig betryggende, og blir litt sikrere på at den karakteren jeg setter er rett.

For Lærer1 betyr altså «rettferdig vurdering» at en elevs prestasjoner vurderes til samme karakter uavhengig av hvilken lærer som vurderer, og uavhengig av hvilken skole eleven går på. Lærer1 ser på sensorskoleringen som en arena for å måle sin egen vurderingspraksis opp mot andres, og som en anledning til å justere sin egen vurderingspraksis dersom hun opplever

at den er i strid flertallets praksis. Hun trekker frem oppmannssystemet under fellessensuren som et velegnet tidspunkt for slik justering: «Hvis (en tekst) kommer ut som en firer, og jeg har litt hardnakket stått på en treer, så må jeg på en måte jekke meg, tenker jeg». Her trekker hun også frem begrepet «tolkningsfellesskapet», og antyder at hun vil føye seg for å passe inn i et slikt fellesskap.

Det er likevel stor forskjell på å skulle sette en karakter som sensor og som faglærer, og på spørsmål om hun kan «definere» hva hun legger i å «vurdere rettferdig» trekker hun derfor i tillegg inn at en rettferdig vurdering også handler om å ikke hensynta det eleven har prestert tidligere, og heller ikke ta elevens person med i betraktningen når en karakter skal settes. I praksis hender det likevel at det er vanskelig å avgjøre hva som er den «riktige» karakteren, sier hun, og i de tilfellene oppgir hun at hun kan legge bort elevteksten og hente den frem igjen på et senere tidspunkt:

Hvis jeg har en bunke på 28 tekster og den tredje ble litt trøblete, så kan jeg godt legge den litt til siden, og når jeg har kommet til (nummer) 25, så kan jeg bli litt tryggere på hvor denne teksten hører hjemme i dette landskapet, da.

Erfaringsbakgrunnen hennes gjør at hun selv opplever å gi rettferdige vurderinger, men hun understreker at det er vanskelig å foreta en slik vurdering av seg selv

4.3.2 Lærer2 sin forståelse av begrepet «rettferdig vurdering» og beskrivelse av egen praksis

I tillegg til noe erfaring som sensor på skriftlig eksamen i norsk har Lærer2 lang erfaring med å vurdere tekster i fellesskap med kollegaer lokalt. På spørsmål om hun kan «definere» hva hun legger i å «vurdere rettferdig» trekker hun frem de ulike vurderingsformene eksamen og standpunkt:

Jeg vet ikke om det er helt rettferdig å sette en karakter på én besvarelse. Det er nokså skummelt, og det er jo det som skjer på eksamen. Mens dette løpet med flere besvarelser på samme karakter, altså standpunkt, gjør meg trygg på at dette er for eksempel «høy måloppnåelse», «dette er en treer» og så videre.

Hun understreker med dette et poeng om at elevene kan ha gode og dårlige dager, og hun opplever at det er mer rettferdig å gi en sluttvurdering basert på et bredt grunnlag heller enn en enkeltprestasjon fordi resultatet da vil være mindre preget av et øyeblikksbilde.

Når det gjelder sin egen praksis oppgir hun at det «som regel går [...] ganske greit» å sette det hun oppfatter som en rettferdig karakter, men de gangene hun er usikker på hvilken karakter som er den «riktige» spør hun kollegaer om de kan lese gjennom teksten og gi en vurdering. Et potensielt hinder for en rettferdig vurdering er derimot, ifølge Lærer2, elevenes forventninger, tidligere prestasjoner og lærerens kunnskap om disse: «Da er det vanskelig å være nøytral». For å illustrere dette forteller hun en historie om en elev som «alltid jobber masse, skriver lange tekster og er flink og kjekk på alle måter». Denne eleven er vant til å få gode karakterer, og Lærer2 er også vant til å gi gode karakterer til denne eleven. Når eleven da på et tidspunkt ikke leverer på forventet nivå er det lett å gi en karakter som er for god, og dermed også per definisjon er urettferdig. I slike tilfeller oppgir Lærer2 å støtte seg på kollegaer i vurderingsarbeidet, og hun opplyser at det skjer «ofte» at lærerne på skolen bytter tekster på denne måten.

4.3.3 Lærer3 sin forståelse av begrepet «rettferdig vurdering» og beskrivelse av egen praksis

For Lærer3 handler rettferdig vurdering først og fremst om at ulike elever må vurderes på samme grunnlag: «Det er at vi bruker de samme kriteriene [...]. Altså om det er Kåre eller Ola eller Kari, så er det de samme tingene vi ser på hos alle». Det er likevel en kjensgjerning, ifølge Lærer3, at det er lett å la seg påvirke av elevenes personer:

Hvis du har hun jenta som du vet sliter med anoreksi, og jobber så godt hun kan, men egentlig ligger og vipper, så kan det hende hun blir vippet opp, ikke sant? For du vil gjerne gi henne den lille boosten. Du vil ikke trykke henne mer ned.

Lærer3 understreker at hun ikke gir ufortjent gode karakterer for å belønne enkeltelever, men erkjenner at det i tvilstilfeller er lett å la tvilen komme eleven litt ekstra til gode. Det kan være for eksempel å gi den høyeste karakteren hvis en elev «ligger på vippen», som illustrert i eksempelet over, eller ved å gi for eksempel 4 +, i stedet for «bare 4». Skolen som Lærer3

jobber ved har gått mer og mer over til anonym retting, og det mener hun er en fordel for å bøte på nettopp denne urettferdigheten. Lærer3 belyser likevel paradokset med at anonym retting gjør det vanskeligere for læreren å gi gode underveis- og fremovermeldinger, og at det på sin side er urettferdig for elever med særlige behov for tett oppfølging, «særlig (elever) med dysleksi»⁴.

Lærer3 reflekterer videre over egen praksis og understreker at man «bør ikke la den anoreksien eller den dysleksien påvirke hvordan man setter karakterer», og de gangene hun er usikker på sin egen vurdering opplyser hun om at hun ber andre kollegaer om råd. Det var likevel enklere for Lærer3 å be om slike råd tidligere. Som nyutdannet lærer fikk hun god opplæring i vurderingsarbeid av mer erfarne lærere, men disse lærerne er nå pensjonister, og læreren har byttet arbeidsplass. På nåværende tidspunkt har hun i større grad en følelse av at hun «maser» viss hun prøver å få til faste rammer for å vurdere elevtekster i fellesskap, og hun gir derfor uttrykk for å savne «disse gamle karene» som hun en gang samarbeidet særlig godt med.

4.3.4 Sammenfatting og oppsummering

De tre lærerne legger vekt på tre ulike aspekter ved vurdering når de blir bedt om å forklare hva de legger i begrepet «rettferdig vurdering»: Lærer1 er særlig opptatt av at et resultat ikke skal være tilfeldig eller avhengig av hvem som vurderer. Lærer2 er på sin side mer opptatt av størrelsen på vurderingsgrunnlaget, og trekker frem at det ikke nødvendigvis er rettferdig å sette en karakter basert på en enkelt prestasjon, sånn som man gjør på eksamen. Lærer3 mener at man vurderer rettferdig viss man bruker de samme målepunktene for alle elever, og dermed representerer hun et tredje synspunkt. Her er det likevel naturlig å understreke at ingen har fått oppfølgingsspørsmål som ber dem om å komme med flere eksempler på hva som kan ligge til begrepet «rettferdig vurdering». Det er heller ikke slik at det ene av disse tre aspektene utelukker de to andre, og dersom de hadde blitt bedt om å komme med flere eksempler er det derfor rimelig å anta at de hadde gitt mer sammenfallende svar. Valget om å ikke stille oppfølgingsspørsmål ble likevel tatt som følge av at alle tre fikk tilsendt

⁴ Av hensyn til de ytre rammene for denne studien vil jeg ikke gå nærmere inn i diskusjonen om rettferdighet i underveisvurdering.

intervjuguiden med dette spørsmålet i god tid før intervjuet, og de hadde derfor, i større eller mindre grad, reflektert over dette i forkant. Det første svaret sier derfor noe om hva de oppfatter som «kjernen», og jeg valgte derfor bevisst å ikke utfordre dem på å gi flere svar på dette spørsmålet.

Til tross for at de velger å trekke frem ulike aspekter ved rettferdig vurdering, så rapporterer de om en tilnærmet lik praksis for å sikre at den vurderingen de selv bedriver skal være mest mulig rettferdig: De støtter seg til profesjonsfellesskapet og ber kollegaer om råd. Lærer1, som har lang sensorerfaring, trekker særlig frem sensorkollegiet, mens både Lærer2 og Lærer3 i hovedsak trekker frem kollegaer på samme skole.

5.0 Drøfting

I dette kapitlet blir resultatene fra analysen av intervjuene drøftet opp mot styringsdokumentene, vitenskapsteorien og den tidligere forskningen som presenteres i kapittel 2.0 *Kunnskapsgrunnlag*. Gjeldende kapittel innledes med en drøfting av hvilke faktorer som skal inngå i en sluttvurdering. Her er det først og fremst styringsdokumentene som danner det sentrale bakteppet, og som derfor naturlig inngår i drøftingen. Videre blir spørsmålet om rettferdig vurdering på tekstnivå behandlet gjennom å drøfte vurderingsmatrisenes og sensorskoleringens verdi opp mot teori og tidligere forskning. I det delkapitlet brukes styringsdokumentene først og fremst for å belyse vurderingene som lærerne gjør av den aktuelle elevteksten. Til slutt behandles rettferdig vurdering på profesjonsnivå ved å drøfte profesjonsfellesskapets rolle i lys av teorier og tidligere forskning.

5.1 Faktorer som skal inngå i en sluttvurdering

Både Lærer1, Lærer2 og Lærer3 trekker inn *språk, struktur og innhold* som tre hovedområder når de begrunner karakteren de har satt på den aktuelle elevteksten. Dette er i tråd med de kjennetegnene på måloppnåelse som er utarbeidet av Utdanningsdirektoratet, og som er å finne i eksamensveiledningen for norsk vg3 (Utdanningsdirektoratet, 2018a). De tre lærerne trekker også frem disse tre hovedområdene når de beskriver vurderingspraksisen sin på mer generelt grunnlag. Berge og Hertzberg (2005) belyser norskeksamens washback-effekt, og Bach (2015, s. 31) understreker at eksamensoppgavene i norsk fungerer som et styringsdokument for lærernes egen praksis. I dette konkrete tilfellet ser vi at det ikke bare er

eksamensoppgavene, men også eksamensveiledningen som gir en slik washback-effekt. Man må likevel hensynta at både undervisningen og eksamen bygger på det samme overordnede dokumentet: læreplanen i faget. Et samsvar mellom undervisning og vurderingspraksis i skolen og innholdet i eksamensoppgavene og vurderingen av disse kan derfor også leses som et uttrykk for en forholdsvis lik tolkning av læreplanen blant lærerne i skolen og de som utarbeider eksamen.

Å vektlegge språk, struktur og innhold i vurderingen av en elevtekst må altså sies å være i tråd med intensjonen i læreplanen, og det er heller en styrke enn en svakhet dersom det er eksamens washback-effekt som har skapt denne bevisstheten blant lærerne. Det er derimot en del som tyder på at andre aspekter ved vurderingspraksisen i mindre grad harmonerer med intensjonene i styringsdokumentene, og disse er forsøkt belyst i de følgende delkapitlene. Resultatet i studien viser nemlig at selv om deltakerne har god kunnskap om vurdering og god forståelse av vurderingsbegrepet, så har de også en praksis som delvis strider mot § 3-3, § 3-14 og § 3-22 i Forskrift til opplæringslova (2020).

5.1.1 Å vurdere elevenes forutsetninger og innsats

Når Lærer2 skal begrunne karakteren hun setter på den aktuelle elevteksten trekker hun frem elevens innsats og det hun oppfatter som et «seriøst forsøk på å skrive et godt svar», som en av de avgjørende faktorene. § 3-3 i Forskrift til opplæringslova slår derimot fast at elevenes innsats ikke skal være en del av vurderingsgrunnlaget i fag, med mindre det følger av læreplanen i faget, hvilket det ikke gjør i læreplanen for norskfaget. Dersom man går til sensorveiledningen (Utdanningsdirektoratet, 2018c) ser man likevel at sensor oppfordres til å være «eksaminandens 'advokat'». Det gjør man, ifølge sensorveiledningen, ved å se etter det eksaminanden faktisk mestrer, og ved å ha en positiv vurdererholdning. Til tross for at det å legge elevens innsats til grunn for sluttvurderingen er i strid med prinsippet for vurdering i fag i *Forskrift til opplæringslova*, kan man derfor, med oppfordringen til sensorene i minne, forstå Lærer2 sitt ønske om å honorere eleven der hun opplever at han fortjener honnør.

Hva angår spesielle tilpasninger gjort med hensyn til elevenes forutsetninger åpner *Forskrift til opplæringslova* for at elever som oppfyller visse kriterier kan søke om fritak fra vurdering med karakter i norsk sidemål (Forskrift til opplæringslova, 2020, §3-19). Utover dette har man

likevel ikke anledning til å gi særlige tilpasninger i sluttvurderingen, for elever med ulike vansker. Det betyr at en elev med dysleksi som har fritak fra vurdering med karakter i norsk sidemål, etter loven skal vurderes på samme grunnlag som elever uten fritak, når karakteren i norsk hovedmål skal settes. Det er tydelig at de tre intervjuede lærerne kjenner godt til dette regelverket, men både Lærer2 og Lærer3 trekker dette frem som en vanskelig øvelse, og det blir tydelig at lærerne slites mellom det som i forskningslitteraturen omtales som universell og differensiell vurdering (Prøitz & Borgen, 2010). Lærer3 understreker at man har ikke lov til å gjøre særlige tilpasninger for elevene når karakteren settes, men når man vet at enkelte elever har slitt hardt for å forbedre seg så er det fristende å sette et pluss bak karakteren «for å gi en boost» (Lærer3, intervju). Lærer2 trekker i tillegg frem at lærerens kunnskap om særlig flinke elevers forventninger ubevisst kan føre til at læreren gir en god karakter også de gangene eleven ikke leverer på sedvanlig høyt nivå. Dette vitner om at de ulike lærerne har et bevisst forhold til sin egen vurderingspraksis, og at de også som faglærere navigerer etter prinsippet om å være elevens advokat.

5.1.2 Å sammenligne elever

Prinsippet om at elevens innsats ikke skal vurderes, videreføres i Forskrift til opplæringslova § 3-14 og § 3-22. Der kan man lese at elevene skal måles opp mot kompetansemålene i læreplanen når henholdsvis standpunktkarakter og eksamenskarakter fastsettes. På dette punktet oppfatter jeg at noen av de intervjuede lærerne delvis har en praksis som strider mot styringsdokumentene, uten at de har et bevisst forhold til det.

Elevteksten har blitt vurdert av to uavhengige sensorpar, og fått henholdsvis karakterene 3 og 4. På spørsmål om hva hun tror sensorparet i klagesensuren har vektlagt som det første sensorparet ikke vektla, trekker Lærer1 frem de ulike kontekstene. Under fellessensuren er det, ifølge Lærer1, svært mange tekster som holder omtrent samme nivå som denne elevteksten. Som klagesensor leser man derimot mange tekster som opprinnelig har fått karakteren 1, sier hun, og hun reflekterer derfor over om denne teksten kan ha utmerket seg positivt blant mange dårlige tekster. Her ligger det altså implisitt en forventning om at sensorkorpset sammenligner de ulike besvarelsene med hverandre.

Både Lærer1 og Lærer2 rapporterer dessuten om en tilsvarende praksis når de skal sette karakterer som faglærere. Dersom de opplever en tekst som vanskelig å vurdere, opplyser begge to at de kan legge den bort, for så å hente den frem igjen etter at de har lest flere besvarelser til samme oppgave. Lærer1 gjør det for å «bli litt tryggere på hvor denne teksten hører hjemme i dette landskapet», mens Lærer2 oppgir dette som en fast praksis i vurderingsarbeidet:

Jeg leser hele bunken først, og så går jeg over besvarelsene en gang til der jeg sammenligner litt de som jeg tenker ligger på 5 og så også de som ligger på 4 og 3. Jeg sammenligner dem litt med hverandre.

Her ser vi altså at både Lærer1 og Lærer2 impliserer at de måler elevenes tekster opp mot hverandre når de setter karakter. Det er likevel viktig å understreke at lærerne ikke sier noe i intervjuene som antyder at de sammenligner elevenes personer, og de tilpasser karakterene etter hvem de mener fortjener eller ikke fortjener gitte karakterer. Jeg oppfatter det snarere som en strategi for å finne ut hvor skillet mellom de ulike karakterene faktisk går, og det kan således leses som en beviselig svakhet ved ordinalskalaen (Berge, 2009). Det er i denne sammenheng nødvendig å presisere at Lærer1 har lang sensorerfaring, og at det er mye som tyder på at hun har Utdanningsdirektoratets kjennetegn på måloppnåelse «under huden». Lærer2 opplyser på sin side om at hun bruker disse kjennetegnene på måloppnåelse fra Utdanningsdirektoratet når hun leser hver enkelt tekst, altså før hun sammenligner de ulike firerne og femmerne. Denne sammenligningspraksisen kan derfor leses som et uttrykk for at vurderingskriteriene er for lite spesifikke, og at lærerne opplever det som krevende å vite hvor skillet mellom ulike karakterer går. Dette funnet er i tråd med funnene til Black et al. (2011), og fenomenet belyses ytterligere i kapittel 5.2.1 *Hvor går skillet mellom de ulike karakterene?*

5.1.3 Forholdet mellom skriveferdighet og leseferdighet

Alle oppgavene som gis til norskeksamen har minst ett tekstvedlegg, og oppgavene som den aktuelle elevteksten svarer på er ikke noe unntak. Kortsvaret har ett, mens langsvaret har tre tekstvedlegg. Alle de tre lærerne leter etter elevens *bruk* av tekstvedleggene når de skal begrunne karakteren de setter, men bare Lærer3 understreker viktigheten av elevens *forståelse* for innholdet i dem. Her ser vi altså et språk mellom lærernes praksis, og hva de gir

uttrykk for å legge vekt på i vurderingen. At leseferdigheten skal inngå som en del av vurderingsgrunnet understrekes derimot i eksamensveiledningen: «Det overordnede formålet med eksamen er at elevene skal vise norskfaglig kompetanse. De skal lese (...), og de skal skrive (...)» (Utdanningsdirektoratet, 2018a). Hva angår vurdering av lesekompetanse under skriftlig eksamen er det dessuten verdt å merke seg at de ulike styringsdokumentene sender motstridende signaler.

I norskfaget på videregående skole er det to skriftlige eksamener: obligatorisk eksamen i hovedmål, og eksamen i sidemål som elevene kan bli trukket ut til. Oppgavetekstene er i begge tilfeller skrevet på begge målformer, for å passe til elever som har både bokmål og nynorsk som sitt hovedmål. Tekstvedleggene, derimot, kommer i bare en målform. Det innebærer at elevene ofte må lese tekstvedlegg på sidemålet sitt, selv om de altså skal skrive på hovedmålet sitt. I henhold til *Forskrift til opplæringslova § 3-17 og § 3-19*, har enkelte elever i videregående skole rett til fritak fra vurdering med karakter i norsk sidemål. Disse elevene skal opp til en obligatorisk eksamen i norsk hovedmål, og de får tilgang til det samme oppgaveheftet som elever uten karakterfritak i sidemål. I prinsippet betyr det at elever som har rett på fritak fra vurdering med karakter i sidemålet, skal opp til en eksamen i norsk der de formelle vurderingskriteriene er knyttet til elevens hovedmål. I praksis er oppgavene med tekstvedlegg likevel utformet på en måte som gjør at elevens lesekompetanse i sidemålet utgjør en del av vurderingsgrunnet, til tross for at de altså er fritatt fra vurdering med karakter i sidemål.

De tre lærerne kan ikke vite om denne aktuelle eleven har fritak fra vurdering med karakter i sidemål eller ikke, og ettersom alle de tre tekstvedleggene er skrevet på elevens hovedmål, bokmål, er det heller ikke av interesse for vurderingen av akkurat denne teksten. Det er likevel verdt å merke seg at dersom eleven hadde hatt nynorsk til hovedmål, hadde alle tekstvedleggene vært skrevet på sidemålet hans, og han hadde dermed blitt vurdert på et betydelig bredere grunnlag enn elever med bokmål som hovedmål⁵.

⁵ Presisering: Elevene kan, som illustrert her, i enkelte tilfeller velge oppgaver der målformen i tekstvedlegget / alle tekstvedleggene er skrevet på deres eget hovedmål, men da vil i tilfelle de elevene som har rett til fritak fra vurdering med karakter i sidemål i praksis ha færre oppgaver å velge mellom enn elever som ikke har et slikt fritak. Ved å lese gjennom tidligere oppgavehefter fra eksamen ser man dessuten at nynorskelever i færre tilfeller enn bokmålelever har anledning til å velge oppgaver med tekstvedlegg som er skrevet utelukkende på elevens hovedmål.

5.2 Rettferdig vurdering på tekstnivå

Før jeg i dette kapitlet går over til å drøfte vurderingsmatrisenes verdi vil jeg understreke at det både i forskningslitteraturen og i intervjuene brukes mange overlappende begreper: «vurderingsmatrise», «vurderingskriterier» og «kjennetegn på måloppnåelse», for å nevne noen. Disse kan selvsagt ta ulike former avhengig av hvem som har utarbeidet dem, og for hvilket skoleslag de er tiltenkt. I det følgende brukes likevel disse begrepene først og fremst om en samlebetegnelse for en oversiktig fremstilling av hvilke krav som stilles til en elevprestasjon for at den skal nå gitte karakterer. Utdanningsdirektoratets «Kjennetegn på måloppnåelse» er eksempel på en slik fremstilling.

Vurderingsmatriser brukes aktivt av både Lærer2 og Lærer3, men mens Lærer2 nøyer seg med å fortelle om at hun har brukt det i vurderingen av den aktuelle elevteksten så refererer Lærer3 til dem aktivt gjennom intervjuet. Hun har dem med seg på papir, og peker i tabellen når hun blir bedt om å redegjøre for hvor i landskapet hun mener den aktuelle elevteksten befinner seg. Lærer1, som altså er den av deltakerne med desidert mest sensorerfaring, sier derimot at hun har sluttet å bruke vurderingsmatriser fordi hun opplevde dem som kunstige. Når hun redegjør for vurderingen sin av den aktuelle elevteksten er det likevel tydelig at hun navigerer etter de samme kjennetegnene på måloppnåelse som de andre to, som altså bruker Utdanningsdirektoratets kjennetegn på måloppnåelse aktivt. Her kan man anta at Lærer1, på grunn av sin lange sensorerfaring, ikke lenger har behov for å ha et fysisk skjema foran seg, men at hun ubevisst har innlemmet innholdet i skjemaet i sin egen vurderingspraksis.

5.2.1 *Hvor går skillet mellom de ulike karakterene?*

Det lave bedømmersamsvaret i norsk så vel som utenlandsk skole har blitt trukket frem og behørig belyst i kapittel 2. Bedømmersamsvaret rapporteres tidvis å være så lavt (Berge, 1996; McNamara, 1996; Weigle, 1998; Bjørnson & Skar, 2021) at det ikke innfrir kravet om reliabilitet. Ved overgangen til tallkarakterer i norsk skole understreket Vagle (2005) behovet for å utarbeide detaljerte signalementer for de forskjellige karakterene, og forskutterte dermed det vi i dag omtaler som vurderingskriterier. Alle de tre lærerne i denne studien er enige i at skillet mellom en god og en mindre god karakter går «midt på middels», altså mellom karakteren 3 og karakteren 4, og vi kan derfor si at de har en viss felles forståelse for hvor skillet mellom de ulike karakterene går. Det er likevel interessant at Lærer1 slites mellom

nettopp karakterene 3 og 4 når hun skal karaktersette den aktuelle elevteksten. Elevteksten ble dessuten først vurdert til karakteren 3 og deretter til karakteren 4 under sensuren. Det er også interessant at Lærer1 forestiller seg at argumentasjonen til begge sensorparene kan ha vært veldig lik, men at de «falt ned på hver sin karakter». Dersom hun har rett i disse antakelsene kan man argumentere for at systemet i for stor grad baserer seg på tilfeldigheter. Ettersom alle de tre lærerne er enige i at skillet mellom en god og en mindre god karakter går nettopp mellom 3 og 4 er det dessuten et paradoks at både Lærer1 og Lærer2 opplever skillet mellom karakterene 3 og 4 som det vanskeligste med vurderingsarbeidet.

Gjennom eksamensveiledningen forstår man at de vedlagte kjennetegnene på måloppnåelse er ment som felles referanserammer for elev og sensor, men Lærer1 har gått bort fra å bruke dem i sitt eget vurderingsarbeid fordi elevene ikke forstår innholdet i dem. Hun nevner eksplisitt «for eksempel det som står der om setningsoppbygging». På akkurat dette punktet har det skjedd en endring i eksamensveiledningen mellom 2018 og 2021, og jeg referer derfor til den nyeste versjonen fra 2021:

Tabell 5: Utdrag fra kjennetegn på måloppnåelse fra 2021. Hentet fra eksamensveiledningen for norskfaget på videregående skole.

| | Karakter 2: Eksamenssvaret | Karakter 3-4: Eksamenssvaret | Karakter 5-6: Eksamenssvaret |
|--------------------------------------|---|--|--|
| Språk og formelle ferdigheter | har enkel setningsstruktur og sammenbinding mellom setninger og avsnitt | har stort sett variert setningsstruktur og nokså god / god sammenbinding mellom setninger og avsnitt | har formålstjenlig og variert setningsstruktur og god sammenbinding mellom setninger og avsnitt. |

(Utdanningsdirektoratet, 2021b, s. 8)

Lærer1 peker altså på at det for en elev vil være vanskelig å forstå hva som ligger i for eksempel ordlyden «formålstjenlig og variert setningsstruktur». Det er i tillegg verdt å merke seg at formuleringene egentlig er like på alle de tre nivåene, og den eneste forskjellen er at de er graderte ved hjelp av formuleringene «enkel», «stort sett variert» og «formålstjenlig og

varierte». I intervjuene ble særlig skillet mellom karakterene 3 og 4 problematisert av deltakerne, og i dette eksempelet er det altså formuleringen «nokså god / god» som indikerer skillet mellom de to karakterene.

Formuleringer som «nokså god / god» i vurderingskriterier er ikke et særnorsk fenomen, og denne form for ordlyd problematiseres blant annet av Black et al. (2011) og belyses også gjennom SKRIV-prosjektet i norsk kontekst (Solheim & Matre, 2014, s. 77). Når man som lærer skal vurdere om et arbeid er «nokså godt» eller «godt», eller om setningsstrukturen er «formålstjenlig og variert» eller bare «varierte» må man altså selv fylle disse ordene med innhold, og det er, ifølge Black et al. (2011, s. 458), her problemene oppstår. I deler av denne studien kommer disse problemene tydelig til uttrykk. Til tross for at Lærer3 vurderte den samlede elevteksten til å være 4, mener hun at deler av den er høy måloppnåelse, og hun bruker vurderingskriteriene aktivt når hun argumenterer for hvilke deler det gjelder. Lærer2, som også har kontrollert vurderingen sin opp mot et vurderingsskjema, mener derimot at teksten er en svak firer. De samme vurderingskriteriene ble altså brukt av to ulike sensorpar, som vurderte teksten til henholdsvis karakteren 3 og karakteren 4.

Omfanget av denne studien gjør at man ikke kan konkludere med at funnene er representative for de deler av populasjonen som ikke selv er deltakere, men man kan likevel slå fast at funnene føyer seg inn i rekken av tilsvarende funn i tidligere og større studier. Det er derfor rimelig å anta at funnene har en viss form for overføringsverdi, også til de norsklærerne og skolene som ikke er direkte representerte i utvalget. På bakgrunn av dette kan man hevde at det er en viss sannsynlighet for at åpne vurderingskriterier med et stort rom for skjønnsmessig vurdering fører til lavt vurderingssamsvar mellom ulike vurderere.

5.2.2 IRR og IRA – vurderingssamsvar og vurderingsenighet

Både Lærer1, Lærer2 og Lærer3 ender opp med å vurdere elevteksten til karakteren 4, men de gir også tre ulike begrunnelser for denne karakteren. For å illustrere dette skillet opererer Tinsley og Weiss (2000, s. 98) med begrepsparet *interrater reliability* og *interrater agreement*, eller sensorsamsvar (IRR) og sensorenighet (IRA), på norsk. I denne studien kan man dermed si at sensorenigheten (IRA) er høy, men siden de tre intervjuede lærerne begrunner karakteren

på tre forskjellige måter kan man likevel argumentere for at sensorsamsvaret (IRR) er betydelig lavere.

Ifølge Berge (1993a, sitert i Berge 2009, s. 45) er det et faktum at lærerne opplever det som vanskelig å bli enige om hvordan man skal rangere tekstenes ulike aspekter opp mot hverandre. Lærer2 opplever kreative tekster som særlig vanskelige å vurdere, og understreker at dette også var inntrykket hun satt igjen med etter å ha deltatt på sensorskolering:

Den siste gangen jeg var sensor på skolering så varierte karakterene på en sånn kreativ tekst [...] Noen ville gi 2, noen ville 3, noen ville gi 4 og noen ville gi 5. Og det er galt at det skal være sånn. Det er veldig, veldig ugreit.

Inntrykket hennes finner støtte i forskningen, og ifølge Berge (2005b, s. 112-113) er det betydelig lavere korrelasjoner for fortellende skrijving enn for resonnerende skrijving. Elevteksten som ligger til grunn for denne studien er derimot en resonnerende tekst, og det er derfor kanskje naturlig at de tre lærerne ender opp med å gi samme karakter til tross for at de gir uttrykk for å være ganske usikre i forkant.

Eckes (2008) presenterer også i sin studie en potensielt medvirkende faktor for at lærerne gir ulike begrunnelser for den samme karakteren. Han tester, og han får langt på vei bekreftet, en hypotese om at ulike lærere har ulik oppfatning av hvilke vurderingskriterier som er særlig viktige, når de vurderer en elevprestasjon. Basert på hvilke vurderingskriterier hver enkelt lærer opp- og nedjusterer, mener Eckes (2008) at han kan plassere lærerne i ulike «vurderingsbåser». Han finner også at lærere som tilhører ulike slike vurderingsbåser mangler felles oppfatning av de ulike kriteriene. Formålet med denne studien har ikke vært å kunne plassere de intervjuede lærerne i slike båser, og det foreligger derfor heller ikke empiri til å foreta en slik kategorisering. Gjennom svarene de gir i intervjuene er det likevel tydelig at de delvis mangler en felles oppfatning av de ulike kriteriene. Det ser vi for eksempel ved at Lærer1 og Lærer3 mener strukturen er god, mens Lærer2 mener den ikke er det, og ved at Lærer2 og Lærer3 mener det er en styrke at eleven trekker inn saktekster i langsvaret sitt, mens Lærer1 mener det er en svakhet.

I kapittel 2 belyses forholdet mellom eleven, teksten og den tenkte leseren, ved å vise til Bakhtin, og i lys av hans teorier om dialogisme er det naturlig å definere elevteksten som en dialogisk tekst utarbeidet for en fremtidig modelleser. Leserene i dette tilfellet har felles bakgrunn i form av høyere utdanning, samt både undervisningskompetanse og undervisningserfaring fra norskfaget. Men de har også erfaringsbakgrunn fra forskjellige arenaer, for eksempel fra ulike skoler og skoleslag, og de har ulik grad av sensorerfaring. I tillegg hadde de, som alle andre, sine egne unike opplevelser som elev, og gjennom jobben som lærer har de møtt elever med ulike historier og ulikt ambisjonsnivå. Dette gjør at de møter den aktuelle elevteksten med ulike blikk, og med dette blir de unike medprodusenter av teksten. Ettersom vurderingskriteriene gir stort rom for tolkning og skjønnsmessige vurderinger, er det dermed naturlig at det også er et visst avvik mellom de vurderingene som gjøres.

5.3 Rettferdig vurdering på profesjonsnivå

Både denne studien og tidligere, mer omfattende, forskningsstudier gir uttrykk for at å ha felles vurderingskriterier ikke er synonymt med å ha et tolknings- eller vurderingsfellesskap. Ecclestone (2001) argumenterer for at man må ha en mer strategisk tilnærming når man skal sosialisere ansatte inn i et profesjonsfellesskap: Vurderingskriterier kan være et nyttig verktøy, men bare hvis de inngår som del av en større prosess. Denne prosessen bør, ifølge Ecclestone, bestå av regelmessige diskusjoner med kollegaer.

Denne studien har delvis sprikende funn hva angår profesjonsfellesskapets rolle i videregående skole. På den ene siden er Lærer3 opptatt av å understreke at hun ikke i like stor grad som tidligere opplever å være del av et velfungerende profesjonsfellesskap. En opplevelse hun har, til tross for at hun etterstreber og etterspør slike regelmessige diskusjoner med kollegiet som Ecclestone peker på at man trenger for å oppnå et tolkningsfellesskap:

Jeg kan når som helst spørre en norsklærer om å hjelpe meg hvis jeg er veldig, veldig usikker. Men det er blitt mindre og mindre samarbeid, for det er ikke en fast greie lenger. Det er noe jeg føler jeg plager mine kollegaer med, [...]

Erfaringene til Lærer3 kan minne om de som beskrives av Ecclestone (2001). Gjennom sin studie fant Ecclestone at erfarne lærere ved enkelte studiesteder har en tendens til å avfeie forsøk på samarbeid og diskusjoner om vurdering, fordi de oppfatter vurderingsarbeid som «intuitivt».

På den andre siden har derimot begrepet profesjonsfellesskap blitt formalisert i skolens styringsdokumenter i senere år, og gjennom den nye overordnede delen av læreplanverket finner man en forventning til at det er gjennom profesjonsfellesskapet man skal utvikle sin praksis og finne svar på pedagogiske spørsmål. Dette kan se ut til å ha manifestert seg i skolen ved at alle de tre lærerne, også Lærer3, oppgir at de søker råd og veiledning fra sine kollegaer dersom de er usikre i vurderingsarbeidet. Når de refererer til kollegaer, snakker Lærer2 og Lærer3 først og fremst om andre lærere på samme arbeidsplass, mens Lærer1 fremhever profesjonsfellesskapet med sensorkollegaer. Alle tre er særlig opptatt av å støtte seg på lærerfellesskapet for å sikre en rettferdig vurdering av elevene.

Det er altså en del som tyder på at historien til Lærer3 om et profesjonsfellesskap i tilbakegang skiller seg fra tendensen i vår tid. Til tross for at hun har opplevd det motsatte, mener jeg derfor det er rimelig å påstå at profesjonsfellesskapet i dag står sterkere enn det tidligere har gjort. Tiden for klisjéen om den suverene norsklektoren som planlegger, gjennomfører og evaluerer sin egen undervisning uten å verken påvirke eller la seg påvirke av lærerfellesskapet ser ut til å nærme seg slutten. Tiden ser derimot ut til å være kommet for å, i tråd med både Bakhtins og Gadammers teorier, bedrive aktivt utviklingsarbeid gjennom dialog.

5.3.1 *Sensorskoleringens effekt*

Etter å ha analysert det innhentede datamaterialet i denne studien fremstår Eckes' (2008) påstander om opplæringsaktiviteter, som paradoksale. Han hevder nemlig at opplæringsaktiviteter har mye mindre effekt enn forventet, dersom målet er å redusere spriket mellom lærernes vurderinger. I norsk kontekst er sensorskoleringen i forbindelse med norskeksamen, særlig i videregående skole, en av de største og mest systematiske opplæringsaktivitetene som bedrives med hensyn til å utvikle et tolkningsfellesskap, og sensorskoleringen nevnes i positive ordelag av alle de intervjuede lærerne. Lærer1 og Lærer2

trekker begge frem denne opplæringsaktiviteten som noe av det mest nyttige de har deltatt på.

At lærerne opplever sensorskoleringen som nyttig, er likevel ikke synonymt med at tolkningsfellesskapet mellom de deltagende lærerne blir større. Studien til Anderson og Hertzberg (2005) peker likevel i den retning. De konkluderer med at sensorskoleringen ser ut til å gjøre lærerne trygge i vurderingen av elevtekstens «stilistiske virkemidler (s. 201). I den aktuelle elevteksten i denne studien spiller ikke stilistiske virkemidler en avgjørende rolle, men kanskje kan vi trekke paralleller til Lærer1 og Lærer2 sine negative reaksjoner på det faktum at eleven bruker begrep fra retorikken i omtale av skjønnlitterære tekster. Dette framstår som en reaksjon på at eleven bryter med en etablert norm om at begrep fra retorikken som hovedregel bare skal brukes i omtale av saktekster.

Nilsen (2005, s. 118-119) understreker likevel at enighet om for eksempel sjangerforståelse eller hvor god den formelle språkkompetansen er, i større grad er avhengig av en felles tolkning av vurderingskriteriene, og det bekreftes i denne studien. Gjennom lærernes begrunnelser for å gi den aktuelle elevteksten karakteren 4 ser ikke sensorskolering ut til å ha innvirkning på tolkningsfellesskapet i nevneverdig grad. Både Lærer1 og Lærer3, med henholdsvis mye og ingen sensorerfaring mener for eksempel at elevteksten har god struktur, mens Lærer2 mener at strukturen er en av tekstens svake sider.

På andre områder er det likevel tegn som tyder på at sensorskoleringen i forholdsvis stor grad har betydning for utviklingen av et tolkningsfellesskap. Lærer1 og Lærer2 er for eksempel mer enige, både seg imellom og med det faktiske eksamensresultatet, enn Lærer3 er. De to lærerne som har erfaring fra sensorskolering, synes altså å ha et tolkningsfellesskap som Lærer3 ikke i like stor grad er en del av. Sitatet som utgjør tittelen for denne studien er hentet fra nettopp Lærer3 og hennes refleksjoner omkring hvilken karakter hun mener den aktuelle elevteksten fortjener: «Jeg tenker det er en sterk firer, jeg da. Men det kan hende at det egentlig er stryk». Det kommer tydelig frem at disse refleksjonene springer ut av sensorhistorier hun har hørt. Lærer3 ser altså ut til å ha et bevisst forhold til at det finnes et tolkningsfellesskap som hun ikke er en del av, fordi hun ikke har vært sensor og derfor heller ikke har deltatt på sensorskolering.

Videre er det interessant at Lærer1 ikke bruker et vurderingsskjema som hjelpemiddel når hun begrunner karakteren hun setter, men at hun likevel er den av de tre lærerne som i størst grad gir en vurdering som sammenfaller med den sensuren som ble gitt til både opprinnelig sensur og klagesensur. Lærer3 bruker derimot vurderingsskjema aktivt, som et verktøy i vurderingsarbeidet, men når hun begrunner karakteren hun gir kommer det likevel frem at hun er den som er lengst unna elevtekstens faktiske eksamensresultat. Her kan det derfor se ut som om sensorskolering, og sensorerfaring som sådan, i større grad gir et tolkningsfellesskap enn det vurderingskriteriene makter å gi, slik de er utformet i dag.

Dersom det er tilfelle, er det ikke uproblematisk. Det er for det første ikke alle norsklærere som melder seg som sensorer, og for det andre er det heller ikke alle de som melder seg som får tildelt sensoroppdrag. Videre stilles det ikke krav om deltakelse på sensorskolering for de som har fått tildelt oppdrag. Skulle det vise seg at deltakelse på sensorskolering, eller tilsvarende aktiviteter, er en forutsetning for å etablere et tolkningsfellesskap, kan man risikere å havne i en situasjon der det oppstår et A-lag og et B-lag blant norsklærerne. Dersom det videre viser seg at tolkningsfellesskapet som etableres under en sensorskolering, har en washback-effekt til skolehverdagen, må man dessuten forvente at det etableres uuttalte «krav» til elevtekstene som ikke alle lærere har lik tilgang til.

5.3.2 De signifikante andre i profesjonsfellesskapet

Det kanskje aller mest entydige funnet i denne studien er hvor viktig lærerfellesskapet, profesjonsfellesskapet, synes å være i vurderingsarbeidet. Ifølge Bakhtin er respons fra *den andre* det aktiverende prinsippet for å utvikle mening (Dysthe, 2012, s. 56), og i lys av den sosiokulturelle læringsteorien kan man forstå lærernes ønske om, og behov for, et nært samarbeid med kollegiet som et uttrykk for å se verdien av å utvikle seg i samspill med andre. Ifølge Mead er det dessuten sånn at selvet gjør justeringer basert på den responsen man mottar fra de signifikante andre. Gjennom denne studien kommer det til uttrykk ved at lærerne i klartekst oppgir at de justerer sin egen vurderingspraksis basert på responsen til kollegaer de ser opp til, de signifikante andre.

Sensorer og oppmenn

Lærere med sensorerfaring ser ut til å være en gruppe som har særlig sterk etos i vurderingsarbeidet. Lærer1 og Lærer2 referer flere ganger til samtaler de har hatt med andre sensorer, og det er tydelig at de med lang sensorerfaring fremstår som *signifikante andre* for samtlige av de tre intervjuede lærerne. Lærer3 har ikke selv vært sensor, men referer til andres opplevelser og justerer seg etter disse. Hun oppgir for eksempel at hun har hørt «fra disse kollegaene mine, som er sensorer, at kortsvaret er kun med på å vippe opp eller ned», og derfor gjør hun nettopp det når hun skal sette karakter på den aktuelle elevteksten. Lærer1 har derimot lang sensorerfaring selv, og man kan således anta at hun har en viss etos i vurderingsspørsmål ved den skolen hun selv jobber på. For henne ser derimot oppmannskorpset og sensorfellesskapet til å fungere som signifikante andre. Hun trekker frem oppmannssystemet som en styrke, og opplyser om at hun må «jekke seg» dersom hun ikke er på linje med sensorfellesskapet når en karakter skal settes. Her ser vi et eksempel på det som i Meads teorier omtales som å gjøre justeringer basert på responsen til de signifikante andre. I dette tilfellet er det dessuten slik at justeringen får en helt konkret konsekvens: Læreren setter en annen karakter enn opprinnelig tenkt. Med dette bekreftes altså Elder et al. (2005) sine funn om at en lærer kan bli «overcautious or fearful» (s. 190) i møte med andre vurdere. Ettersom man ikke har noen garanti for at flertallet, eller de mest erfarne, nødvendigvis er de som i størst grad vurderer i tråd med kravene som til enhver tid stilles i læreplan og lovverk kan en slik praksis, hvis utbredt, gå på bekostning av vurderingens gyldighet.

Praksislærere

Lærer2 har begrenset sensorerfaring, men er den av de tre intervjuede lærerne med desidert lengst erfaring som lærer i videregående skole. I løpet av karrieren har hun vært praksislærer, og vi kan derfor anta at hun selv har bekledd rollen som signifikant andre for praksisstudentene. Lærer2 trekker i den forbindelse inn viktigheten av at de får delta i autentiske vurderingssituasjoner: «det er jo kjempeviktig at de får prøvd seg i vurderingssituasjoner [...]. Det vil de lære av». Lærer2 ønsker altså å bruke sin egen rolle for å bidra til at praksisstudentene får utviklet seg, og man kan ane ekkoet av John Deweys pragmatiske kunnskapssyn (Dysthe et al., 2012, s. 49). Hos Lærer2 ser vi her en holdning om at samhandling er avgjørende for praksisstudentenes utvikling, og ved å inkludere studentene

i slikt arbeid er hun også med på å innlemme dem i et tolkningsfellesskap der hun selv, som den signifikante andre, kan gi dem respons som de kan bruke for å justere egen praksis.

Kollegaer på egen arbeidsplass

De tre lærerne oppgir tre forskjellige forklaringer når de blir bedt om å beskrive hva de legger i begrepet «rettferdig vurdering», men i sum er alle tre opptatt av at rettferdig vurdering handler om at de karakterene elevene får, gir et riktig bilde av hvilken kompetanse eleven besitter. Forskningslitteraturen gir mange eksempler på at et lavt tolkningsfellesskap i enkelte tilfeller gir så store sprik i lærernes vurderinger av elevtekster at man i realiteten må betvile vurderingsarbeidets gyldighet. I internasjonal sammenheng blir problematikken behørig belyst av for eksempel McNamara (1996) og Weigle (1998). I norsk kontekst finner Berge (1996) lignende resultat, og eksamensrapporten *Sensorreliabilitet på skriftlig eksamen i videregående opplæring* (Bjørnson & Skar, 2021) bekrefter at tolkningsfellesskapet blant norsklærere fortsatt er svakt i 2021. Gjennom denne studien har jeg derimot avdekt et unisont ønske, blant de intervjuede lærerne, om å inngå i et tolkningsfellesskap. De er dessuten opptatt av å sette «rettferdige karakterer», og i dette arbeidet trekker de frem profesjonsfellesskapet som den aller viktigste støttespilleren.

Både Lærer2 og Lærer3 er opptatt av at det er lett å være forutinntatt i vurderingen, og begge trekker i den forbindelse frem muligheten til å be en kollega om råd for å sikre at vurderingen blir rettferdig. En praksis som synes å være felles på alle skolene som lærerne jobber ved er å «bytte tekster». Lærer1 opplyser om at hun jobber i et kollegium som gjør dette systematisk:

Hvis vi er litt i tvil, så bruker vi hverandre, leser gjennom, får litt second opinion. Og det kan det være blank retting, sånn at den som skal lese gjennom teksten ikke nødvendigvis vet hvilken karakter jeg tenker på. Og så får jeg en kjapp vurdering på hva de tenker. Det gjør vi egentlig ganske mye. (Lærer1, intervju)

Her ser vi at i kollegiet til Lærer1 legger de vekt på å ikke «avsløre» karakterene for hverandre de gangene de skal hjelpe hverandre med å vurdere hva som er den riktige karakteren. Dette kan forstås nesten som en «test» på kollegiets tolkningsfellesskap. Lærer2 oppgir også at praksisen med å bytte tekster er forholdsvis systematisk på hennes arbeidssted:

Vi har ganske tett samarbeid vi som har samme trinnene, så vi samarbeider og samleser tekster. Ikke hele bunker, men kanskje 5-6 stykker, så vi er flere som går sammen og leser og diskuterer nivået, at vi er sikre på hva som kjennetegner en firer og en femmer. (Lærer2, intervju)

Praksisen mellom arbeidsstedet til Lærer1 og Lærer2 er altså ganske lik, men Lærer2 og hennes kollegium kan i større grad se ut til å forsøke å utvikle et tolkningsfellesskap, heller enn å teste om de allerede har et. Både Bakhtin og Gadamer er opptatt av dialogen, og for å utvikle profesjonsfellesskapet er man, ifølge dem, avhengig av en kontinuerlig dialog. Ecclestone (2001) understreker også viktigheten av regelmessige diskusjoner med kollegaer i arbeidet med å utvikle et tolkningsfellesskap. I så henseende kan man argumentere for at praksisen som Lærer2 beskriver kan tenkes å være mer fruktbar enn den Lærer1 beskriver. Det er derfor et paradoks at den praksisen som Lærer1 beskriver, sammenfaller med den praksisen som i hovedsak benyttes under sensorskoleringen:

Når vi har denne sensorskoleringen [...] da er det jo en stor sal full av sensorer som skal rette. Da skal vi bli enige i sensorgruppene våre om hvilken karakter en tekst vil få. Og så har vi håndsopprekning i lokalet om hvem som vil gi den karakteren, den karakteren, den karakteren, den karakteren. (Lærer1, intervju)

Av dette sitatet kommer det riktignok frem at det foreligger en diskusjon i forkant, men av konteksten i intervjuet er det tydelig at opplevelsen av å skulle avsløre karakteren man har satt gjennom en håndsopprekning gjør sterkere inntrykk. Det kan dessuten, i lys av praksisen fra egen skole som Lærer1 forteller om, virke som denne aktiviteten har potensiale til å bli tatt med tilbake til skolene og bli benyttet i et forsøk på å oppnå et vurderingsfellesskap.

Lærer3 sin opplevelse av kollegiets rolle skiller seg noe ut fra opplevelsene som både Lærer1 og Lærer2 beskriver. Hun har gode samarbeidserfaringer fra tidligere arbeidsforhold og trekker også frem erfaringer med å bytte elevtekster: «På hver eneste rettebunke så byttet vi 3 tekster». Med årene har derimot kollegiet endret seg, og selv om hun opplyser om at «jeg kan når som helst spørre en norsklærer om å hjelpe meg hvis jeg er veldig, veldig usikker», er det tydelig at hun opplever at det ikke egentlig er et ønske om å bedrive slikt samarbeid i

kollegiet: «Det er noe jeg føler jeg plager mine kollegaer med, når jeg ikke er helt sikker selv. 'Kan du ta den her i tillegg til dine?', liksom». Her kommer dessuten Lærer3 inn på et moment som ikke tidligere er belyst: Kollegaene hennes har allerede svært mye å gjøre, og hun vil ikke legge stein til byrden. Man kan derfor stille spørsmål ved i hvilken grad norsklærernes totale arbeidsbelastning er til hinder for å utvikle fungerende tolknings- og profesjonsfelleskap som kan bidra til rettferdig vurdering av elever.

Lærer1 og Lærer2 jobber altså på skoler med kultur for samarbeid, og de opplyser om at de «ofte» ber om hjelp fra en kollega dersom de er usikre i vurderingsarbeidet. Lærer3 er ikke tilfreds med samarbeidet på skolen, og hun er opptatt av at hun selv må være «ekstra nøye» i vurderingen for å sikre at den blir rettferdig, og deretter kan hun gå til en kollega hvis det er behov for det. Dette viser at profesjonsfelleskapets rolle er helt sentralt i utviklingen av et tolkningsfelleskap.

6.0 Avslutning

Denne studien handler om sluttvurdering i norskfaget i den videregående skolen, og datainnsamlingen er gjort ved semistrukturerte intervju med tre lærere. Lærerne ble bedt om å lese og vurdere en elevtekst, og under intervjuet gav alle lærerne uttrykk for at de var spente på å få vite «fasiten», altså hvilken karakter den aktuelle elevteksten hadde fått under sensuren. De gav dessuten uttrykk for å være bekymret for at karakteren de satte skulle avvike fra den karakteren som ble gitt på eksamen. Det, mener jeg, må forstås som et uttrykk for et genuint ønske om å være rettferdige. Samtidig er det også et uttrykk for at de verktøyene lærerne har til rådighet i vurderingsarbeidet ikke i tilstrekkelig grad gir den støtten som lærerne har behov for. I dette avsluttende kapittelet vil jeg presentere studiens hovedfunn, og skissere mulige veier videre i dette forskningsfeltet.

6.1 Hovedfunn og konklusjon

Studiens overordnede problemstilling er forsøkt besvart gjennom to forskningsspørsmål, hvorav det siste er delt i to ledd:

1. Hvilke faktorer peker lærerne selv på som viktige i sin vurdering?
2. Hvilke faktorer avgjør om en vurdering er rettferdig, ifølge lærerne?

- Rettferdig vurdering på tekstnivå
- Rettferdig vurdering på profesjonsnivå

Hovedfunnene som presenteres i det følgende er derfor først organisert etter forskningsspørsmål, før den overordnede problemstillingen besvares i *6.1.4 Konklusjon*.

6.1.1 Hvilke faktorer peker lærerne selv på som viktige i sin vurdering?

For det første viser analysen at lærerne peker på de samme faktorene ved elevteksten når de vurderer: *språk, struktur og innhold*. Disse faktorene finner man også i Utdanningsdirektoratets kjennetegn på måloppnåelse for norskeksamen etter vg3, og det er mye som tyder på at disse er styrende også for den vurderingen som foregår utenom eksamen. Lærerne er dessuten samstemte når de hevder at man må sette en karakter basert på en helhetsvurdering. Hva som karakteriserer god struktur eller godt innhold er det derimot større uenighet om, og til tross for at alle tre vurderer elevteksten til karakteren 4, trekker alle frem ulike sider ved teksten når de blir bedt om å si noe om hva som er tekstens styrker og svakheter. Det første hovedfunnet er dermed at det i denne studien er høy vurderingsenighet (IRA), men betydelig lavere vurderingssamsvar (IRR).

6.1.2 Hvilke faktorer avgjør om en vurdering er rettferdig, ifølge lærerne?

For det andre viser analysen at det er høyere vurderingssamsvar mellom de to lærerne som har erfaring fra sensorskolering og sensorarbeid, enn det er mellom dem og den læreren som ikke har erfaring som sensor. Læreren uten sensorerfaring er derimot den som mest aktivt bruker Utdanningsdirektoratets kjennetegn på måloppnåelse, når hun begrunner karakteren hun setter. Det andre hovedfunnet er dermed knyttet til spørsmålet om rettferdig vurdering på tekstnivå, og det viser at sensorskolering som opplæringsaktivitet synes å ha større effekt i arbeidet med å utvikle et tolkningsfellesskap, enn aktiv bruk av vurderingskriterier alene har.

For det tredje bekreftes det gjennom studien at det er ulike oppfatninger av hva som kjennetegner rettferdig vurdering. Alle tre trekker frem ulike sider ved det å vurdere rettferdig, og dilemmaet mellom den universelle og differensielle dimensjonen i spørsmålet om rettferdig karaktersetning viser seg å være høyst reelt. Alle lærerne har god kjennskap til lovverkets krav om å vurdere elevene opp mot kompetansemålene i faget, men i praksis synes

dette å være en krevende øvelse. Det er tydelig at den omsorgen lærerne viser for elevene i undervisningssammenheng i en del tilfeller gjenspeiles i vurdering- og karaktersettingsammenheng, et fenomen som også belyses av Prøitz og Borgen (2010). Det tredje hovedfunnet er dermed knyttet til spørsmålet om rettferdig vurdering på profesjonsnivå, og det viser at selv om lovverket stiller krav til universell tilnærming til vurdering, er en differensiell tilnærming forholdsvis utbredt i praksisfeltet.

6.1.4 Konklusjon

Den overordnede problemstillingen for prosjektet er: *Hvilke faktorer påvirker tre norsklæreres vurdering av en elevtekst?*, og oppsummert kan denne besvares slik: Utdanningsdirektoratets kjennetegn på måloppnåelse legger åpenbare føringer for hvilke faktorer ved elevteksten lærerne vektlegger i vurderingen, og dette er også en faktor som de tre lærerne selv er bevisste på. En annen faktor, som de i mer sprikende grad gir uttrykk for å ha et bevisst forhold til, er rettferdighetsspørsmålet. To av lærerne gir uttrykk for å sammenligne ulike elevers prestasjoner, og særlig to av lærerne er opptatt av at det er lett å la seg påvirke av forhold rundt eleven i vurderingsarbeidet. Når de er usikre i vurderingsarbeidet, har alle tre strategier for å unngå urettferdig vurdering og karaktersetting, og disse strategiene går i hovedsak ut på å søke råd i profesjonsfellesskapet. Responsen de får fra kollegaene sine, er den tredje hovedfaktoren som synes å påvirke lærernes vurderinger. Lærerne gir uttrykk for å ville «føye seg» etter kollegaenes vurderinger, og det kan forstås som et genuint ønske om å ville sette den riktige, eller den rettferdige, karakteren. Det er dessuten tydelig at profesjonsfellesskapets arbeid med å utvikle et vurderings- og tolkningsfellesskap er en viktig faktor for at hver enkelt lærer skal utvikle sin egen evne til å gi rettferdige vurderinger som er i tråd med læreplanmålene.

6.2 Videre forskning

Denne studien baserer seg på intervjuer med bare tre av svært mange norsklærere i videregående skole, og viser til dels lavt vurderingssamsvar til tross for at vurderingsenigheten er høy. I lys av tidligere vurderingsforskning som har konkludert med at lærerne mangler fullgode tolkningsfellesskap, er det nærliggende å tenke at vurderingsenigheten hadde vært lavere ved å studere et bredere utvalg. Med dette som utgangspunkt hadde det vært interessant å gjennomføre denne studien i større skala, som en mixed methods-studie. En

større survey-undersøkelse der mange lærere ble bedt om å vurdere og karaktersette den samme elevteksten, hadde vært aktuelt, og denne kunne man ha fulgt opp med dybdeintervjuer med et større antall lærere. For å få et innblikk i utviklingen av et vurderingsfellesskap og felles forståelse for vurderingskriteriene i et kollegium, kunne man også ha foretatt observasjon av for eksempel fagmøter med vurderingspraksis på agendaen.

Det er gjort mange, både store og små studier, som omhandler karaktersetting og sprikvurdering i norsk skole. Gjennom kunnskapsgrunlaget som ligger til grunn for denne studien, avdekkes det likevel at det i nyere tid er gjort få omfattende studier av hvilke mekanismer som ligger til grunn for de sluttvurderingene som gjøres i norsk skole, og her er det fortsatt forskningshull å fylle. Slik forskning vil kunne gi kunnskap som kan bidra i arbeidet med å utvikle et større tolkningsfellesskap blant de lærerne som foretar sluttvurderinger med karakter i skolen. Tittelsitatet «Jeg tenker det er en sterk firer, jeg da. Men det kan hende at det egentlig er stryk.» viser at slik kunnskap er nødvendig for å sikre rettferdig vurdering i skolen.

7.0 Referanser

Aftenposten (2018, 26. juni). Sanner vil ha slankere og mer praktiske skolefag.

Alvesson, M., & Sköldbberg, K. (2018). *Reflexive Methodology: New Vistas for Qualitative Research* (3. utg.). SAGE.

Anderson, S & Hertzberg, F. (2005). Å ta en risk – «utfordrerne» blant eksamensskriverne. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse: B.2: Norskeksamen som tekst* (s. 275-301). Oslo: Universitetsforlaget.

Bach, C. (2015). Eksamensoppgaver i norsk etter 13 år: sensorstemmer fra våren 2012 og våren 2014. I H. Otnes (Red.), *Å invitere elever til skiving: ulike perspektiver på skriveoppgaver* (s. 29–44).

Bakhtin, M. M. & Slaattelid, R. T. (1998). *Spørsmålet om talegenrane*. Ariadne forlag.

Berge, K. L. (1996). Norsksensorenes tekstnormer og doxa: en kultursemiotisk og sosioteknologisk analyse. Doktoravhandling, Norges teknisk-naturvitenskapelige universitet, Trondheim.

Berge, K. L. & Hertzberg, F. (2005). Etertanker – og tanker om framtidens utfordringer. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse: B.2: Norskeksamen som tekst* (s. 387-393). Oslo: Universitetsforlaget.

Berge, K. L. (2005a). Skiving som grunnleggende ferdighet og nasjonal prøve – ideologi og strategier. I A.J. Aasen & S. Nome (Red.). *Det nye norskfaget* (s. 161-187). Vigmostad & Bjørke AS.

Berge, K. L. (2005b). Skriveprøvenes pålitelighet. I K. L. Berge, L. S. Evensen, F. Hertzberg & W. Vagle (Red.), *Ungdommers skrivekompetanse: B.1: Norsksensuren som kvalitetsvurdering* (s. 101-113). Oslo: Universitetsforlaget.

- Berge, K. L. (2009). Er tolkningsfellesskap mulig å oppnå i skriveprøver?. I O.K. Haugaløkken, F. Hertzberg, L.S. Evensen & H. Otnes (Red.), *Tekstvurdering som didaktisk utfordring* (s. 44-54). Universitetsforlaget.
- Berge, K. L. (2014). Fra eksamensskrivning i norsk til skriving som grunnleggende ferdighet - to vurderingsutfordringer. I R. Hvistendahl & A. Roe (Red.), *Alle tiders norskdiraktiker – Festskrift til Frøydis Hertzberg på 70-årsdagen den 18. november 2014* (s. 259-277). Oslo: Novus Forlag.
- Berger, P. & Luckmann, T. (1966) *The Social Construction of Reality. A Treatise in the Sociology of Knowledge*. London: Penguin Books.
- Bjørnsson, J. & Skar, G. (2021). *Sensorreliabilitet på skriftlig eksamen i videregående opplæring*. Institutt for lærerutdanning og skoleforskning. Universitetet i Oslo
- Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education: Principles, Policy & Practice*, 18(4), 451–469.
<https://doi.org/10.1080/0969594X.2011.557020>
- Czarniawska, B. (2004). *Narratives in social science research*. Sage.
- Dahl, T., Grut, G. & Østerås, A. K. (2005). Kan vi stole på karakterene? Innhold og vurdering i samfunnsfag i norsk skole. *Bedre skole*, (3), 46-51.
<https://utdanningsforskning.no/artikler/2015/kan-vi-stole-pa-karakterene-innhold-og-vurdering-i-samfunnsfag-i-norsk-skole/>
- Dysthe, O., Bernhardt, N., Esbjørn, L., & Strømsnes H. (2012). *Dialogbasert undervisning: Kunstmuseet som læringsrom*. Fagbokforlaget.
- 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: <https://doi.org/10.1080/03098770126527>

- Eckes, T. (2008). Rater types in writing performance assessments: A classification approach to rater variability. *Language Testing*, 25(2), 155–185.
<https://doi.org/10.1177/0265532207086780>
- Elder, C., Knoch, U., Barkhuizen, G., & von Randow, J. (2005). Individual Feedback to Enhance Rater Training: Does It Work? *Language Assessment Quarterly*, 2(3), 175–196.
https://doi.org/10.1207/s15434311laq0203_1
- Emerson, C. & Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics: Mikhail Bakhtin; edited and translated by Caryl Emerson; introduction by Wayne C. Booth*. University of Minnesota Press.
- Engelsen, B. U. (2020). Læreplanens generelle del - et historisk perspektiv. *Norsk Pedagogisk Tidsskrift*, 104(2), 206–217. <https://doi.org/10.18261/issn.1504-2987-2020-02-10>
- Engh, R., Dobson, S. & Høihilder, E. K. (2007). *Vurdering for læring*. Høyskoleforlaget.
- Evensen, L. S. (2009). Vurdering av skrivekompetanse: En kompleks utfordring. I O.K. Haugaløkken, F. Hertzberg, L.S. Evensen & H. Otnes (Red.), *Tekstvurdering som didaktisk utfordring* (s. 15-23). Universitetsforlaget.
- Findlay, I. R., & Purves, A. C. (1993). The IEA Study of Written Composition II: Education and Performance in Fourteen Countries [Review of The IEA Study of Written Composition II: Education and Performance in Fourteen Countries]. *British Journal of Educational Studies*, 41(3), 313–314. Blackwell Publishing Ltd. <https://doi.org/10.2307/3122297>
- Forskrift til opplæringslova (2020). Forskrift til opplæringslova (FOR-2020-06-29-1474). Lovdata. <https://lovdata.no/LTI/forskrift/2020-06-29-1474>
- Fuglseth, K. (2018). Vitskapsteori og hermeneutikk. I M. Krogtoft & J. Sjøvoll (Red.), *Masteroppgaven i lærerutdanninga* (s. 245-264). Cappelen Damm Akademisk.
- Gadamer, H.-G. (2001): Opphøyelsen av forståelsen historitet til hermeneutisk prinsipp. I S. Lægroid., T. Skorgen & E. B. Hagen (Red.), *Hermeneutisk lesebok* (s. 115-136). Spartacus.

- Gardner, J. (2007). Is teaching a 'partial' profession?. *Making the Grade*, Summer: 18–21.
https://www.researchgate.net/publication/324165229_Is_teaching_a_%27partial%27_profession
- Grønmo, S. (2016). *Samfunnsvitenskapelige metoder*. 2.utgave. Bergen: Fagbokforlaget.
- Holand, A. (2018) Oversiktsstudier og spørreskjema. I M. Krogtuft & J. Sjøvoll (Red.),
Masteroppgaven i lærerutdanninga. (2. utg., s. 95-115). Cappelen Damm Akademisk.
- Igland, M.A., Dysthe, O. (2001): Michael Bakhtin og sosiokulturell læringsteori. I O. Dysthe
(Red.), *Dialog, samspel og læring* (s. 107-128). Abstrakt forlag.
- Jencks, C. (1988). Whom must we treat equally for educational opportunity to be equal?
Ethics, 98(3), 518-533. <https://doi.org/10.1086/292969>
- Kaiser, K. (2012). Protecting confidentiality. I J. F. Gubrium, J. A. Holstein, A. B. Marvasti & K.
D. McKinney (Red.), *Handbook of interview research: The complexity of the craft*. (2.
utg.), (s. 457-464). Thousand Oaks, CA: Sage.
- Kjørup. (2008). *Menneskevidenskabene: 2: Humanistiske forskningstraditioner* (2. utg.).
Roskilde Universitetsforlag.
- Kleven, T. A., & Hjordemaal, F. (2018). *Innføring i pedagogisk forskningsmetode: en hjelp til
kritisk tolking og vurdering* (3. utg.). Fagbokforlaget.
- Kunnskapsdepartementet (2013). *Læreplan i norsk (NOR1-05)*. Fastsatt som forskrift.
Læreplanverket for Kunnskapsløftet 2013. <https://www.udir.no/kl06/nor1-05?lplang=http://data.udir.no/kl06/nob>
- Kunnskapsdepartementet (2019). *Læreplan i norsk (NOR01-06)*. Fastsatt som forskrift.
Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/nor01-06?lang=nob>
- Kvale, S., Brinkmann, S., Anderssen, T. M., & Rygge, J. (2015). *Det kvalitative
forskningssintervju* (3. utg.). Gyldendal akademisk.

- Kvistad, A. H. & Otnes, H. (2019). Mottakerinstansen i skoleskriving: En studie av skriveoppgaver fra Normprosjektet. *Nordic journal of literacy research*, 5(2), 100–119. <https://doi.org/10.23865/njlr.v5.1644>
- L93: Læreplan for grunnskole, videregående opplæring, voksenopplæring. Generell del. 1993. Oslo: Kirke- utdannings- og forskningsdepartementet.
- LK06: Læreplanverket for Kunnskapsløftet. Utdannings- og forskningsdepartementet (opprinnelig; senere: Kunnskapsdepartementet; nå: Kunnskaps- og integreringsdepartementet). Hentet fra <https://www.udir.no/laring-og-trivsel/lareplanverket/>
- Lortie, D. (1975). *Schoolteacher: A Sociological study*. University of Chicago Press.
- Marzano, M. (2012). Informed consent. I J. F. Gubrium, J. A. Holstein, A. B. Marvasti & K. D. McKinney (Red.), *Handbook of interview research: The complexity of the craft*. (2. utg.), (s. 443-456). Thousand Oaks, CA: Sage.
- Matre, S. & Solheim, R. (2014). Lærersamtalar om elevtekstar: Mot eit felles fagspråk om skrivning og vurdering. I R. Hvistendahl & A. Roe (Red.), *Alle tiders norskdidaktiker: Festskrift til Frøydís Hertzberg på 70-årsdagen den 18. november 2014* (s. 219-243). Oslo: Novus forlag.
- Matre. (2017). *Bruk av forventningsnormer i skriveopplæringa: Erfaringer fra en intervensjonsstudie om skrivning og vurdering*. The Swedish Association for Applied Linguistics.
- McNamara, T. F. (1996). *Measuring second language performance*. New York: Longman.
- Mead, H. G. (1934). *Mind, Self, & Society from the Standpoint of a Social Behaviorist*. London: The University of Chicago Press.
- Mead, G. H. & Reck, A. J. (1964). *Selected writings*. The University of Chicago Press

- Meld. St. 28 (2015-2016). *Fag – Fordypning – Forståelse – En fornyelse av Kunnskapsløftet*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/>
- 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* (s. 115-160). Oslo: Universitetsforlaget.
- NOU 2014: 7 (2014). *Elevenes læring i fremtidens skole – Et kunnskapsgrunnlag*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/NOU-2014-7/id766593/>
- NOU 2015: 8 (2015). *Fremtidens skole – Fornyelse av fag og kompetanser*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2015-8/id2417001/>
- Overordnet læreplandel (2017). *Overordnet del – verdier og prinsipper*. Oslo: Utdanningsdirektoratet. Hentet fra <https://www.udir.no/lk20/overordnet-del/>
- Postholm, M. B. & Jacobsen, D. I. (2018). *Forskningsmetode for materstudenter i lærerutdanning*. Cappelen Damm Akademisk.
- Prøitz, T. S. & Borgen, J. S. (2010). *Rettferdig standpunktvurdering - det (u)muliges kunst?: læreres setting av standpunkt karakter i fem fag i grunnopplæringen* (NIFU STEP Rapport 16/2010). <https://nifu.brage.unit.no/nifu-xmlui/bitstream/handle/11250/279131/NIFUrapport2010-16.pdf?sequence=1&isAllowed=y>
- Resh, N. (2009). Justice in grades allocation: teachers' perspective. *Social Psychology of Education*. 12(3), 315-325. <https://doi.org/10.1007/s11218-008-9073-z>
- Rye, J. F. (2013), Mead, Berger & Luckmann og de signifikante andre. *Sosiologisk tidsskrift*, 21(2), 169–189. <https://doi.org/10.18261/ISSN1504-2928-2013-02-04>
- Sadler, R. (2010). Beyond feedback: developing student capability in complex appraisal. I: *Assessment & Evaluation in Higher Education*, Vol. 35 (5), s. 535–550.

- Skar, G. B. U. & Aasen, A. J. (2018). Å måle skrivning som grunnleggende ferdighet. *Acta didactica Norge*. 12(4). <https://doi.org/10.5617/adno.6280>
- Solheim, R. & Matre, S. (2014). Forventninger om skrivekompetanse. Perspektiver på skrivning, skriveopplæring og vurdering i «Normprosjektet». *Viden om Læsning* 2014 (15), s. 76-89. https://skrivesenteret.no/wp-content/uploads/2021/02/videnom_15_8.pdf
- Thagaard, T. (2018). *Systematikk og innlevelse: en innføring i kvalitative metoder* (5. utg.). Fagbokforlaget.
- Tinsley, H.E.A & Weiss, D.J. (2000). Interrater reliability and agreement. I H.E.A. Tinsley & S.D. Brown (Red.), *Handbook of Applied Multivariate Statistics and Mathematical Modelling* (s. 95-124). Academic Press.
- Utdanningsdirektoratet (2018a). *Eksamensveiledning – om vurdering av eksamenssvar. Norsk vg3.*
- Utdanningsdirektoratet (2018b). *Eksamen NOR1211-NOR1231 Norsk hovedmål. 25.05.2018.*
- Utdanningsdirektoratet (2018c). *Sensorveiledning NOR1211-NOR1231 Norsk hovedmål.*
- Utdanningsdirektoratet (2020, 16. juni). *Vurderinger og anbefalinger om fremtidens eksamen.* <https://www.udir.no/eksamen-og-prover/eksamen/vurderinger-og-anbefalinger-fremtidens-eksamen/>
- Utdanningsdirektoratet. (2021a, 26. august). *Endringer i norskeksamen etter ny læreplan for vg3.* <https://www.udir.no/eksamen-og-prover/eksamen/eksempeloppgaver/eksempeloppgaver-norsk-vg3/>
- Utdanningsdirektoratet. (2021b). *Eksamensveiledning – om vurdering av eksamenssvar. Norsk vg3.*
- Utdanningsdirektoratet. (2022, 02. mars). *Standpunktvurdering.* <https://www.udir.no/laring-og-trivsel/vurdering/sluttvurdering/standpunktvurdering/>

- 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* (s. 35-99). Oslo: Universitetsforlaget.
- Vassenden, A. & Andrews, T. (2007). Snøballen som ikke ruller. Utvalgsproblemer i kvalitativ forskning. *Sosiologisk tidsskrift*, 15(2), 151-162 <https://doi.org/10.18261/ISSN1504-2928-2007-02-02>
- Vygotsky, L., Cole, M., John-Steiner, V., Scribner, S. & Souberman, E. (1978). *Mind in society: the development of higher psychological processes*. Harvard University Press.
- Vygotsky, L. (2001). *Tenkning og tale*. Gyldendal akademisk.
- Weigle, S. C. (1998). Using FACETS to model rater training effects. *Language Testing*, 15(2), 263-287. Doi: <https://doi.org/10.1177/026553229801500205>

Vedlegg

| | |
|--|-----|
| Vedlegg I – Informasjonsskriv og samtykkeskjema for bruk av elevtekst..... | I |
| Vedlegg II – Informasjonsskriv og samtykkeskjema for deltakere i intervju..... | III |
| Vedlegg III – Intervjuguide..... | V |
| Vedlegg IV – Godkjenning fra NSD (skjerm bilde) | VI |

Vedlegg I – Informasjonsskriv og samtykkeskjema for bruk av elevtekst

Vil du delta i forskningsprosjektet

«Vurderingssamsvar mellom norsklærere i videregående skole?»

Dette er et spørsmål til deg om å delta i et forskningsprosjekt om vurderingssamsvar mellom norsklærere i videregående skole. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Prosjektet er en masteroppgave i studieprogrammet «Erfaringsbasert master i undervisning og læring», ved Høgskulen i Volda. Formålet med prosjektet er å utvikle ny kunnskap om summativ vurdering i videregående skole.

Hvem er ansvarlig for forskningsprosjektet?

Høgskulen i Volda er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Du får denne henvendelsen fordi jeg ønsker å bruke din eksamensbesvarelse fra hovedmåleeksamen i 2018 som utgangspunkt for deler av intervjuene jeg skal gjennomføre i forbindelse med masteroppgaven.

Hva innebærer det for deg å delta?

Dersom du samtykker til at din eksamensbesvarelse kan brukes i prosjektet må du gi meg digital tilgang til besvarelsen og eksamensresultat. Ved å signere dette samtykkeskjemaet gir du meg rettigheter til å bruke eksamensbesvarelsen din og informasjon om eksamensresultatet til det formål som er skissert over. Det vil også innebære at besvarelsen blir delt med de som skal delta i intervjuene. Besvarelsen blir anonymisert og det vil ikke være mulig for de som skal delta i intervjuene å identifisere deg.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

Navn og kontaktinformasjon vil erstattes med en kode som er lagret i en egen liste og oppbevares på en egen disk, adskilt fra øvrige data. Opplysningene behandles konfidensielt, og det vil ikke være mulig å gjenkjenne enkeltpersoner i den ferdige oppgaven.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene, både besvarelse og karakter, slettes når prosjektet er avsluttet og oppgaven er godkjent, noe som etter planen er august 2022.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke. På oppdrag fra Høgskulen i Volda har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke opplysninger vi behandler om deg, og å få utlevert en kopi av opplysningene
- å få rettet opplysninger om deg som er feil eller misvisende
- å få slettet personopplysninger om deg
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger

Hvis du har spørsmål til studien, eller ønsker å vite mer om eller benytte deg av dine rettigheter, ta kontakt med:

- Høgskulen i Volda ved student Synnøve Skare Baade på e-post synnove@baadeweb.com eller telefon 92 21 19 61, eller veileder Pernille Fiskerstrand på e-post pernille.fiskerstrand@hivolda.no eller telefon 70 02 54 61
- Vårt personvernombud: Cecilie Røeggen på e-post cecilie.roeggen@hivolda.no eller telefon 700 75 073

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

- NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller på telefon: 53 21 15 00.

Med vennlig hilsen

Pernille Fiskerstrand
Veileder

Synnøve Skare Baade
Student

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet *Vurderingssamsvar mellom norsklærere i videregående skole* og har fått anledning til å stille spørsmål. Jeg samtykker til:

- at en anonymisert versjon av min eksamensbesvarelse kan brukes som utgangspunkt for intervju
- at en anonymisert versjon av min eksamensbesvarelse kan deles digitalt med de som skal delta i intervjuet
- at intervjudeltakerne får vite (muntlig) hvilken karakter besvarelsen opprinnelig ble vurdert til
- at intervjudeltakerne får vite (muntlig) hvilken karakter besvarelsen ble vurdert til etter klage på opprinnelig eksamenskarakter

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet

(Signert av prosjektdeltaker, dato)

Vil du delta i forskningsprosjektet

«Vurderingssamsvar mellom norsklærere i videregående skole»?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt om vurderingssamsvar mellom norsklærere i videregående skole. I dette skrevet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Prosjektet er en masteroppgave i studieprogrammet «Erfaringsbasert master i undervisning og læring», ved Høgskulen i Volda. Formålet med prosjektet er å utvikle ny kunnskap om summativ vurdering i videregående skole.

Hvem er ansvarlig for forskningsprosjektet?

Høgskulen i Volda er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Du får denne henvendelsen fordi du har undervisningserfaring fra norskfaget i videregående skole.

Hva innebærer det for deg å delta?

Hvis du velger å delta i prosjektet, innebærer det at du stiller til et intervju, og at du i forkant av intervjuet leser og vurderer en elevbesvarelse fra en tidligere gitt eksamen. Spørsmålene vil handle om hva du vektlegger når de vurderer elevbesvarelser i norsk, og elevbesvarelsen du har vurdert i forkant vil være tema under deler av intervjuet. For at jeg skal kunne gjengi dine uttalelser korrekt vil jeg ta notater og lydopptak under intervjuet. Omfanget av intervjuet vil ikke gå over 60 minutt, og i tillegg kommer tid til å vurdere en elevbesvarelse i forkant av intervjuet.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrevet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

Navn og kontaktinformasjon vil erstattes med en kode som er lagret i en egen liste og oppbevares på en egen disk, adskilt fra øvrige data. Opplysningene behandles konfidensielt, og det vil ikke være mulig å gjenkjenne enkeltpersoner i den ferdige oppgaven.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene, både notater og lydfiler, slettes når prosjektet er avsluttet og oppgaven er godkjent, noe som etter planen er august 2022.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke. På oppdrag fra Høgskulen i Volda har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke opplysninger vi behandler om deg, og å få utlevert en kopi av opplysningene
- å få rettet opplysninger om deg som er feil eller misvisende
- å få slettet personopplysninger om deg
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger

Hvis du har spørsmål til studien, eller ønsker å vite mer om eller benytte deg av dine rettigheter, ta kontakt med:

- Høgskulen i Volda ved student Synnøve Skare Baade på e-post synnove@baadeweb.com eller telefon 92 21 19 61, eller veileder Pernille Fiskerstrand på e-post pernille.fiskerstrand@hivolda.no eller telefon 70 02 54 61
- Vårt personvernombud: Cecilie Røeggen på e-post cecilie.roeggen@hivolda.no eller telefon 700 75 073

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

- NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller på telefon: 53 21 15 00.

Med vennlig hilsen

Pernille Fiskerstrand
Veileder

Synnøve Skare Baade
Student

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet *Vurderingssamsvar mellom norsklærere i videregående skole* og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i intervju

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet

(Signert av prosjektdeltaker, dato)

Vedlegg III – Intervjuguide

Vurdering i norskfaget

Spørsmålene er nummererte fordi det skal være lett for intervjuer å navigere, ikke fordi spørsmålene må komme i akkurat denne rekkefølgen. Det er likevel naturlig at man holder seg til denne rekkefølgen til en viss grad da noen av spørsmålene bygger på hverandre. Spørsmål nummer 17 må for eksempel komme etter spørsmål 13-16. De spørsmålene som er skrevet i blått er spørsmål som deltakerne vil få tilsendt på forhånd.

Del 1. Innledning

- Hvem er jeg? Hvorfor er jeg her?
- Hva er hensikten med intervjuet?
- Hvordan vil materialet bli behandlet?

Del 2. Bakgrunnsinformasjon om deltakeren

1. Hvor lenge har du jobbet som lærer?
2. Hvilke fag underviser du i? Har du undervist i andre fag tidligere?
3. Hvor lenge har du undervist i norsk?
4. Fra hvilke trinn har du undervisningserfaring i norsk? Er det noen trinn du har mer erfaring fra enn andre?
5. Hvilken utdanning/fagbakgrunn har du (både norsk og andre fag)?
6. Når var du ferdig utdannet lærer? Har du tatt mer utdanning på senere tidspunkt?
7. Hvilke erfaringer har du med å sette karakterer sammen med andre?
 - Under utdanning - på skolen og/eller i praksis?
 - Fra tidligere/nåværende arbeidssted?
 - Fra sensorarbeid? Har du i tilfelle deltatt på sensorskolering?

Del 3. Vurdering i norsk hovedmål

PRESISERINGER:

- Med mindre noe annet er presisert så menes «norsk hovedmål» når jeg i det følgende snakker om «norsk» eller «norskfaget».
 - I forkant av intervjuet har du blitt bedt om å vurdere en elevtekst. Vi skal komme til denne etter hvert, men jeg vil først høre dine tanker om vurdering på et mer generelt nivå.
8. Hva mener du kjennetegner en god elevbesvarelse i norsk?
 9. Hva vektlegger du når du setter karakter på en elevtekst i norsk?
 10. Hvordan går du frem når du vurderer en elevbesvarelse? Bruker du hjelpemidler og/eller systemer når du setter karakter i norsk?
 11. Hva legger du i «å vurdere rettferdig»?
 12. Synes du det er «lett» eller «vanskelig» å gi rettferdige karakterer?
 - Hvis «lett»: hva er det som gjør det lett?
 - Hvis «vanskelig»: hva er det som gjør det vanskelig? Har du tanker om hva som kunne gjort det lettere for deg?

Del 4. Vurdering av elevbesvarelse

13. I forkant av intervjuet har jeg bedt deg om å vurdere en elevbesvarelse. Hvilken karakter har du vurdert denne teksten til?
14. Hvorfor har du satt denne karakteren?
15. Dersom du skulle gitt en fremovermelding til eleven som har skrevet denne teksten: Hva ville du ha trukket frem som særlig godt, og hva ville du ha trukket frem som det viktigste forbedringspunktet?
16. Har du diskutert den aktuelle elevbesvarelsen med kollegaer / andre lærere?
 - Hvis ja: kan du fortelle litt om hva som ble sagt?
 - Hvis nei: hvordan tror du at kollegaene dine hadde vurdert denne teksten? Hva tror du de hadde trukket frem som tekstens styrker og svakheter?
17. Om eksamensresultatet:
 - Denne teksten ble opprinnelig vurdert til karakteren 3 på eksamen. Hva tenker du om det?
 - Kandidaten klaget og fikk medhold. Den nye karakteren ble 4. Hva tenker du om det? Hva tror du det nye sensorparet har sett/lagt vekt på som det opprinnelige sensorparet ikke vektla, eller omvendt?

Del 5. Avslutning

18. Er det noe du ønsker å presisere eller føye til, utover det vi allerede har snakket om?

Takk for at du har stilt opp!

Vedlegg IV – Godkjenning fra NSD (skjerm bilde)

Vurdering

 Skriv ut

Referansenummer

818821

Prosjektittel

Vurderingssamsvar mellom norsklærere i videregående skole

Behandlingsansvarlig institusjon

Høgskulen i Volda / Avdeling for humanistiske fag og lærarutdanning / Institutt for språk og litteratur

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Pernille Fiskerstrand, pernille.fiskerstrand@hivolda.no, tlf: 70075461

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Synnøve Skare Baade, synnove@baadeweb.com, tlf: 92211961

Prosjektperiode

01.11.2021 - 31.08.2022

Vurdering (1)**19.11.2021 - Vurdert**

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 19.11.2021, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

DEL PROSJEKTET MED PROSJEKTANSVARLIG

For studenter er det obligatorisk å dele prosjektet med prosjektansvarlig (veileder). Del ved å trykke på knappen «Del prosjekt» i menylinjen øverst i meldeskjemaet. Prosjektansvarlig bes akseptere invitasjonen innen en uke. Om invitasjonen utløper, må han/hun inviteres på nytt.

TYPE OPPLYSNINGER OG VARIGHET

Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 31.08.2022.

LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake.

Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER

NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), og dataportabilitet (art. 20).

NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER

NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

Ved bruk av databehandler (spørreskjemalieferandør, skylagring eller videosamtale) må behandlingen oppfylle kravene til bruk av databehandler, jf. art 28 og 29. Bruk leverandører som din institusjon har avtale med.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

MELD VESENTLIGE ENDRINGER

Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

<https://www.nsd.no/personverntjenester/fylle-ut-meldeskjema-for-personopplysninger/melde-endringer-i-meldeskjema>

Du må vente på svar fra NSD før endringen gjennomføres.

OPPFØLGING AV PROSJEKTET

NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!