NTNU
Det skapende universitet

Flervalgsoppgaver og digitale læringsplattformer

En kvalitativ studie av forberedelse til den første delen av skriftlig eksamen i Biologi 2

Tonje Tomine Seland Olsen

Master i lærerutdanning med realfag
Oppgaven levert: Mai 2009
Hovedveileder: Alex Strømme, BIO

Norges teknisk-naturvitenskapelige universitet
Institutt for biologi

Forord

Nå er masteroppgaven skrevet ferdig, og arbeidet har væt svært tidkrevende, spennende, frustrerende, interessant og lærerikt. Totalt sett har arbeidet vært en positiv prosess, som har gitt meg en stadig større grad av innsikt og forståelse for spørsmålene i problemstillingene. Masteroppgaven er et sluttprodukt som jeg har skrevet alene. Til tross for at oppgaven ble skrevet alene, har det vært viktig for meg å diskutere innholdet med noen. Jeg skylder noen mennesker takk enten for at de har lest eller bidratt med innvendinger og kommentarer underveis eller for at de har støttet meg underveis for at jeg skulle nå målet.

Først og fremst vil jeg takke min veileder, Alex Strømme, som har vært en viktig støttespiller gjennom hele arbeidet med masteroppgaven. Han har gitt meg gode, konstruktive og kritiske veiledninger som har hjulpet meg gjennom hele prosessen for å få til dette sluttproduktet. Takk til mine respondenter, lærer og elever, for at de stilte opp. Uten deres innsats ville ikke denne oppgaven fått noe resultat. Takk til Berit Bungum, som har gitt meg mange tilbakemeldinger innenfor metodedelen.

Jeg har hatt et liv på siden av studiene også, selv om det til tider har blitt lite av det. I den forbindelse må jeg takke mine medstudenter. Dere har vært helt fantastiske til få meg ut av mastergradsboblen som jeg har tilbrakt mye tid inni. Takk til min kjære mamma, som alltid har vært der for meg. Hun har alltid tid til en telefon fra sin datter som trenger litt trøst og gode ord på veien, eller bare har behov for å få ut litt frustrasjon. Og til sist, takk til min fantastiske kjæreste, Torgeir Strat, som har fulgt interessert med gjennom hele prosessen. Han har vært en strålende motivator og støttet meg når det har stått på som verst. Han har også brukt mye tid til på å korrekturlese oppgaven. Han har rett og slett vært til uvurderlig hjelp!

Trondheim, 28.mai.2009

Tonje Tomine Seland Olsen

Sammendrag

Denne masteroppgaven ble skrevet som et avsluttende ledd i studiet Lærerutdanning i realfag ved NTNU. Hensikten med oppgaven var å forberede elever på bruk av flervalgsoppgaver som vurderingsform. Det er en av vurderingsformene som blir benyttet på skriftlig eksamen i Biologi 2 fra våren 2008. I denne studien ble det fokusert på læringsverktøyet Test 2.0 i it's learning som metode for å forberede elevene til denne eksamenen. Det ble benyttet en testgruppe bestående av en lærer og 18 elever som gjennomførte undervisning i Biologi 2. I studien skulle elevene utføre to prøver: en flervalgsprøve i Test 2.0 og en kortsvarsprøve på papir. Studien ble gjennomført for å finne ut i hvilken av vurderingsmetodene elevene får vist sin kompetanse fra læreplanmålet *Genetikk* i Kunnskapskapsløftet best.

Den teoretiske forankringen baserer seg på litteratur innenfor flervalgsoppgaver med utgangspunkt i fordeler og utfordringer ved bruk av flervalgsoppgaver som vurderingsform. Det ble også fremlagt teori rundt bruk av digitale tester i læringsplattformer. Metoden i undersøkelsen ble basert på en kasusstudie med bruk av kvantitative og kvalitative metoder. Til datainnsamlingen ble det utført spørreundersøkelser med avkrysningspåstander og åpne spørsmål, samt intervju av lærer og elever. Tendensene fra datamaterialet viste at elevene fikk vist sin kompetanse godt ved bruk av flervalgsoppgaver som vurderingsform, men flere av elevene mente at de fikk vist mer av sin kompetanse gjennom bruk av kortsvarsoppgaver.

Læreren i denne studien hadde ikke benyttet Test 2.0 tidligere, og han var positiv til dette læringsverktøyet etter flervalgsprøven. Han så mange muligheter for å benytte dette verktøyet for underveisvurdering og for å forberede elevene til skriftlig eksamen. Sluttkonklusjonen av denne studien er at flervalgsprøver ikke må betraktes som en erstatning for-, men som et verdifullt supplement til de andre prøveformene som benyttes i undervisningen. Det gjelder også for Test 2.0. Den kan tilby elevene en ny læringsform, og den kan utgjøre et viktig supplement i lærerens pedagogiske verktøykasse.

Nøkkelord: Vurdering *av* læring, vurdering *for* læring, skriftlig eksamen i Biologi 2, flervalgsoppgaver, it's learning, Blooms taksonomi, kasusstudie.

Innholdsfortegnelse

1. Introduksjon 9
1.1. Bakgrunn for valg av oppgave 9
1.2. Problemstillinger 11
1.3. Oppbygning av oppgaven 11
2. Flervalgsoppgaver 12
2.1. Produksjon av flervalgsoppgaver 13
2.2. Fordeler 14
2.3. Utfordringer 15
2.4. Analyse av enkeltoppgaver 15
2.5. Reliabilitet og validitet 16
2.6. Gjetting på flervalgsprøver 19
2.7. Karaktersetting av flervalgsprøver 22
3. Digitale læringsplattformer 23
3.1. Kunnskapsløftet 23
3.2. Digitale læringsplattformer 24
3.3. Digitale tester og bruk av Test 2.0 i it's learning 26
4. Blooms taksonomi og flervalgsoppgaver 29
4.1. Blooms taksonomi 30
4.2. Bruk av Blooms taksonomi i flervalgsoppgaver 32
5. Metode 35
5.1. Innledning 35
5.2. Kvalitativ metode 36
5.3. Utvalg 36
5.4. Metodevalg 39
6. Datainnsamling og resultater 45
6.1. Flervalgsoppgaver 45
6.2. Kortsvarsprøven 61
6.3. Flervalgsoppgaver og kortsvarsoppgaver på eksamen 67
6.4. Bruk av Test 2.0 for å forberede elevene til skriftlig eksamen 76
7. Diskusjon 77
7.1. Oppsummering av resultatet 77
7.2. Fikk elevene vist sin faglige kompetanse ved bruk av flervalgsoppgaver som vurderingsform? 78
7.3. Hvordan fungerte testverktøyet Test 2.0 i it's learning til å forberede elevene til skriftlig eksamen i Biologi 2? 81
7.4. Erfaringer fra datainnsamlingen 83
7.5. Erfaringer fra metodevalget 91
7.6. En kritisk oppsummering av studien 95
8. Etterord 99
9. Litteraturliste 101

Vedlegg

Vedlegg 1: Kontrakt mellom forskeren og utvalget

Vedlegg 2: Tematiske kategorier for å analysere datamaterialet

Vedlegg 3: Spørreundersøkelsen

Vedlegg 4: Intervjuspørsmål til elever

Vedlegg 5: Intervjuspørsmål til lærer

Vedlegg 6: Flervalgsprøven

Vedlegg 7: Resultater på flervalgsprøven uten og med korrigering for gjetting

Vedlegg 8: Kortsvarsprøven

Vedlegg 9: Resultater på flervalgsprøven, kortsvarsprøven og total karakter

Vedlegg 10: Postintervjuspørsmål til læreren

1. Introduksjon

1.1. Bakgrunn for valg av oppgave

Vurdering har vært et omstridt tema i norsk skole, og begrepene *vurdering av læring* og *vurdering for læring* dukket opp i internasjonal sammenheng for å forklare vurderingens formål (Birenbaum et al., 2006). Vurdering *av* læring og vurdering *for* læring er to ulike to måter å gi både lærere og elever innsikt i elevenes læringsprosess. Vurdering *av* læring blir ofte omtalt som en lite fleksibel vurderingsform som baserer seg på resultatbaserte prøver som har til formål å kartlegge læringsresultater. I følge Birenbaum et al. (2006) kan en overfokusering på vurdering *av* læring føre til en *teach for the test – undervisning*. Læreren legger da opp til at elevene lærer kun de læringsmål som skal være gjenstand for vurderingen. Vurdering *for* læring er derimot å samle inn informasjon om elevenes læringsutbytte eller prosess med formålet å forbedre læring. Den tar også hensyn til elevenes individuelle behov, og den er en fleksibel og lite kontekstsavhengig vurderingsform (Birenbaum et al., 2006).

I denne masteroppgaven er det tatt utgangspunkt i den skiftlige eksamensformen fra våren 2008 i Biologi 2 for å undersøke hvordan en lærer kan forberede elever på denne eksamensformen. Skriftlig eksamen er et typisk eksempel på vurdering *av* læring, og for at læreren skal kunne veilede elevene best mulig til denne skriftlige eksamenen, må det forekomme vurdering *for* læring i undervisningen ved underveisvurdering.

To former for skriftlig eksamen

I følge Utdanningsdirektoratet (2008c) er det to former for skriftlig eksamen. Den første formen er den tradisjonelle metoden for skriftlig eksamen. Denne metoden er en eksamen med eller uten forberedelsesdel, der eksamenstiden er satt til fem timer. På en eventuell forberedelsesdag er alle hjelpemidler tillatt, også bruk av internett. På eksamen er også alle hjelpemidler tillatt, med unntak av internett og andre verktøy som tillater kommunikasjon (Utdanningsdirektoratet, 2008c). Den andre formen for skriftlig eksamen ble tatt i bruk for første gang våren 2008, og denne eksamenen har en todelt oppbygning. Den første delen er uten hjelpemidler, og det er kun tillatt å bruke skrivesaker, passer, linjal og vinkelmåler. Den totale eksamenstiden er fem timer, den første delen innleveres etter to timer. Den andre delen varer i tre timer og blir utlevert når den første delen er innlevert. Her er det lov å benytte seg av alle hjelpemidler med unntak av internett og andre kommunikasjonsverktøy (Utdanningsdirektoratet, 2008c). Våren 2008 var det kun biologi og matematikk på

videregående skole som benyttet seg av denne formen for skriftlig eksamen. Fra våren 2009 vil også matematikk i grunnskolen og fysikk og kjemi på videregående skole ta denne formen i bruk (Utdanningsdirektoratet, 2008d).

Skriftlig eksamen i Biologi 2 våren 2008

Den første delen av skriftlig eksamen som ble gitt våren 2008 i Biologi 2 bestod av flervalgsoppgaver og kortsvarsoppgaver (Utdanningsdirektoratet, 2008a). Flervalgsoppgavene hadde ulikt nivå i forhold til resonnering, og kortsvartoppgavene skulle besvares kort på det nivået elevene hadde sin biologikompetanse (Utdanningsdirektoratet, 2008c). Den andre delen bestod av oppgaver med mer åpne problemstillinger, der elevene måtte besvare oppgavene ut fra sitt kompetansenivå (Utdanningsdirektoratet, 2008b). Elevene måtte anvende og analysere åpne problemstillinger, og de måtte aktivt benytte kilder for å besvare oppgavene. Flere ulike oppgaveformer bidro til at elevene med ulik *læringsstil* fikk vist sin kompetanse. Småoppgavene bidro til et variert oppgavesett ved at de gikk i dybden innenfor smalere deler av læreplanen[1]. Alle elever er ulike, og de vil lære på grunnlag av hver enkelt elevs egenskaper og forutsetninger. Dermed vil en eksamensform med flere ulike oppgaver bidra til at flere elever får vist hele sin kompetanse (Imsen, 2005, side 41). Hensikten med denne inndelingen av oppgavene var at elevene skulle vise en bredere og dypere kompetanse i forhold til pensum.

Flervalgsoppgaver har sine begrensninger som vurderingsform, men slike oppgaver vil også ha fordeler i forhold til åpne oppgaver (McDonald, 2002, side 83). Det er forsket mye på både flervalgsoppgaver og åpne oppgaver som vurderingsform. McDonald (2002) beskriver hvordan man skal lage gode flervalgsoppgaver, samt fordeler og ulemper ved denne vurderingsformen (McDonald, 2002, side 83-121). Horgen (2007) beskriver pedagogisk bruk av flervalgsoppgaver, og hva som er viktig å legge på ved slike oppgaver. I forhold til åpne oppgaver har Gronlund (1965) beskrevet hvordan man skal lage en åpen oppgave, hva som kreves av en student for å besvare slike oppgaver og fordeler og utfordringer ved en slik vurderingsform (Gronlund, 1965, side 180-193). Det har ikke blitt utført direkte forskning på eksamensformen fra våren 2008 i Biologi 2, og det er denne eksamenen som har vært bakgrunnen for denne studien.

[1] E-post fra Utdanningsdirektoratet ved Siv Hilde Lindstrøm om begrunnelse av valg av oppgaveform på skriftlig eksamen i Biologi 2 våren 2009. Denne e-posten ble mottatt 17.02.2009.

1.2. Problemstillinger

I denne studien ble det fokusert på den første delen av skriftlig eksamen fra våren 2008 i Biologi 2. Denne delen omhandlet flervalgsoppgaver og kortsvarsoppgaver, der hovedfokuset var flervalgsoppgaver. Det ble i denne studien benyttet en testgruppe bestående av en lærer og 18 elever som gjennomførte undervisning i faget Biologi 2. Det ble undersøkt om elevene får vist sin kompetanse i Biologi 2 ved hjelp av flervalgsoppgaver, og hvordan læreren best mulig kan forberede sine elever på en slik vurderingsform. Undersøkelsen ble gjort ved å benytte læringsplattformen *it's learning* for å se hvordan *læringsverktøyet Test 2.0* kan være et effektivt hjelpemiddel for lærerens undervisning og elevenes læring. Som utgangspunkt for datainnsamlingen hadde elevene i utvalget to prøver som omhandlet de samme kompetansemålene fra læreplanen i Biologi 2 (Utdanningsdirektoratet, 2006a). Den første prøven bestod av flervalgsoppgaver, mens den andre prøven bestod av kortsvarsoppgaver. Det ble samlet inn kvantitative og kvalitative resultater fra lærer og elever som omhandler ulike elementer fra de to prøvene, der hovedfokuset var bruk av flervalgoppgaver som vurderingsform. Ved å ta utgangspunkt i dette hovedfokuset ble det redegjort for følgende problemstillinger:

1. Får elever i tilstrekkelig grad vist sin faglige kompetanse i Biologi 2 ved bruk av vurderingsformen *flervalgsoppgaver* i den første delen av skriftlig eksamen i Biologi 2?
2. Hvordan kan en lærer forberede elevene sine på den første delen av skriftlig eksamen i Biologi 2, ved å benytte *læringsverktøyet Test 2.0* i læringsplattformen *it's learning* i undervisningen?

1.3. Oppbygning av oppgaven

Denne oppgaven er delt inn i fire hoveddeler. Den første hoveddelen handler om teori rundt flervalgsoppgaver, digitale læringsplattformer og Blooms taksonomi. Det blir beskrevet hva som kreves ved produksjon av flervalgsoppgaver og fordeler og utfordringer ved slike oppgaver. Teorien fortsetter med å fremlegge teori rundt digitale læringsplattformer og it's learning. Den første hoveddelen avsluttes med teori rundt Blooms taksonomi. Det blir beskrevet hva som kjennetegner Blooms taksonomi og disse kjennetegnene kobles opp mot å vurdere elevenes faglige kompetanse ved bruk av flervalgsoppgaver.

Den neste hoveddelen handler om metodevalg for datainnsamlingen. Metodene som ble benyttet for datainnsamling var kvantitative og kvalitative metoder. Denne hoveddelen inneholder klassisk teori rundt disse metodene før det blir beskrevet hvilke metodevalg som har foregått i denne datainnsamlingen og gitt begrunnelser for disse valgene.

Deretter inneholder oppgaven hvordan datainnsamlingen foregikk og resultater fra datainnsamlingen. Resultater og analyser som er viktig for å besvare problemstillingene fra kapittel 1.2 blir fremlagt.

I den siste hoveddelen av oppgaven blir resultatene fra datainnsamlingen diskutert opp mot teorien som ble fremlagt i den første hoveddelen. Problemstillingene blir besvart, og det blir også diskutert erfaringer fra dataanalysen og datainnsamlingen. Oppgaven avsluttes ved å gi en oppsummering av oppgaven og resultater fra et postintervju av læreren. Postintervjuet ble gjennomført for å få kjennskap til i hvilke grad læreren har benyttet flervalgsoppgaver som vurderingsform og Test 2.0 for å forbedre elevene til skriftlig eksamen etter at datainnsamlingen til denne studien ble gjennomført.

Den første hoveddelen av oppgaven som omhandler teori, og er inndelt følgende kapitler:

- Flervalgsoppgaver
    - Produksjon av flervalgsoppgaver
    - Fordeler
    - Utfordringer
    - Analyse av flervalgsoppgaver
    - Reliabilitet og validitet
    - Gjetting på flervalgsprøver
    - Vurdering av flervalgsprøver
- Digitale læringsplattformer
    - Kunnskapsløftet
    - It's learning
    - Digitale tester og bruk av Test 2.0 i it's learning
- Blooms taksonomi og flervalgsoppgaver

2. Flervalgsoppgaver

Flervalgsoppgaver kan spille en viktig rolle når man skal måle mange viktige aspekter rundt kompetansenivået til elevene (Haladyna, 2004, side 5). En av de store utfordringene ved å benytte flervalgsoppgaver i undervisningen er at det krever mye forarbeid av lærerne å lage slike oppgaver. Det er mye man må ta hensyn til ved produksjon av flervalgsoppgaver. Det er

viktig å lage oppgaver med små feilkomponenter, slik at testens *reliabilitet* og *validitet* blir ivaretatt på en tilfredsstillende måte. Sirnes (2005) definerer reliabilitet som et mål som angir i hvilken grad en test en elev tar, vil gi samme resultat dersom testen gjentas (Sirnes, 2005, side 95). Validitet handler om i hvilken grad de slutningene man kan trekke fra resultatene på en flervalgsprøve, vil være reelle for situasjoner utenfor prøven (Kleve, 1994). Reliabiliteten og validiteten blir beskrevet og diskutert i kapittel 2.5.

2.1. Produksjon av flervalgsoppgaver

Den tradisjonelle formen for en flervalgsoppgave består av et spørsmål eller et ikke-avsluttet utsagn med flere svaralternativer, der kun ett svaralternativ er korrekt (Haladyna, 2004, side 68). Spørsmålet eller utsagnet blir kalt *stammen*, og den er stimulus for responsen. For de som vet det korrekte svaralternativet på flervalgsoppgaven, bør stammen inneholde nok informasjon til at man ikke behøver å lese svaralternativene for å kunne besvare oppgaven. I produksjon av flervalgsoppgaver bør man også unngå negasjoner og negative uttrykk som *ikke, aldri, med unntak av* og *unntatt*. Slike ord kan lett overses av elevene, og de vil dermed kunne bomme når de skal velge et svaralternativ. Dersom slike ord er med i stammen, bør de fremheves på en eller annen måte: ordene kan ha store bokstaver, understrekninger eller fet skrift (Haladyna, 2004, side 132; Sirnes, 2005, side 31-33).

Etter stammen følger svaralternativene, og det vanligste er å ha tre til fem svaralternativer for hver flervalgsoppgave. Jo flere svaralternativer, desto vanskeligere blir flervalgsoppgaven. Flervalgsoppgavene som ble gitt på skriftlig eksamen i Biologi 2 våren 2008 hadde fem svaralternativer. I 2009 vil flervalgsoppgavene på skriftlig eksamen i Biologi 2 kun ha fire svaralternativer, siden det er veldig utfordrende å produsere gode flervalgsoppgaver[2]. I den tradisjonelle formen for en flervalgoppgave er det kun ett kun korrekt svar, og det svaret kalles *nøkkelen*. De andre svaralternativene kalles *distraktorer*. En stor utfordring ved å lage gode flervalgsoppgaver, er å lage gode distraktorer i forhold til nøkkelen (Sirnes, 2005, side 33-35). For å lage gode flervalgsoppgaver, må svaralternativene være gode og realistiske. Jo flere svaralternativer, desto vanskeligere er det å finne gode distraktorer. Det at det er vanskelig å finne gode distraktorer er hovedårsaken til at antall svaralternativer på skriftlig

[2] E-post fra Utdanningsdirektoratet ved Siv Hilde Lindestrøm om begrunnelse av endring av antall svaralternativer på flervalgsoppgavene i skriftlig eksamen i Biologi 2 fra våren 2009. Denne e-posten ble mottatt 20.05.2009.

eksamen i Biologi 2 blir endret fra våren 2009[2]. En god distraktor vil være mer attraktiv for faglig svake elever enn faglig sterke elever (Sirnes, 2005, side 68). Nøkkelen og distraktorene kan være et ord, en frase eller en setning, formen er avhengig av stammens spørsmålsformulering. Vanskelighetsgraden i flervalgsoppgaven kan øke dersom svaralternativene ligner på hverandre. Likevel bør alle svaralternativene ha samme presisjonsnivå og tilnærmet lik syntaktisk struktur og lengde for ikke å gi hint til elevene av svaralternativene som er nøkkelen (Haladyna, 2004, side 68-69; Sirnes, 2005, side 33-35).

Det finnes også alternative former for flervalgsoppgaver, som hull-i-setning, matche-ord-sammen og lignende (Haladyna, 2004, side 69-93; Sirnes, 2005, side 43-44). Å benytte alternative former for flervalgsoppgaver kan være en bra vurderingsmetode ved vurdering *for* læring. Slike alternative former for flervalgsoppgaver kan gjennomføres ved bruk av it's learning, og det blir diskutert i kapittel 3.3. Flervalgsoppgavene som ble gitt på eksamen i Biologi 2 våren 2008 hadde den tradisjonelle formen, der det var ett korrekt svaralternativ for hver flervalgsoppgave og det var fire distraktorer i tillegg til nøkkelen.

2.2. Fordeler

Det er mange fordeler ved å benytte flervalgsoppgaver både som vurdering *for* læring og vurdering *av* læring. I denne teoridel blir det beskrevet teori som omhandler bruk av flervalgsoppgaver ved vurdering *av* læring. Elevene skal forberede seg mot skriftlig eksamen i Biologi 2, og der vil det bli benyttet flervalgsoppgaver i den første delen. En av fordelene ved å benytte flervalgoppgaver som vurderingsform er at det tar kortere tid for elever å besvare oppgaven, og en flervalgsoppgave krever ikke skrivearbeid og faglige formuleringer. Det gjør at elevenes tid under prøven blir benyttet til tenking, og ikke tid å formulere og skrive svarene (Sirnes, 2005, side 10). Det å kunne bruke naturfaglige begreper og uttrykksformer er en viktig del i Biologi 2, og mange elever kan oppleve det å bruke disse begrepene som problematisk. Derfor kan flervalgsoppgaver gi elever som har problemer med å formulere svar en økt opplevelse av mestring (Lykknes & Smidt, 2009). På bakgrunn av at det tar kortere tid for elever å besvare flervalgsoppgaver, kan en prøve bestående av slike oppgaver dekke større deler av pensum enn kortsvarsoppgaver og langsvarsoppgaver. Den økte dekningen av pensum kan øke reliabiliteten og gi større innholdsvaliditet i prøven (Jansen, 1971, side 4; Sirnes, 2005, side 10-11).

Flervalgsoppgaver er en objektiv vurderingsform, siden sensureringen ikke påvirkes av elevenes håndskrift og evne til å uttrykke seg skriftlig. Det å benytte flervalgsoppgaver i prøver, vil gi elevene en mer rettferdig vurdering. Det tar også kortere tid for læreren å vurdere flervalgsoppgaver enn åpne oppgaver for læreren (Sirnes, 2005, side 10-11).

2.3. Utfordringer

Det er også flere utfordringer ved å benytte flervalgsoppgaver som vurdering *av* læring. Denne vurderingsformen vil ikke måle skriveevne hos elevene. Oppgavene kan fort bli en test som spør etter fakta og detaljkunnskap, noe som bare til en viss grad måler evnen til gjenkjenning av pensum. Slike spørsmål kan være en umotiverende faktor for faglig sterke elever, siden de ikke får vist utdypende kompetanse innen pensum (Jansen, 1971, side 5; Sirnes, 2005, side 27). En av de store utfordringene ved en flervalgsoppgave er muligheten for eliminering av distraktorer og gjetting av nøkkelen når elevene skal besvare oppgavene. Denne utfordringen blir beskrevet i kapittel 2.6.

2.4. Analyse av enkeltoppgaver

Det er mulig å analysere enkeltoppgaver for å undersøke hvor godt oppgavene fungerte i en flervalgsprøve. Som et utgangspunkt for en slik analyse tar man høyde for at faglig sterke elever vil velge nøkkelen, mens faglig svake elever vil trekkes mot distraktorene (Haladyna, 2004, side 203). Som nevnt i kapittel 2.1, er det en stor utfordring å lage gode distraktorer til stammen i flervalgsoppgaver. En analyse av enkeltoppgavene vil gi informasjon om deres vanskelighetsgrad og distraktorenes effektivitet. Analysen vil altså si noe om i hvilken grad oppgavene var for lette eller for vanskelige og om distraktorene fungerte. En slik analyse kan påvise eventuelle svakheter, slik at man kan forbedre oppgavene ved senere bruk (Sirnes, 2005, side 65). Sirnes (2005) har beskrevet en formel for å beregne vanskelighetsgraden, *P-verdi*, til enkeltoppgaver. Denne formelen er:

$$P = \left(R / T\right) \cdot 100\% \qquad \text{(Formel 1)}$$

I denne formelen sier P noe om hvor mange prosent av elevene som svarte korrekt på oppgaven, R er antall elever som svarte korrekt og T er total antall elever som besvarte

oppgaven (Sirnes, 2005, side 66). Boks 1 viser et eksempel på hvordan man kan benytte denne formelen.

Boks 1: Et eksempel som viser hvordan man kan benytte formel 1 for å beregne vanskelighetsgraden til en flervalgsoppgave der 17 elever hadde prøven og 13 av disse elevene svarte korrekt på oppgaven.

En flervalgsoppgave på en prøve der 17 elever var tilstedet. Det var 13 elever som svarte korrekt på oppgaven og 4 elever som svarte ukorrekt.

Oppgave: *Antall hydrogenbindinger mellom adenin og tymin i DNA er:*
Svaralternativ 1: Det er ingen hydrogenbindinger mellom adenin og tymin i DNA
Svaralternativ 2: 1 binding
Svaralternativ 3: 2 bindinger
Svaralternativ 4: 3 bindinger
Svaralternativ 5: Det varierer mellom 2 og 3 bindinger
Korrekt svar var svaralternativ 3.

Beregning av oppgavens vanskelighetsgrad ut fra formel 1:
$P = R/T \cdot 100\ \%$
$\mathbf{P} = 13/17 \cdot 100\ \% = \mathbf{76\ \%}$

Det er flere elementer som kan påvirke denne P-verdien. Det kan være at stammen er for enkel, distraktorene er for dårlige eller at elevene er faglig sterke. Dersom det er en gruppe med flinke elever som besvarer oppgaven, kan det godt hende at nesten alle elevene svarer korrekt. En generell regel er at dersom P-verdien er 90 % eller høyere, er det meget lett å få korrekt svar på flervalgsoppgaven (Sirnes, 2005, side 66). Da bør man endre stammen og/eller distraktorene for å forbedre oppgaven ved senere bruk.

2.5. Reliabilitet og validitet

Det er viktig å ivareta testens reliabilitet og validitet på en tilfredsstillende måte. For å ivareta dette, må man lage flervalgsoppgaver med så små feilkomponenter som mulig, og det kan være en stor utfordring. Ved å benytte flervalgsoppgaver både som vurdering *av* læring og vurdering *for* læring, er det viktig å ta hensyn til reliabiliteten og validiteten i enkeltoppgavene og alle oppgavene samlet sett.

Reliabilitet

Som nevnt tidligere, definerer Sirnes (2005) reliabilitet som et mål som angir i hvilken grad en test som en elev tar, vil gi samme resultat dersom testen gjentas. Denne reliabiliteten oppnås gjennom korrelering av råskårer[3] på to versjoner av samme test (Sirnes, 2005, side 95). Det er mulig å skille mellom *ekstern konsistens reliabilitet* og *intern konsistens reliabilitet*.

Ved ekstern konsistens reliabilitet kan man stille følgende spørsmål: dersom målingene kunne gjentas under like forutsetninger, ville man da fått samme resultat? Ved å definere en korrelasjon mellom skåren på disse to parallelle testene, vil man få en *reliabilitetskoeffisient*. At to tester er parallelle, betyr at en elev skal ha samme skår på de to testene og at de to testene skal ha samme varians[4]. Matematisk kan denne reliabilitetskoeffisienten defineres som forholdet mellom variansen til den sanne skår[5] og variansen til den observerte skår. Dette forholdet sørger for et estimat for målingens reliabilitet. Reliabilitetskoeffisienten mellom to helt parallelle tester er et teoretisk begrep, og kalles for *presisjonskoeffisienten*. Den er et uttrykk for korrelasjoner mellom prøveresultater dersom eleven skal ta prøven flere ganger, gitt at eleven glemmer alt mellom hver gang. Det at eleven skal glemme alt mellom hver gang er en tenkt teoretisk situasjon, og er ikke mulig å få til i praksis (Kleve, 1994).

Ved intern konsistens reliabilitet vil tilfeldige feil ikke bare variere over tid, men også fra et spørsmål til et annet på den samme prøven. Dersom vi antar at alle spørsmålene på en prøve måler det samme temaet og de samme begrepene, vil mangel på korrelasjon mellom oppgavene kun være et resultat av tilfeldige feil. Dersom tilfeldige feil har høy innflytelse på en prøve, vil det vise seg ved at det på noen oppgaver blir skåret høyt, mens det på andre blir skåret lavt. Dersom den tilfeldige feilkomponenten er liten, betyr det at hver oppgave på en prøve måler det samme temaet (Kleve, 1994).

---

[3] Sirnes (2005) definerer råskårer som antall korrekte oppgaver når det ikke forekommer korrigering for gjetting (Sirnes, 2005, side 96).

[4] Sirnes (2005) definerer varians som et mål på variasjonen av en fordeling rundt gjennomsnittet. Den oppnås ved å ta hver skårs avvik fra gjennomsnittlig skår, og kvadrere avviket, summere disse kvadrerte avvikene og dividere dem med antall personer (Sirnes, 2005, side 99).

[5] Sirnes (2005) definerer sann skår en skår som er helt uten målefeil. Sanne skårer er hypotetiske verdier som aldri oppnås i faktisk testing, der en alltid må regne med noen grad av målefeil (Sirnes, 2005, side 96).

Kleve (1994) nevner også flere faktorer som vil påvirke reliabiliteten til en prøve, og blant annet vil en lengre prøve være mer reliabel enn en kort prøve. Det er også viktig å ta hensyn til homogeniteten i elevgruppen for å øke reliabiliteten. Dersom en prøve blir brukt i en annen elevgruppe, som enten er mer eller mindre homogen, vil reliabiliteten bli en helt annen, så variasjonsbredden[6] av elevene vil påvirke reliabiliteten. Instruksjoner som er gitt før prøven vil også påvirke reliabiliteten, slik at fravær av distraksjoner og misforståelser vil gi større reliabilitet. Det er også viktig å la alle som tar prøven få god nok tid til å fullføre den. I tillegg til de nevnte faktorene, vil også sensureringen påvirke reliabiliteten til en prøve. Det betyr at om det er samsvar mellom to eller flere sensorers retting av og karakterforslag til en prøve, vil det øke reliabiliteten. Ved en flervalgsprøve vil sensorreliabiliteten være høy, siden flervalgsoppgaver er en objektiv vurderingsform. Dersom det blir benyttet åpne oppgaver, vil sensorreliabiliteten være mer utsatt (Kleve, 1994).

Validitet

Validitet handler om i hvilken grad de slutningene man kan trekke fra resultatene på en flervalgsprøve, vil være reell for situasjoner utenfor prøven. *Innholdsvaliditet, kriterievaliditet* og *begrepsvaliditet* er tre ulike tilnærminger for å teste validiteten. Det er ikke snakk om å enten bruke en tilnærming eller en annen, men disse tilnærmingene overlapper hverandre. For å studere validiteten til en flervalgsprøve er alle tilnærmingene ofte nødvendige, men begrepsvaliditeten er ofte den viktigste (Kleve, 1994).

Innholdsvaliditet er en viktig kilde for å bevise validitet (Haladyna, 2004, side 185). Det vil si at de spørsmålene som man stiller til elevene, vil måle det de er ment til å måle (Sirnes, 2005, side 99). Innholdsvaliditeten vil besvare spørsmålet om i hvilken grad oppgavene som testen innholder, er representative for det faget eller emnet som elevene skal testes i (Sirnes, 2005, side 81). Spørsmålet blir om graden av måloppnåelse i forhold til de faglige mål, kan måles med den aktuelle prøven (Kleve, 1994).

Den kriterierelaterte validiteten er en studie av forholdet mellom resultater på en prøve og ferdigheter som vanskelig eller umulig lar seg måle. For å kunne benytte resultater fra en prøve til å si noe om elevenes prestasjoner innenfor et område, er det viktig å ha visshet om det er en sammenheng mellom prøven og prestasjonene til elevene på prøven (Kleve, 1994). Det vil si at den kriteriebaserte validiteten tar utgangspunkt i spørsmålet om elevenes

[6] Sirnes (2005) definerer variasjonsbredde som forskjellen mellom de laveste og høyeste skårene som er oppnådd på en test som en gruppe har tatt (Sirnes, 2005, side 100).

resultater på en test kan si noen om fremtidige presentasjoner eller danne grunnlag for en vurdering av et annet mål (Sirnes, 2005, side 81). Det er mulig å skille kriteriebasert validitet i to typer: *forutsiende validitet* og *samtidig validitet*. Med forutsiende validitet menes i hvilken grad presentasjoner på en prøve kan si noen om presentasjoner på kriteriet i fremtiden. Samtidig validitet refererer til relasjoner mellom prøveresultat og kriteriemålinger foretatt på samme tid. Den kan referere til sammenheng mellom et prøveresultat og standpunktkarakteren eleven ligger på (Kleve, 1994).

Den siste tilnærmingen for validitet er begrepsvaliditet. Den er en indikasjon på om prøven som elevene har utført, virkelig tester elevens kunnskap i forhold til kompetansemålene fra Kunnskapsløftet, og om det er mulig på bakgrunn av resultatene fra denne prøven å trekke slutninger om elevenes måloppnåelse i faget (Kleve, 1994).

2.6. Gjetting på flervalgsprøver

Gjetting er en stor utfordring ved denne oppgavetypen, men det viktig å ta hensyn til at gjetting også forekommer i åpne spørsmål (Sirnes, 2005, side 10). Flervalgsoppgaven som ble gitt på eksamen i Biologi 2 våren 2008 hadde fem svaralternativer for hver oppgave. Sannsynligheten for at elevene traff korrekt svar ved tilfeldig gjetting var 20 %, men dersom elevene kunne eliminere enkelte svaralternativer, ville sannsynligheten øke for å velge nøkkelen. På skriftlig eksamen våren 2009 vil sannsynligheten for riktig svar ved tilfeldig gjetting på en flervalgsoppgave øke til 25 %, siden hver oppgave vil ha fire svaralternativer.

For å bøte på problemet med gjetting er det utviklet flere metoder som skal korrigere eller justere skårer som følge av tilfeldig gjetting. Slike metoder er ikke benyttet i vurderingen av skriftlig eksamen. Der vil helhetsinntrykket angi sluttkarakteren til elevene (Utdanningsdirektoratet, 2008c), men læreren kan anvende disse metodene ved vurdering av flervalgsoppgaver som underveisvurdering. Dersom man anvender en korrigeringsmetode ved flervalgsprøver, skal det fremgå i testinstruksjonene at elevene ikke skal gjette og kun besvare de spørsmålene elevene mener at de vet svaret på (Sirnes, 2005, side 45). I denne oppgaven er det fremlagt tre ulike metoder for å korrigere poengsummen til elevene på en flervalgsprøve.

Den første metoden består av den vanligste justeringsformelen som benyttes ved korrigering av resultater for en flervalgsprøve.

Formelen er:

$$Skåre = \mathrm{Re}tt - \left({}^{Galt}\!/_{n-1} \right) \qquad \text{(Formel 2)}$$

I denne formelen er den endelige skåren beregnet ved å summere antall oppgaver som eleven har besvart korrekt og antall oppgaver som eleven har besvart ukorrekt. Disse verdiene plasseres inn i formel 2 og *n* er antall svaralternativer det er i flervalgsoppgaven (Gronlund, 1998, side 121). Boks 2 viser resultater fra en flervalgsprøve til to elever. Den første metoden, med formel 2, er benyttet for å korrigere resultatet.

Boks 2: Et eksempel som viser hvordan resultatene fra en flervalgsprøve til Elev 1 og Elev 2 kunne ha blitt, dersom man hadde benyttet den første metoden for gjetting ved bruk av formel 2.

En prøve har 35 flervalgsoppgaver, der hver oppgave gir 1 poeng for hvert korrekte svaralternativ, men man vil benytte formel 2 for å korrigere for gjetting. Toppskåren for prøven er 35 poeng, og hver oppgave har 5 svaralternativer.

Elev 1 har 25 korrekte svar, 5 ukorrekte svar og 5 ubesvarte oppgaver. Elevens skår vil bli utregnet slik:
Total skår: 25 – 5/(5-1) = 25 – 1,15 = **23,85 poeng**

Elev 2 har også 25 korrekte svar, men 10 ukorrekte svar. Elevens skår vil bli utregnet slik:
Total skår: 25 – 10/ (5-1) = 25 – 2,5 = **22,5 poeng**

Dersom man benytter den første korrigeringsmetoden ved bruk av formel 2, vil det lønne seg for elevene å la være å svare på flervalgsoppgavene dersom de er usikre på det korrekte svaret.

Dale (2008) beskriver en annen metode for korrigering av gjetting. Hun gjennomførte flere flervalgstester ved Universitetet i Agder, der blant annet en test ble gjennomført som midtsemestereksamen i engelsk for studenter. Flervalgsoppgavene var bygget opp med én nøkkel og to distraktorer på hver oppgave. Det ble gitt 3 poeng for hvert korrekte svar, - 2 poeng for hvert ukorrekte svar og 0 poeng for ubesvarte oppgaver. På denne måten var det to måter å få lav skår på, enten ved å ha få korrekte svar og mange ubesvarte oppgaver eller å ha mange besvarte oppgaver, hvorav mange var korrekte svar og mange var gale svar (Dale, 2008). Boks 3 viser hvordan man benytter den andre metoden ved korrigering for gjetting ved en flervalgsprøve.

Boks 3: Et eksempel som viser hvordan resultatene fra en flervalgsprøve til Elev 1 og Elev 2 ville blitt, dersom man benytter den andre metoden for gjetting.

En prøve har 35 flervalgsoppgaver, der hvert korrekte svar gir 3 poeng, hvert ukorrekte svar gir -2 poeng og hver ubesvarte oppgave gir 0 poeng. Det vil si at toppskåren er 90 poeng. Hver oppgave har 5 svaralternativer.

Elev 1 har 25 korrekte svar, 5 ukorrekte svar og 5 ubesvarte oppgaver. Elevens skår vil bli utregnet slik:
Korrekte svar: $25 \cdot 3 = 75$ poeng
Ukorrekte svar: $5 \cdot (-2) = -10$ poeng
Ubesvarte oppgaver: $5 \cdot 0 = 0$ poeng
Total skår: $75 - 10 + 0 =$ **65 poeng**

Elev 2 har også 25 korrekte svar, men 10 ukorrekte svar. Elevens skår vil bli utregnet slik:
Korrekte svar: $25 \cdot 3 = 75$ poeng
Ukorrekte svar: $10 \cdot (-2) = -20$ poeng
Ubesvarte oppgaver: $0 \cdot 0 = 0$ poeng
Total skår: $75 - 20 + 0 =$ **55 poeng**

Av boks 3 ser man at det vil være en større forskjell mellom totalskåren til Elev 1 og Elev 2 i forhold til eksempelet i boks 2, selv om de i begge eksamplene har samme sammensetning av korrekte og ukorrekte svar og ubesvarte oppgaver.

En tredje og siste metode er å gi delvis skår for flervalgsoppgaver som er ubesvarte i steden for å trekke skårer ved ukorrekte svar. Et eksempel på poenggiving kan være å gi 2 poeng for hvert korrekte svar, 0 poeng for ukorrekte svar og ½ poeng for ubesvarte oppgaver. Denne metoden har en psykologisk fordel i forhold til de to første metodene ved at den premierer elever som ikke gjetter, isteden for å straffe ved gjetting (Sirnes, 2005, side 46). Boks 4 viser et eksempel på hvordan man kan beregne total skår med den siste metoden.

Boks 4: Et eksempel som viser hvordan resultatene fra en flervalgsprøve til Elev 1 og Elev 2 ville ha blitt, dersom man benytter den siste metoden for gjetting ved å premiere elever som ikke gjetter isteden for å straffe ved gjetting.

En prøve har 35 flervalgsoppgaver, der hver oppgave gir 2 poeng for hvert korrekte svar, 0 poeng for hvert ukorrekte svar og ½ poeng for hver ubesvarte oppgave. Det vil si at toppskåren er 60 poeng. Hver oppgave har 5 svaralternativer.

Elev 1 har 25 korrekte svar, 5 ukorrekte svar og 5 ubesvarte svar. Elevens skår vil bli utregnet slik:
Korrekte svar: $25 \cdot 2 = 50$ poeng
Ukorrekte svar: $5 \cdot 0 = 0$ poeng
Ubesvarte oppgaver: $5 \cdot ½ = 2{,}5$ poeng
Total skår: $50 + 0 + 2{,}5 =$ **52,5 poeng**

Elev 2 har også 25 korrekte svar, men 10 ukorrekte svar. Elevens skår vil bli utregnet slik:
Korrekte svar: $25 \cdot 2 = 50$ poeng
Ukorrekte svar: $10 \cdot 0 = 0$ poeng
Ubesvarte oppgaver $= 0 \cdot ½ = 0$ poeng
Total skår: $50 + 0 + 0 =$ **50 poeng**

Boks 4 viser også at det vil lønne seg for elever å la være å besvare flervalgsoppgaver dersom de er usikre på svaret. Uansett hvilken metode man velger å benytte ved korrigering for gjetting i en flervalgsprøve, er det viktig å informere elevene før prøvestart. Det er for å hindre misforståelser ved beregning av skår. Sirnes (2005) mener at det er mest forsvarlig å oppfordre elevene til å svare på spørsmål hvor ett eller flere svaralternativer kan elimineres. Bare i de tilfeller der svaret vil være ren gjetting blant svaralternativene, bør eleven hoppe over oppgaven dersom formelskåring anvendes (Sirnes, 2005, side 46).

Jansen (1971) drøfter om det er nødvendig å korrigere for gjetting. Dersom man korrigerer for gjetting, kan det gi *spill-typen* en fordel. Slike elever vil vite om det lønner seg å dra nytte av den minste kunnskap. De mer forsiktige elevene, men kanskje like kunnskapsrike, vil unngå å gjette. På grunnlag av dette mener Jansen (1971) ar vi ikke burde korrigere for gjetting (Jansen, 1971, side 45-46).

2.7. Karaktersetting av flervalgsprøver

Det er viktig å tenke gjennom hvordan man skal sette karakter på en flervalgsprøve. I følge vurderingsveiledningen for skriftlig eksamen i Biologi 2 våren 2008 baserte karaktersettingen

på eksamen seg på et helhetsinntrykk av elevenes besvarelser. Det betyr at sensor vurderte i hvilken grad eleven viste faglig forståelse, så sammenhenger i faget, var oppfinnsom, kunne anvende fagkunnskap i nye situasjoner, gjennomførte logiske resonnementer, forklarte fremgangsmåter og begrunnet svar og skrev oversiktlige og nøyaktige fremstillinger (Utdanningsdirektoratet, 2008c). Denne vurderingsveiledningen nevnte ingenting om hvor mye flervalgsoppgavene telte og hvordan flervalgsoppgavene ble vurdert i forhold til de andre oppgavene i eksamenssettet.

Benytter man flervalgsoppgaver i en prøve, er det mye læreren må ta hensyn til. For det første må læreren fastsette om han vil benytte bestått/ikke-bestått-skala, og da er det viktig å ta hensyn til hvor grensen går mellom karakterene bestått og ikke-bestått. Det er også mulig å gi tall- eller bokstavkarakterer ved en flervalgsprøve, og her vil det også være viktig å ta hensyn til hvordan skalaen vil være inndelt innenfor de ulike skårene, og om strykgrensen skal være høyere enn ved en prøve med åpne oppgaver. Læreren må også gjøre et valg på om han ønsker å korrigere for gjetting, og hvilken metode han ønsker å benytte seg av for å korrigere. Det er mange faktorer å ta hensyn til, men det viktigste er å gi tydelig beskjed til elevene om korrigeringen før prøven slik at det blir en korrekt og rettferdig karaktersetting.

3. Digitale læringsplattformer

3.1. Kunnskapsløftet

Stortingsmelding nr. 30 (2003-2004) *Kultur for læring* la grunnlag for de nye læreplanene i Kunnskapsløftet (LK06). I denne stortingsmeldingen ble det definert en *basiskompetanse* for grunnskolen som består av lese- og skriveferdigheter, regneferdigheter og tallforståelse, ferdigheter i engelsk, digital kompetanse, læringsstrategier og sosial kompetanse. Hovedmålet etter endt skolegang er at alle elever skal ha opparbeidet et nødvendig kompetansenivå i de mest sentrale ferdighetene for å kunne ta del i kunnskapssamfunnet. Disse ferdighetene kalles for *de grunnleggende ferdighetene*, og de er viktige for at elevene skal kunne tilegne seg og utvikle kunnskap og viten i fag. De er også viktige for at elevene skal tilegne seg ferdigheter for å kunne kommunisere og samhandle med andre i utdanning, yrke, samfunn og/eller på det personlige plan. For å sikre en kontinuerlig utvikling av elevens grunnleggende ferdigheter

gjennom hele opplæringsløpet, er disse ferdighetene integrert i læreplanene for alle fag på ulike nivåer (St.meld.nr.30, 2003-2004).

Disse ferdighetene er:

- Å kunne uttrykke seg muntlig
- Å kunne uttrykke seg skriftlig
- Å kunne lese
- Å kunne regne
- Å kunne bruke digitale verktøy

De fire første ferdighetene var definert i tidligere læreplaner, men det å kunne bruke digitale verktøy var først definert i Stortingsmelding nr. 30 (2003-2004) *Kultur for læring*. Grunnleggende digitale ferdigheter er å kunne velge, vurdere og bruke informasjon ved bruk av informasjons- og kommunikasjonsteknologi (IKT). Det innebærer at elevene ved bruk av IKT skal kunne søke, lokalisere, behandle, produsere, gjenbruke, presentere og evaluere informasjon, samt kommunisere og samhandle med andre. Digitale ferdigheter er nødvendig for at elever skal kunne fungere i en stadig mer digitalisert hverdag og kunne delta aktivt i arbeids- og samfunnslivet. Dermed vil denne ferdigheten være en forutsetning i et livslangt læringsperspektiv (St.meld.nr.30, 2003-2004).

Med utgangspunkt i denne stortingsmeldingen er det lagt opp til økt bruk av IKT ved vurdering. Bruk av IKT skal bidra til å gjenspeile sammenhengen mellom læring/undervisning og vurdering. Elevene som bruker digitale verktøy for å tilegne seg kunnskap og ferdigheter under læringsprosessen skal også kunne bruke de samme verktøyene når de skal vurderes (Djupedal, 2006). Det er utviklet mange webbaserte verktøy for kunnskapsdeling. De har fått sitt uttrykk gjennom bruk av *digitale læringsplattformer* som er ment å koble pedagogiske prosesser sammen med institusjonenes administrative og organisatoriske prosesser i en og samme webbaserte plattform (Utdanningsdirektoratet, 2006b).

3.2. Digitale læringsplattformer

Learning Management System (LMS) kalles på norsk digitale læringsplattformer, og de er mye brukt i skolen og ved andre institusjoner i Norge (UninettABC, 2006b). I henhold til Utdanningsdirektoratets kartleggingsrapport om utstyrs- og driftssituasjonene (2007) benytter

98 % av de videregående skolene i Norge, som omfatter 99 % av elevene, en eller flere digitale læringsplattformer. Dette tallet forteller kun i hvilken grad systemet er innkjøpt, men det sier ikke noe om i hvilken grad den enkelte skole faktisk utnytter mulighetene i læringsplattformen (Utdanningsdirektoratet, 2007).

Markedet for læringsplattformer i norsk skole domineres av produsentene it's learning[7] og Fronter[8], og over 90 % av videregående skoler i Norge benytter seg av en av disse to læringsplattformene (Utdanningsdirektoratet, 2007). Disse to leverandørene legger vekt på brukerinvolvering og har etablert ulike fora for at brukernes interesser og behov skal bidra i produktutviklingen, både i forhold til pedagogiske og administrative verktøy i læringsplattformen (Utdanningsdirektoratet, 2006b). Digitale læringsplattformer presenterer flere nyttige verktøy for å administrere, utnytte, synliggjøre og utvikle kompetanse på en effektiv og gjerne målbar måte (Håland, 2007), og det er mange fordeler ved å benytte seg av en læringsplattform i skolen. I denne masteroppgaven har det blitt fokusert på bruk av it's learning i forhold til å lage en test med læringsverktøyet Test 2.0, for å forberede elever til skriftlig eksamen i Biologi 2.

It's learning

It's learning er et av de ledende e-læringsselskaper i Skandinavia, med hovedvirksomhet i Norge, Sverige, Danmark, England og Nederland. De har i 2008 mer enn 600 000 brukere og er nå en av de etablerte markedslederne i Norge. It's learning er fokusert på utvikling og distribusjon av en brukervennlig og kraftig nettbasert læringsplattform for skoler og andre undervisningsinstitusjoner. Deres virksomhet leverer blant annet et verktøy som kan støtte læringsprosesser, Test 2.0 (It'sLearning, 2008, side 70). Man kan lage ulike tester til elevene ved bruk av dette verktøyet, og elevene kan få umiddelbar tilbakemelding på hvilke områder de har prestert bra og på hvilke områder de kan øke kompetansen innefor ulike læreplanmål (UninettABC, 2006a, 2006b). Disse testene kan være en god måte å forberede elever til skriftlig eksamen som involverer flervalgsoppgaver og kortsvarsoppgaver.

[7] http://www.itsolutions.no/

[8] http://no.fronter.info/

3.3. Digitale tester og bruk av Test 2.0 i it's learning

En digital test som Test 2.0 kan benyttes til ulike formål. For eksempel kan den bruktes til å teste elevenes kunnskapsnivå innen tidligere kunnskap, teste om elevene har fått med seg viktige deler av pensum eller den kan brukes direkte i undervisningssammenheng for å lære elevene ulike prinsipper fra ulike læreplanmål. Dersom man benytter en digital test kan en slik test være selvrettende og vise resultatet til deltakerne umiddelbart etter testen, der de korrekte svarene vises som en prosentverdi av den maksimale skåren. Det er også mulig å legge inn tilbakemeldinger til elevene som kan inneholde kommentarer, hint eller ledetråder som elevene kan lese umiddelbart etter innlevering av testen. En slik tilbakemelding kan informere elevene om hvordan de ligger an i forhold til måloppnåelse i faget. Det vil også gi en god pekepinn på hvordan innsatsen bør legges fremover for å oppnå høyere grad av måloppnåelse.

Det er mange variasjonsmulighetene for oppgavene som lages i Test 2.0. Figur 1 viser en enkel oversikt over de ulike generelle og interaktive oppgaveformene som finnes i Test 2.0. De fleste oppgaveformene har mulighet til å gi selvrettende oppgaver som inkluderer enten/eller, flervalgsoppgave, flere svar, velg fra liste – spørsmål, finn par, rekkefølge og klikkpunkt – klikk på bildet. Felles for disse oppgavene er at de har ett eller flere objekter som er riktige svar, og det registreres i hvilken grad eleven identifiserer av korrekt svar. Andre oppgaver kan inkludere mer åpne svar hvor man får poeng ut ifra hvilke nøkkelord som er tatt med. Det er mulig å inkludere spørsmål der man leser inn en tekst, spiller av en lydfil eller legger ved en tekst eller et bilde. Oppgaver i Test 2.0 kan spare læreren for tid, være en effektiv måte å gjøre forarbeid til et tema på, og være til hjelp i bearbeiding av elevens kunnskap etter gjennomgang av et tema (Strømme, 2008, side 159-160).

Figur 1: En oversikt over de ulike variasjonsmulighetene i de generelle og interaktive oppgaveformene som man kan benytte i Test 2.0.

Digitale tester ved vurdering *for* læring

Vurdering *for* læring har hovedfokuset på selve læringsprosessen, og at det er mulig å gi løpende tilbakemelding underveis i læringsprosessen. Denne tilbakemeldingen skal peke på elevens styrker og svakheter i faget (Engh, Dobson, & Høihilder, 2007, side 13; Sirnes, 2005, side 14). Den har også til hensikt å hjelpe eleven til innsikt i sin egen måte å løse en oppgave på, samtidig som den skal hjelpe eleven til å bli bevisst på sine egne *læringsstrategier*. Woolfolk (2004) definerer læringsstrategier som hvordan elevene selv kan nå læringsmål ved å benytte ulike læringsteknikker (Woolfolk, 2004, side 235). Vurdering *for* læring vil altså gi en pekepinn på hva elevene trenger å arbeide mer med. Selve hovedpoenget er at elevene skal vite hva de bør gjøre videre for å oppnå en større grad av måloppnåelse (Engh et al., 2007, side 28).

En digital test har en rekke muligheter for å støtte opp vurdering *for* læring og læringsprosessen til elevene. Den kan brukes til å øke læringen og motivere elevene til videre læring. Det er flere fordeler ved en slik formativ bruk av digitale tester, og blant annet vil den gi mulighet for en umiddelbar og automatisert tilbakemelding til elevene som kan inneholde kommentarer, hint og ledetråder. En slik test vil også gi en god pekepinn på hvor elever står faglig, og dermed kan den virke veiledende på hvor innsatsen bør legges fremover for å oppnå høyere grad av måloppnåelse. En viktig fordel i forhold til å øke motivasjon for læring er at digitale tester åpner for varierte, spennende og utfordrende spørsmål der både tekst, lys og bilde kan inngå i oppgavene. Slike variasjoner kan åpne for mer avanserte muligheter for å

lage vanskeligere spørsmål på et høyere kognitivt nivå til eleven. Det er også mulig å benytte digitale tester før, under og/eller etter en undervisningsøkt. Ved en pretest før en læringsøkt kan man øke fokuset til elevene før selve læringsøkten og motivere dem til økt engasjement i undervisningen. En test gitt underveis i en læringsøkt, kan sette fokus på de viktige temaene, skape refleksjon og gi økt bevissthet rundt elevenes læring. En posttest i etterkant av en læringsøkt vil oppsummere stoffet fra undervisningsøkten slik at elevene får bearbeidet temaene. Ved en posttest er det viktig at elevene får bekreftelser på at de har jobbet tilstrekkelig med lærestoffet. En slik bekreftelse kan gjøres ved at læreren kan legge inn tilbakemeldinger at spørsmålene er besvart. En god tilbakemelding vil være veiledende for eleven, og vil henvise hvordan eleven kan jobbe videre for å lære det som skal læres for å klare å svare korrekt på oppgaven. Slike digitale tester som brukes formativt vil skape refleksjon, repetisjon og motivasjon for elevene, samtidig vil det gi en visshet om eleven er klar for å gå videre med fagstoffet (Horgen, 2007, 2009).

Læreren kan på sin side kontrollere en rekke egenskaper ved digitale tester, som blant annet antall forsøk, grad av tilbakemelding, visning av svar og karantenetid. Ved å sette karantenetid kan en unngå prøv-og-feil-metoden, og stimulere elevene til arbeid med lærestoffet før testen kan tas på nytt. Det er viktig at gode utarbeidede tester også er gode på selve testoppsettet (Horgen, 2009).

Digitale tester ved vurdering *av* læring

Vurdering *av* læring har hovedfokus på karaktersetting og sluttvurdering. En slik vurdering kan enten være et mål på hvor mye læring som har forekommet etter hvert tema eller fagområde som et fag består av, eller et mål på total grad av måloppnåelse. Denne vurderingen skal gis uten hensyn til hvordan elevene har arbeidet med temaet, deres forutsetninger, undervisning eller andre ytre forhold. Mange ganger vil den oppfattes som en kontroll på at elevene har utviklet tilstrekkelig kunnskap eller ferdigheter, og den kan også brukes som en tilbakemelding til læreren på hvor god eller virkningsfull undervisningen har vært (Engh et al., 2007, side 29).

For læreren vil bruk av digitale tester for vurdering *av* læring ha flere fordeler. Læreren kan velge automatisert retting ved at systemet setter poeng avhengig av hva som er riktig, og læreren selv kan velge beståttgrense eller karakterskala. En digital test vil ikke ta hensyn til håndskrift ved retting, og det gir en mer objektiv retting. Denne fordelen vil uansett være tilstede ved bruk av flervalgsoppgaver, uavhengig av om de gjennomføres på papir eller

digitalt. Dersom testen består av andre oppgaver enn fritekst-spørsmål vil rettingen heller ikke ta hensyn til evnen til å formulere seg skriftlig, men det kan både være en fordel og en ulempe for ulike elever. Siden testen foregår digitalt, kan den digitale informasjonen lett redigeres og gjenbrukes, og både for hele tester og enkeltspørsmål. Dersom en samler opp spørsmålene i en større base kan man bygge opp stadig bedre og mer rikholdige samlinger av spørsmål. En digital test vil også tilby systemet statistikk over resultatene og det kan gjøre det lettere å luke ut dårlige spørsmål, samtidig som en kan bygge opp statistikkgrunnlag over tid (Horgen, 2009).

Utfordringer ved bruk av digitale tester

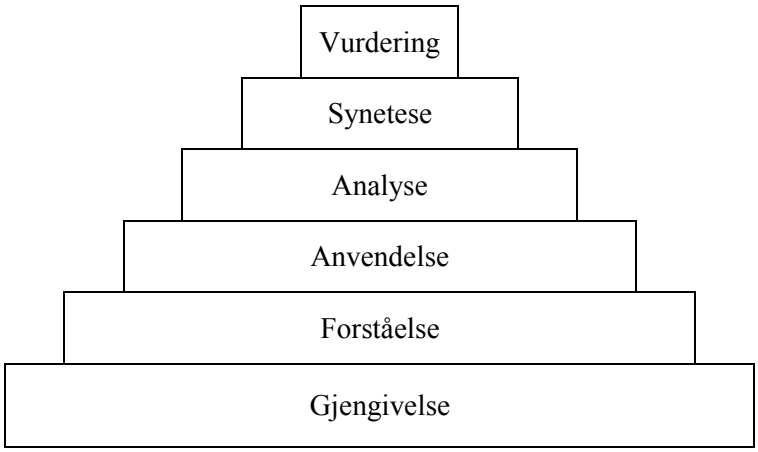
Det er flere utfordringer ved å bruke flervalgsoppgaver, og flere av dem er nevnt i kapittel 2.3. Mange av disse utfordringene vil være de samme for flervalgsoppgaver som utføres i en digital læringsplattform, og en av de største utfordringene er å lage gode spørsmål til elevene ved bruk av gode distraktorer. Det å lage gode distraktorer gjelder både ved bruk av vurdering *for* læring og vurdering *av* læring. Ved bruk av digitale tester ved vurdering *for* læring, bør man som lærer strebe etter å lage gode tester for å gi best mulig læringsutbytte til elevene. Digitale tester kan tilby elevene en ny spennende læringsform og vurderingsform, og de kan utgjøre et viktig supplement i lærerens pedagogiske verktøykasse. Det er viktig å påpeke at både summativ og formativ vurdering kan brukes om hverandre og samtidig i samme fag (Horgen, 2009).

4. Blooms taksonomi og flervalgsoppgaver

En av de store utfordringene ved bruk av flervalgsoppgaver i undervisningssammenheng er å lage oppgaver innenfor de høyere nivåene i Blooms taksonomi for å utfordre elevene og øke deres kompetansenivå. I rammene fra LK06 er ikke Blooms taksonomi så sentralt innenfor kompetansevurdering som i tidligere læreplaner, men med å benytte Blooms taksonomi ved flervalgsoppgaver kan den være med på å gi en god pekepinn på hvordan elevene ligger an i forhold til kompetansemålene i læreplanen.

4.1. Blooms taksonomi

Benjamin Bloom (1913-1999) satt opp en rekke kjennetegn for læring. Disse kjennetegnene kan visualiseres som en pyramide, der hvert nivå forutsetter kvalitetene fra et lavere nivå (Imsen, 2006). Figur 2 viser en oversikt over de ulike kjennetegnene for læring. I denne pyramiden for kunnskap starter man på det laveste nivået, der eleven bare skal kunne gjengi enkeltkunnskap. For hvert nivå øker det kognitive nivået for læring og man blir i stand til å formulere kunnskap med egne ord, anvende tillærte kunnskaper, analysere, syntetisere og vurdere kunnskap (Sirnes, side 23).



Figur 2: En oversikt over de ulike nivåene i Blooms taksonomi. Denne oversikten fremstilles ofte som en trapp eller pyramide for å synliggjøre hvordan læring går fra enkel gjengivelse til forståelse og anvendelse og opp til vurdering. Ethvert nivå forutsetter kvalitetene fra de lavere nivåene. (Bearbeidet etter Imsen (2006)).

Nivå 1: Gjengivelse

Det første nivået i Blooms taksonomi er gjengivelse, her skal elevene kunne gjengi enkeltkunnskaper i den samme formen som kunnskapen er undervist (Sirnes, 2005, side 23). Det innebærer kjennskap og evne til gjenkalling av enkle og spesifikke fakta. Flervalgsprøver blir ofte relatert til dette nivået (Jansen, 1971, side 16). I følge Bloom (1956) kan gjengivelsesnivået deles inn i flere undernivå, blant annet *kunnskap om prinsipper og generaliseringer*. Her skal elevene kunne kjenne til de prinsipper som ligger til grunn ved vurderingen av en viss situasjon, eller når den riktigste eller mest rasjonelle handling skal bestemmes eller beskrives (Bloom, 1956, side 75-76).

Nivå 2: Forståelse

Forståelse er det andre nivået i Blooms taksonomi. Her skal elevene kunne gjengi kunnskaper med egne ord eller symboler (Sirnes, 2005, side 24). Mange ferdigheter og egenskaper som skolen legger stor vekt på omhandler at elevene skal kunne tilegne seg kompetanse som innebærer forståelse. Undervisningen går på mange måter inn for at det som kommuniseres skal bli forstått av elevene. Dersom elevene forstår det som blir fremstilt, vil de ha lettere for å kunne bruke det i sine egne former. (Jansen, 1971, side 18). Bloom (1956) deler også dette nivået inn i flere underkategorier, en av underkategoriene er *tolkning*. For at elevene skal kunne tolke lærerens fremstilling av temaet må eleven kunne omdanne de ulike elementene i denne fremstillingen. Elevene må dermed kunne identifisere og forstå de ulike elementene og forbindelsen mellom dem (Bloom, 1956, side 93).

Nivå 3: Anvendelse

For at elevene skal kunne anvende et prinsipp kreves det forståelse av metoden eller prinsippet som skal anvendes. Disse kjennetegnene er det tredje nivået i Blooms taksonomi. Ved anvendelse skal elevene kunne se at kunnskapen som allerede er lært kan brukes i en ny og ukjent situasjon, elevene må kunne omstrukturere det innlærte. Det er viktig at elevene blir testet i en ny situasjon, ellers blir det bare snakk om evne til gjenkalling (Jansen, 1971, side 19; Sirnes, 2005, side 24).

Nivå 4: Analyse

Analyse er et noe høyere trinn enn forståelse og anvendelse. I dette fjerde nivået skal elevene ha så grundig forståelse av kunnskap at de skal kunne bryte kunnskapen ned i enkeltdeler, identifisere årsak og virkning, trekke slutninger og finne bevis som kan underbygge generaliseringer (Sirnes, 2005, side 24). Bloom (1956) har også delt dette nivået inn i flere underkategorier. På den laveste kategorien i nivå 4 er *analyse av enkelte elementer*, der elevene skal kunne skille det vesentlige fra det uvesentlige, og elevene skal kunne avsløre usaklige konklusjoner i en fremstilling. I den neste kategorien, *analyse av relasjoner*, kreves det at elevene, etter å ha identifisert elementene i en fremstilling, skal se sammenhengen mellom de ulike elementene. På det høyeste trinnet, *analyse av organisatoriske prinsipper*, forventes det av elevene at de skal kunne analysere helheten i de ulike elementene. Helhetsanalysen er ofte nødvendig for at elevene skal få fullt utbytte av en fremstilling (Bloom, 1956, side 146-148).

Nivå 5: Syntese

Det nest siste og femte nivået kalles syntese. Her skal elevene kunne sette sammen informasjon på en annerledes måte ved å kombinere elementer i et nytt system eller ved å foreslå alternative løsninger (Sirnes, 2005, side 24). Selve prosessen innebærer at elevene må omstrukturere elementer og kombinere dem slik at de danner en ny helhet, og summen må være noe mer enn summen av de enkelte delene. Disse kjennetegnene er noe mer omfattende enn foregående trinn, som innebar at elevene skulle studere en helhet for å forstå den bedre. På synetesenivået må elevene trekke inn elementer fra ulike kilder og sette dem sammen til en struktur som ikke var der tidligere (Jansen, 1971, side 20).

Nivå 6: Vurdering

På det siste og sjette nivået i Blooms taksonomi skal elevene kunne presentere og gjøre rede for synspunkter og meninger ved å forta vurderinger av informasjon, gyldigheten til ideer og kvaliteten av arbeid på grunnlag av et sett kriterier (Sirnes, 2005, side 25). Dette nivået kalles vurdering, og det er plassert til slutt fordi det antas at det i noen grad betyr en kombinasjon av alle de foregående trinn. I tillegg vil også verdisetting av ideer, påstander, metoder, konklusjon og lignende være tilstede. Vurderingen kan være både kvantitativt og kvalitativt og kriteriene kan være gitte eller selvvalgte (Jansen, 1971, side 21).

Bloom (1956) deler dette nivået inn i to underkategorier, *indre* og *ytre kriterier*. Ved de indre kriteriene skal elevene kunne vurdere om fremstillingen er konsistent og om den inneholder uttrykk og tanker som er selvmotsigende. Det vil også være naturlig for elevene å prøve å finne logiske feilaktigheter i fremstillingen. For å kunne ta hensyn til de ytre kriteriene må elevene kunne kjenne til det området som det arbeidet som skal vurderes omfatter. Elevene må kjenne de kriterier som vanligvis legges til grunn ved vurdering og elevene bør også være vant til å anvende disse kriteriene. Det siste trinnet i Blooms taksonomi krever altså at eleven er godt kjent med stoffet (Bloom, 1956, side 188-192).

4.2. Bruk av Blooms taksonomi i flervalgsoppgaver

Det blir ofte hevdet at flervalgsoppgaver bare kan måle enkeltstående fakta og detaljspørsmål fra pensum. Denne kunnskapen er kjennetegn på de laveste nivåene i Blooms taksonomi, men flervalgsoppgaver kan utfordre elevene på høyere nivå enn gjengivelsesnivået (Haladyna,

2004, side 137). Ved å bruke flervalgsoppgaver i avanserte regneoppgaver eller bruk av tabeller eller lignende, må elevene forstå de ulike elementene i fremgangsmåten for å finne det korrekte svaralternativet. Dersom man tester elevene på et høyere nivå i Blooms taksonomi, vil testen bli mer utfordrende for både elevene og lærerne. For lærerne sin del vil det i slike flervalgsoppgaver være mer utfordrende å produsere stammen, nøkkelen og distraktorene. Et praktisk tips til lærerne for å lage gode distraktorer er først å lage kortsvarsspørsmål til elevene slik at man kan bruke de ukorrekte svarene fra kortsvarsoppgaven som distraktorer i flervalgsoppgaven. Boks 5 viser et eksempel på hvordan man kan benytte svar fra en kortsvarsoppgave som nøkkel og distraktorer i en flervalgsoppgave. Her vil kortsvarsoppgaven være på nivå 4 *analyse,* i kategorien *analyse av relasjoner* i Blooms taksonomi*,* siden elevene må kunne bryte kunnskapen om stoffkretsløpet til karbon i planter ned i enkeltdeler og trekke slutninger for å kunne formulere et svar på kortsvarsoppgaven. Elevene skal da kunne se sammenhengen mellom elementene i svaret. Flervalgsoppgaver vil være på nivå 4 *analyse*, på den laveste kategorien *analyse av enkelte elementer*, siden elevene må kunne skille nøkkelen fra distraktorene, og elevene må kunne avsløre usakelige konklusjoner av de ulike svaralternativene.

Boks 5: Et eksempel på hvordan man kan benytte svar fra en kortsvarsoppgave som nøkkel og distraktorer i en flervalgsoppgave[9].

Kortsvarsspørsmål: *Forklar kort hvordan stoffkretsløpet til karbon fungerer i planter. Bruk 5-7 setninger.*
Svar 1: Ved hjelp av celleånding tar dyr opp oksygen og puster ut karbondioksid. Karbondioksid blir tatt opp hos planter som bruker karbondioksidmolekylet til å spalte glukose til energi i form av ATP, vann og oksygen. Oksygenet slipper ut igjen via spalteåpninger i bladet og tilbake til luften, mens noe blir tatt i selve planten.
Svar 2: Plantene tar opp karbondioksid for å bryte det ned, slik at de får energi. Karbondioksid er blant annet med i fotosyntesen, der platene produserer oksygen. Vi mennesker puster ut karbondioksid.
Svar 3: Produsentene tar opp solenergi. I kloroplastene blir det utført fotosyntese. Organiske stoffer og oksygen blir dannet. I mitokondriet blir det utført celleånding og det blir dannet karbondioksid og vann. Energien blir frigitt som varme.
Svar 4: Karbon kommer inn i et kretsløp fra blant annet eksos fra bilen. Dyr som spiser planter får i seg karbon, fordi plantene bruker karbon i fotosyntese. Da blir karbon overført til neste ledd i kjeden. Når dyret med karbon dør eller blir spist, kommer karbon ut i luften når det døde dyret blir brutt ned av mikroorganismer, eller blir overført til neste ledd i kjeden.

Flervalgsspørsmål: *I karbonkretsløpet…*
Svaralternativ 1: blir karbonet i plantene brukt til å bygge opp organiske stoffer.
Svaralternativ 2: får plantene den energien de trenger ved å bryte ned karbondioksid.
Svaralternativ 3: tar dyrene opp karbon fra plantene i form av karbondioksid.

En åpen oppgave kan enten være en kortsvarsoppgave eller langsvarsoppgave. I følge Woolfolk (2004) er en kortsvarsoppgave en skriveoppgave som kan teste mye faktakunnskap på forholdsvis kort tid, og det kan gjøre at det blir vanskelig å måle kompleks kunnskap. En langsvarsoppgave kan teste denne komplekse kunnskapen, og læreren kan vurdere resonneringsevne og kreativitet. Både kortsvarsoppgaver og langsvarsoppgaver blir gitt på den skriftlige eksamenen i Biologi 2 fra våren 2008. I alle åpne oppgaver er det viktig at spørsmålet er et klart og presist, og det bør indikere hva svaret skal dekke. Elevene må også vite hvor utdypende svaret skal være. En stor utfordring ved åpne oppgaver er å vurdere dem. Woolfolk (2004) beskriver ulike metoder for å unngå subjektivitetsproblemet med vurderingen. Det er viktig å konstruere et *modellsvar*. Det betyr at læreren avgjør hva som bør være med i besvarelsen. Etter at man har laget et modellsvar, må man vektlegge de ulike delene i oppgaven i form av poeng, og på dette grunnlaget kan læreren sette karakter (Woolfolk, 2004, side 400-403).

[9] Elevsvarene fra dette kortsvarspørsmålet er fra en spørreundersøkelse som ble utført i september 2008 for et annet fag ved NTNU.

Det er lettere å benytte Blooms taksonomi for å undersøke kompetansen til elevene på en åpen oppgave i forhold til en flervalgsoppgave. En kortsvarsoppgave kan teste elevene på de fire første nivåene: gjengivelse, forståelse, anvendelse og analyse. Alt er avhengig av spørsmålsformuleringen som elevene skal besvare. I en langsvarsoppgave kan det være lettere å teste den komplekse kunnskapen, og dermed undersøke elevenes nivå innenfor syntese og vurdering i Blooms taksonomi.

Den neste delen av oppgaven omhandler mer om klassisk forskning innenfor kvalitativ forskning, metodevalg og begrunnelser for å besvare problemstillingene som ble presentert i kapittel 1.2.

5. Metode

5.1. Innledning

Klassisk forskning kan deles inn i to hovedkategorier; *kvantitativ* og *kvalitativ forskning*. I kvantitativ forskning er hensikten å avdekke sammenheng mellom ulike variable og ha et fokus på det representative og det generaliserbare. Hensikten med kvalitativ forskning er derimot å forstå og beskrive et fenomen der fokuset er mer spesialisert (Robsen, 2002, side 86-93). Som nevnt i introduksjonen var mine to problemstillinger for denne oppgaven:

1. Får elever i tilstrekkelig grad vist sin faglige kompetanse i Biologi 2 ved bruk av vurderingsformen *flervalgsoppgaver* i den første delen av skriftlig eksamen i Biologi 2?
2. Hvordan kan en lærer forberede elevene sine på den første delen av skriftlig eksamen i Biologi 2, ved å benytte *læringsverktøyet Test 2.0* i læringsplattformen *it's learning* i undervisningen?

Hovedfokuset for datainnsamlingen var å benytte kvalitative metoder, og metodene var spørreundersøkelse med åpne svar og intervju av elever og lærer for å få svar på de to problemstillingene. Før datamaterialet ble samlet inn, gjennomførte elevene to prøver for å teste grad av måloppnåelse innen kompetansemålet *Genetikk* fra læreplanen i Biologi 2 (Utdanningsdirektoratet, 2006a). Datainnsamlingen med spørreundersøkelse og intervjuer ble

utført etter disse to prøvene. Det ble brukt totalt åtte skoletimer à 45 minutter for å samle inn datamaterialet.

Kapittelet om metode og metodevalg vil bli inndelt på følgende måte:

- Kvalitativ metode
- Utvalget
- Metodevalg
    - Spørreundersøkelse
    - Intervju

5.2. Kvalitativ metode

Postholm (2005) beskriver kvalitativ forskning som noe som innebærer å utforske menneskelige prosesser eller problemer i en virkelig setting. Forskeren skal være åpen for hva deltagerne gjør og sier og videre løfte fram deres perspektiver (Postholm, 2005, side 9). Kvalitativ forskning kan deles inn i flere undergrupper, der en av dem er *kasusstudier* (Postholm, 2005, side 41). En kasusstudie involverer forskning på et fenomen eller en aktivitet som er viktig i sin kontekst (Robsen, 2002, side 89, 178). Det betyr å gi en detaljert beskrivelse av fenomenet som det er. Ved å rette fokus mot dette fenomenet kan man avdekke interaksjonen mellom ulike faktorer som er karakteristiske for denne aktiviteten. Det kan gi en mer helhetlig beskrivelse av studiet, som er målet for kvalitativ forskning (Postholm, 2005, side 50).

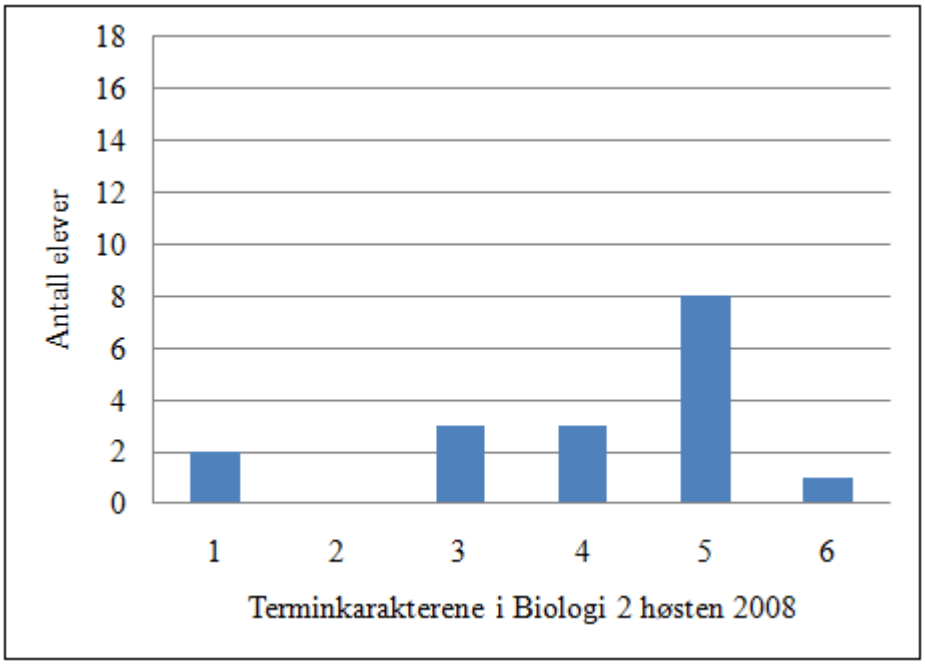
I denne oppgaven er det utførte en studie som har likhetstrekk med en kasusstudie, der utvalget var en lærer og hans klasse i Biologi 2. I følge Robsen (2002), er denne studien et sett med individuelle kasusstudier og alle individene hadde noen felles kjennetegn (Robsen, 2002, side 181). Kjennetegnene for denne studien er at alle elevene skal ha samme skriftlig eksamen i Biologi 2 våren 2009, der de samme kravene vil bli stilt til dem på eksamenen. Disse kjennetegnene bekreftet at denne studien var et sett med individuelle kasusstudier.

5.3. Utvalg

Utvalget for denne oppgaven ble delvis valgt selektivt. Siden Test 2.0 er en del av it's learning, var det viktig å utføre kasusstudiet på en skole som benyttet seg av denne læringsplattformen. I nærmiljøet var det bare to skoler som benyttet it's learning. Et annen

viktig element i seleksjonen av utvalget var valg av lærer. Læreren måtte ha mulighet og tid til å være med på denne studien, i tillegg til å undervise i Biologi 2 på videregående skole. Etter disse kriteriene, var det bare ett alternativ igjen i forhold til skolene i nærmiljøet. Denne begrunnelsen kan kanskje tolkes som at utvalget ble valgt ved *convenience sampling*, som Robsen (2002) kaller det. Robsen (2002) beskriver convenience sampling som en veldig enkel måte å skaffe utvalget sitt på, siden man bare velger det letteste alternativet (Robsen, 2002, side 265). Det å finne en videregående skole i Norge som benytter seg av it's learning var ikke noe problem, og beskrivelsen av convenience sampling er derfor ikke dekkende for utvalget. Som nevnt i kapittel 3.2. er it's learning et av de ledende e-læringsselskapene i Skandinavia. Utfordringen var å finne en lærer som ikke benyttet flervalgsoppgaver som vurderingsmetode og som samtidig var motivert til å være med på denne studien. Læreren fra skolen som ble plukket ut informerte om at han hadde svært lite erfaring med bruk av flervalgsoppgaver som vurderingsform, og han ville derfor bli med på studien. Denne studien å velge utvalg på er en *quota sampling*, som er en sampling etter spesifikke karakteristikker (Robsen, 2002, side 264). En studie der utvalget er plukket ut gjennom en quota sampling gjør at man er mer sensitiv for *researchers bias*. Researchers bias betyr at forskeren selv er med på å påvirke resultatet, noe som videre kan påvirke validiteten i oppgaven (Robsen, 2002, side 172). Validitet handler om gyldighet i studien, og det betyr at de funnene man har observert i datainnsamlingen, er det som kommer fram i resultatet i studien, upåvirket av forskerens meninger (Robsen, 2002, side 101).

Den valgte læreren hadde 18 elever som hadde valgt Biologi 2 skoleåret 2008/2009, og elevene hans ble elevutvalget for denne studien. Skolen elevene gikk på var en offentlig videregående skole, og den bestod av flere utdanningsprogram som studiespesialiserende fag og andre utdanningstilbud innen yrkesutdannelse. Alle elevene i utvalget gikk på studiespesialiserende utdanning. I spørreundersøkelsen ble elevene spurt om terminkarakteren som de fikk høsten 2008 i Biologi 2, og figur 3 viser en oversikt over disse terminkarakterene.



Figur 3: En oversikt over fordelingen av terminkarakterene som elevene i utvalget fikk høsten 2008 i Biologi 2.

Figur 3 viser at flertallet av elevene som var faglig sterke elever. 56 % av elevene fikk terminkarakter 5 eller 6 høsten 2008 i Biologi 2[10]. Det kan være en faktor som kan påvirke resultatet og troverdigheten i studien. I kvantitativ forskning er det viktig å oppnå høy reliabilitet i oppgaven, og det handler om pålitelighet i studien (Robsen, 2002, side 101). I kvalitativ forskning er det også viktig å oppnå høy reliabilitet, men det handler mer om å oppnå høy troverdighet i studien. Det er vanskelig å etterprøve en kvalitativ studie, og derfor vil det være viktig å arbeide grundig og forklare hva man har gjort i studiene for å øke troverdigheten i oppgaven (Robsen, 2002, side 176). I denne studien var det mange faglig sterke elever i utvalget, men så lenge det er begrunnet og forklart grundig, vil troverdigheten i oppgaven være høy.

Av etiske hensyn er også datamaterialet anonymisert. Det ble underskrevet en kontrakt mellom meg, veilederen min, læreren i studien og elevene i klassen hans før datainnsamlingen startet. Denne kontrakten ligger vedlagt som vedlegg 1. Det var for å fastslå for elevene at alle resultatene fra datainnsamlingen ble behandlet konfidensielt mellom meg og dem, samt at elevenes deltagelse var basert på frivillighet. I spørreundersøkelsen, som vil bli beskrevet i kapittel 5.4, skulle elevene velge et femsifret tall som skulle identifisere dem. De kunne velge de fem første eller siste sifrene i mobilnummeret sitt, fødselsdagen sin og lignende. Det

[10] Jfr. figur 3.

viktigste var at tallet skulle være anonymt. Alle spørreundersøkelsene ble lagt i stigende rekkefølge i forhold til tallene, og hvert individ fikk et fiktivt navn. Det laveste tallet i spørreundersøkelsen var ei jente, og dermed ble det fiktive navnet Anna etter første bokstaven i alfabetet. Den nest laveste tallet i spørreundersøkelsen var også ei jente, og det fiktive navnet ble Berit. Slik fortsetter det, og fremlegg av det kvalitative resultatet vil bare inneholde fiktive navn som kun identifiserer kjønnet fra spørreundersøkelsen. Det er også oppgitt terminkarakter etter hvert fiktivt navn i de kvalitative resultatene. Det er for å se om det er noen ulikheter mellom resultatene fra de faglig svake og de faglig sterke elevene.

5.4. Metodevalg

Postholm (2005) beskriver at datainnsamling for kasusstudier kan være omfattende, og studiens varighet vil være avhengig av fenomenet som skal studeres. Et essensielt trekk for en slik studie er at det blir samlet inn tilstrekkelig med data for at forskeren skal kunne utforske viktige trekk og tolke det som blir studert (Postholm, 2005, side 53). En bra måte å få inn tilstrekkelige mengder data på, er å benytte *triangulering*. Det er en prosedyre for å sikre datakvaliteten, og det vil gjøre at resultatet vil bli mer detaljert, slik at man kan sikre validiteten i studien (Postholm, 2005, side 132; Robsen, 2002, side 174-175). I denne oppgaven ble det benyttet flere innsamlingsmetoder for å få svar på problemstillingene, se kapittel 1.2. Det ble benyttet spørreundersøkelser som var inndelt i en kvantitativ og en kvalitativ del, samt intervju av lærer og elever. Denne kombinasjonen av å benytte både kvantitative og kvalitative metode kalles *metodetriangulering* (Robsen, 2002, side 174).

En faktor som er viktig i forhold til å opprettholde en god troverdighet, som ble nevnt tidlig, og det er transkripsjon og translasjon av datamaterialet. Robsen (2002) nevner tre forhold som er viktig å ta hensyn til ved disse aspektene. Det første forholdet er *beskrivelsen*. Det er viktig at beskrivelsene av fenomenet i studiene er detaljert og fullstendig. For å hindre ufullstendige beskrivelser, bør man ta opp intervjuer på bånd eller filme fenomenet for å få en fullstendig transkripsjon av datamaterialet. Det andre forholdet Robsen (2002) nevner er *tolkningen*. For å kunne få en valid tolkning av studien, er det viktig ikke å synse og søke etter tendenser i datamaterialet som ikke er tilstede. Ved en grundig transkripsjon av datamaterialet, vil det være lettere å translere datamaterialet. Det som skal fremlegges er en nøyaktig tolkning av fenomenet. Det er mulig å øke validiteten i datamaterialet ved enten å vedlegge eksempler på hvordan tolkningen mellom transkripsjon og translasjon har skjedd, eller ved å benytte direkte

sitater fra transkripsjon i resultatet. Det siste forholdet er *teori*, altså forklaringen på fenomenet som har blitt studert. Dette forholdet omfatter å analysere translasjonen i forhold til teori. Her er det viktig å ta hensyn til *researchers bias*, altså at forskeren kan påvirke tolkningen av translasjonen (Robsen, 2002, side 171-172).

Det er også to andre elementer som man må ta hensyn til ved innsamling av datamaterialet: *reactivity* og *respondent bias*. Robsen (2002) beskriver reactivity som det at forskerens tilstedeværelse kan påvirke elevenes oppførsel i klasserommet (Robsen, 2002, side 172). Det ble ikke foretatt noen formelle observasjoner av elever i klasserommet under datainnsamlingen, så dette elementet var ikke avgjørende for analysen i denne oppgaven. Det andre elementet som Robsen (2002) nevner er respondent bias, og det er mer avgjørende for resultatene og analysene i denne oppgaven. Respondent bias betyr at elevene kan holde tilbake informasjon som kan være viktig for resultatet og analysen av datamaterialet eller at de svarer det forskeren gjerne vil høre. Robsen (2002) kaller det siste for *good bunny syndrome*. Det kan være med på å senke validiteten i studien, og man kan aldri ha garanti for at respondent bias ikke vil forekomme ved datainnsamling (Robsen, 2002, side 172).

Spørreundersøkelse

Spørreundersøkelsen bestod av både avkrysningspåstander og åpne spørsmål. Avkrysningspåstander er en kvantitativ metode, siden man vil få ut et statistisk resultat fra avkrysningene. Metoden ble benyttet for å få en enkel oversikt og for å se trendene i datamaterialet. For å få en dypere forståelse av elevenes synspunkter ble det benyttet åpne spørsmål, som er en kvalitativ metode.

Svarinnholdet fra de åpne spørsmålene hadde ingen restriksjoner. Elevene hadde mulighet til å svare i dybden på spørsmålene, og det kunne resultere i detaljer i resultatet. Slike detaljer ville også gi en mer nøyaktig vurdering av elevenes synspunkter, og det vil øke den indre validiteten i oppgaven. Som nevnt tidligere handler validitet om gyldighet i resultatene, og indre validitet betyr at de slutningene man har kommet til, er gyldige innenfor studien (Robsen, 2002, side 103-106). Ved å benytte åpne spørsmål kan man også få uforutsette svar, og disse kan utgjøre viktige poenger for oppgaven. En annen stor fordel med å benytte spørreundersøkelser med åpne spørsmål, er at spørsmålene i seg selv kan angi kategoriene for analysen av datamaterialet. Ved at kategoriene allerede er klare, vil en slik metode være tidsbesparende i forhold til analysen av datamaterialet. (Robsen, 2002, side 274-276).

Det er også viktig å vurdere utfordringer ved et metodevalg. En av utfordringene ved en spørreundersøkelse med åpne spørsmål, er at det er mer tidkrevende å analysere kvalitative resultater enn kvantitative. I kvantitative resultater arbeider man mye med å beskrive fenomenet for det utvalget en har, og man ønsker å trekke slutninger som kan fremstilles statistisk. I kvalitative resultater blir det lagt mer vekt på analysen av resultatene og forankringen av disse i teori. En annen ulempe ved å stille åpne spørsmål i en spørreundersøkelse, er at man ikke kan klare opp i misforståelser i forhold til spørsmålsformuleringen eller i forhold til svarene på spørsmålene. Åpne spørsmål blir ofte benyttet ved et intervju, og da vil det være mulig å klare opp i slike misforståelser. Kategorisering av datamaterialet kan også være en annen ulempe ved åpne spørsmål. I løpet av en kvalitativ forskningsprosess kan forskeren ende opp med mengder av like fenomener og tendenser i datamaterialet. Disse begrepene må grupperes for at datamaterialet skal bli håndterlig, og denne prosessen for å få oversikt over tendensene og samle dem i grupper, kalles *kategorisering* (Postholm, 2005, side 88). Kategoriene i denne studien var tematiske kategorier, og det vil si at kategoriene blir dannet ut fra hvilke tema elevene uttaler seg i, ikke hva de sier om selve temaene. Kategoriene var i denne studien forhåndsbestemte før datainnsamlingen, og det kunne føre til problemer dersom svarene til elevene ikke passet inn i de tematiske kategoriene. Det ble løst ved å danne nye kategorier eller underkategorier underveis i dataanalysen. Vedlegg 2 viser en oversikt over de tematiske kategorier og om de var forhåndsbestemt eller oppstod underveis i analysen av datamaterialet. En slik analyse av datamaterialet, kan være preget av mye synsing, og det vil være en ulempe. Synsingen kan også være med på å påvirke den indre validiteten i oppgaven (Robsen, 2002, side 274-276).

Spørreundersøkelsen som elevene måtte besvare etter flervalgsprøven og kortsvarsprøven var delt inn i fire hoveddeler, og den er vedlagt som vedlegg 3. Spørreundersøkelsen startet først med litt bakgrunnsinformasjon om eleven, som gav en kontekst som var til hjelp for å tolke resultatet i analysen av spørreundersøkelsen. Den andre hoveddelen omhandlet flervalgsprøven. De første spørsmålene var avkrysningspåstander med likertsskala med alternativene enig, litt enig, litt uenig, uenig og vet ikke. Deretter var det flere åpne spørsmål der elevene kunne greie ut i dybden om avkrysningspåstandene og fordeler og ulemper ved flervalgsprøven. Den neste hoveddelen handlet om kortsvarsprøven, og denne delen var også inndelt i en avkrysningsdel og åpne spørsmål. Den siste hoveddelen inneholdt felles spørsmål om flervalgsprøven og kortsvarsprøven, der de skulle sammenligne i hvilken grad de fikk vist sin kompetanse på de to prøvene. Alle elevene besvarte denne spørreundersøkelsen, selv om

noen av dem bare hadde vært tilstede på den ene av de to prøvene. Dersom en elev ikke var tilstede på den ene prøven, besvarte de kun den hoveddelen som omhandlet den prøven de var tilstede på.

I beskrivelsen og tolkningen av spørreundersøkelsen, ble de kvantitative resultatene fremlagt grafisk med enkle oversikter for å kunne se tendenser i datamaterialet. De kvalitative resultatene ble nøyaktig fremlagt i forhold til hva elevene svarte på de ulike spørsmålene. Elevbesvarelsene ble fremlagt i bokser i resultatet, for å opprettholde en god troverdighet i tolkningen av datamaterialet.

Målet med denne spørreundersøkelsen var å få svar på flere elementer rundt den første problemstillingen: Får elever i tilstrekkelig grad vist sin faglige kompetanse i Biologi 2 ved bruk av vurderingsformen *flervalgsoppgaver* i den første delen av skriftlig eksamen i Biologi 2? Alle besvarelsene fra de kvalitative spørsmålene måtte begrunnes, for å få et bedre grunnlag for å kartlegge elevenes synspunkter rundt denne problemstillingen.

Intervju

Det er ulike former for intervju, og Postholm (2005) deler dem inn i tre former, *strukturert intervju*, *halvstrukturert* og *ustrukturert intervju* (Postholm, 2005, side 69-74). Det er forskningens hensikt som styrer valget av de ulike intervjuformene. I denne oppgaven ble det benyttet strukturert og halvstrukturert intervju, og siden denne oppgaven er en kasusstudie vil spørsmålene dannes ut fra forskerens erfaringer, med utgangspunkt i teori og observasjoner (Postholm, 2005, side 77).

Det planlagte og formelle intervjuet er betegnelsen for strukturert intervju. Det betyr at intervjueren stiller alle respondentene den samme serie av spørsmål som er utformet i forkant av intervjuet. En stor fordel ved å benytte strukturert intervju er at det går lett å kategorisere datamaterialet, siden kategoriene er forhåndsbestemt før datainnsamlingen. Informasjonen som kommer fram i intervjuet, vil bli sortert i forhold til disse forhåndsbestemte kategoriene. Som nevnt tidligere blir problemet med forhåndsbestemte kategorier at dersom svarene til elevene ikke passet inn i disse kategoriene, kan det gi problemer i forhold til analysen av datamaterialet. Andre utfordringer ved strukturert intervju er at det vanligvis er lite rom for variasjon i svarene. Alle respondentene vil få de samme spørsmålene i den samme rekkefølge av intervjueren, og det vil gi lite fleksibilitet i måten spørsmålene stilles og svarene gis på (Postholm, 2005, side 69). Et strukturert intervju kan føre til at respondentene prøver å gjøre

intervjueren til lags ved gi det svaret de tror intervjueren ønsker (Postholm, 2005, side 71), og det er en stor kilde til responders bias. All informasjon som samles inn, vil gi forskeren hjelp til å forstå forskningsstedet fra respondentenes perspektiv. Forskeren kan ha noen kategorier klare før intervjuene, men en kvalitativ forsker er også åpen for at hver av disse kategoriene kan bli avkreftet og at kategorier som man ikke har definert på forhånd, kan dukke opp. Et planlagt, formelt intervju hindrer dermed ikke forskeren i å være induktiv i forskningsprosessen samlet sett (Postholm, 2005, side 71).

Det ble også utført et halvstrukturert intervju, som er den andre formen for intervju. Her vil intervjueren ha noen spørsmål klare på forhånd, samtidig som han er åpen for at respondentene har ulike forhold som de vil ta opp. Dersom intervjuet dreier i en annen retning enn på forhånd planlagt, vil intervjueren være med på respondentenes refleksjonsreise i stedet for å holde seg til spørsmålsrekken (Postholm, 2005, side 72).

Intervjuene av elevene og lærer ble tatt opp på bånd for å kunne beskrive datamaterialet nøyaktig, og det ga en god troverdighet i oppgaven. Beskrivelsen og tolkningen av disse intervjuene, ble fremlagt i bokser i resultatet. Det var kun tatt med utdrag av intervjuene som var viktig for å besvare problemstillingene, se kapittel 1.2. Det forekom til en viss grad translasjon av utdragene som ble fremlagt, men det var kun for å få et bedre språk på utsagnene til elevene og læreren. En slik beskrivelse og tolkning av intervjuene vil også være med på å opprettholde en god troverdighet i oppgaven.

Intervju av elever

Intervjuet av elevene ble bygget opp som et strukturert intervju, og disse spørsmålene som ble stilt dem er vedlagt i vedlegg 4. Elevene måtte fortelle om hvordan de forberedte seg til en flervalgprøve og kortsvarsprøve, hvordan vanskelighetsgraden på flervalgsprøven og kortsvarsprøven var, om de gjettet på flervalgsprøven og om de benyttet eliminasjonsmetoden, om de fikk uttrykt kunnskapen sin på flervalgsprøven og kortsvarsprøven og hva slags fordeler dette opplegget som ble gjennomført ville gi dem på en eksamen. Intervjuet ble avsluttet med et åpent spørsmål om de hadde noen siste kommentarer.

Navnsettingen til elevene som ble intervjuet ble laget med utgangspunkt i prestasjonen på flervalgsprøven. Her er elevene satt opp hierarkisk i forhold til prestasjon fra flervalgsprøven og navnsetting ble videre gjennomført med utgangspunkt i alfabetet, der den beste eleven som fikk 22 poeng, fikk navnet Elev A. Den nest beste eleven fikk navnet Elev B, og slik fortsetter

det. De fire elevene som ble intervjuet, ble valgt med utgangspunkt i prestasjonene fra flervalgsprøven og i forhold til hvordan de i følge læreren presterte i Biologi 2.

De fire elevene som ble intervjuet var:

- Elev A, som fikk 22 korrekte svar på flervalgsprøven. Hun var faglig sterk i Biologi 2.
- Elev G, som fikk 19 korrekte svar på flervalgsprøven. Hun var også faglig sterk i Biologi 2.
- Elev P, som fikk 12 korrekte svar på flervalgprøven. Han presterte middels faglig sett i Biologi 2.
- Elev Q, som fikk 11 korrekte svar på flervalgprøven. Hun presterte godt i Biologi 2.

Det var også ønskelig å intervjue en elev som gjorde det bedre enn forventet på flervalgprøven i forhold til kortsvarsprøven og vanlig underveisvurdering i Biologi 2, men de elevene det dreide seg om, møtte ikke opp på kortsvarsprøven. Siden de bare hadde vært tilstede på den ene av de to prøvene ble det vanskelig å utføre et intervju med dem. Hensikten med å utføre intervjuer av elevene var å få en dypere forståelse av elevenes subjektive, individuelle meninger om bruk av flervalgsoppgaver som vurderingsform. Det ville gi et bedre grunnlag for å kunne besvare den første problemstillingen min om i hvilke grad elevene får i tilstrekkelig grad vist sin faglige kompetanse ved bruk av flervalgsoppgaver, se kapittel 1.2.

Intervju av lærer

Læreren ble også intervjuet etter gjennomføring av flervalgsprøven og kortsvarsprøven. Intervjuet var et halvstrukturert intervju der en del spørsmål var klare på forhånd, men læreren kunne også bestemme hva slags temaer som skulle tas opp underveis i intervjuet. De forhåndsbestemte intervjuspørsmålene er fremstilt i vedlegg 5. I spørsmålene måtte læreren fortelle litt om hva han synes om hvordan flervalgsoppgaver fungerte ved vurdering *av* læring og vurdering *for* læring, nytteverdien av bruk av flervalgoppgaver på it's learning, den tekniske bruken av Test 2.0, måling av kunnskapsnivå ved bruk av flervalgsoppgaver i forhold til kompetansemål og Blooms taksonomi, fordeler og ulemper ved bruk av flervalgoppgaver som vurderingsform og hvordan læreren vil benytte flervalgsoppgaver videre i undervisningen. Hensikten med dette intervjuet var å supplere elevenes synspunkter med lærerens for å kunne besvare den første problemstillingen, se kapittel 1.2, samt samle datamateriale for å kunne besvare den andre problemstillingen min, se kapittel 1.2.

6. Datainnsamling og resultater

Datamaterialet ble samlet inn fra en spørreundersøkelse som bestod av avkrysningspåstander og åpne spørsmål og intervju av fire elever og læreren. Alle elevene i utvalget måtte utføre spørreundersøkelsen, og de fire elevene som ble intervjuet ble selektert ut fra faglig nivå i Biologi 2 og i forhold til prestasjonen deres på flervalgsprøven. I dette kapittelet presenteres de ulike elementene fra datainnsamlingen for å besvare problemstillingene, som ble beskrevet i kapittel 1.2.

Kapittelet deles inn i fire hoveddeler, og hoveddelene vil bestå av følgende elementer:

- Flervalgsprøven
    - Var flervalgsprøven for lett?
    - Gjettet elevene på flervalgsprøven?
    - Fikk elevene vist sin faglige kompetanse gjennom flervalgsprøven?
- Kortsvarsprøven
    - Var kortsvarsprøven for lett?
    - Fikk elevene vist sin faglige kompetanse gjennom kortsvarsprøven?
- Flervalgsoppgaver og kortsvarsoppgaver
    - Foretrekker elever flervalgsoppgaver ved underveisvurdering?
    - Foretrekker elever flervalgsoppgaver ved sluttvurdering?
    - Skriftlig eksamen i Biologi 2
- Bruk av Test 2.0 for å forbedre elevene til skriftlig eksamen i Biologi 2

6.1. Flervalgsoppgaver

Flervalgsprøven ble gjennomført tirsdag 13.01.09, og den er vedlagt som vedlegg 6. Denne prøven ble utført i Test 2.0. Prøven bestod av 24 spørsmål med ulik vanskelighetsgrad, og den ble laget med utgangspunkt i 3 av 5 punkter i kompetansemålet *Genetikk* fra læreplanen for Biologi 2 (Utdanningsdirektoratet, 2006a). I dette kompetansemålet skal elevene blant annet kunne:

- Forklare strukturen til DNA og hvordan DNA blir kopiert før cellene deler seg.
- Gjøre greie for transkripsjon og translasjon av gen (…).
- Sammenligne mitose og meiose med vekt på fordeling av genmaterialet i cellene som blir dannet.

Oppgavene på flervalgprøven inneholdt dermed følgende temaer:

- DNA-oppbygning
- DNA-replikasjon
- DNA-transkripsjon
- DNA-translasjon
- Mitose og meiose

Alle korrekte svar gav 1 poeng per oppgave, og det forekom ingen korrigering for gjetting, se kapittel 2.6. Noen av oppgavene utfordret elevene lavt på Blooms taksonomi med nivå 1 *gjengivelse* og nivå 2 *forståelse*, som ble beskrevet i kapittel 4.1. Andre oppgaver utfordret elevene på høyere plan ved bruk av nivå 3 *anvendelse* og nivå 4 *analyse*. Tabell 1 viser en oversikt over hvor de ulike flervalgsoppgavene kan plasseres på Blooms taksonomi. For å lese stammen og svaralternativene, se vedlegg 6.

Tabell 1: En oversikt over de ulike oppgavene som var i flervalgsprøven, og hvor de kan plasseres i forhold til Blooms taksonomi. For å lese stammen og svaralternativene, se vedlegg 6.

Flervalgsoppgave	**Blooms taksonomi**
Oppgave 1	Nivå 1
Oppgave 2	Nivå 2
Oppgave 3	Nivå 2
Oppgave 4	Nivå 1
Oppgave 5	Nivå 2
Oppgave 6	Nivå 4
Oppgave 7	Nivå 3
Oppgave 8	Nivå 2
Oppgave 9	Nivå 3
Oppgave 10	Nivå 1
Oppgave 11	Nivå 2
Oppgave 12	Nivå 1
Oppgave 13	Nivå 1
Oppgave 14	Nivå 2
Oppgave 15	Nivå 2
Oppgave 16	Nivå 4
Oppgave 17	Nivå 1
Oppgave 18	Nivå 1
Oppgave 19	Nivå 3
Oppgave 20	Nivå 2
Oppgave 21	Nivå 2
Oppgave 22	Nivå 1
Oppgave 23	Nivå 2
Oppgave 24	Nivå 4

Siden prøven foregikk på it's learning, var det forhåndsbestemt at oppgavene skulle bli gitt i ulik rekkefølge til elevene og det skulle ikke være mulig med fri navigering mellom spørsmålene. Det var for å hindre juks. Det var også en stor utfordring å lage mange flervalgsoppgaver fra et kapittel, der begrepene fra svaret i en flervalgsoppgave ikke ble brukt som stamme i en annen oppgave. Denne utfordringen ble også eliminert ved å hindre fri navigering mellom flervalgsoppgavene.

De fleste elevene skåret høyt på flervalgsprøven, der gjennomsnittlig poengsum var 18,31 poeng av totalt 24 poeng. Den laveste skåren var 11 poeng og den høyeste skåren var 22 poeng. Det ble analysert ulike elementer ved denne flervalgsprøven ut fra datainnsamling fra spørreundersøkelsen og intervjuene av elever og lærer. I denne delen av kapittel 6 er det presentert elementer rundt i hvilken grad flervalgsprøven var for lett, om elevene gjettet mye, om de benyttet elimineringsmetoden og om de klarte å vise sin faglige kompetanse fra kompetansemålet *Genetikk* fra LK06 gjennom denne flervalgsprøven.

Var flervalgsprøven for lett?

Det var 17 av 18 elever som var tilstede på flervalgsprøven. Det var høye skårer på flervalgsprøven, og gjennomsnittlig poengsum var, som nevnt, 18,31 poeng. Flere av spørsmålene fra flervalgsprøven hadde en P-verdi i intervallet 80-100 %. P-verdi ble beskrevet i kapittel 2.4, der en generell regel er at om P-verdien er 90 % eller høyere, er oppgaven meget lett for å få korrekt svar. Det er samtidig avhengig av om elevgruppen er en gruppe med meget flinke elever. Derfor er det viktig å diskutere homogeniteten i elevgruppen. Utvalget for denne studien bestod av mange faglig sterke elever, og 56 % av elevene fikk karakter 5 eller 6 som terminkarakter i Biologi 2 høsten 2008, se kapittel 5.3. Det er viktig å vurdere flervalgsoppgavene opp mot vanskelighetsnivå i forhold til elevgruppen, og en P-verdi kan være til hjelp for å analysere om stammen var for enkel og/eller om distraktorene var for dårlige (Sirnes, 2005, side 66). Boks 6 viser hvordan P-verdien ble utregnet med utgangspunkt i resultatene fra oppgave 2 på flervalgsprøven.

Boks 6: Et eksempel som viser utregning av P-verdi med utgangspunkt i resultatet fra oppgave 2 på flervalgsprøven.

***Flervalgsoppgave nr 2**: Hva er komponentene i et DNA-nukleotid?*
Svaralternativ 1: Par av komplementære baser som holdes sammen av hydrogenbindinger.
Svaralternativ 2: En dobbelhelix som holdes sammen av hydrogenbindinger.
Svaralternativ 3: En dobbelhelix som holdes sammen av elektronparbindinger.
Svaralternativ 4: Sukkeret deoxyribose og fosfat.
Svaralternativ 5: Sukkeret deoxyribose, en base og fosfat.

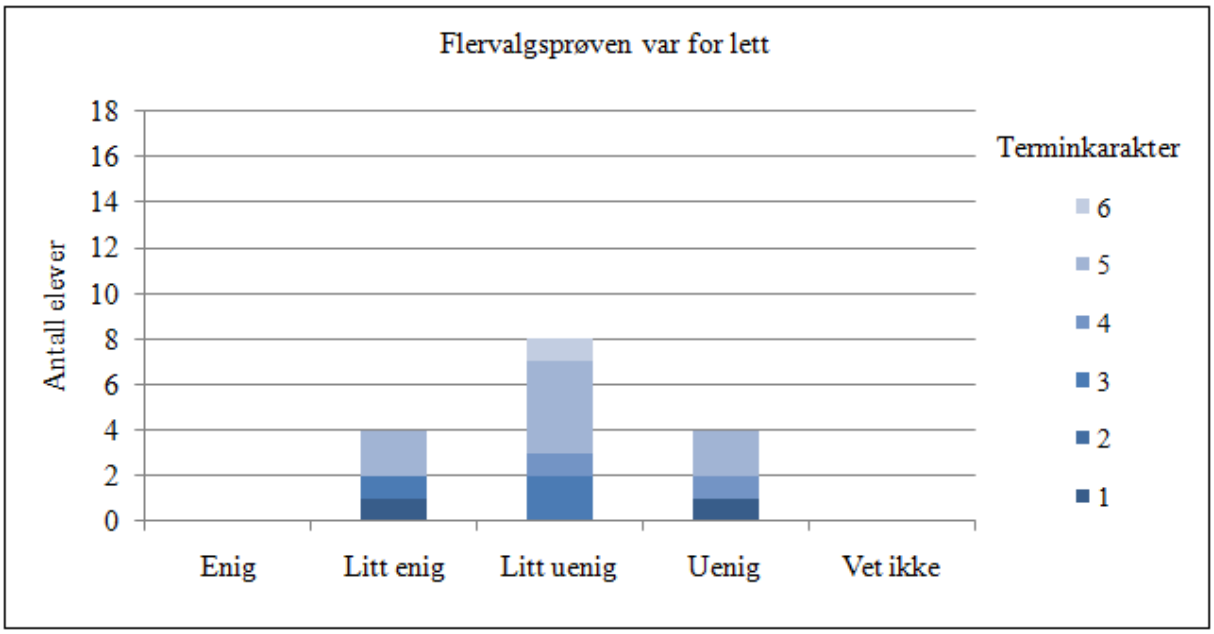
Nøkkelen var svaralternativ 5, og det var 16 elever som svarte korrekt på oppgaven. 1 elev som svarte feil.
$P = R / T \cdot 100\ \%$ (Formel 1, fra kapittel 2.4)
$P = 16 / 17 \cdot 100\ \%$
P = 94 %

I eksempelet fra boks 6 var P-verdien 94 %, og dersom denne prøven skulle gjennomføres en gang til, måtte man ha vurdert stammen og/eller distraktorene for å gjøre denne oppgaven mer utfordrende for elevene. Tabell 2 viser en oversikt over P-verdiene for alle oppgavene som ble gitt på flervalgsprøven, og ut fra den oversikten var det 42 % av oppgavene som hadde en P-verdi i intervallet 80-100 %.

Tabell 2: En oversikt over P-verdiene for alle oppgavene som ble gitt på flervalgsprøven. P-verdiene ble beregnet ut fra formel 1. For å lese stammen og svaralternativene, se vedlegg 6.

Flervalgsoppgave	**Antall elever som svarte korrekt, R**	**Totalt antall elever som besvarte oppgaven, T**	**Utregnet P-verdi ut fra formel 1**
Oppgave 1	14	17	82 %
Oppgave 2	16	17	94 %
Oppgave 3	16	16	100 %
Oppgave 4	12	17	71 %
Oppgave 5	12	17	71 %
Oppgave 6	12	17	71 %
Oppgave 7	3	17	18 %
Oppgave 8	13	17	76 %
Oppgave 9	10	17	59 %
Oppgave 10	13	17	76 %
Oppgave 11	12	17	71 %
Oppgave 12	14	17	82 %
Oppgave 13	14	17	82 %
Oppgave 14	8	17	47 %
Oppgave 15	15	17	88 %
Oppgave 16	17	17	100 %
Oppgave 17	10	17	59 %
Oppgave 18	14	17	82 %
Oppgave 19	10	17	59 %
Oppgave 20	12	17	71 %
Oppgave 21	11	17	65 %
Oppgave 22	17	17	100 %
Oppgave 23	17	17	100 %
Oppgave 24	13	17	76 %

En påstand i spørreundersøkelsen var *Flervalgsprøven var for lett*, se vedlegg 3. Resultatet fra denne påstanden er fremstilt i figur 4, og den viser at 75 % av elevene var *litt uenig* eller *uenig* til denne påstanden. Det betyr at de fleste elevene *ikke* synes at prøven var for lett, selv om P-verdien var høy på flere av oppgavene. Denne figuren viser også en oversikt over påstanden i forhold til terminkarakteren som elevene fikk i Biologi 2 høsten 2008. Det var ikke noe særlig differensiering i datamaterialet i forhold til karakteren.



Figur 4: En oversikt over resultatet fra påstanden "Flervalgsprøven var for lett" i spørreundersøkelsen. Denne påstanden er sett i forhold til terminkarakterene som elevene fikk høsten 2008 i Biologi 2.

For å observere tendensene fra figur 4 tydeligere, må man gå inn i datamaterialet i de åpne spørsmålene fra spørreundersøkelsen. I det åpne spørsmålet *Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven?*, se vedlegg 3, kom det tydeligere fram at noen elever synes at flervalgsprøven var for lett, i forhold til oversikten i figur 4. Boks 7 viser et utdrag fra tre elevbesvarelser på spørreundersøkelsen i forhold til hva de besvarte på påstanden *Flervalgsprøven er for lett* og det åpne spørsmålet *Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven*.

Boks 7: En oversikt over synspunktene til tre elever om hvordan de svarte ved påstanden "Flervalgsprøven var for lett" og det åpne spørsmålet "Fikk du vist din kunnskap fra kapittel 6 gjennom flervalgsprøven?".

Påstand 2a ii: *Flervalgsprøven var for lett.*
Heidi, terminkarakter 5: Litt uenig.
Anna, terminkarakter 4: Litt uenig.
Erik, terminkarakter 3: Litt enig.

Spørsmål 2b ii: *Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven? Begrunn svaret.*
Heidi, terminkarakter 5: Ja, synes jeg fikk vist min kunnskap. Var bra spørsmål, ikke for lette. For å svare rett måtte man ha en del kunnskap (eventuelt veldig flaks).
Anna, terminkarakter 4: Ja, (…) selv om en del spørsmål var litt enkle, var det også vanskelige spørsmål som krevde kunnskap fra kapittelet.
Erik, terminkarakter 3: Nei, jeg synes at spørsmålene som vi fikk var litt mer oversiktlige enn jeg forventet. På eksamen vil spørsmålene være litt vanskeligere.

Boks 7 viser at det var stor spredning i synspunktene hos de tre elevene. Heidi var *litt uenig* i påstanden om at flervalgsprøven var for lett, og i det åpne spørsmålet kom det også fram at oppgavene ikke var for lette. Heidi begrunnet det med at man måtte ha kunnskap for å kunne besvare dem. Anna var også *litt uenig* i påstanden om flervalgsprøven var for lett, men i det åpne spørsmålet kom det fram at noen av oppgavene var litt enkle, samtidig som det var noen mer utfordrende spørsmål som krevde mer kunnskap fra kapittelet. Erik, derimot, var *litt enig* i påstanden om at flervalgsprøven var for lett, og hans begrunnelse på det åpne spørsmålet var at spørsmålene på prøven var litt mer oversiktlige enn det som var forventet. Synspunktene fra disse tre elevene viser at det var stor spredning i datamaterialet, men de generelle tendensene fra det åpne spørsmålet *Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven?* var at mange elever synes at de fikk vist sin faglige kompetanse gjennom flervalgsprøven. Flere av elevene forklarte videre at de fikk vist den faglige kompetansen sin bedre gjennom kortsvarsoppgaver, siden de kunne formulere svarene sine selv ved åpne oppgaver.

Læreren i klassen observerte[11] at de faglig sterke elevene skåret høyt på flervalgsprøven og at de faglig svake elevene gjorde det generelt bedre enn ved en vanlig prøve som inneholdt åpne oppgaver. Det var forventet at elevene ville skåre bedre på flervalgsprøven enn på en prøve med åpne oppgaver, siden elevenes tid under prøven ikke ble benyttet for å formulere svarene på oppgavene (Sirnes, 2005, side 10). Det man må undersøke videre er om elevene kom fram

[11] Det ble ikke benyttet observasjon som en formell metode for å samle inn resultatene, men jeg var med i undervisningen under prøvene og pratet underveis med læreren om hva han observerte. Mye at dette ble formelt tatt opp under intervjuet med læreren etter prøvene.

til det riktige svaret på flervalgsprøven på grunn av sin faglige kunnskap eller ved bruk av elimineringsmetoden og gjetting. Dersom elevene var usikre på svaret, men klarte å eliminere bort flere distraktorer i flervalgsoppgavene, kan det vise til at distraktorene var for dårlige. Dersom elevene ikke gjettet, men likevel klarte å få korrekt svar, vil det være en indikator på at stammen var for enkel eller at det var mange faglig sterke elever som gjennomførte prøven.

Gjettet elevene på flervalgsprøven?

Det kom fram i både det kvantitative og kvalitative datamaterialet fra spørreundersøkelsen og intervjuene at flere elever gjettet på en eller flere oppgaver på flervalgsprøven. Før prøvestart ble elevene informert om at det ikke ville forekomme en korrigering av resultatet ved gjetting, se kapittel 2.6, og elevene ble oppfordret til å svare på alle spørsmålene. Den ene eleven hvor ble intervjuet, Elev Q, hadde gjettet ved flere av oppgavene. På de flervalgsoppgavene som hun gjettet uten å eliminere svaralternativer først, ble oppgavene som regel ukorrekt, men dersom hun klarte å eliminere bort noen distraktorer og sto da mellom færre svaralternativer, ble oppgavene oftere korrekt. Boks 8 viser et utdrag fra intervjuet med Elev Q.

Boks 8: Utdrag av intervjuet med Elev Q, utdraget viser om hun gjettet mye og om hun benyttet eliminasjonsmetoden sammen med gjetting.

Intervjuer: *Gjettet du mye?*
Elev Q: Gjettet bare på dem jeg ikke hadde peiling på.

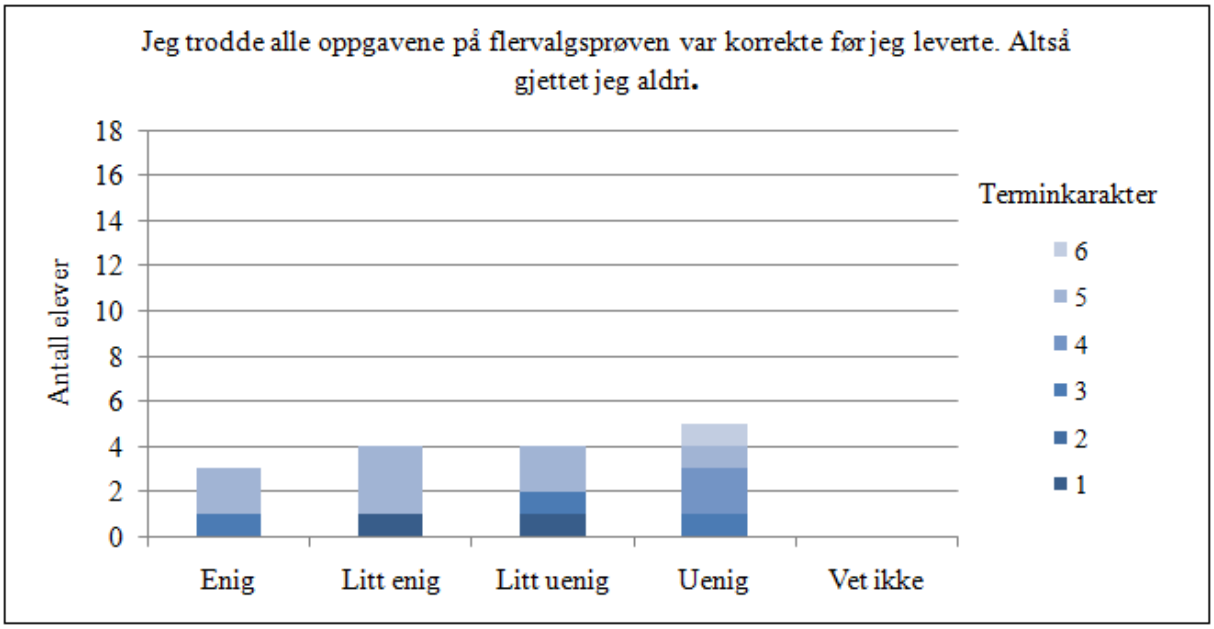
Intervjuer: *Klarte du å eliminere noen svaralternativer når du gjettet?*
Elev Q: Satt et vilkårlig kryss på noen, og noen eliminerte jeg på.

Intervjuer: *Ble det korrekte eller ukorrekte svar?*
Elev Q: Feil på de fleste svarene. De oppgavene jeg eliminerte, fikk jeg flere korrekte på.

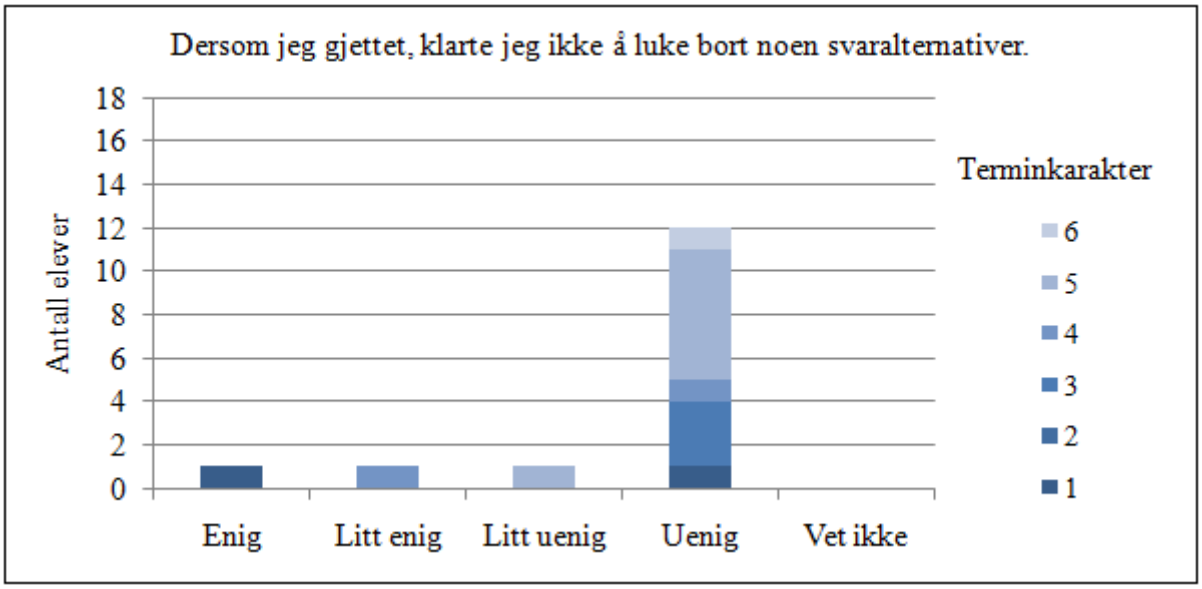
Boks 8 viser at bruk av gode distraktorer er viktig. Dersom elevene klarer å eliminere bort distraktorer, vil det være større sannsynlighet at elevene klarer å gjette nøkkelen, og ved bruk av mange og gode distraktorer vil det være vanskeligere for elevene å velge nøkkel ved vilkårlig gjetting.

Det var også to påstander i spørreundersøkelsen som handlet om i hvilken grad elevene gjettet på flervalgsoppgavene, og om elevene benyttet eliminasjonsmetoden før de gjettet. Påstandene var *Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte.*

Altså gjettet jeg aldri og *Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer*. Resultatene fra disse to påstandene er vist i figur 5 og 6.



Figur 5: En oversikt fra påstanden "Jeg trodde alle oppgavene på flervalgsoppgaven var korrekte før jeg leverte. Altså gjettet jeg aldri". Denne oversikten er sett i forhold til terminkarakteren som elevene fikk i Biologi 2 høsten 2008.



Figur 6: En oversikt over resultatene fra påstanden "Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer". Oversikten er sett i forhold til terminkarakterene som elevene fikk i Biologi 2 høsten 2008.

Resultatene fra figur 5 viste at elevene var forholdsvis jevnt fordelt i forhold til om de gjettet eller ikke på flervalgsprøven, men i følge figur 6 var det mange elever som benyttet

eliminasjonsmetoden. Dette kan virke som motsigende resultater, siden flere elever var sikre på resultatet sitt før de leverte inn flervalgsprøven, samtidig som de benyttet eliminasjonsmetoden ved gjetting. For å få oppklaring i disse motsigende resultatene, måtte man gå inn i besvarelsene fra det åpne spørsmålet *Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mellom alle svaralternativene?* i spørreundersøkelsen. Det var flere elever som var *litt enig* i påstanden om at de trodde alle oppgavene var korrekte før de leverte oppgavene, men de gjettet på enkelte oppgaver. Det gjelder blant annet for Frida, og boks 9 viser et utdrag fra spørreundersøkelsen til Frida.

Boks 9: Utdrag av spørreundersøkelsen til Frida, som viser hva hun svarte ved påstandene "Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte. Altså gjettet jeg aldri" og "Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer". Utdraget viser også hvordan hun begrunnet sitt svar i det åpne spørsmålet "Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mye mellom alle svaralternativene?".

Påstand 2a iv: *Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte, altså gjettet jeg aldri.*
Frida, terminkarakter 1: Litt enig.

Påstand 2a v: *Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer.*
Frida, terminkarakter 1: Enig.

Spørsmål 2b v: *Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mellom alle svaralternativene?*
Frida, terminkarakter 1: Gjettet ikke mye, men der jeg gjettet, eliminerte jeg alternativer.

Boks 9 viser at det kvantitative og kvalitative datamaterialet ikke er samstemt på alle punkter, og det gjelder også for flere elever som har sagt at de er *litt enig* i påstanden om at de aldri gjettet selv om de benyttet eliminasjonsmetoden for å besvare oppgaver, blant annet for Nina og Oda. De var *litt enig* i påstanden at om de aldri gjettet, som ble presentert i figur 5. I resultatet fra det åpne spørsmålet *Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mellom alle svaralternativene* kom det fram at Nina og Oda gjettet noen ganger og benyttet eliminasjonsmetoden for å komme fram til det korrekte svaret. Boks 10 viser utdrag fra spørreundersøkelsen til Nina og Oda.

Boks 10: Utdrag fra spørreundersøkelsen til Nina og Oda, som viser hva de svarte ved påstandene "Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte. Altså gjettet jeg aldri" og "Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer". Utdraget viser også hvordan de begrunnet sine svar i det åpne spørsmålet "Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mye mellom alle svaralternativene?".

Påstand 2a iv: *Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte, altså gjettet jeg aldri.*
Nina, terminkarakter 5: Litt enig.
Oda, terminkarakter 5: Litt enig.

Påstand 2a v: *Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer.*
Nina, terminkarakter 5: Litt uenig.
Oda, terminkarakter 5: Uenig.

Spørsmål 2b v: *Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mellom alle svaralternativene?*
Nina, terminkarakter 5: Gjettet ikke så mye, men jeg kunne være usikker på en eller to svaralternativer noen ganger.
Oda, terminkarakter 5: Ja, på noen kunne jeg gjettet, men jeg var ganske sikker på de fleste spørsmålene.

I besvarelsene fra det åpne spørsmålet om i hvilken grad elevene gjettet mye, viste det seg at flere elever ikke var så sikre på at alle oppgavene på flervalgsprøven var korrekte før de leverte. Det kunne ha resultert en større forskjell i figur 5, der flere elever ville ha vært *litt uenig* og *uenig* i påstanden om at de trodde at oppgavene var korrekte før de leverte dem inn. Samtidig forekom det at flere elever var sikre på svaret sitt, selv om at de benyttet eliminasjonsmetoden. Det kom fram i figur 6 og i det åpne spørsmålet *Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mellom alle svaralternativene*? Boks 11 viser hvordan ulike elever benyttet eliminasjonsmetoden ulikt for å besvare oppgavene på flervalgsprøven. Erik og Rita benyttet eliminasjonsmetoden for å velge det korrekte svaret, mens Petra og Berit benyttet eliminasjonsmetoden for å få færre alternativer å gjette mellom.

Boks 11: Utdrag av resultatene fra spørreundersøkelsen, som viste hvordan Erik, Rita, Petra og Berit besvarte påstandene "Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte. Altså gjettet jeg aldri" og "Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer". Utdraget viser også hvordan de begrunnet sine svar i det åpne spørsmålet "Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mye mellom alle svaralternativene?".

Påstand 2a iv: *Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte, altså gjettet jeg aldri.*
Erik, terminkarakter 3: Enig.
Rita, terminkarakter 5: Litt uenig.
Petra, terminkarakter 1: Litt uenig.
Berit, terminkarakter 4: Uenig.

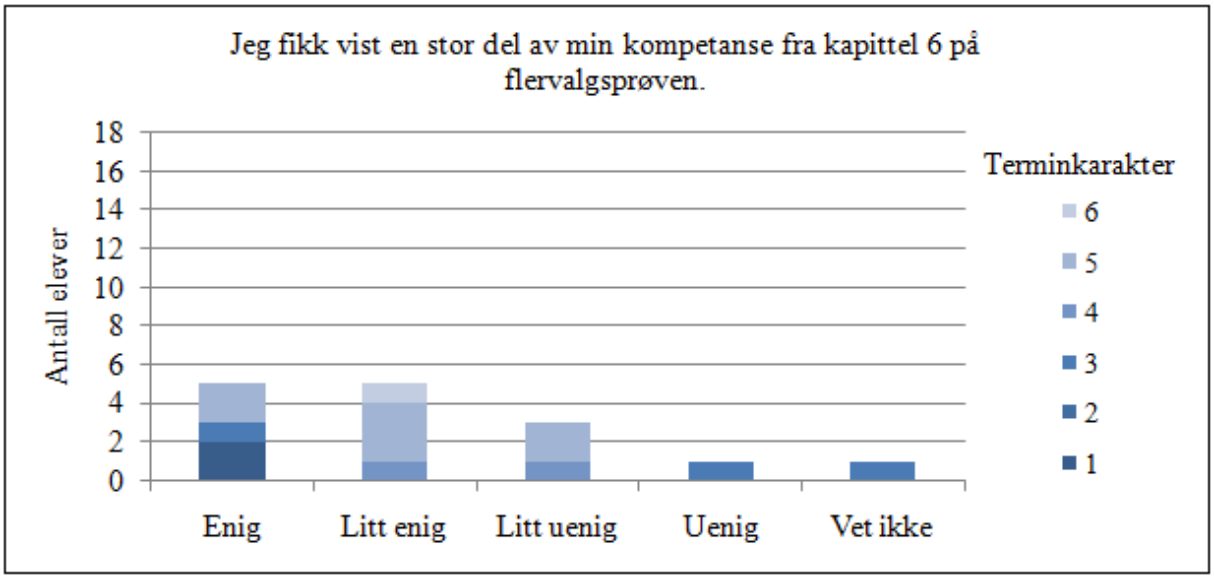
Påstand 2a v: *Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer.*
Erik, terminkarakter 3: Uenig.
Rita, terminkarakter 5: Uenig.
Petra, terminkarakter 1: Uenig.
Berit, terminkarakter 4: Litt enig.

Spørsmål 2b v: *Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mellom alle svaralternativene?*
Erik, terminkarakter 3: Nei, jeg kunne eliminere noen svaralternativer slik at det ble lettere å velge det riktige svaret.
Rita, terminkarakter 5: Jeg gjettet ikke så mye, var sikker på de meste. Jeg benyttet elimineringsmetoden som hjalp meg til å komme frem til det korrekte svaret.
Petra, terminkarakter 1: Jeg gjettet på noen spørsmål, men da hadde jeg benyttet eliminasjonsmetoden på de fleste av oppgavene.
Berit, terminkarakter 4: Jeg gjettet på noen. På noen av oppgavene fikk jeg til å eliminere noen svaralternativer slik at det ble lettere å gjette.

Det ble beskrevet i kapittel 2.6 hvordan man kan korrigere for gjetting ved hjelp av tre forskjellige metoder. Sirnes (2005) mente at det er mest forsvarlig å oppfordre elevene til å svare på alle spørsmålene når ett eller flere svaralternativer kan elimineres som gale dersom man skal korrigere skåren (Sirnes, 2005, side 46). I denne studien forekom det ingen korrigering for gjetting på flervalgsprøven. Dersom det hadde forekommet korrigering ville sluttresultatet og gjennomsnittlig poengsum vært annerledes. Vedlegg 7 viser en oversikt over resultatene fra flervalgsprøven, samt en oversikt over resultatene dersom de tre korrigeringsmetodene som ble fremstilt i kapittel 2.6 hadde blitt benyttet. Siden elevene ble oppfordret til å svare på alle oppgavene, vil ikke korrigering for gjetting gi realistiske skårer.

Fikk elevene vist sin faglige kompetanse gjennom flervalgsprøven?

Den første delen av skriftlig eksamen i Biologi 2 består som nevnt av flervalgsoppgaver, og det er derfor viktig å undersøke om elevene i tilstrekkelig grad får vist sin faglige kompetanse gjennom flervalgsoppgaver som vurderingsform. Det ble samlet inn både kvantitativt og kvalitativt datamaterialet fra elevene om i hvilken grad de fikk vist sin faglige kompetanse gjennom flervalgsprøven. Resultatet fra påstanden *Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på flervalgsprøven* i spørreundersøkelsen er fremlagt i figur 7.



Figur 7: En oversikt over resultatet fra påstanden "Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på flervalgsprøven". Oversikten viser også fordelingen i forhold til terminkarakterene som elevene fikk høsten 2008 i Biologi 2.

I resultatene fra figur 7 var det et flertall av elevene som var *enig* og *litt enig* i påstanden om at de fikk vist en stor del av kompetansen sin på flervalgsprøven. De faglig svake elevene svarte *enig*, og flesteparten av de faglig sterke og middels faglig sterke elevene svarte også *enig* og *litt enig* til denne påstanden. For å undersøke dette resultatet nærmere, måtte man gå inn i det kvalitative datamaterialet. I resultatet fra det åpne spørsmålet fra spørreundersøkelsen *Fikk du vist din kompetanse gjennom flervalgsprøven?* kom det fram at flere elever synes at de fikk vist tilstrekkelig kompetanse, men de følte samtidig at de fikk vist mer og bedre faglig kompetanse i åpne oppgaver. Flere av elevene var også veldig fornøyde med bruk av flervalgsoppgaver som vurderingsform, og de skulle ønske at det ble brukt mer av en slik vurderingsform. Boks 12 viser resultatet fra hvordan flere elever, besvarte påstanden *Jeg fikk*

vist en stor del av min kompetanse fra kapittel 6 på flervalgsprøven og det åpne spørsmålet *Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven?*

Boks 12: Utdrag som viser resultatet fra hvordan elevene besvarte påstanden "Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på flervalgsprøven" og det åpne spørsmålet "Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven?".

Påstand 2a vi: *Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på flervalgsprøven.*
Daniel, terminkarakter 3: Vet ikke.
Berit, terminkarakter 4: Litt uenig.
Cecilie, terminkarakter 5: Litt uenig.
Kaia, terminkarakter 5: Litt uenig.
Rita, terminkarakter 5: Litt enig.
Petra, terminkarakter 1: Enig.
Lars, terminkarakter 3: Enig.
Ine, terminkarakter 5: Enig.

Spørsmål 2b ii: *Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven? Begrunn svaret.*
Daniel, terminkarakter 3: Med tanke på at jeg gjettet, må jeg si nei.
Berit, terminkarakter 4: Egentlig ikke. Jeg følte at jeg fikk vist kunnskapen min bedre gjennom kortsvarsoppgavene.
Cecilie, terminkarakter 5: Jeg vet ikke helt jeg, både og. Det kom spørsmål der jeg ikke kunne så mye, men klarte å svare riktig allikevel, og spørsmål på tema som jeg kunne mer om hvor jeg ville ha vist større kompetanse ved å skrive utfyllende på en langsvarsoppgave.
Kaia, terminkarakter 5: Egentlig ikke, for jeg fikk ikke beskrive selv hva jeg mente hadde vært et "bedre" svar.
Rita, terminkarakter 5: Jeg fikk vist en god del siden det meste ble rett, men det var mange detaljer jeg kunne som jeg ikke fikk vist.
Petra, terminkarakter 1: Ja, jeg synes det… Det gikk mye bedre enn på vanlige prøver (…). Jeg skulle ønske alle prøver var flervalgsprøver.
Lars, terminkarakter 3: Ja, jeg slapp å utforme svaret, dermed klarte jeg bedre å vise min kunnskap.
Ine, terminkarakter 5: Ja, det gjorde jeg. Vi måtte ha en viss forståelse av stoffet for å svare riktig, hvis du ikke gjettet.

Resultatet fra boks 12 viser sammenheng mellom påstanden *Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på flervalgsprøven* og det åpne spørsmålet *Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven?* fra spørreundersøkelsen om i hvilken grad elevene fikk vist faglig kompetanse gjennom flervalgsprøven. Elevene begrunnet svarene i det åpne spørsmålet, og det kom fram fra flere elever at de får vist mer kompetanse gjennom kortsvarsoppgaver enn ved flervalgsoppgaver, men det var stor spredning i datamaterialet. I besvarelsene, som vises i boks 12, var det også oppgitt terminkarakteren som elevene fikk høsten 2008 i Biologi 2. Lars, som hadde terminkarakter 3, svarte at han fikk vist

kompetansen sin, siden han slapp å formulere svarene sine selv. Det å benytte tiden under en prøve til tenking, og ikke på å formulere svarene er en fordel som Sirnes (2005) nevner (Sirnes, 2005, side 10). Faglig svake elever kan ha en større utfordring med å formulere svarene sine selv enn faglig sterke elever. Flere faglig sterke elever var *litt uenig* eller *uenig* i påstanden om at de fikk vist en stor del kunnskapen sin på flervalgsprøven. I boks 12 nevnte Cecilie, som hadde terminkarakter 5, at hun bedre fikk vist sin kompetanse på en åpen oppgave enn på flervalgsprøven, siden hun på den åpne oppgaven kunne skrive utfyllende svar. Dette synspunktet kom også fram under intervjuet med Elev A, som var en faglig sterk elev, og boks 13 viser en utdrag fra dette intervjuet.

Boks 13: Utdrag fra intervjuet med Elev A om i hvilken grad hun fikk uttrykt sin kompetanse gjennom flervalgsprøven.

Intervjuer: *Fikk du uttrykt din kompetanse på flervalgsprøven?*
Elev A: Har du lest litt, så får du vist en del på flervalgsprøven av hva du kan siden du kan svare på alt. Samtidig får man ikke vist alt, siden du sitter inne med kunnskap som man ikke blir spurt om. Derfor er det bedre med kortsvarsprøver hvor man kan skrive på litt mer enn det som blir spurt om. På flervalgsprøven må man bare krysse av på et svaralternativ, og på kortsvarsprøven kan man skrive mer om emnet.

Det var mange elever som fikk høye skårer på flervalgsprøven, se vedlegg 7, og det er viktig å tenke gjennom hvordan man ønsker å karaktersette elevene etter en flervalgsprøve. Elevene fikk oppgitt en samlet poengsum automatisk etter innlevering av flervalgsprøven i Test 2.0, men hvordan karakterskalaen for karaktersetting av prøven skulle læreren selv vurdere. Læreren hadde informert elevene på forhånd om at flervalgsprøven skulle sees i samsvar med prestasjonene fra kortsvarsprøven. Det var en høy gjennomsnittlig skår på flervalgsprøven med 18,31 poeng av 24 mulig. I følge læreren presterte de faglig sterke elevene bra på flervalgsprøven, men de faglig svake elevene presterte bedre enn det de gjorde til vanlig ved kortsvarsprøver. Da kortsvarsprøven ble gjennomført, utelot de faglig svake elevene å møte opp. De ville ikke ødelegge det gode utgangspunktet som de hadde fra før på flervalgsprøven, se boks 14. Boks 14 viser et utdrag av intervjuet med læreren, der han tar opp denne problematikken, og hva han vil gjøre videre for å unngå denne utfordringen.

Boks 14: Utdrag fra intervjuet med læreren, der han kommenterte hvorfor elever unngikk å møte opp på kortsvarsprøven.

Lærer: Elever tenker taktisk i forhold til karakterer. I denne gruppen har det skjedd en taktisk vurdering. Det var to elever, som før prøven startet opp ville komme seg unna denne kortsvarsprøven. En av de som ville komme seg unna, hadde gjort det bra på flervalgsprøven. De fire andre som var borte hadde også skåret relativt høyt på flervalgsprøven i forhold til hvordan de normalt gjør det. Når de skal prøves i det samme, foregår det en taktisk vurdering. Selv om jeg hadde sagt at prøvene skulle sees som en helhet, valgte de å holde seg borte. De hadde kanskje håpet at det ville gi dem et godt resultat allikevel. Så man må tenke seg veldig godt om rundt hvordan man vil gjøre i forhold til elevene for å forebygge taktiske valg. Nå fungerte det mot sin hensikt. I stedet for oppkjøring mot å måle en viktig del av kompetansen deres, fungerte det slik at de synes det var best å holde seg borte. Dette var stikk motsatt av det som var intensjonen. Slike effekter må man prøve å forebygge.

Jeg prøvde å forbygge det, men jeg var ikke tydelig nok til elevene. Jeg fikk en melding av en elev som uttrykte akkurat dette: ”Jeg gjorde det veldig bra på flervalgsprøven, og nå har jeg vært syk, så jeg er redd for å gjøre det dårligere. Jeg lar være å komme til den kortsvarsprøven.” Eleven skrev det svart på hvitt, en dyktig elev. Det er ikke så enkelt å lage en prøve og vurdere elevene. Elevene gjør taktiske valg, og det taper de på. De ser ikke helheten i det, og at det er sluttevalueringen som er viktig. Nå er vi i starten av en termin, og det som er viktig er kompetansen deres i slutten av terminen i følge vurderingssystemet. Elevene tar det ikke inn over seg.

Intervjuer: *Hvordan vil du vurdere dem som bare møtte opp på flervalgsprøven?*
Lærer: Vet ikke helt, men elevene må nok ha en prøve i kapittelet i en annen form. Dette kapittelet vil bli tatt med på neste prøve, for å måle kompetansen deres. Vi lærere må gjøre noen taktiske valg vi også… Jeg må sikre meg at jeg får målt kompetansen deres i de ulike læreplanmålene og målt sluttkompetansen og satt en karakter.

I den neste delen av resultatet presenteres resultatene fra kortsvarsprøven og ulike elementer ved denne prøven som kom fram i spørreundersøkelsen og intervjuene med elevene.

6.2. Kortsvarsprøven

Kortsvarsprøven ble gjennomført 11 dager etter flervalgsprøven, fredag 23.01.09, kortsvarsprøven er vedlagt som vedlegg 8. Prøven bestod av fem kortsvarsoppgaver, der elevene skulle svare maks en A4-side per spørsmål. Oppgavene skulle besvares på papir, ikke via it’s learning eller tekstredigeringsverktøy som Microsoft Word, siden det var lettere for elevene å kunne tegne relevante figurer på papir enn på datamaskinen. Ved åpne oppgaver kan det være en større utfordring å fastsette hvor i Blooms taksonomi oppgaven befinner seg. Noen elever vil besvare oppgavene ut fra nivå 2 *forståelse* og andre elever vil besvare

oppgavene ut fra nivå 3 *anvendelse* og nivå 4 *analyse*. Hva som kjennetegner de ulike nivåene, ble nevnt i kapittel 4.1. Oppgavene i kortsvarsprøven omhandlet det samme kompetansemålet *Genetikk* fra LK06 som flervalgsprøven, for at elevene kunne sammenligne de to vurderingsformene flervalgsoppgaver og kortsvarsoppgaver under datainnsamlingen.

Kortsvarsprøven bestod av følgende temaer:

- DNA-oppbygning
- DNA-transkripsjon
- DNA-translasjon
- Mitose og meiose
- Genregulering i eukaryote celler

Kortsvarsprøven var litt for omfattende, siden flere elever for fikk liten tid å besvare alle oppgavene fullt ut, selv om temaet DNA-replikasjon ikke var med på prøven. Noe av årsaken var at læreren hadde bestemt å ta med temaet *genregulering i eukaryote celler* på kortsvarsprøven, siden dette temaet ikke ble gitt på flervalgsprøven. Læreren rakk ikke å gjennomgå dette temaet før flervalgsprøven, men det ble gjennomgått i perioden mellom flervalgsprøven og kortsvarsprøven. Læreren ønsket å undersøke elevenes måloppnåelse innen dette temaet i kortsvarsprøven. Det var læreren selv som vurderte kortsvarsprøven, og karaktersettingen skulle ta utgangspunkt i hva elevene besvarte på de ulike oppgavene, samt relevante figurer som var tatt med. Læreren bestemte at vurderingen fra flervalgsprøven og kortsvarsprøven skulle utgjøre en total karakter, som viste en samlet måloppnåelse innen kompetansemålet *Genetikk* fra LK06. Det er samme metode som blir benyttet ved skriftlig eksamen i Biologi 2. I følge vurderingsveiledningen for eksamen i Biologi 2 våren 2008 ble karaktersettingen på eksamen basert på helhetsinntrykket (Utdanningsdirektoratet, 2008c). Karaktersettingen for dem samlede karakteren på de to prøvene ble gjort på samme måte, der karaktersettingen ble basert på helhetsinntrykket. Læreren bestemte at hovedvekten av total karakter ble basert på resultatene fra kortsvarsprøven og den ble justert i forhold til resultatene fra flervalgsprøven. Vedlegg 9 viser en oversikt over karakterene fra flervalgsprøven, kortsvarsprøven og total karakter til elevene.

Ut fra vedlegg 9 ser man at gjennomsnittskarakteren fra kortsvarsprøven var ¾. Læreren korrigerte dette resultatet opp mot resultatene fra flervalgsprøven. De fleste elevene gjorde det bedre på flervalgsprøven, og det gjorde at gjennomsnittskarakteren for total karakter fra

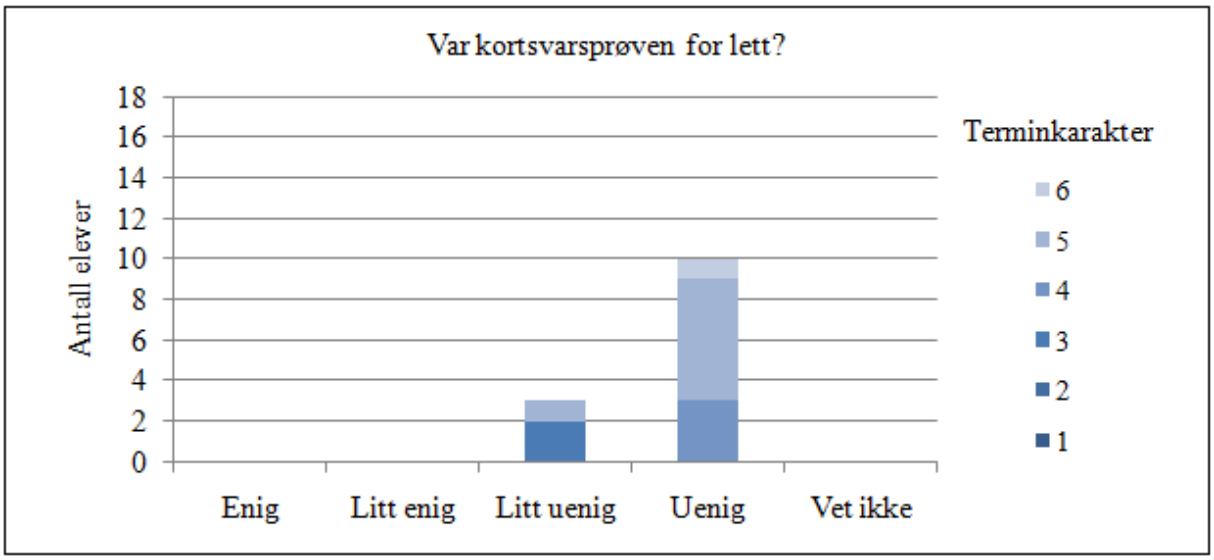
flervalgsprøven og kortsvarsprøven ble en svak 4. Elever som kun møtte opp på kortsvarsprøven, fikk total karakter, siden hovedvekten for karaktersettingen lå på kortsvarsprøven. Elever som kun møtte opp på flervalgsprøven, fikk ikke total karakter.

Kortsvarsoppgavene, som ble gitt på eksamen i Biologi 2 våren 2008, var mer detaljerte enn spørsmålene som elevene fikk på kortsvarsprøven i denne studien (Utdanningsdirektoratet, 2008a). Oppgavene på eksamen skulle besvares direkte i oppgaveheftet, men det var åpnet for at elevene kunne skrive svaret på eget ark og tegne figurer som var relevante for oppgavene (Utdanningsdirektoratet, 2008c). Spørsmålene som elevene fikk på kortsvarsprøven var mer omfattende enn de på skriftlig eksamen i Biologi 2. Dersom oppgavene på kortsvarsprøven skulle være like spesifikke som på kortsvarsoppgavene på eksamen, hadde det blitt flere spørsmål som var identiske til flervalgoppgavene. For å hindre at elevene kun leste på spørsmålene og svaralternativene fra flervalgsprøven som lå på it's learning, ble det bestemt at kortsvarsoppgavene på prøven skulle være mer omfattende enn de normale kortsvarsoppgavene som ble gitt på skriftlig eksamen. Denne avgjørelsen var også en faktor som resulterte i at flere elever fikk for lite tid til å besvare alle spørsmålene på kortsvarsprøven.

Det ble analysert ulike elementer ved denne kortsvarsprøven ut fra datainnsamlingen fra spørreundersøkelsen og intervjuene av elever og lærer. I denne delen av kapittel 6 presenteres det elementer som sier noe om i hvilken grad kortsvarsprøven var for lett og om elevene fikk vist sin faglige kompetanse fra kompetansemålet *Genetikk* fra LK06 gjennom denne kortsvarsprøven.

Var kortsvarsprøven for lett?

Det var 14 av 18 elever som var tilstedet på kortsvarsprøven, og de faglig svake elevene møtte ikke opp. De ville ikke ødelegge det gode utgangspunktet som de fikk på flervalgsprøven, se boks 14 i kapittel 6.1. Det gjorde det vanskelig å sammenligne resultatet fra flervalgsprøven og kortsvarsprøven etter at begge prøvene var gjennomført. Ut fra resultatene fra kortsvarsprøven, se vedlegg 9, presterte flere elever dårligere enn forventet. Derfor var det viktig å undersøke hva elevene synes om vanskelighetsgraden på kortsvarsprøven. Figur 8 viser resultatet fra spørreundersøkelsen, og viser om elevene var *enig*, *litt enig*, *litt uenig* eller *uenig* i påstanden *Kortsvarsprøven var for lett*. Resultatet viste at det var ingen elever som var *enig* eller *litt enig* i påstanden. Flertallet av elevene var *uenig* i denne påstanden, og resultatet var jevnt fordelt i forhold til terminkarakteren som de fikk i Biologi 2 høsten 2008.



Figur 8: En oversikt over resultatet fra påstanden "Var kortsvarsprøven for lett?". Oversikten viser også fordelingen i forhold til terminkarakterene som elevene fikk høsten 2008 i Biologi 2.

For å observere tendensene fra figur 8 tydeligere, måtte man gå inn i intervjuene med de fire elevene. I et intervjuspørsmål ble disse elevene spurt om vanskelighetsgraden på kortsvarsprøven. Boks 15 viser utdrag fra disse intervjuene.

Boks 15: Utdrag fra svar på intervjuspørsmålet "Hvordan var vanskelighetsgraden på kortsvarsprøven?" fra de fire elevene som ble intervjuet, Elev A, Elev G, Elev P og Elev Q.

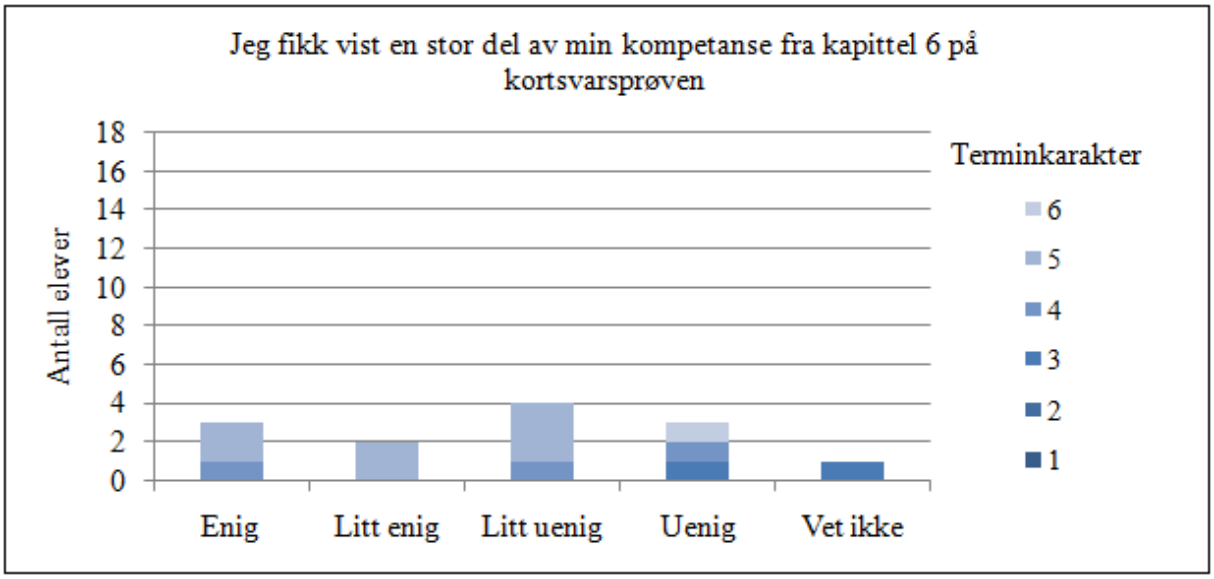
Intervjuer: *Hvordan var vanskelighetsgraden på kortsvarsprøven?*
Elev A: Den var vanskeligere enn flervalgsprøven. Jeg hadde ikke fått lest så mye som jeg hadde håpt, men den var ikke så veldig vanskelig den heller.
Elev G: Det var det samme som på flervalgsprøven. Det var forventede spørsmål, men utfordringen var å formulere seg.
Elev P: Har du lest, var den enkel, men for meg var den ganske vanskelig siden jeg ikke hadde lest. Det var jo spørsmål fra læreboka, så det var bare å lese på de emnene.
Elev Q: Den var bedre enn flervalgprøven. Du får vist litt mer hva du kan. Du kan skrive ned det du vil i stedet for å krysse av.

Dette utdraget viser at elevene synes det var ulike utfordringer ved kortsvarsprøven. Elev P sin utfordring var å svare på oppgavene, mens Elev G sin utfordring var formuleringene av svarene på kortsvarsoppgavene. Slike utfordringer kan være med på å bestemme om elevene synes at kortsvarsprøven var vanskelig eller ikke. Sirnes (2005) beskriver utfordringene ved flervalgsoppgaver, blant annet at de ikke vil måle skriveevne hos elevene (Sirnes, 2005, side 27). En kortsvarsoppgave vil bedre måle skriveevne hos elevene, og det kan være en utfordring for elever som har problemer med å formulere svaret. Det å kunne uttrykke seg

skriftlig i skolen er en av de fem grunnleggende ferdighetene, som nevnt i kapittel 3.1. I naturvitenskapelige fag må elevene kunne bruke naturfaglige begreper og uttrykksformer. Mange elever synes at å lære seg og mestre dette kan oppleves som problematisk (Lykknes & Smidt, 2009). Det er viktig å undersøke om elevene fikk vist sin kunnskap gjennom kortsvarsprøven, selv om prøven ble oppfattet som vanskelig av elevene.

Fikk elevene vist sin faglige kompetanse gjennom kortsvarsprøven?

Det ble samlet inn datamateriale fra elevene om i hvilken grad de fikk vist en stor del av sin kompetanse gjennom kortsvarsprøven. Resultatet fra påstanden *Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på kortsvarsprøven* er fremlagt i figur 9.



Figur 9: En oversikt over resultatene fra påstanden "Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på kortsvarsprøven". Oversikten viser også resultatene i forhold til terminkarakterene som elevene fikk høsten 2008 i Biologi 2.

Figur 9 viser hvordan resultatet fra påstanden om at elevene fikk vist en stor del av sin kompetanse på kortsvarsprøven fordeler seg. Det var stor spredning i resultatet, men elevene var jevnt fordelt i forhold til terminkarakterene som de fikk høsten 2008 i Biologi 2. For å observere tendensene tydeligere i datamaterialet, måtte man gå inn i datamaterialet fra det åpne spørsmålet *Fikk du vist din kompetanse fra kapittel 6 gjennom kortsvarsprøven?* I dette resultatet kan en lettere observere elevenes synspunkter. Boks 16 viser resultatet for hva flere elever svarte på påstanden *Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på kortsvarsprøven* og det åpne spørsmålet *Fikk du vist din kompetanse fra kapittel 6 gjennom kortsvarsprøven?*

Boks 16: Utdrag fra svarene på påstanden "Jeg fikk vist en stor del av min kunnskap fra kapittel 6 på kortsvarprøven" og det åpne spørsmålet "Fikk du vist din kompetanse fra kapittel 6 gjennom kortsvarsprøven?".

Påstand 3a iii: *Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på kortsvarsprøven.*
Gunn, terminkarakter 6: Uenig.
Jorunn, terminkarakter 4: Uenig.
Lars, terminkarakter 3: Uenig.
Ine, terminkarakter 5: Litt uenig.
Oda, terminkarakter 5: Litt uenig.
Nina, terminkarakter 5: Litt enig.
Berit, terminkarakter 4: Enig.
Heidi, terminkarakter 5: Enig.
Mona, terminkarakter 5: Enig.

Spørsmål 3b ii: *Fikk du vist din kompetanse fra kapittel 6 gjennom kortsvarsprøven? Begrunn svaret.*
Gunn, terminkarakter 6: Nei, pga dårlig forberedelse. Men kunne gjort det under andre omstendigheter.
Jorunn, terminkarakter 4: Nei, fordi jeg leste på feil kapittel. Hadde jeg lest riktig, ville jeg nok gjort det. Det var gode spørsmål.
Lars, terminkarakter 3: Nei, hadde dårlig konsentrasjonsevne når jeg hadde prøven.
Ine, terminkarakter 5: Fikk vist en del av min kunnskap når vi fikk skrive en hel side.
Oda, terminkarakter 5: Ikke helt, men litt. Det var jo selvfølgelig en del ting det ikke ble spurt om, men sånn er det jo til vanlig også. Jeg skulle ha ganske stor oversikt over kapittelet dersom jeg skulle klare å svare på spørsmålene.
Nina, terminkarakter 5: Tja, både og. Fikk skrevet ned det meste jeg kunne, men det ble litt dårlig tid, så jeg rakk ikke å lese over eller lage noen tegninger.
Berit, terminkarakter 4: Ja, for da kan man skrive det man vil, i stedet for å bare krysse av på et svaralternativ.
Heidi, terminkarakter 5: Fikk vist en del kunnskap. Var noen vanskelige spørsmål som man måtte ha pugget for å kunne besvare.
Mona, terminkarakter 5: Ja, fordi jeg fikk skrevet mye mer i detalj om et tema. Dette hadde jeg ikke mulighet til på flervalgsprøven.

I boks 16 ser man tydelig at de elevene som var *uenig* i påstanden om at de fikk vist en stor del av kompetansen sin på kortsvarsprøven, hadde lest på feil kapittel, forberedt seg dårlig eller hadde dårlig konsentrasjon under prøven. Det kan forklare at resultatet fra kortsvarsprøven gav et dårligere gjennomsnitt enn forventet. De elevene som var *litt enig* eller *enig* med påstanden begrunnet dette synspunktet med at de kunne skrive mer detaljer i prøven og de kunne skrive det de ville under hvert spørsmål i forhold til å krysse av på et svaralternativ som forekom på flervalgsprøven. I intervjuet med Elev A kom det også fram at man kan uttrykke mer kunnskap i en åpen oppgave. Boks 17 viser utdraget, der elev A uttrykte om hun fikk vist sin faglige kompetanse på kortsvarsprøven.

Boks 17: Utdrag fra intervjuet med Elev A, der hun svarer på om hun fikk uttrykt kompetansen sin på kortsvarsprøven.

Intervjuer: *Fikk du uttrykt din kompetanse på kortsvarsprøven?* **Elev A:** Jeg synes at du får vise mer hva du kan på en kortsvarsprøve. På en flervalgsoppgave får du så mange svaralternativer hvor man krysser av på det som er riktig. På en kortsvarsprøve kan man skrive mer, og man har en hel side som gjør at man får uttrykt mye mer av hva man kan.

Den neste delen av kapittel 6 inneholder resultater fra spørsmål om hva elevene synes om eksamensformen fra våren 2008, og hvilke vurderingsform de foretrekker ved underveisvurdering og sluttvurdering.

6.3. Flervalgsoppgaver og kortsvarsoppgaver på eksamen

Det er viktig å undersøke om hva elevene synes om bruk av flervalgsoppgaver som vurderingsform. Det er viktig å få kartlagt hvilken vurderingsmetode, flervalgsoppgaver eler kortsvarsvarsoppgaver, elevene foretrekker ved vurdering *for* læring og vurdering *av* læring. Det var disse elementene som ble tatt opp i den siste delen av spørreundersøkelsen, se vedlegg 3. Boks 18 viser en oversikt over det resultatet fra det åpne spørsmålet *Hva likte du best av flervalgsoppgaver og kortsvarsoppgaver, i forhold til de to ulike metodene?*

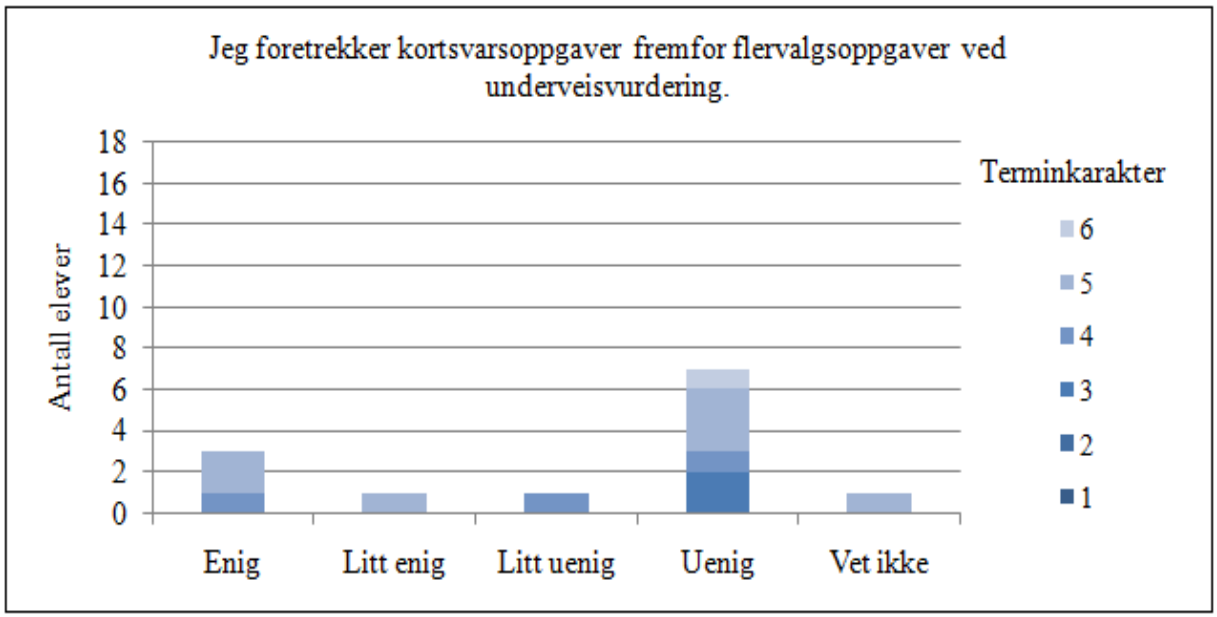
Boks 18: Oversikt over resultatene fra det åpne spørsmålet "Hva likte du best av flervalgsoppgaver og kortsvarsoppgaver, i forhold til de to ulike metodene?".

Spørsmål 4a: *Hva likte du best av flervalgsoppgaver og kortsvarsoppgaver, i forhold til de to ulike metodene? Begrunn svaret.*
Daniel, terminkarakter 3: Flervalgsprøven. Mindre stress, kan sitte og tenke igjennom hvilke svar som er rette. Er du helt tom, kan du tippe, i motsetning til kortsvarsprøven der svaret vil være helt tomt.
Lars, terminkarakter 3: Flervalgsoppgavene**,** slipper å utforme svaret.
Anna, terminkarakter 4: Flervalgsprøven, fordi jeg fikk vist mer kunnskap.
Berit, terminkarakter 4: Jeg likte kortsvarsprøven best, fordi jeg fikk vist kunnskapen min. Grunnen til at jeg likte kortsvarsprøven best, kan være på grunn av at vi er vant til denne prøveformen.
Cecilie, terminkarakter 5: Flervalgsprøven, selv om kortsvar var enklere, fordi jeg er mer vant til dette.
Heidi, terminkarakter 5: Jeg likte best flervalgsprøven. Synes det er enklere å få rett svar.
Ine, terminkarakter 5: Synes flervalgsprøven var best, men begge metodene var helt greie.
Oda, terminkarakter 5: Jeg likte flervalgsprøven best, fordi da måtte jeg ha mye kunnskap og jeg fikk vist det også.
Rita, terminkarakter 5: Flervalgsoppgaver er nok mindre gunstig for meg, men metoden er jo egentlig slik at det blir mer moro å ta prøven. Ikke like mye stress.
Kaia, terminkarakter 5: Kortsvarsoppgaven var best. Flervalgsprøven var enkel å "lese" feil og ikke lese oppgavene godt nok, samtidig som at jeg fikk brukt mine egne ord i kortsvarsprøven.
Mona, terminkarakter 5: Likte vel egentlig begge like godt, men fikk vist min kunnskap mye bedre på kortsvar, så sånn sett likte jeg denne litt bedre.
Nina, terminkarakter 5: Likte kortsvarsprøven best for da fikk jeg skrevet omtrent alt jeg kunne om det temaet. På flervalgsprøven var det så lett å gjøre "dumme" feil, som man egenlig visste svaret på.
Gunn, terminkarakter 6: Likte kortsvarsprøven best, vet ikke hvorfor.

Resultatene fra boks 18 er systematisert i forhold til terminkarakterene som elevene fikk høsten 2008 i Biologi 2. Det var stor spredning i elevgruppen om hvilke av metodene elevene likte best, i forhold til vurderingsformen. Både Lars og Anna foretrakk flervalgsprøven, og begrunnet det med at de fikk vist kompetansen sin bedre, samt at de slapp å utforme svaret. Mona derimot foretakk kortsvarsprøven, siden hun fikk vist kompetansen sin bedre gjennom denne vurderingsformen. Dette synspunktet stemmer overens med teorien til Sirnes (2005) og Jansen (1971), flervalgsoppgaver vil ikke måle skriveevne hos elever, og det kan virke umotiverende for faglig sterke elever, siden de ikke får vist utdypende kunnskap om emnet (Jansen, 1971, side 5; Sirnes, 2005, side 27). Spørsmålet fra boks 18 handlet om hvilke av vurderingsformene elevene likte best, men det er viktig å undersøke om elevene hadde ulike synspunkter i forhold til å benytte flervalgsoppgaver ved underveisvurdering og ved sluttvurdering.

Flervalgsprøver ved underveisvurdering

Boks 18 viser at det var stor spredning i om elevene likte flervalgsoppgaver eller kortsvarsoppgaver best som vurderingsform. Figur 10 viser en oversikt over hva elevene foretrakk ved underveisvurdering. Denne figuren viser at flertallet av elevene var *litt uenig* eller *uenig* i den fremsatte påstanden. Det betyr at de fleste elevene foretrakk flervalgsoppgaver som underveisvurdering.



Figur 10: En oversikt over resultatet fra påstanden "Jeg foretrekker kortsvarsoppgaver fremfor flervalgsoppgaver ved underveisvurdering". Oversikten viser også en fordeling over resultatet i forhold til terminkarakterene som elevene fikk høsten 2008 i Biologi 2.

I intervjuet med læreren ble det også tatt opp bruk av flervalgsoppgaver ved underveisvurdering. Boks 19 viser et utdrag fra dette intervjuet.

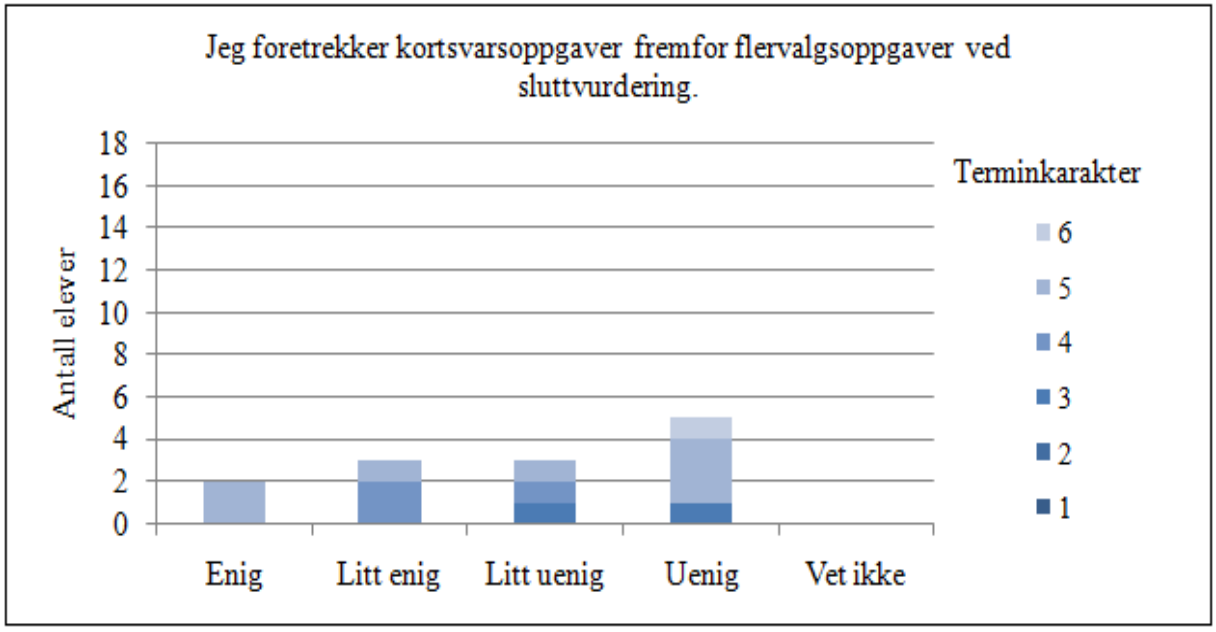
Boks 19: Utdrag fra intervjuet med læreren, der det blir tatt opp hva han synes om å benytte flervalgsoppgaver som underveisvurdering.

Intervjuer: *Hva tenker du om å benytte flervalgsoppgaver ved underveisvurdering?*
Lærer: Ja, flervalgsoppgaver kan fungere ved underveisvurdering. Du kan lage oppgaver som elevene kan prøve seg og ha egenvurdering på, og samtidig få tilbakemelding fra lærer. Det er underveisvurdering som i størst grad kan motivere for videre innsats. Det er en viss fare ved underveisvurdering med karakter. Får elevene dårlig karakter, kan elevene miste motet. Elevene må altså få en vurdering og tilbakemelding med karakter, slik at de vet hva de må arbeide videre for å få bedre kompetanse.

Boks 19 viser viktige elementer ved å benytte flervalgsoppgaver ved underveisvurdering. Det kan motivere elevene til videre innsats i faget, men det er også viktig å gi tilbakemelding til elevene om hva de kan gjøre for å oppnå høyere kompetanse i faget og forbedre kompetansen til sluttvurderingen (Engh et al., 2007).

Flervalgsoppgaver ved sluttvurdering

Figur 11 viser en oversikt over i hvilke grad elevene foretrakk kortsvarsprøver fremfor flervalgsprøver ved sluttvurdering. Ved første øyekast kan det virke som at det er større spredning i disse resultatene enn de i figur 10, som viser oversikten over i hvilken grad elevene foretrakk kortsvarsprøver fremfor flervalgsprøver som underveisvurdering. Dersom man gjør en generalisering av resultatene, vil man få mer nøyansert inntrykk. Man kan tenke seg følgende generalisering for resultatene i figur 10 og figur 11: Elever som har svart at de er *enig* eller *litt enig* i påstanden anses å være positive til den aktuelle oppgaveformen. Likeledes vil elever som har svart at de er *litt uenig* eller *uenig* anes å være negative til den aktuelle oppgaveformen. Da viser figur 10 at fire elever er positive til kortsvarsprøver som underveisvurdering. Åtte elever var negative til kortsvarsprøver som underveisvurdering. I figur 11 var det fem elever som er positive til kortsvarsprøver som sluttvurdering, og fremdeles åtte elever som var negative til dette. Det betyr at det samlet sett ikke er det noen endring i datamaterialet, dersom dette generaliseres.



Figur 11: En oversikt over resultatet fra påstanden "Jeg foretrekker kortsvarsoppgaver fremfor flervalgsoppgaver ved sluttvurdering". Oversikten er vist i forhold til terminkarakterer som elevene fikk høsten 2008 i Biologi 2.

I de åpne spørsmålene fra spørreundersøkelsen ble det spurt om hva elevene synes om at det er flervalgsoppgaver på skriftlig eksamen i Biologi 2. Boks 20 viser en oversikt over noen av synspunktene fra det åpne spørsmålet *Hva synes du om at det er flervalgsoppgaver på eksamen i Biologi 2?*

Boks 20: Et utdrag fra resultatet fra det åpne spørsmålet "Hva synes du om at det er flervalgsoppgaver på eksamen i Biologi 2?".

Spørsmål 2a vi: *Hva syns du om at det er flervalgsoppgaver på eksamen i Biologi 2?*
Daniel, terminkarakter 3: Vil trolig bli vanskelig, når de tar detaljerte spørsmål fra hele boken. Men er likevel greit.
Anna, terminkarakter 4: Greit. Gjør det enklere å stå, men samtidig må jeg da pugge mye mer detaljer enn hvis det ikke var det.
Cecilie, terminkarakter 5: Synes det er unødvendig. Kompetansen en viser i disse oppgavene burde en få vist i de andre skriftlige oppgavene. Synes flervalgsoppgaver handler mer om å "lure" eleven.
Heidi, terminkarakter 5: Jeg synes flervalgsoppgaver kan være bra på eksamen. Man får litt variasjon, og jeg likte denne prøvemetoden godt.
Ine, terminkarakter 5: Synes det er helt greit, siden man da alltid har en mulighet til å svare riktig.
Mona, terminkarakter 5: Både og, ettersom at det har fordeler og ulemper. Samtidig synes jeg stort sett at det er helt greit, ettersom det er mye lettere å få til noe rett på denne måten (vanskeligere å stryke). Samtidig er det kanskje ikke like lett for lærere/sensor å se hva eleven virkelig kan ut ifra disse oppgavene.
Nina, terminkarakter 5: Litt dumt, fordi jeg føler jeg får vist kunnskapen min best ved å ha langsvarsoppgaver som "Grei ut om…".
Rita, terminkarakter 5: Det er greit når det er i kombinasjon med andre prøveformer, men ikke nok til å vise kompetanse i faget.
Gunn, terminkarakter 6: Helt greit, noe annet enn bare skriving for en gangs skyld. Det gjør kanskje at man skjerper seg litt mer.

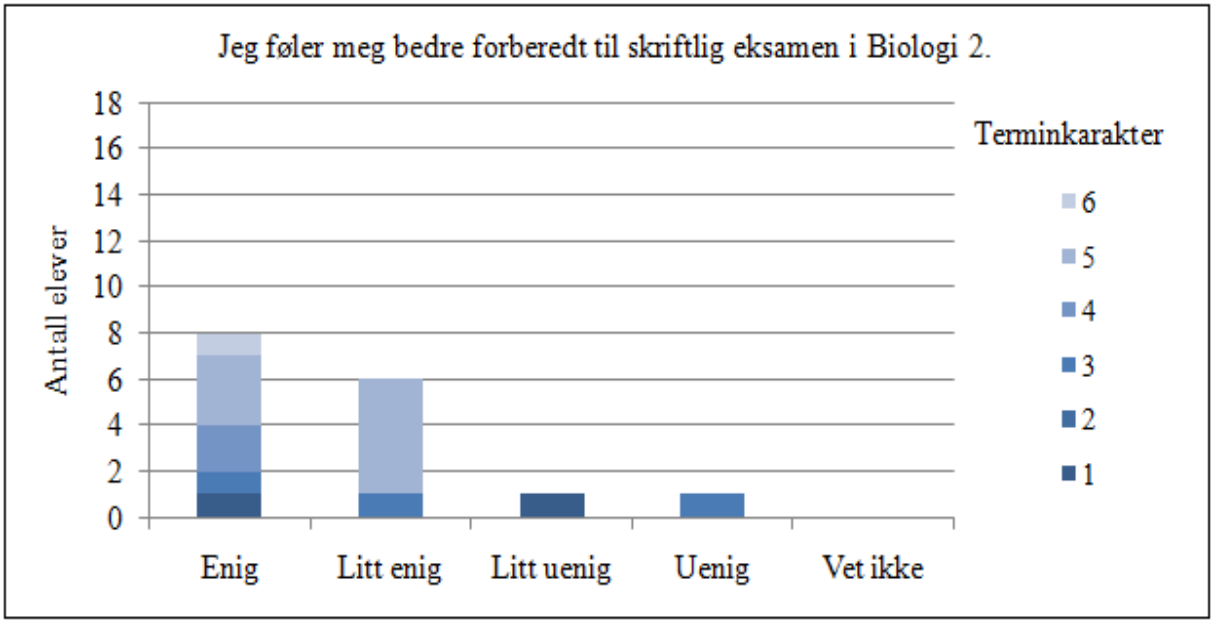
Boks 20 viste at det var mange ulike synspunkter blant elevene. Flere elever synes at det er helt greit med flervalgsoppgaver på eksamen. Blant annet Heidi har begrunnet det med at det skaper variasjon i vurderingsformen, og i følge Utdanningsdirektoratet er variasjon i vurderingsformene er en av årsakene for at flervalgsoppgaver blir benyttet på skriftlig eksamen, se fotnote 1 i kapittel 1.1. Utdanningsdirektoratet mener også at ulike oppgaver vil bidra til at elever med ulike lærestiler vil få vist sin kompetanse, og denne nye inndelingen på oppgavene vil føre til at elevene får vist en bredere og dypere kompetanse, i forhold til pensum. Det kommer også fram i begrunnelsen til Rita i boks 20. Hun mener at flervalgsoppgaver i seg selv ikke er nok til å vise kompetanse i faget, men en kombinasjon med flervalgsoppgaver og andre oppgaver kan føre til hun får vist sin faglige kompetanse i Biologi 2. Det viktigste er at elevene får vist sluttkompetansen sin gjennom skriftlig eksamen, og eksamensformen fra våren 2008 var slik sett en grei måte å få målt sluttkompetansen sin på. Det kom også fram i intervjuet med læreren, boks 21 viser et utdrag fra dette intervjuet.

Boks 21: Utdraget fra intervjuet med læreren, der han forteller om eksamensformen fra våren 2008 og hvordan den vil måle sluttkompetansen til elevene.

Lærer: Jeg må sikre meg at jeg får målt kompetansen til elevene mine i de ulike læreplanmålene, målt sluttkompetansen og satt en karakter. Det viktigste er at elevene føler at de får vist sluttkompetansen sin gjennom alle evalueringsmåtene. Sånn sett er den nye eksamensformen grei for at de får målt sluttkompetansen sin. Den gamle måten er nok ikke noe dårligere. De fikk kartlagt kompetansen der også. Kanskje det er lettere å stå med den nye eksamensformen, men så lenge elevene får kartlagt kompetansen sin, så er jeg fornøyd.

Skriftlig eksamen i Biologi 2

Det ble også undersøkt om elevene etter studien følte seg bedre forberedt til den første delen på den skriftlige eksamen som de skulle ha våren 2009. Elevene hadde som nevnt to prøver, der den ene bestod kun av flervalgsoppgaver og den andre bestod av kortsvarsoppgaver. For å få svar på problemstillingene, se kapittel 1.2, var det viktig at elevene følte at disse to prøvene var relevante for skriftlig eksamen i Biologi 2. Figur 12 viser en oversikt over påstanden *Jeg føler meg bedre forberedt til eksamen i Biologi 2, i forhold til å benytte flervalgsoppgaver som prøveform, enn før jeg hadde flervalgprøven*. Det var et stort flertall av elevene som var *enig* og *litt enig* i denne påstanden. Det betyr at den tiden som ble benyttet for å samle inn datamaterialet har/har hatt en etisk verdi i forhold til forberedelse til skriftlig eksamen i Biologi 2.



Figur 12: Oversikt over resultatet fra påstanden "Jeg føler meg forberedt til skriftlig eksamen i Biologi 2". Oversikten er vist i forhold til terminkarakterene som elevene fikk høsten 2008 i Biologi 2.

I forhold til skriftlig eksamen er det viktig at elevene får vist sin sluttkompetanse, og derfor var det viktig å kartlegge i hvilken av metodene, flervalgsoppgaver eller kortsvarsoppgaver, elevene fikk vist sin faglige kunnskap best. Boks 22 viser et utdrag fra resultatene, der elevene uttrykker i hvilke av vurderingsmetodene de føler at de får vist sin faglige kompetanse best.

Boks 22: Et utdrag fra resultatet fra det åpne spørsmålet "I forhold til eksamen i Biologi 2, i hvilken av flervalgsoppgaver eller kortsvarsoppgaver føler du at du får vist din kompetanse best?".

Spørsmål 4b: *I forhold til eksamen i Biologi 2, i hvilken av flervalgsoppgaver eller kortsvarsoppgaver føler du at du får vist din kompetanse best? Begrunn svaret.*
Lars, terminkarakter 3: Flervalgsoppgave.
Daniel, terminkarakter 3: Kortsvarsprøven, kan ikke tippe. Må vise at du kan stoffet.
Berit, terminkarakter 4: Kortsvarsprøven. Da kan du skrive det du kan, istedenfor å krysse av på et svaralternativ.
Jorunn, terminkarakter 4: Kortsvarsprøven, for der skriver man jo tross alt ned det man kan og ikke bare krysser av et svaralternativ i hytt og pine.
Cecilie, terminkarakter 5: Liker flervalgsprøven best, fordi jeg er mest vant til det, og jeg har innøvd hvordan jeg skal vise mitt beste ved dette.
Heidi, terminkarakter 5: Får kanskje vist mest kunnskap via kortsvar, siden man kan dra inn annet stoff. Dette kan man ikke med en flervalgsprøve.
Ine, terminkarakter 5: Tror jeg egentlig får vist min kunnskap best gjennom en kortsvarsprøve.
Kaia, terminkarakter 5: Kortsvarsoppgaven, fordi der er det jeg selv som må skrive det jeg mener er mest relevant og riktig i forhold til spørsmålet. Samt på flervalgsprøven kan man gjette og fortsatt få full score.
Mona, terminkarakter 5: Kortsvarsprøven, fordi du får skrevet mye mer i detalj på hvert tema.
Nina, terminkarakter 5: Jeg likte kortsvarsprøven best, for der fikk jeg vist kunnskapen min bedre enn på flervalgsprøver.
Oda, terminkarakter 5: Flervalgsoppgaver, fordi da er det ikke slik at du må være så presis når du skriver. Man må bare vite det riktige svaret og krysse av på det.
Rita, terminkarakter 5: Kortsvarsprøven, fordi jeg som regel kan en del detaljer rundt hvert tema som er på oppgavene.
Gunn, terminkarakter 6: Det er vel bare å vente og se egentlig, men kortsvarsprøven kanskje? Kanskje jeg kan stoffet godt nok, så klarer jeg vel å svare godt på den.

Boks 22 viste at flertallet av elevene følte at de fikk vist sin kompetanse best ved bruk av kortsvarsoppgaver, siden elevene der fikk skrevet ned flere detaljer rundt temaene som de ble gitt i oppgaveteksten. Som nevnt i kapittel 5.3. bestod denne klassen av mange faglig sterke elever, og mange av dem vil sannsynligvis ikke ha problemer med å formulere svar på en kortsvarsoppgave.

I intervjuet med læreren ble det spurt om hva han synes om at det er flervalgsoppgaver på skriftlig eksamen i Biologi 2. Boks 23 viser et utdrag fra intervjuet med læreren, der det kommer fram hva han synes om at det er flervalgsoppgaver på skriftlig eksamen i Biologi 2.

Boks 23: Utdrag fra intervjuet med læreren om hva han synes om at det er flervalgsoppgaver på skriftlig eksamen i Biologi 2.

Intervjuer: *Hva synes du om at det er flervalgsoppgaver på skriftlig eksamen i Biologi 2?*
Lærer: Det synes jeg er helt greit. Hovedoppgaven til vurdering i et fag, i forhold til læreplanmål er for å se hvor stor kompetanse elevene har. Desto mer kompetanse man klarer å få kartlagt, jo bedre grunnlag har man for å sette karakter.

Intervjuer: *Er flervalgsoppgaver en god metode for å måle kompetansemål?*
Lærer: Ja, sammen med andre metoder. Jeg får litt problemer med flervalgsoppgaver i forhold til Blooms taksonomi. Mye er avhengig av utformingen av spørsmålet. Det stiller store krav til utforming hvis du skal klare å få fram Blooms taksonomi. Jeg ville ha brukt flervalgsoppgaver for å måle de laveste trinnene i Blooms taksonomi. Det er krevende å få frem de øverste trinnene. Jeg mener at hvis du gir elevene åpne oppgaver, ser man hvordan de formulerer seg og hvilke vurderinger de gjør. På flervalgsoppgaver kan de krysse rett, men du vet ikke om de har vurdert rett. Ved åpne oppgaver står vurderingen på papiret, og derfor mener jeg at flervalgsoppgaver er best for de laveste trinnene i Blooms taksonomi.

I boks 23 kommer det fram at læreren mente at flervalgsoppgaver måler kompetansen til elevene på de laveste nivåene i Blooms taksonomi. Flervalgsoppgaver kan fort bli en test som spør om rene faktaspørsmål og detaljkunnskap, og det er kjennetegn for de laveste nivåene i Blooms taksonomi, se kapittel 4.1. Flervalgsoppgaver kan også utfordre elevene på et høyere nivå enn gjengivelsesnivået (Haladyna, 2004, side 137). Ved bruk av regneoppgaver og tabeller, må elevene forstå de ulike elementene i fremgangsmåten for å vite det korrekte svaralternativet. Tabell 1 i kapittel 6.1 viser en oversikt over hvor i Blooms taksonomi de ulike flervalgsoppgavene på prøven ble plassert. Flere av oppgavene var på nivå 1 *gjengivelse* og nivå 2 *forståelse*, men flere oppgaver var også på nivå 3 *anvendelse* og nivå 4 *analyse*. Det gjelder blant annet for oppgave 6 og 24 i flervalgsprøven som ligger i vedlegg 6. I disse to oppgavene måtte elevene forstå figuren og tabellen for bryte kunnskapen ned i enkeltdeler for å kunne se sammenhengen mellom de ulike elementene. Dette var nødvendig for å kunne besvare oppgaven.

6.4. Bruk av Test 2.0 for å forberede elevene til skriftlig eksamen

I denne siste delen av kapittelet om datainnsamlingen og resultatet presenters det elementer i forhold til den andre problemstillingen: Hvordan kan en lærer forberede elevene sine på den første delen av skriftlig eksamen i Biologi 2, ved å benytte *læringsverktøyet Test 2.0* i læringsplattformen *it's learning* i undervisningen? Datainnsamlingen foregikk ved intervju av lærer, gjennomført fredag 23.01.09. De formelle spørsmålene som var planlagt før intervjuet, ligger i vedlegg 5.

Under intervjuet kom det fram flere synspunkter som læreren hadde rundt bruk av Test 2.0 i it's learning. Boks 24 viser utdrag fra dette intervjuet.

Boks 24: Utdrag fra intervju med læreren, der han forteller om nytteverdi og teknisk bruk av Test 2.0.

Intervjuer: *Hva slags nytteverdi ser du ved bruk av flervalgsoppgaver på it's learning?*
Lærer: Det er den læringsplattformen vi bruker, og den er lett tilgjengelig for elevene. De vil få respons veldig fort, de får den der og da. Det er veldig viktig for å motivere dem videre i faget. Det er også lett å legge inn oppgavene på læringsplattformen, og det er lett å variere oppgaveformen. Du kan kombinere flervalgsoppgaver på andre måter, med bilder og spørsmål på figurer. Kanskje mer variasjon få bedre vil få fram kompetansen dem har enn bare med flervalgsoppgaver.

Intervjuer: *Hvordan er den tekniske bruken av Test 2.0?*
Lærer: Den synes jeg helt grei.

Intervjuer: *Foretrekker du papir eller læringsplattformer?*
Lærer: Jeg foretrekker digital.

Intervjuer: *Hvorfor det?*
Lærer: Fordi utformingen tar like lang om det er papir eller digital, men responsen kommer mye fortere, og den er det viktigste. Det er ikke problematisk å rette for lærer, men det er viktig med respons til elever. Elevene kan også gjenbruke testen. De kan kjøre testen og gang, og de kan trykke på knappen og kjøre testen en gang til like etterpå. Eller jeg kan gå igjennom ting som jeg ser er vanskelig for dem, før jeg lar dem ta testen en gang til. Det er også mulig å ta ut en avgrenset del av en test, la elevene ta denne delen og la det være en kartleggingsprøve for å legge opp undervisningen etter resultatet og å lage undervisning for å øke kompetansen til elevene.

Fordelene som kom fram i boks 24, nevnes også av Engh et al., (2007) og Horgen (2007, 2009) ved underveisvurdering. En slik vurdering skal gi elevene en pekepinn på hva de trenger å arbeide mer med, og hovedpoenget er at elevene skal vite hva de trenger å gjøre videre for at sluttresultatet skal bli bra, og oppnå en større grad av måloppnåelse (Engh et al., 2007, side 28). En digital test har en rekke muligheter for å støtte opp under vurdering *for*

læring og læringsprosessene til elevene. Slike tester vil skape refleksjon, repetisjon og motivasjon for elevene (Horgen, 2007, 2009).

Flervalgsoppgaver er en av vurderingsformene som blir gitt på skriftlig eksamen i Biologi 2, og det er viktig at lærere gir slike oppgaver underveis i faget for å forberede elevene på denne vurderingsformen. Test 2.0 kan være en god metode for å forberede elevene og den har mange fordeler som både læreren nevnte i boks 24 og som Horgen (2007, 2009) nevner. Test 2.0 kan tilby elevene en ny lærings- og vurderingsform, og den kan utgjøre et viktig supplement i lærerens pedagogiske verktøykasse.

7. Diskusjon

Denne siste hoveddelen av oppgaven inneholder diskusjon av de kvantitative og kvalitative resultatene som ble fremlagt i kapittel 6 for å besvare problemstillingene fra kapittel 1.2. Resultatene støttes opp av teori fra kapittel 2, 3 og 4. Det fremlegges også erfaringer fra datainnsamlingen og metodevalg. Det diskuteres hvordan reliabiliteten og validiteten ble ivaretatt i studien, og hva slags endringer som måtte ha blitt utført for å øke reliabiliteten og validiteten.

Kapittelet vil bli inndelt i følgende punkter:

- Oppsummering av resultatet
- Fikk elevene vist sin faglige kompetanse ved bruk av flervalgsoppgaver som vurderingsform?
- Hvordan fungerte testverktøyet Test 2.0 i it's learning til å forberede elevene til skriftlig eksamen i Biologi 2?
- Erfaringer fra datainnsamlingen
    - Flervalgsprøven
    - Kortsvarsprøven
- Erfaringer fra metodevalg
- Kritisk oppsummering av studien
    - Avslutning

7.1. Oppsummering av resultatet

Denne masteroppgaven er en kvalitativ studie av en lærer som underviste i Biologi 2, og hans elever i klassen. Studien har tatt utgangspunkt i skriftlig eksamen i Biologi 2 fra våren 2008, og den eksamenen inkluderte flervalgsoppgaver som vurderingsform. Det ble benyttet både

kvantitative og kvalitative metoder som spørreundersøkelse med avkrysningspåstander og åpne spørsmål, samt intervju av elever og lærer for å få svar på to problemstillinger som ble beskrevet i kapittel 1.2.

Den første problemstillingen, se kapittel 1.2, handlet om elevene fikk i tilstrekkelig grad vist sin faglige kompetanse i Biologi 2 ved å benytte flervalgsoppgaver som vurderingsform. Ut fra datamaterialet kom det fram at de fleste elevene i klassen fikk vist den faglige kompetansen sin fra kompetansemålet *Genetikk* fra LK06 ved å benytte flervalgsoppgaver som vurderingsform. Det kom også fram i de åpne spørsmålene i spørreundersøkelsen at elevene fikk vist kompetansen bedre ved bruk av åpne oppgaver. For læreren sin del, så mente han at det viktigste for ham var at han fikk målt den faglige kompetansen til elevene og satt en rettferdig karakter. Han hevdet også at eksamensformen som inkluderer flervalgsoppgaver er en grei metode for at elvene skal få målt sluttkompetansen sin.

Den andre problemstillingen, se kapittel 1.2, handlet om hvordan en lærer kan forberede elevene sine på den første delen av skriftlig eksamen i Biologi 2, ved å benytte *testverktøyet Test 2.0* i læringsplattformen *it's learning*. Læreren hadde ikke benyttet Test 2.0 før denne studien, og han var positiv til dette læringsverktøyet etter flervalgsprøven. Han så mange muligheter for å benytte dette verktøyet ved vurdering *for* læring. For det første var det enkelt å legge inn oppgavene i læringsplattformen, og det var mulig å kombinere flervalgsoppgaver med andre oppgaver i Test 2.0. En slik kombinasjon kan gi variasjon i undervisningen, og det kan øke den faglige kompetansen til elevene. For det andre var det mulig å gi automatisk tilbakemelding til elevene etter innlevering av testen som kan inneholde kommentarer, hint eller ledetråder. En automatisk tilbakemelding fra Test 2.0 kan gi en pekepinn på elevenes måloppnåelse, og det kan være lettere å veilede elevene. Det siste hovedpunktet læreren tok opp var at det er mulig å lage kartleggingstester *før* undervisningen for å legge opp undervisningen etter resultatet fra testen.

7.2. Fikk elevene vist sin faglige kompetanse ved bruk av flervalgsoppgaver som vurderingsform?

Den første problemstillingen, som ble beskrevet i kapittel 1.2, var om elever fikk i tilstrekkelig grad vist sin faglige kompetanse ved å benytte flervalgsoppgaver som vurderingsform. Gjennom datainnsamling fra spørreundersøkelsene og intervju med fire

elever var det generelle inntrykket fra elevene at de fikk vist den faglige kompetansen sin ved en slik vurderingsform. I resultatet fra påstanden *Jeg fikk vist en stor del av min kunnskap på flervalgsprøven* fra spørreundersøkelsen i vedlegg 3, som ble grafisk fremstilt i figur 7 i kapittel 6.1, kom det fram at flertallet av elevene var *enig* eller *litt enig* til denne påstanden. Dette kvantitative resultatet ble koblet opp mot besvarelsene til elevene fra det åpne spørsmålet *Fikk du vist din kunnskap fra kapittel 6 gjennom flervalgsprøven?* i spørreundersøkelsen. Disse besvarelsene ble fremstilt i boks 12, og det kom fram at mange av elevene fikk vist sin faglige kompetanse gjennom flervalgsprøven, men flere av dem forklarte videre at de fikk vist mer kompetanse gjennom åpne oppgaver. Den ene eleven, Rita, forklarte i boks 12 at hun fikk vist en god del av kompetansen sin, siden de fleste av oppgavene ble korrekte, men det var mange detaljer som hun ikke fikk vist. Rita fikk 5 som terminkarakter i Biologi 2 høsten 2008, og hun vil sannsynligvis ha lettere for å beskrive flere detaljer fra pensum i en åpen oppgave enn en faglig svak elev.

Flertallet av elevene i utvalget var faglig sterke, og 56 % av elevene fikk terminkarakter 5 eller 6 i Biologi 2 høsten 2008. Prosenttallet ble beregnet ut fra figur 1 i kapittel 5.3. Flervalgsoppgaver vil ikke måle skriveevnen hos elever, og for enkelte elever kan det å kunne formulere seg skriftlig i naturfag med faglige begreper og uttrykksformer være problematisk (Lykknes & Smidt, 2009; Sirnes, 2005, side 27). For faglig sterke elever kan slike begreper og uttrykksformer være en mindre utfordring å mestre enn for faglig svake elever. Siden flertallet av elevene i utvalget var faglig sterke, kan det være en forklaring på at det generelle inntrykket av elevene var at de ville vise mer faglig kompetanse i åpne oppgaver. I en åpen oppgave kan elevene ta med detaljer og figurer som er relevant for oppgaven, det er ikke mulig i en flervalgsoppgave. En slik ulempe ble nevnt av flere elever, se boks 22, der elevene skulle uttrykke seg om ved hvilken vurderingsform de føler at de får vist kompetanse sin best. Blant annet Kaia, Mona og Rita foretrakk kortsvarsoppgaver. Disse tre elevene fikk terminkarakter 5 i Biologi 2 høsten 2008. De kan sannsynligvis formulere seg bedre i en åpen oppgave enn en faglig svak elev.

Det var to faglig svake elever i utvalget, Frida og Petra, som begge fikk terminkarakter 1 høsten 2008 i Biologi 2. Ut fra det kvantitative og kvalitative datamaterialet, som ble fremlagt i kapittel 6, var det generelle inntrykket at disse to elevene var fornøyde med bruk av flervalgsoppgaver som vurderingsform. De benyttet gjettet i kombinasjon med eliminasjonsmetoden på flere av oppgavene, og det kom fram i boks 9 og 11. Både Frida og Petra benyttet eliminasjonsmetoden for å få færre svaralternativer å gjette mellom på

flervalgsoppgavene. I figur 7, som ble fremlagt i kapittel 6.1, kom det fram at begge disse elevene var *enige* i at de fikk vist den faglige kompetansen sin på flervalgsprøven, og i boks 12 kom det fram fra Petra at det var fordi det gikk bedre enn ved vanlige prøver. Det blir antatt at *vanlige prøver* bestod av åpne oppgaver. Lars, som fikk terminkarakter 3 i Biologi 2 høsten 2008, forklarte i boks 12 at siden han slapp å formulere svaret sitt på flervalgsprøven, klarte han å vise sin kompetanse bedre på denne prøven. I følge Sirnes (2005) vil en slik fordel som Lars beskrev, være viktig for flere elever. Elevenes tid blir benyttet til tenking, og ikke skriving for å formulere svarene i forhold til en åpen oppgave (Sirnes, 2005, side 10). Det kan være en viktig faktor som gjør at faglige svake elever foretrekker flervalgsoppgaver som vurderingsform. I Lykknes og Smidt (2009) kom det også fram at enkelte elever kan ha problemer med å uttrykke seg gjennom faglige begreper og uttrykksformer i naturvitenskapelige fag, og det kan være en annen faktor som kan gjøre at faglig svake elever kan foretrekke flervalgsoppgaver som vurderingsform.

Det ble også samlet inn kvalitativt og kvantitativt datamateriale om elevenes synspunkter om at det er flervalgsoppgaver på skriftlig eksamen i Biologi 2. Det var stor spredning i resultatet fra det åpne spørsmålet *Hva synes du om at det er flervalgsoppgaver på eksamen i Biologi 2?*, se boks 20. Flertallet av elevene syntes at det var helt greit med slike oppgaver på skriftlig eksamen, men det var en stor variasjon på begrunnelser fra elevene. Daniel mente at oppgavene ville bli vanskeligere enn på flervalgsprøven, siden oppgaven vil være fra hele læreplanen, og ikke bare fra kompetansemålet *Genetikk* fra LK06. Ina synes at det var helt greit, siden man uansett vil ha en mulighet til å svare korrekt på oppgavene. Dette nevnte også Mona. Heidi mente at flervalgsoppgaver vil skape variasjon på en skriftlig eksamen, og en slik variasjon er en av årsakene til at flervalgsoppgaver blir benyttet på skriftlig eksamen i Biologi 2, ifølge Utdanningsdirektoratet, se fotnote 1 i kapittel 1.1. Utdanningsdirektoratet mener at ulike oppgaveformer vil bidra til at ulike elever med ulik læringsstil vil få vist sin kompetanse, og denne nye inndelingen av oppgavene, med flervalgsoppgaver, kortsvarsoppgaver og langsvarsoppgaver vil føre til at elevene får vist en bredere og dypere kompetanse i forhold til pensum, se fotnote 1 i kapittel 1.1. Det kom også fram av Ritas svar, se boks 20. Hun mente at flervalgsoppgaver er helt greit som vurderingsform på skriftlig eksamen, men at slike oppgaver alene ikke vil være nok til å vise sluttkompetansen i faget. Dette synspunktet kom også fram i intervjuet med læreren. Han nevnte at jo mer kompetanse man klarer å få kartlagt, desto bedre grunnlag har man for å sette en rettferdig karakter til elevene på skriftlig eksamen. Flervalgsoppgaver kan være en god vurderingsmetode for å

kartlegge kunnskap på en ny måte, og eksamensformen fra våren 2008 var en grei måte å få målt sluttkompetansen til elevene, men den tradisjonelle formen for skriftlig eksamen *ikke* nødvendigvis var noe dårligere. Flervalgsoppgaver må ikke betraktes som en erstatning for, men som et verdifullt supplement til andre vurderingsformer som kan benyttes ved vurdering *for* læring og vurdering *av* læring.

Flervalgsoppgaver er blitt en av vurderingsformene på avsluttende skriftlig eksamen i Biologi 2, og det er en stor utfordring for lærere å forberede elevene til denne eksamensformen. Lærere må benytte flervalgsoppgaver aktivt i undervisningen for å gi elevene erfaring med denne vurderingsformen, og lærere må kunne produsere gode flervalgsoppgaver med gode stammer, nøkler og distraktorer. Det å produsere gode flervalgsoppgaver krever mye tid, og dersom oppgavene man produserer er dårlige, må man endre stammen og/eller distraktorene for å gjøre dem mer utfordrende for elevene ved senere bruk. Som Haladyna (2004) og Sirnes (2005) nevner, vil vanskelighetsgraden i flervalgsoppgaver øke dersom svaralternativene ligner på hverandre. Vanskelighetsgraden vil også øke dersom de har samme presisjonsnivå og tilnærmet lik syntaktisk struktur og lengde. Det er for ikke å gi hint til elevene på hva som er nøkkelen av svaralternativene i flervalgsoppgaven (Haladyna, 2004, side 68-69; Sirnes, 2005, side 33-35). Det kan tid før lærerne får bygget opp en spørsmålsbank som inneholder flervalgsoppgaver som utfordrer elevenes faglige kompetanse. Ved å legge inn flervalgsoppgavene i Test 2.0 etter hvert som man produserer dem, vil lærerne lett få en oversikt på hvordan oppgavene fungerte i forhold til hverandre etter at de har blitt gitt til elevene. Lærerne vil også få en statistisk oversikt over resultatene, og det kan da være lettere å luke ut flervalgsoppgaver med dårlige stammer og/eller distraktorer (Horgen, 2009). Testverktøyet Test 2.0 kan dermed være et bra hjelpemiddel for både lærere og elever for å forberede elevene til skriftlig eksamen i Biologi 2.

7.3. Hvordan fungerte testverktøyet Test 2.0 i it's learning til å forberede elevene til skriftlig eksamen i Biologi 2?

Den andre problemstillingen, se kapittel 1.2, handlet om hvordan en lærer kan forberede elevene sine på vurderingsformen flervalgsoppgaver ved å benytte testverktøyet Test 2.0 i læringsplattformen it's learning. Å benytte Test 2.0 var en ny erfaring for læreren, og han var veldig positiv til dette læringsverktøyet etter flervalgsprøven. Han synes det var veldig lett å

legge oppgavene inn i Test 2.0 og at verktøyet hadde mange fordeler i forhold til vurdering *for* læring.

Det kom fram mange fordeler for læreren ved bruk av Test 2.0 i intervjuet, og et utdrag fra dette intervjuet ble fremlagt i boks 24. En av fordelene læreren nevnte, var at Test 2.0 inneholdt mange variasjonsmuligheter. Å kombinere de generelle og interaktive oppgaveformene i Test 2.0, som er vist i figur 1, vil gi variasjon i undervisningen som både kan øke motivasjonen og den faglige kompetansen til elevene. En slik variasjon ble nevnt i kapittel 3.3, der Strømme (2008) også nevner at det er mulig å inkludere spørsmål der man leser inn en tekst, spiller av en lydfil eller legger ved en tekst eller et bilde (Strømme, 2008, side 159-160). Læreren hevdet også at Test 2.0 vil gi elevene mulighet til å få en umiddelbar tilbakemelding etter innlevering. En slik tilbakemelding kan være et godt hjelpemiddel for å veilede eleven, siden tilbakemeldingene kan peke på elevenes styrker og svakheter i faget. En slik tilbakemelding kan gi elevene beskjed om hva de trenger å arbeide mer med for at de skal oppnå en større grad av måloppnåelse i faget (Engh et al., 2007, side 28). En siste fordel som læreren tok opp var muligheten for å lage kartleggingstester *før* undervisningen for å legge opp undervisningen etter resultatet fra testen. Det kan øke fokuset hos elevene før selve læringsøkten og motivere til økt engasjement hos elevene (Horgen, 2007, 2009). Horgen (2007, 2009) nevner også at en test gitt *underveis* i en læringsøkt, kan sette fokus på de viktige temaene, skape refleksjon og gi økt bevissthet rundt elevenes læring. En test *etter* undervisningen vil kunne oppsummere temaene, slik at elevene får bearbeidet dem. Læreren vil da også få en tilbakemelding på om elevene har arbeidet tilstrekkelig med lærestoffet (Horgen, 2007, 2009).

Under intervjuet ble læreren spurt om hvordan han ønsket å benytte seg av Test 2.0 etter datainnsamlingen. Boks 25 viser et utdrag fra dette intervjuet.

Boks 25: Utdrag fra intervjuet med læreren, der han uttaler seg om hvordan han ønsker å benytte seg av Test 2.0 etter datainnsamlingen.

Intervjuer: *Hvordan vil du benytte Test 2.0 videre i undervisningen?*
Lærer: Jeg vil ta det mer og mer i bruk. Jeg ser mange muligheter ved underveisvurdering, og der vil jeg vurdere det mest til kartlegging, motivering og variasjon for elevene. Jeg har brukt slike oppgaver ganske lenge, men jeg har brukt dem som finnes på nett. De er mange læreverk i biologi og naturfag på nett med flervalgsoppgaver med tre svaralternativer, ikke fem.

I boks 25 kom det fram at læreren vil benytte seg av Test 2.0 videre i undervisningen. Han så mange muligheter i læringsverktøyet ved vurdering *for* læring, og det kan brukes ved kartlegging av elevenes faglige kompetanse, variasjon og motivasjon for elevene. Dette er viktige aspekter ved digitale tester som også ble nevnt av Horgen (2007, 2009) i kapittel 3.3.

7.4. Erfaringer fra datainnsamlingen

I dette avsnittet diskuteres sterke og svake elementer ved flervalgsprøven og kortsvarsprøven som kan ha påvirket analysen av datamaterialet. Det diskuteres også reliabilitet og validitet i oppsettet og prøveresultatet fra flervalgsprøven og kortsvarsprøven.

Flervalgsprøven

I kapittel 2.3 ble det nevnt at en av de store utfordringene ved å produsere en flervalgsprøve er det å lage oppgavene. Det å produsere gode oppgaver med en god stamme og gode distraktorer kan være krevende for lærere. I tillegg har oppgavene ofte en tendens å bli veldig detaljorienterte og kan inneholde rene faktaspørsmål, og det vil kun kartlegge kunnskapen til elevene på de lave nivåene i Blooms taksonomi: nivå 1 *gjengivelse* og nivå 2 *forståelse* (Sirnes, 2005, side 33-35).

Flervalgsprøven, som ligger i vedlegg 6, inneholdt 24 oppgaver som testet elevene på ulike nivåer i Blooms taksonomi. Det ble foretatt en analyse av flervalgsprøven om i hvilken grad elevene kom fram til det korrekte svaret på grunn av høy faglig kompetanse eller om de gjettet og/eller benyttet eliminasjonsmetoden. Dersom flere elever var usikre på svaret, men klarte å eliminere bort flere distraktorer i flervalgsoppgavene, kan det indikere at distraktorene var for dårlige. Dersom flere elever gjettet uten å eliminere bort distraktorer, og likevel klarte å få korrekt svar, kan det indikere at stammen var for enkel eller at det var mange faglige sterke elever som gjennomførte prøven.

Ut fra tabell 2 i kapittel 6.1 ser man at 42 % av oppgavene hadde en P-verdi mellom 80 og 100 %. I kapittel 2.4 ble det nevnt at dersom P-verdien var over 90 %, var flervalgsoppgaven for lett. I denne studien ble det bestemt at grensen skulle være 80 %, siden det kun var 17 elever som gjennomførte flervalgsprøven. Dersom det var 16 elever som svarte korrekt, var P-verdien på 94 % og dermed 15 elever som svarte korrekt, var P-verdien på 88 %. Dersom 14 elever svarte korrekt, hadde flervalgsoppgaven en P-verdi på 82 %, og det var forholdsvis mange elever i forhold til av total deltagelse på flervalgsprøven. Derfor ble P-verdien ansett som høy dersom den oversteg 80 %.

Det kan være flere årsaker til at P-verdien var høy på flere av flervalgsoppgavene. Hovedårsaken var sannsynligvis en kombinasjon av at stammen var for enkel og at distraktorene var for dårlige. Mange av elevene gjettet på en eller flere av flervalgsoppgavene, og det kom fram i figur 5 og 6 som viste resultatet for påstandene *Jeg trodde alle oppgavene var korrekte før jeg leverte* og *Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer* i spørreundersøkelsen. Det at elevene gjettet og eliminerte en eller flere av alternativene, gav en pekepinn på at stammen var enkel og/eller distraktorene var for dårlige. En annen årsak til at P-verdien var høy på flere av flervalgsoppgavene var at flertallet av elevene var faglig sterke, der 56 % av elevene fikk terminkarakter 5 eller 6 høsten 2008 i Biologi 2, se kapittel 5.3. Det var en kvalitetssjekk på flervalgsprøven før den ble gitt til elevene. Både veileder og lærer godkjent oppgavene, og flere av oppgavene var også laget med utgangspunkt i eksempeloppgavene som er tilgjengelig på www.naturfag.no[12]. Disse oppgavene er laget for elever som tar Biologi 1 og Biologi 2, og det var med utgangspunkt i disse oppgavene, samt pensumboka til elevene (Sletbakk et al., 2008) man definerte vanskelighetsgraden på flervalgsoppgavene. For å hindre høy P-verdi kunne det ha blitt utført en *pilottest* på et annet utvalg elever før flervalgsprøven ble gitt i denne studien. I følge Robsen (2002) er en pilottest er en test av metoden som skal benyttes for å samle inn datamaterialet. Den kan peke på elementer som kan virke manglende eller misforstående for elevene (Robsen, 2002, side 185). Ved å benytte pilottest før flervalgsprøven, kunne man ha forutsett dårlige stammer og distraktorer, slik at færre flervalgoppgaver fikk en P-verdi mellom 80 og 100 %.

I et utdrag fra intervjuet med læreren, som ble fremlagt i boks 23 i kapittel 6.3, kom det fram at læreren synes at flervalgsoppgaver kun måler kompetansen til elevene på de laveste nivåene i Blooms taksonomi. Sirnes (2005) nevnte at flervalgsoppgaver fort kan ha en tendens til å spørre om rene faktaspørsmål og detaljkunnskaper, og det er kjennetegn for de laveste nivåene i Blooms taksonomi. Det er mulig å utfordre elevene på høyere nivåer i Blooms taksonomi ved å bruke figurer, tabeller eller regneoppgaver i flervalgsoppgavene, der elevene må forstå og vite fremgangsmåten i oppgaven for å finne det korrekte svaralternativet (Haladyna, 2004, side 137).

Tabell 1 i kapittel 6.1 viser en oversikt over hvor i Blooms taksonomi de ulike flervalgsoppgavene som ble gitt på flervalgsprøven lå. Flere oppgaver utfordret elevene lavt

[12] Oppgavene ble lastet ned 03.01.09 fra http://www.naturfag.no/binfil/download.php?did=3884 og http://www.naturfag.no/binfil/download.php?did=3886.

på Blooms taksonomi, blant annet oppgave 1 og 22, se vedlegg 3. I disse to oppgavene må elevene kunne gjengi spesifikke fakta i den samme formen som kunnskapen er undervist, og dette er kjennetegn ved nivå 1 *gjengivelse* i Blooms taksonomi, se kapittel 4.1 (Jansen, 1971, side 16; Sirnes, 2005, side 23). De fleste oppgavene på flervalgsprøven var på nivå 2 *forståelse*, som for eksempel oppgave 15, se vedlegg 3. I denne oppgaven måtte elevene kunne forstå likhetene og ulikhetene mellom mitose og meiose, og de måtte kunne identifisere og forstå de ulike elementene ved celledeling. Dette er kjennetegnet på underkategorien *tolkning* i nivå 2 *forståelse* som ble beskrevet i kapittel 4.1. Oppgave 7, 9 og 19 i vedlegg 3 ble kategorisert i nivå 3 *anvendelse* i tabell 1. I disse tre oppgavene måtte elevene kunne omstrukturere det innlærte og bruke det i en ny situasjon. Figurene i oppgave 7 og 9 var ikke identiske figurer fra pensumboka, og en slik omstrukturering av kunnskapen er kjennetegn for nivå 3 *anvendelse* i Blooms taksonomi (Jansen, 1971, side 19; Sirnes, 2005, side 24). Det var også tre oppgaver som utfordret elevene på nivå 4 *analyse* i Blooms taksonomi; oppgave 6, 16 og 24. I oppgave 6 og 24 måtte elevene kunne bryte kunnskapen ned i enkeltdeler og trekke slutninger for å kunne forstå figuren i oppgave 6 og tabellen i oppgave 24. Oppgave 16 var en flervalgsoppgave, der elevene måtte velge *det mest korrekte* svaralternativet. Alle svaralternativene hadde korrekte elementer i seg, men det var kun *ett* av dem som skilte seg ut. Dette er en måte å gjøre flervalgsoppgaver mer utfordrene for elevene. En slik oppgave kategoriseres i nivå 4 *analyse*, siden elevene må kunne trekke slutninger og analysere helheten i svaralternativene.

Det var også en oppgave i flervalgsprøven som viste seg å være for vanskelig for elevene, oppgave 7 i vedlegg 6, som handlet om hvilke av basesekvensene som ved normal replikasjon kunne kobles til den *klebrige enden* på DNA-figuren. Den har blitt kategorisert som nivå 3 *anvendelse* i tabell 1, se kapittel 6.1, og det var kun tre elever som besvarte oppgaven korrekt. Det gav oppgaven en P-verdi på 18 %. Dersom denne oppgaven hadde blitt utført en gang til, måtte man ha sett på stammen og svaralternativene for å gjøre oppgaven litt enklere. I oppgaven ble det oppgitt hvilken retning som var 3'-enden og 5'-enden, og begge disse basesekvensene fra 3'-enden og 5'-enden var med i svaralternativene. 3'-enden og 5'-enden på figuren var lett å overse for elevene, og de bommet dermed på svaralternativet. De to endene burde ha blitt fremhevet i stammen på en annen måte.

Reliabilitet i flervalgsoppgaven

I kapittel 2.5 er det lagt frem teori om reliabilitet i flervalgsoppgaver. Kleve (1994) nevner at man kan skille reliabilitet i ekstern konsistens reliabilitet og intern konsistents reliabilitet. Ekstern konsistens reliabilitet handler om i hvilken grad elevene ville få de samme resultatene på en prøve, dersom den skulle gjennomføres en gang til. Dette er kun en teorietisk situasjon, siden elevene ikke vil glemme alt fra en gang til den neste (Kleve, 1994). Intern konsistens reliabilitet kan man derimot beregne. Det handler om tilfeldige feil på prøven. Som utgangspunkt må alle spørsmålene på prøven handle om det samme temaet. Dersom tilfeldige feil har en høy innflytelse på en prøve, vil det vise seg gjennom at på noen oppgaver vil bli skåret lavt, mens det på andre oppgaver vil bli skåret høyt (Kleve, 1994). Flervalgsprøven som elevene fikk i denne studien, handlet om DNA, og begreper relatert til DNA som DNA-oppbygning, DNA-replikasjon, DNA-transkripsjon, DNA-translasjon, mitose og meiose. Siden spørsmålene i prøven testet ulike begreper, ble det vanskelig å beregne intern konsistens reliabilitet.

Det er andre faktorer Kleve (1994) nevner, som kunne påvirke reliabiliteten, og disse faktorene var det tatt hensyn til i denne studien. En faktor som nevnes er at en lang prøve vil være mer reliabel enn en kort prøve. Flervalgsprøven elevene fikk, varte i 60 minutter. Det er lengre tid enn det som er vanlig ved en underveisprøve, som vanligvis varer i 45 minutter. Siden prøven varte i 60 minutter, kan det være med på å øke reliabiliteten i prøven. Man måtte også ta hensyn til homogeniteten i elevgruppen for å øke reliabiliteten. Flertallet av elevene var faglig sterke, der 56 % av elevene fikk terminkarakter 5 eller 6 høsten 2008 i Biologi 2. Den lille variasjonsbredden i utvalget er tatt hensyn til i diskusjonen av resultatet, se kapittel 7.2. En annen faktor som man måtte ta hensyn til, var instruksjonene som ble gitt i forkant av før prøven. Slike instruksjoner var veldig viktige dersom det forekom korrigering av resultatet, som nevnt i kapittel 2.6. I dette tilfelle forekom det ingen korrigering, og alle elevene fikk de samme beskjedene før prøvestart angående prøvetiden, korrigering av resultatet og at det ikke var mulig med fri navigering på it's learning. Enda en viktig faktor for å øke reliabiliteten, var at alle elevene fikk god nok tid til å besvare flervalgsprøven. Tiden som var satt av til prøven, var som nevnt 60 minutter, og alle elevene ble ferdige før tiden var ute. Det vil også øke reliabiliteten for denne prøven. Den siste faktoren Kleve (1994) nevner er at sensureringen vil påvirke reliabiliteten. Ved en flervalgsprøve vil sensorreliabiliteten være høy, siden flervalgsoppgaver er en objektiv måte å rette på. Etter å ha tatt hensyn til alle

disse faktorene som Kleve (1994) nevner, kan man fastslå at reliabiliteten på flervalgsprøven ble godt ivaretatt.

Validitet i flervalgsprøven

Kleve (1994) nevner tre ulike tilnærminger for å teste validiteten i kapittel 2.5. Disse handler om i hvilken grad man kan trekke slutninger fra resultatene på en flervalgsoppgave. De tre tilnærmingene er innholdsvaliditet, kriterievaliditet og begrepsvaliditet. Alle disse tilnærmingene overlapper hverandre, men for å studere validiteten til en flervalgsprøve er alle tilnærmingene nødvendige, selv om begrepsvaliditeten er den viktigste (Kleve, 1994). Innholdsvaliditet kan være en viktig kilde for å øke validitet. Det handler om i hvilken grad testen målte det den var ment til å måle. Spørsmålene på testen var laget med utgangspunkt i kompetansemålet *Genetikk* fra LK06 og kapittelet fra Sletbakk et al. (2008) som omhandlet DNA og begreper rundt dette kompetansemålet (Sletbakk et al., 2008, side 151-188). Derfor kan det påstås at det var en god innholdsvaliditet på flervalgsprøven. Kriterievaliditet handler om elevenes prestasjon på prøven, og i hvilken grad om resultatet kan si noe om fremtidige prestasjoner i faget. De faglig sterke elevene oppnådde høy skår på flervalgsprøven, og det var forventet. De faglig svake elevene oppnådde høyere skårer på flervalgsprøven enn forventet, og det kan hende at de ikke klarte å opprettholde et slikt resultat senere i faget. Siden de faglig svake elevene ikke møtte opp til kortsvarsprøven, var det vanskelig å gjøre en sammenligning mellom resultatene fra flervalgsprøven og kortsvarsprøven i forhold til om de klarte å opprettholde resultatet fra flervalgsprøven. Begrepsvaliditet er den siste tilnærmingen for validitet. Denne tilnærmingen ga en indikasjon på om oppgavene som elevene utførte, virkelig målte kompetansemålene og om man kunne trekke slutninger om elevenes måloppnåelse i faget. Ut i fra kompetansemålet *Genetikk* fra læreplanen i Biologi 2 skal elevene kunne:

- Forklare strukturen til DNA og hvordan DNA blir kopiert før cellene deler seg.
- Gjøre greie for transkripsjon og translasjon av gen (…).
- Sammenligne mitose og meiose med vekt på fordeling av genmaterialet i cellene som blir dannet.

De fleste elevene kunne besvare de fleste oppgavene, og gjennomsnittlig skår på flervalgsprøven var 18,31 poeng av 24 mulige. I boks 12, som er fremstilt i kapittel 6.1, kommer det fram at flere elever fikk vist kompetansen sin på flervalgsprøven. Det kommer også fram i figurene 5 og 6, som fremvises i kapittel 6.1, at flere elever gjettet en eller flere

ganger. De benyttet også eliminasjonsmetoden for enten å gjette blant færre svaralternativer eller som hjelpemiddel for å velge det korrekte svaralternativet. Begrepsvaliditeten ble ivaretatt, men den ble også svekket, siden flere elever gjettet og benyttet eliminasjonsmetoden for å besvare en eller flere oppgaver.

Ut i fra disse elementene kan man si at innholdsvaliditeten ble ivaretatt på en tilfredsstillende måte, men at det ikke er grunnlag for å avgjøre om kriterievalditeten ble ivaretatt. Begrepsvaliditeten er den viktigste i forhold til Kleve (1994), og den ble ivaretatt til en viss grad. Mange av elevene fikk vist sin kunnskap fra kompetansemålet *Genetikk* fra LK06, og det kom fram i boks 12. Samtidig benyttet flere elever gjetting og eliminasjonsmetoden for å besvare en eller flere av flervalgsoppgavene, og det kan svekke begrepsvaliditeten.

Kortsvarsprøven

Resultatene fra kortsvarsprøven ble ikke helt som forventet. De faglig svake elevene utelot å møte opp på kortsvarsprøven, fordi de ikke ville ødelegge det gode utgangspunktet som de fikk fra flervalgsprøven. Siden kun de faglig sterke elevene møtte opp på flervalgsprøven, ble det da forventet at gjennomsnittskarakteren for kortsvarsprøven skulle være forholdsvis høy. Det skjedde ikke. Gjennomsnittskarakteren fra kortsvarsprøven var ¾, se vedlegg 9, og det var flere grunner til at gjennomsnittet var lavere enn forventet. Som nevnt i kapittel 6.2. var kortsvarsprøven for omfattende, siden flere elever fikk for kort tid til å besvare alle oppgavene. Grunnen var at læreren hadde bestemt å ta med emnet genregulering på prøven, siden dette temaet ikke var gjennomgått før flervalgsprøven. Det resulterte at noen elever ikke rakk å lese gjennom besvarelsene før de måtte levere dem inn. I boks 16, som er fremlagt i kapittel 6.2, kom det også fram at flere elever gjorde det dårligere på kortsvarsprøven enn forventet, siden de hadde lest på feil kapittel, forberedt seg dårlig eller hadde dårlig konsentrasjon på prøven. Disse begrunnelsene var også elementer som gjorde at gjennomsnittskarakteren ble lavere enn forventet.

Oppgavene på kortsvarsprøven var også mer omfattende enn kortsvarsoppgavene på skriftlig eksamen i Biologi 2. På kortsvarsprøven kunne elevene skrive maks en A4-side per oppgave, inkludert figurer. Årsaker til at oppgavene på kortsvarsprøven var mer omfattende, var at de ikke skulle være identiske med oppgavene fra flervalgsprøven. Det at kortsvarsoppgavene var mer omfattende enn kortsvarsoppgavene som blir gitt på skriftlig eksamen i Biologi 2, kunne resultere i at elevene fikk dårlig tid på kortsvarsprøven. Ut fra boks 16 i kapittel 6.2, kom det fram at vurderingsformen med større kortsvarsoppgaver gjorde at mange elever følte at de

fikk vist mye av sin faglige kompetanse på denne prøven. Elevene fikk beskrevet flere detaljer rundt emnet og tegnet ned relevante figurer. Som nevnt tidligere, var flertallet av elevene faglig sterke, og disse elevene møtte opp på kortsvarprøven. For faglig sterke elever vil kortsvarsoppgaver gjøre at de får vist utdypende kunnskap om emnet det blir spurt om på prøven. Siden de faglige svake elevene utelot å møte opp på kortsvarsprøven, resulterte det også i at det ble vanskelig å sammenligne resultatene fra flervalgsprøven og kortsvarsprøven.

Reliabilitet i kortsvarsprøven

Som nevnt tidligere i kapittel 2.5, kan man skille reliabilitet i ekstern konsistens validitet og intern konsistens reliabilitet (Kleve, 1994). Tidligere i dette kapittelet kom det fram at det ikke var grunnlag for å undersøke ekstern konsistens reliabilitet, siden det kun er en teoretisk situasjon. Det gjaldt også for kortsvarsprøven, siden elevene ikke vil glemme alt fra en gang til den neste. Det var heller ikke grunnlag for å beregne intern konsistens reliabilitet, siden kortsvarsprøven ikke bare handlet om ett tema, men den bestod av fem spørsmål som omhandlet flere begreper under hovedtemaet DNA.

Av de andre faktorene som Kleve (1994) nevner i kapittel 2.5, så var kortsvarsprøven mer omfattende enn flervalgsprøven, og det kunne gi høyere reliabilitet. Samtidig så ble ikke alle elevene ferdige i tide, og det kunne senke reliabiliteten. Alle elevene fikk de samme beskjedene før prøvestart, og det vil også øke reliabiliteten. Det er også tatt hensyn til testgruppens homogenitet i diskusjonen, se kapittel 7.2, med tanke på at det kun var de faglig sterke elevene som møtte opp på kortsvarsprøven. Den siste faktoren som Kleve (1994) nevner, handler om sensurering. Ved en flervalgsprøve vil sensorreliabiliteten være høy, siden en slik prøve har en objektiv retting. Ved åpne oppgaver, som i dette tilfelle med kortsvarsprøven, vil sensorreliabiliteten være mer utsatt. Ut fra disse faktorene, vil reliabiliteten for kortsvarsprøven være middels ivaretatt. Den kunne ha blitt høyere om alle de deltagende elevene hadde fått muligheten til å bli ferdig med kortsvarsprøven. Den kunne også ha blitt høyere om de faglig svake elevene også hadde møtt opp på prøven.

Validitet i kortsvarsprøven

Kleve (1994) nevner i kapittel 2.5 at det finnes tre ulike tilnærminger for å teste validiteten, som handler om i hvilken grad man kan trekke slutninger fra resultatene på en prøve. De tre

tilnærmingene er innholdsvaliditet, kriterievaliditet og begrepsvaliditet, og de overlappet hverandre, men begrepsvaliditet er den viktigste (Kleve, 1994).

Innholdsvaliditet var en viktig kilde for å bevise validiteten i oppgavene. I kortsvarsprøven var det tatt utgangspunkt i kompetansemålet *Genetikk* fra LK06 og kapittelet fra Sletbakk et al. (2008) som omhandlet DNA og begreper rundt dette kompetansemålet (Sletbakk et al., 2008, side 151-188). Man kan derfor anta at det kan være en god innholdsvaliditet på kortsvarsprøven. Kriterievaliditet handler om elevenes prestasjon på prøven, og om resultatet kan si noe om fremtidige prestasjoner i faget. I følge boks 16 i kapittel 6.2, var det flere av elevene som presterte dårligere på kortsvarsprøven enn forventet, på grunn av dårlig forberedelser og dårlig konsentrasjon under prøven. Derfor vil resultatene fra denne kortsvarsprøven ikke si mye om fremtidig prestasjon i faget. Begrepsvaliditet er den siste tilnærmingen for validitet. Denne tilnærmingen vil gi en indikasjon på om oppgavene som elevene utførte, virkelig målte kunnskapen innen kompetansemålene og om man kunne trekke slutninger om elevenes måloppnåelse i faget. Ut i fra kompetansemålet *Genetikk* fra læreplanen i Biologi 2 (Utdanningsdirektoratet, 2006a) skulle elevene kunne:

- Forklare strukturen til DNA og hvordan DNA blir kopiert før cellene deler seg.
- Gjøre greie for transkripsjon og translasjon av gen og forklare hvordan regulering av gen kan styre biologiske prosesser.
- Sammenligne mitose og meiose med vekt på fordeling av genmaterialet i cellene som blir dannet.

Gjennomsnittskarakteren fra kortsvarsprøven var ¾, se vedlegg 9. Det var dårligere enn forventet, siden de faglig svake elevene utelot å møte opp på kortsvarsprøven. Flere elever hadde forberedt seg dårlig på denne prøven og noen hadde dårlig konsentrasjon under prøven. Samtidig kom det fram i boks 16 at flere elever fikk vist kompetansen sin gjennom kortsvarsprøven. Ut fra disse elementene kan man si at begrepsvaliditeten ble ivaretatt til en viss grad. Innholdsvaliditeten ble også ivaretatt på en tilfredsstillende måte på denne kortsvarsprøven, siden oppgavene ble laget ut fra kompetansemålet *Genetikk* fra LK06 og kapittelet fra Sletbakk et al. (2008) som omhandlet DNA og temaer rundt det kompetansemålet (Sletbakk et al., 2008, side 151-188). Oppgavene som ble gitt til elevene, var også godkjent av veileder og læreren i studien før prøvestart.

7.5. Erfaringer fra metodevalget

Som nevnt i kapittel 5.2 var denne studien en kvalitativ studie som hadde likhetstrekk med en kasusstudie, der utvalget var en lærer og hans klasse i Biologi 2. En kasusstudie involverer forskning på et fenomen som er viktig i sin kontekst. Det betyr å gi en detaljert beskrivelse av dette fenomenet, og man måt da samle inn tilstrekkelig med data (Postholm, 2005, side 41, 50; Robsen, 2002, side 89, 178). Datainnsamlingen ble utført ved hjelp av spørreundersøkelser og intervju av lærer og elever.

En stor utfordring med denne datainnsamlingen var antall elever i utvalget. Det ble tatt utgangspunkt i en lærer og klassens hans i Biologi 2. Klassens hans bestod av 18 elever, der 17 stykker møtte opp til flervalgsprøven og 14 stykker møtte på kortsvarsprøven. Det ble også bestemt at et utvalg på fire elever skulle intervjues etter de to prøvene. Dette utvalget skulle bestå av to faglig sterke elever, hvorav den ene eleven gjorde det bra og den andre ikke fullt så bra på flervalgsprøven, og de to faglig svake elever, hvorav en elev gjorde det bedre enn forventet og den andre eleven hadde et resultat som var samme nivå som utgangspunktet til eleven. Kriteriet for elevene som skulle intervjues var at de måtte ha utført både flervalgsprøven og kortsvarsprøven for at det skulle være mulig å sammenligne disse to vurderingsformene. På kortsvarsprøven var det flere elever som ikke møtte opp, blant annet de faglig svake elevene. Siden de faglige svake elevene utelot å møte opp på kortsvarsprøven, ble det vanskelig å sette sammen utvalget til intervjuene. Dermed måtte det velges ut nye elever til intervjuene, og det ble to faglig sterke elever som gjorde det bra på flervalgsprøven og to elever som presterte middels i faget, der den ene gjorde det dårligere enn forventet og den andre var på samme nivå som forventet. Under intervjuene med elevene besvarte de spørsmålene veldig kort, og det var vanskelig å få ut begrunnelser på valg og synspunkter rundt spørsmålene som ble stilt til dem. I kapittel 5.4. ble det nevnt at elevintervjuene ble bygget opp som strukturerte intervjuer. Postholm (2005) nevner at man kan dele intervju inn i tre ulike former: strukturert intervju, halvstrukturert intervju og ustrukturert intervju (Postholm, 2005, side 69-74). Det kan hende at man fra elevene hadde fått tydeligere synspunkter rundt de to vurderingsformene, flervalgsoppgaver og kortsvarsoppgaver, ved å benytte en annen form for intervju enn strukturert intervju. I et halvstrukturert intervju vil intervjueren ha noen spørsmål klare på forhånd, samtidig som intervjueren er åpen for at respondentene har ulike forhold de vil ta opp (Postholm, 2005, side 72). Det ble ikke tatt hensyn til dette i intervjuene med elevene. I forhold til de kvalitative resultatene i

datamaterialet var det mer informasjon i besvarelsene fra de åpne spørsmålene i spørreundersøkelsen fra alle elevene enn svarene fra de fire elevene i intervjuene.

Som nevnt i kapittel 5.4, ble både det kvantitative og kvalitative datamaterialet transkribert og translert. Svarene fra avkrysningspåstandene og de åpne spørsmålene i spørreundersøkelsen er nøyaktig gjengitt i resultatene i kapittel 6. Intervjuene ble tatt opp på bånd for å få en nøyaktig transkripsjon, slik at translasjonen også skulle bli nøyaktig. De utdragene fra intervjuene som er fremlagt i kapittel 6, var transkribert i liten grad. Det var kun renskrevet fra transkripsjonen, og det var for å opprettholde en høy validitet i dataanalysen. I diskusjonen, se kapittel 7.2 og 7.3, ble det foretatt en dataanalyse av datamaterialet. Det er viktig at man ikke feiltolker datamaterialet i dataanalysen, og utsetter oppgaven for respondent bias, se kapittel 5.4. Siden datamaterialet inneholdt det transkriberte materialet, kan det være vanskeligere å utsette oppgaven for respondent bias.

Det ble også bestemt på forhånd at det ikke var nødvendig med en kontrollgruppe, siden utvalget skulle ha begge prøvene med 11 dagers mellomrom. Det betyr at elevene i utvalget var sin egen kontrollgruppe. I boks 14, som er fremlagt i kapittel 6.1, kom det fram at flere elever ikke møtte opp til kortsvarsprøven for ikke å ødelegge det gode utgangspunktet som de fikk på flervalgsprøven. Siden elever utelot å møte opp på kortsvarsprøven, ble det vanskelig å sammenligne resultatene fra flervalgsprøven og kortsvarsprøven. Dersom denne studien skulle ha blitt utført en gang til, burde man ha benyttet en kontrollgruppe. Denne kontrollgruppen hadde da fått en prøve som var identisk til flervalgprøven, men med åpne oppgaver. Kortsvarsprøven som elevene fikk i denne studien, var oppgaver som elevene skulle besvare på maks en A4-side, og besvarelsene skulle innholde de nødvendige detaljene og figurene for å besvare oppgavene. Dette ble gjort for å hindre at elevene bare leste på spørsmålene og det korrekte svaralternativet fra flervalgsprøven som lå på it's learning. Dersom det hadde blitt benyttet en kontrollgruppe som fikk en prøve som var identisk til flervalgsprøven, hadde det blitt lettere å sammenligne og trekke paralleller om faglig kompetanse mellom de to vurderingsformene. Å sammenligne resultatene fra kortsvarsprøven var vanskelig å gjennomføre i denne studien, siden de faglig svake elevene utelot å møte opp på kortsvarsprøven.

I kapittel 5.4. er det nevnt at respondent bias var et element som man måtte ta hensyn til ved datainnsamlingen. Respondent bias betyr i denne sammenhengen at elevene kan holde tilbake informasjon som kan være viktig for resultatet og analysen av datamaterialet eller at de svarer

det de tror forskeren gjerne vil høre (Robsen, 2002, side 172). Respondent bias kan ha forekommet i denne studien, men det er vanskelig å undersøke det. Ved å benytte metodetriangulering, kan respondent bias bli redusert. Svarene på avkrysningspåstandene, som var en kvantitativ metode, ble sjekket mot svarene på de åpne spørsmålene, som er en kvalitativ metode. Dersom det var store avvik, kunne det ha forekommet respondent bias. Det generelle inntrykket var at det ikke var store avvik mellom avkrysningspåstandene og de åpne oppgavene, men at det forekom i enkelte tilfeller. Blant annet forekom det i avkrysningspåstandene som omhandlet gjetting og bruk av eliminasjonsmetoden for å besvare flervalgsoppgaver, se figur 5 og 6 i kapittel 6.1. Påstanden *Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer* kunne tolkes på to måter, enten kunne eliminasjonsmetoden bli benyttet ved gjetting eller ved valg av korrekt svaralternativ. Det er vist i boks 11 i kapittel 6.1. Det kan også være lettere å unngå respondent bias ved intervju enn ved spørreundersøkelser. Forskeren kan ved intervju spørre elevene tilleggsspørsmål for å få fullstendige svar eller for å klarne opp i misforståelser, dersom noe er uklart. Dersom det forekom uklarheter under intervjuene, ble elevene og læreren spurt om begrunnelser for utsagnene som de kom med.

Reliabilitet i metodevalg

Som nevnt i kapittel 2 er reliabilitet et mål som angir i hvilken grad en test som en elev tar, vil gi det samme resultatet, dersom testen gjentas. I kapittel 5.3. er det nevnt at det i kvalitativ forskning ikke er så viktig å oppnå høy reliabilitet. Det er viktigere å oppnå høy troverdighet i studien. Det vil si at det er viktig å arbeide grundig og forklare hva man har gjort i studiene for å øke troverdigheten i oppgaven (Robsen, 2002, side 101, 176). Ut fra erfaringene fra datainnsamlingen og metodevalgene som har blitt gjort, var troverdigheten godt ivaretatt i denne studien. Det er nøye forklart hva som har blitt utført, og det er fremstilt de viktigste elementene fra datainnsamlingen som er viktig for å få svar på de to problemstillingene. Intervjuene ble tatt opp på bånd, for å få en nøyaktig transkripsjon, og translasjonen var direkte avskrift fra transkriptet. De eneste endringene som forekom var rettskriving. Det kvantitative resultatet er vist i figurer og det kvalitative resultatet ble vist i bokser. De nøye beskrivelsene og direkte avskrift av synspunktene til elevene, vil øke troverdigheten i oppgaven.

Validitet i metodevalg

Validitet handler om at de slutningene man kan trekke er reelle for situasjoner utenfor prøven, se kapittel 2. For at slutningene utenfor prøven skal være reelle, må den indre validiteten i studien være god. Spørreundersøkelsen som elevene utførte bestod av både avkrysningspåstander og åpne spørsmål. Siden svarene i avkrysningspåstandene ble krysskontrollert i de åpne spørsmålene, var det en bra metode å samle inn datamaterialet på. Svarinnholdet i de åpne oppgavene hadde ingen restriksjoner og elevene hadde mulighet til å svare i dybden på spørsmålene. Det gav en mer nøyaktighet i vurderingen av hva elevene mente. En slik metodetriangulering, se kapittel 5.4, økte den indre validiteten i oppgaven. Dersom denne studien skulle ha blitt utført en gang til, ville det ikke ha blitt benyttet intervjuer, men kun spørreundersøkelse med avkrysningspåstander og åpne spørsmål som metode med et større utvalg, siden man ikke kunne trekke så mye informasjon ut fra intervjuene med elevene. Spørreundersøkelsen kunne da i stedet ha inneholdt flere spørsmål. Det kunne også vært mulig å endre intervjuformen for elevintervju, ved å endre til halvstrukturert eller ustrukturert intervju. I analysen av datamaterialet var det lite synsing, siden translasjonen var renskrevet transkripsjon. I kvalitativ forskning skal man beskrive fenomenet, og ikke trekke konklusjoner og synse i resultatet. Det var ikke tilfellet at man trakk konklusjoner og synes i denne studien, og det vil øke den indre validiteten i studien.

For å få tydeligere tendenser i datamaterialet fra læreren, kunne man også ha økt utvalget i antall lærere som var med på studien. Ved å intervjue flere lærere, ville man ha fått tydeligere tendenser, noe som ville gjort det lettere å besvare den andre problemstillingen, om hvordan en lærer kan forberede elever på flervalgsoppgaver som vurderingsform ved å benytte Test 2.0 i it's learning.

Som nevnt, må man ved validitet kunne trekke slutningene som man kom fram i studien i situasjoner utenfor prøven. I denne studien ble det benyttet et forholdsvis lite utvalg, der flertallet av elevene var faglig sterke. Det at utvalget var lite, gjorde det vanskelig å generalisere resultatet i en større skala. Ut fra datamaterialet kom det fram at elevene synes at de fikk vist kunnskap ved bruk av flervalgsoppgaver som vurderingsform, men de synes samtidig at de fikk vist kunnskapen sin bedre gjennom åpne oppgaver, siden de her kunne skrive ned flere detaljer og figurer som var relevante for oppgaven. Dersom utvalget av elevene hadde vært mer spredt i forhold til faglig kompetanse, kunne konklusjonen vært annerledes om i hvilken grad elevene fikk vist sin faglige kompetanse ved bruk av

flervalgsoppgaver. Tendensene fra de faglig svake elevene, var at de var veldig positive til denne vurderingsformen, se kapittel 7.2, men det var kun to elever som var faglig svake i utvalget. Det gjorde det vanskelig å generalisere disse tendensene i større grad.

7.6. En kritisk oppsummering av studien

Denne masteroppgaven er skrevet som et avsluttende ledd i studiet Lærerutdanning i realfag ved NTNU. Utgangspunktet for oppgaven var å benytte læringsverktøyet Test 2.0 i it's learning for å forberede elever til skriftlig eksamen i Biologi 2. Hensikten var å forberede elever på bruk av flervalgsoppgaver som vurderingsform fra 2008, siden denne vurderingsformen blir benyttet ved skriftlig eksamen i Biologi 2. Denne masteroppgaven er på 30 studiepoeng, og den er skrevet innenfor et begrenset tidsrom. Denne tidsbegrensningen har satt sitt preg på ulike elementer ved oppgaven. Dersom det var mulighet til å ha bedre tid under datainnsamlingen, ville enkelte ting ha blitt gjort annerledes. Det er også flere ting, som man ser i ettertid, som en kunne ha gjort annerledes. Det vil nå nevnes flere elementer som kunne ha blitt gjort annerledes for å få tydeliggjøre tendenser i datamaterialet.

Oppgaven ble påbegynt i august 2008, og i løpet av høstsemesteret 2008 ble det bestemt fokus og problemstillinger for masteroppgaven. Det ble fastsatt to problemstillinger for oppgaven:

1. Får elever i tilstrekkelig grad vist sin faglige kompetanse i Biologi 2 ved bruk av vurderingsformen *flervalgsoppgaver* i den første delen av skriftlig eksamen i Biologi 2?
2. Hvordan kan en lærer forberede elevene sine på den første delen av skriftlig eksamen i Biologi 2, ved å benytte *læringsverktøyet Test 2.0* i læringsplattformen *it's learning* i undervisningen?

Det har blitt gjort svært få endringer på disse to problemstillingene underveis i studien, og de har vært viktige for fokuset i datainnsamlingen og analysen av resultatene. I løpet av høsten 2008 ble det også arbeidet en del med bakgrunnslitteratur og metode, slik at dette skulle være mest mulig ferdig for datainnsamlingen startet. Datainnsamlingen ble gjennomført i januar 2009, siden det ikke var tid å gjennomføre den i løpet av høsten 2008. Kravene for gjennomføringen av datainnsamlingen, var at den måtte foregå ved en skole som benyttet læringsplattformen it's learning, og det var nødvendig med en lærer som underviste i Biologi

2 som ikke hadde benyttet læringsverktøyet Test 2.0 tidligere. Blant skolene i nærmiljøet som innfridde disse kravene, var det en lærer som var interessert å bli med på studien. Det var 18 elever i hans klasse som hadde valgt Biologi 2 i skoleåret 2008/2009, og disse elevene ble elevutvalget for denne studien.

Siden det kun var en lærer og 18 elever som var med på studien, ble det bestemt at studien skulle være en kvalitativ studie med bruk av kvalitative og kvantitative metoder. Fenomener rundt bruk av flervalgoppgaver som vurderingsform skulle utforskes og beskrives, og slike beskrivelser av fenomener er kjennetegn på kasusstudier (Postholm, 2005, side 41). Etter datainnsamlingen ser man at kasusstudier var det korrekte valget for valg av metode, men for å få tydeligere tendenser i datamaterialet kunne det ha blitt benyttet et større utvalg av lærere og elever. Som nevnt i kapittel 7.5, ble det bestemt at det ikke var nødvendig med en kontrollgruppe, siden utvalget skulle ha både flervalgsprøven og kortsvarsprøven med 11 dagers mellomrom. I boks 14, som ble fremlagt i kapittel 6.1, kom det fram at flere elever ikke møtte opp til kortsvarsprøven. Dermed ble det vanskelig å sammenligne resultatene fra flervalgsprøven og kortsvarsprøven.

Kortsvarsprøven, som er vedlagt i vedlegg 8, inneholdt 5 spørsmål, der elevene kunne skrive maks en A4-side per spørsmål, inkludert figurer. Disse kortsvarsoppgavene omhandlet de samme begrepene som i flervalgsprøven, men spørsmålsformuleringen og omfanget av elevbesvarelsene krevde mer av elevene. Dersom det var blitt benyttet en kontrollgruppe, kunne både utvalget og kontrollgruppen fått de samme oppgavene. Vurderingsformen for utvalget kunne ha vært flervalgsoppgaver, men vurderingsformen for kontrollgruppen kunne ha vært kortsvarsoppgaver som var identiske med spørsmålene på flervalgsprøven, bortsett fra at de måtte skrive ned svaret selv med egne ord. Ved å benytte veldig spesifikke kortsvarsspørsmål i stedet for de mer omfattende kortsvarsoppgavene, som ble gitt i denne studien, hadde det vært mulig å benytte kortsvarsfunksjonen i Test 2.0, se figur 1 i kapittel 3.3. Ved å benytte mer spesifikke kortsvarsoppgaver, samt å benytte Test 2.0 for kortsvarsprøven ville det ha gitt et bedre utgangspunkt for å kunne sammenligne resultatene mellom flervalgsprøven og kortsvarsprøven.

I ettertid kunne det også ha blitt noen endringer i selve flervalgsprøven, se vedlegg 6. Det var mange elever som fikk mange korrekte svar på denne prøven, og det er jo egentlig hovedmålet ved en prøve. Man ønsker at elevene skal oppnå høy grad av måloppnåelse innen et kompetansemål, men det kan også indikere at oppgavene som ble produsert hadde dårlig

stamme og/eller dårlige distraktorer. Ut fra tabell 2 i kapittel 6.1 ser man at 42 % av flervalgsoppgavene hadde en P-verdi som lå i intervallet 80-100 %. Ut fra figur 5 og 6 i kapittel 6.1, var det mange elever som gjettet og benyttet eliminasjonsmetoden ved en eller flere av flervalgsoppgavene. Det kan indikere at stammen var for enkel og/eller distraktorene var for dårlige. For å hindre en høy P-verdi på oppgavene, kunne det ha blitt utført en pilottest på flervalgsprøven med et annet utvalg. En slik pilottest kan peke på elementer som kan virke manglende eller misforstående for elevene (Robsen, 2002, side 185). Ved å benytte en pilottest, kunne man ha forutsett dårlige stammer og distraktorer, slik at færre flervalgsoppgaver fikk en P-verdi på mellom 80 og 100 %.

Som nevnt tidligere foregikk datainnsamlingen i januar 2009, og den ble gjennomført ved bruk av spørreundersøkelser med avkrysningspåstander og åpne spørsmål, samt intervju av elever og lærer. En slik metodetriangulering var en god måte å samle inn datamaterialet på. Dette gjorde også at validiteten i oppgaven ble ivaretatt, siden resultatet be mer detaljert. Avkrysningspåstandene ble benyttet for å få en enkel oversikt og for å se trender i datamaterialet. For få en dypere forståelse av elevenes synspunkter ble det benyttet åpne spørsmål i spørreundersøkelsen og intervjuer. Det var mye informasjon i datamaterialet fra spørreundersøkelsene, og det resulterte i en god analyse for å kunne besvare den første problemstillingen, se kapittel 1.2. Det kom ikke like mye fram fra intervjuene med elevene. Som nevnt i kapittel 5.4, var elevintervjuene bygget opp som et strukturerte intervjuer. Det kan hende at man hadde fått tydeligere synspunkter fra elever om bruk av flervalgsoppgaver og kortsvarsoppgaver som vurderingsform ved å benytte en annen form enn strukturert intervju. Dersom det hadde blitt benyttet halvstrukturert eller ustrukturert intervju kunne det hende at man hadde fått tydeligere synspunkter fra elevene om de to vurderingsformene. Dersom det hadde blitt benyttet et større utvalg elever, ville man sannsynligvis ha intervjuet faglig svake elever for å få deres synspunkter rundt bruk av flervalgsoppgaver som vurderingsform. Det ble ikke gjennomført i denne studien, siden et av kravene for intervjuet var at elevene måtte ha møtt opp på begge prøvene. De faglig svake elevene utelot å møte opp på kortsvarsprøven, og dermed ble de ikke intervjuet.

Det ble intervjuet kun én lærer i studien, og tendensene som kunne besvare den andre problemstillingen, se kapittel 1.2, var veldig avhengige av denne læreren. Læreren så mange muligheter ved bruk av læringsverktøyet Test 2.0 i underveisvurderingen og for å forberede elevene til skriftlig eksamen, se kapittel 7.3. Dersom studien skulle ha blitt gjennomført en gang til, ville det sannsynligvis ha blitt benyttet flere lærere for å få flere synspunkter rundt

bruk av Test 2.0 for å kunne forbedre elever til flervalgsoppgaver på skriftlig eksamen i Biologi 2.

Som nevnt i kapittel 5, ble datainnsamlingen utført ved bruk av åtte skoletimer à 45 minutter. Grunnet at masteroppgaven ble skrevet innenfor et begrenset tidsrom, ble spørreundersøkelsen og intervjuene utført rett etter at kortsvarsprøven var gjennomført. Det var læreren som skulle vurdere kortsvarsprøvene til elevene, og det resulterte i at datainnsamlingen med spørreundersøkelsene og intervjuene ble gjennomført før elevene hadde fått en vurdering av kortsvarsprøven. Dersom studien hadde vært gjennomført en gang til, ville datainnsamlingen blitt gjennomført etter at elevene hadde fått vurderingen av kortsvarsprøven.

Det er flere elementer rundt datamaterialet og resultatene som ikke ble tatt opp til videre diskusjon opp mot bakgrunnsteori. I dette diskusjonskapittelet, har det kun blitt tatt hensyn til resultater som besvarte problemstillingene, med de andre elementene i datamaterialet og resultatet er tilgjengelig for videre studier. Mange av endringene som ble foreslått her, kunne ha vært mulig å gjennomføre dersom studien skulle ha blitt utført en gang til. Dersom man skulle ha benyttet en pilottest, et større utvalg med kontrollgruppe og flere lærere som ble intervjuet, hadde analysen av datamaterialet tatt lengre tid, og oppgaven måtte da ha omfattet flere studiepoeng. Det hadde ikke vært mulig å gjennomføre disse endringene innenfor en masteroppgave på 30 studiepoeng.

Avslutning

Det har kommet fram i denne studien at flervalgsoppgaver kan benyttes både ved vurdering *for* læring og vurdering *av* læring. Hovedfunnene for denne studien er at elevene i tilstrekkelig grad klarte å vise sin faglige kompetanse gjennom bruk av flervalgsoppgaver, men at det er mye man må ta hensyn til ved å benytte en slik vurderingsform. Flervalgsoppgaver er en av vurderingsformene som blir gitt på skriftlig eksamen i Biologi 2, og det er viktig at lærerne gir slike oppgaver underveis i faget for å forberede elevene på denne vurderingsformen. For å forberede elevene på flervalgsoppgaver, kan det å benytte læringsverktøyet Test 2.0 være en god metode, siden det kan brukes for å øke læringen og motivere elevene til videre læring. En digital test kan gi lærerne og elevene variasjon og motivasjon i undervisningen og læringen. Det er også mulig å variere oppgaveformene i Test 2.0 ved å gi generelle eller interaktive oppgaver, som også kan inkludere bruk av tekst, lyd og/eller bilde. Ved å benytte Test 2.0 vil lærerne få samlet oppgaver i en større base, slik at man stadig får en bedre samling av

oppgaver. Dette kan gjøre det lettere for lærere å forberede elever til flervalgsoppgaver som vurderingsform. Uansett er det viktig at flervalgsoppgaver ikke betraktes som en erstatning for, men som er verdifullt supplement til de andre prøveformene som kan benyttes ved vurdering *for* læring og vurdering *av* læring.

8. Etterord

Flervalgsoppgaver er en av vurderingsmetodene som blir benyttet på skriftlig eksamen i Biologi 2. Det er viktig at lærere forbereder elevene til denne vurderingsformen, og Test 2.0 kan være en god metode for å forbedre elevene på flervalgsoppgaver som vurderingsform. Test 2.0 kan tilby elevene en ny læringsform og vurderingsform, og den kan utgjøre et viktig supplement i lærerens pedagogiske verktøykasse. I boks 25 kommer det fram at læreren ønsket å benytte flervalgsoppgaver og Test 2.0 mer og mer i undervisningen. Læreren ønsket å benytte Test 2.0 å kartlegge elevenes kompetanse før, under og etter en undervisningsøkt, samt benytte det til å motivere elevene og variere undervisnings- og vurderingsmetodene. Torsdag 07.05.09 ble det foretatt et postintervju av læreren for å få kartlagt i hvilken grad han hadde benyttet flervalgsoppgaver og Test 2.0 for å forberede elevene til skriftlig eksamen i løpet av våren 2009. Elevenes eksamen i Biologi 2 nærmet seg, og det var interessant å undersøke hvilke metoder læreren hadde benyttet og hvor mye han hadde benyttet Test 2.0 for å forberede dem til eksamen. Vedlegg 10 viser spørsmålene som ble stilt under postintervjuet.

I postintervjuet kom det fram at læreren har benyttet flervalgsoppgaver i undervisningen. Han hadde benyttet de flervalgoppgavene som finnes i tilknytning til læreboka, tidligere eksamenssett og andre ferdige tester på internett. Læreren hadde ikke laget egne flervalgsoppgaver, siden det er så tidkrevende å lage *gode* oppgaver til elevene. Han hadde ikke benyttet læringsverktøyet Test 2.0 i undervisningen. Han hadde lagt ut pekere på it's learning hvor elevene skulle gå inn og besvare ferdige tester på nettsider og han hadde også skrevet ut tester og gitt dem til elevene på papir. Det ble spurt om i hvilken grad han hadde forberedt elevene på skriftlig eksamen i Biologi 2, i forhold til hvilke metoder han hadde benyttet. Boks 26 viser en oversikt over hva læreren svarte.

Boks 26: En oversikt over hvordan læreren har forberedt elevene til skriftlig eksamen i Biologi 2 etter datainnsamlingen.

Intervjuer: *I hvilken grad har du forberedt elevene til skriftlig eksamen i Biologi 2, i forhold til hvilke metoder du har benyttet?*
Lærer: Det har jeg brukt mye tid på. Jeg har benyttet både flervalgsoppgaver, kortsvarsoppgaver og langsvarsoppgaver. Da jeg laget prøvene har jeg stort sett konsentrert meg om kortsvarsoppgaver og langsvarsoppgaver. Heldagsprøven, som elevene nettopp hadde, hadde samme prosedyre som skriftlig eksamen med først 20 flervalgsoppgaver, 4 kortsvarsoppgaver og 2 langsvarsoppgaver, der elevene kunne bruke alle hjelpemidlene på langsvarsoppgavene. Heldagsprøven ble kjørt sånn som eksamen, der elevene fikk en full gjennomgang av hele settet. Tidligere har jeg sett bare kjørt biter av denne prosedyren på prøven, slik at de har blitt forberedt bitvis først, og deretter hele opplegget på heldagsprøven så elevene fikk prøve seg på hele prosedyren.

Intervjuer: *Men hvordan har du forberedt elevene på flervalgsoppgaver, siden det ikke har blitt gitt på prøver?*
Lærer: Det har blitt gitt mer i undervisningen. Slike oppgaver har blitt brukt som *test deg selv*, underveisvurdering på fagstoff som vi har gått i gjennom for å teste elevene i forhold itl hva som sitter igjen etter undervisningen. Etter at de har gjort det, har jeg pratet med dem for å høre hva som var vanskeligst og hvor elevene oppdaget at de har hull, og deretter gjennomgått dette etterpå. Så flervalgsoppgaver har blitt brukt som underveisevaluering for undervisningen.

Det kom også fram fra intervjuet at lærerens kompetanse for å lage og bruke flervalgsoppgaver har økt. Han nevnte at han har kommet over terskelen for å prøve å bruke flervalgsoppgaver i undervisningen. Det kom også fram at han kommer til å benytte Test 2.0 ved senere anledninger. Han har benyttet flervalgsoppgavene fra eksamenssettene som var gitt ut tidligere som forberedelse til elevene, og jo flere eksamenssett som kommer ut, jo større register vil han få som han kan benytte. Disse oppgavene vil bli lagt inn i Test 2.0, og brukes til å forberede elever på flervalgsoppgaver som vurderingsform.

9. Litteraturliste

Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., Ridgeway, J., et al. (2006). A Learning Integrated Assessment System. from http://www.earli.org/position_papers

Bloom, B. (1956). *Taxonomy of Educational Objectives* (6 ed.). New York: David McKay Company, inc.

Dale, K. (2008). Survival of the "net"est? Experiences with electronic test tools - reduced teacher hours? *Seminar.net - International journal of media, technology and lifelong learning, 4*(2), 1-8.

Djupedal, Ø. (2006). Digital kompetanse er viktigere enn noensinne. *Digital Kompetanse, 1*(1), 5-11.

Engh, R., Dobson, S., & Høihilder, E. K. (2007). *Vurdering for læring* (1 ed.). Kristiansand: Høyskoleforlaget.

Gronlund, N. E. (1965). *Measurement and Evaluation in Teaching* (1 ed.). New York: The MacMillian Company.

Gronlund, N. E. (1998). *Assessment of Student Achievement* (6 ed.). USA: Allyn & Bacon.

Haladyna, T. M. (2004). *Developing and Validating Multiple-Choice Test Items* (3 ed.). USA: Lawrence Erlbaum Associates.

Horgen, S. A. (2007). Pedagogical use of multiple choice tests, Students create their own tests. *Proceeding of the Informations Education Europe II Conferance, South East European Research Center (SEERC)*, 386-392.

Horgen, S. A. (2009). Gjemmesiden, mine digitale spor, sist endret 18.01.09. from http://gjemmesiden.blogspot.com/search/label/Tipsserie1, lastet ned 11.02.2009

Håland, E. (2007). Må ha det, bare må ha det! Om fenomenet Learning Management System (LMS). *Digital Kompetanse, 1*(2), 4-22.

Imsen, G. (2005). *Elevenes verden, innføring i pedagogisk psykologi* (4 ed.). Oslo: Universitetsforlaget.

Imsen, G. (2006). *Lærerens verden, innføring i generell didaktikk* (3 ed.). Oslo: Universitetsforlaget.

It'sLearning. (2008). *Grunnkurs Versjon 3.3*. Bergen: It'sLearning AS.

Jansen, K. T. (1971). *Konstruksjon av flervalgsprøver*. Bergen: Handelslærerstudiet NHH.

Kleve, B. (1994). *En testteoretisk og diagnostisk analyse av flervalgsoppgaver i matematikk fra TIMSS' pilottest*. Oslo: Universitetet i Oslo.

Lykknes, A., & Smidt, J. (2009). *Skriving i arbeidsbok i naturfag på ungdomstrinnet: innhold, form og formål*: Under trykker i konferanserapport.

McDonald, M. E. (2002). *Systematic assessment of learning outcomes, developing multiple-choice exams* (1 ed.). USA: Jones & Bartlett Publishers.

Postholm, M. B. (2005). *Kvalitativ metode, en innføring med fokus på fenomenologi, etnografi og kasusstudier* (1 ed.). Oslo: Universitetsforlaget.

Robsen, C. (2002). *Real World Research* (2 ed.). USA: Blackwell Publishing.

Sirnes, S. M. (2005). *Flervalgsoppgaver, konstruksjon og analyse* (1 ed.). Bergen: Fagbokforlaget.

Sletbakk, M., Håpnes, A., Hessen, D. O., Gjærevoll, I., Røsok, Ø., Borge, O. J., et al. (2008). *Bios - biologi 2* (2 ed.). Aurskog: Cappelen Damm AS.

St.meld.nr.30. (2003-2004). Kultur for læring. from http://www.regjeringen.no/nb/dep/kd/dok/regpubl/stmeld/20032004/Stmeld-nr-030-2003-2004-.html?id=404433, lastet ned 15.04.09

Strømme, A. (2008). Biologifaget og digital kompetanse. In P. van Marion & A. Strømme (Eds.), *Biologididaktikk* (1 ed.). Kristiansand: Høyskoleforlaget.

UninettABC. (2006a). Digital mappevurdering, http://www.uninettabc.no/content.ap?thisId=684&language=0, lastet ned 25.09.2008.
UninettABC. (2006b). LMS - hva og hvordan, http://www.uninettabc.no/content.ap?thisId=675&language=0, lastet ned 25.09.2008.
Utdanningsdirektoratet. (2006a). Læreplan i biologi, programfag i studiespesialiserende utdanning. from http://udir.no/grep/Lareplan/?laereplanid=170703&visning=5&sortering=2&kmsid=170715, lastet ned 29.04.09
Utdanningsdirektoratet. (2006b). Digitale læringsplattformer, en mulig katalysator for digital kompetanse i grunnopplæringen. from http://udir.no/upload/Rapporter/LMS.pdf, lastet ned 12.02.09
Utdanningsdirektoratet. (2007). Utstyrs- og driftssituasjonen i grunnopplæringen 2006-2007, en gjentatt kartlegging og sammenligning med 2004-2005. from http://udir.no/upload/Satsningsomraader/digital_kompetanse/Utstyrs-og_driftssituasjonen_i_grunnopplaringen_06-07.pdf, lastet ned 12.02.09
Utdanningsdirektoratet. (2008a). Eksamen for REA3002 Biologi del 1, http://udir.no/upload/Eksamen/Videregaende/Tidligere_gitte_eksoppg_Kunnskapsl/Programfag_studieforberedende/V08/REA3002_Biologi_2_del_1_V08.pdf, lastet ned 25.09.2008.
Utdanningsdirektoratet. (2008b). Eksamen for REA3002 Biologi del 2, http://udir.no/upload/Eksamen/Videregaende/Tidligere_gitte_eksoppg_Kunnskapsl/Programfag_studieforberedende/V08/REA3002_Biologi_2_del_2_V08.pdf, lastet ned 25.09.2008.
Utdanningsdirektoratet. (2008c). Vurderingsveiledning for REA3002 Biologi, http://udir.no/upload/Eksamen/Videregaende/V2009/Vurderingsv/REA3002_Biologi_2_vurderingsveiledning_H08.pdf, lastet ned 25.09.2008.
Utdanningsdirektoratet. (2008d). Informasjon om eksamen våren 2009, http://udir.no/templates/udir/TM_Artikkel.aspx?id=3883#, lastet ned 27.11.2008.
Woolfolk, A. (2004). *Pedagogisk psykologi* (1 ed.). Trondheim: Tapir Adademiske Forlag.

Vedlegg 1: Kontrakt mellom forskeren og utvalget

I dette vedlegget er det en oversikt over kontrakten som ble inngått mellom forskeren og utvalget i denne studien. Det var for å fastslå for elevene at alle resultatene fra datainnsamlingen ble behandlet konfidensielt mellom meg og dem, samt at elevenes deltagelse var basert på frivillighet.

Formell kontrakt om deltagelse og behandling av datainnsamlingen i masterstudiet til Tonje Olsen

Mitt navn er Tonje Olsen. Jeg er masterstudent ved NTNU og jeg skal utføre en studie om flervalgsoppgaver som vurderingsform og bruk av det nye testvektøyet til it's learning i undervisningen. Dersom du har noen spørsmål eller lurer på noe innenfor prosjektet, kan du gjerne kontakte meg på telefon, nn eller e-mail, nn@nn.

Før du signerer denne kontrakten, vil jeg bare påpeke at:

- Din deltagelse er basert på frivillighet;
- Du kan trekke tilbake deltagelsen når som helst.

Alle resultater fra datainnsamlingen vil bli behandlet konfidensielt mellom meg, min veileder og deg. Utdrag fra datainnsamlingen kan være med i masteroppgaven, men det vil ikke bli oppgitt navn eller noen andre identifiserbare trekk i oppgaven.

Vennligst sett kryss ved ett av alternativene:

☐ Jeg har lest kontrakten, og jeg godkjenner at Tonje Olsen kan bruke mine svar i sin masteroppgave.

☐ Jeg har lest kontakten, og jeg godkjenner **IKKE** at Tonje Olsen kan bruke mine svar i sin masteroppgave.

Elev	Lærer
Veileder	Masterstudent

Vedlegg 2: Tematiske kategorier for å analysere datamaterialet

I dette vedlegget er det en oversikt over kategoriene som ble dannet i analyseprosessen. Vedlegget viser også når kategoriene ble dannet og hva hvilke spørsmål fra spørreundersøkelsen kategoriene ble dannet ut fra, se vedlegg 3.

Kategorier	Når og fra hvor ble kategoriene dannet i analyseprosessen?
Forberedelse til flervalgsprøven	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 2b i) og intervju av elever
Forberedelse til kortsvarsprøven	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 3b i) og intervju av elever
Vise kunnskap gjennom flervalgsprøven	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 2b ii) og intervju av lærer og elever
Vise kunnskap gjennom kortsvarsprøven	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 3b ii) og intervju av lærer og elever
Fordeler med flervalgsprøver	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 2b iii) og intervju av lærer og elever
Fordeler med kortsvarsprøver	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 3b iii)
Ulemper med flervalgsprøver	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 2b iv) og intervju av lærer og elever
Ulemper med kortsvarsprøver	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 3b iv)
Gjetting	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 2b v) og intervju av elever
Eliminasjon	Bestemt underveis i analysen ut fra svar på spørsmål 2b v) og intervju av elever
Flervalgsoppgaver på eksamen	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 2b vi) og intervjuspørsmål av lærer
Prøve - flervalg vs kortsvar	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 4a)
Eksamen - flervalg vs kortsvar	Forhåndsbestemt ut fra spørreundersøkelsen, spørsmål 4a) og intervjuspørsmål av lærer
Fordel eksamen	Forhåndsbestemt ut fra intervju av elever
Vanskelighetsgrad flervalg	Forhåndsbestemt ut fra intervju av elever
Vanskelighetsgrad kortsvar	Forhåndsbestemt ut fra intervju av elever
Andre kommentarer	Bestemt etter analysen av intervjuene. Benyttet denne kategorien for å ikke glemme viktige elementer ved intervjuet siden alt det andre var kategorisert

Vedlegg 3: Spørreundersøkelsen

I dette vedlegget er det oversikt over spørreundersøkelsen som ble gitt til elevene etter flervalgsprøven og kortsvarsprøven.

Flervalgsprøven og kortsvarsprøven

Dette spørreskjemaet inneholder spørsmål om deg, dine erfaringer og meninger i forhold til bruk av flervalgsoppgaver, og hva du synes om slike oppgaver som vurderingsform etter at du har utført to prøver i Biologi 2.

Tenk nøye om gjennom spørsmålene og skriv ned det som du mener.

Spørreskjemaet starter med enkel bakgrunnsinformasjon om deg. Deretter deles spørreskjemaet i to deler, flervalgsprøven og kortsvarsprøven. Ved begge disse delene skal du besvare noen avkrysningspåstander og noen utfyllende spørsmål. Tilslutt er det noen felles utfyllende spørsmål om både flervalgsprøven og kortsvarsprøven. Skriv helst i hele setninger, ikke stikkord.

Hensikten med denne undersøkelsen er å finne ut av hva du synes om flervalgsprøven siden slike spørsmål blir benyttet ved skriftlig eksamen i Biologi 2. Denne informasjonen kan hjelpe meg med å kartlegge din mening om den nye vurderingsformen i Biologi 2.

Dine svar vil være til stor hjelp.

Tusen takk!

Vennlig hilsen

Tonje Olsen

1. Din bakgrunn

a. Kjønn

☐ Jente

☐ Gutt

b. Skriv fem siffer som du benyttet i høst på den forrige spørreundersøkelsen.

For eksempel de 5 siste sifrene i telefonnummeret ditt, eller de fem første tallene i fødselsdatoen din. Ikke skriv 00000 eller 12345 eller lignende. Det kan være mange som vil velge slike tall, og jeg kan ikke få inn to undersøkelser med samme siffer. Dette sifferet skal også brukes i alle spørreundersøkelser som du utfyller for meg, så det er viktig at du husker det.

☐ _____

c. Hva fikk du i terminkarakter i Biologi 2 høsten 2008?

Jeg trenger denne informasjonen for masteroppgaven min. Jeg skal se sammenheng mellom karakterer og hvilke vurderingsform man liker best.

☐ _____

2. Flervalgsprøven.

Du vil nå få flere spørsmål som tar utgangspunkt i flervalgsprøven som du hadde tirsdag 13.01.09.

a. Generelle påstander

Kryss av på en skala fra enig til uenig i forhold til din mening om flervalgsprøven. Sett kun et kryss per påstand. La linjen stå blank hvis du ikke forstår spørsmålet.

PÅSTAND	Enig	Litt enig	Litt uenig	Uenig	Vet ikke
i) Jeg forberedte meg ingenting til flervalgsprøven.					
ii) Flervalgsprøven var for lett.					
iii) Flervalgsprøven var for kort.					
iv) Jeg trodde alle oppgavene på flervalgsprøven var korrekte før jeg leverte. Altså gjettet jeg aldri.					
v) Dersom jeg gjettet, klarte jeg ikke å luke bort noen svaralternativer.					
vi) Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på flervalgsprøven.					
vii) Jeg føler meg bedre forberedt til skriftlig eksamen i Biologi 2, ift å benytte flervalgsoppgaver som vurderingsprøveform.					

b. Utfyllingsspørsmål om flervalgsprøven

i. Hvordan forberedte du deg til denne flervalgsprøven, altså hvilke metode benyttet du?

ii. Fikk du vist din kompetanse fra kapittel 6 gjennom flervalgsprøven? Begrunn svaret.

iii. **Hva slags fordeler har flervalgsprøver i Biologi 2 for deg?**

iv. **Hva slags ulemper har flervalgsprøver i Biologi 2 for deg?**

v. **Gjettet du mye? Kunne du eliminere noen svaralternativer slik at det ble lettere å gjette eller gjettet du mellom alle svaralternativene?**

vi. **Hva syns du om at det er flervalgsoppgaver på eksamen i Biologi 2?**

3. Kortsvarsprøven

Du vil nå få flere spørsmål som tar utgangspunkt i kortsvarsprøven, som du hadde fredag 23.01.09.

a. Generelle påstander

PÅSTAND	Enig	Litt enig	Litt uenig	Uenig	Vet ikke
i) Jeg forberedte meg ingenting til kortsvarsprøven.					
ii) Kortsvarsprøven var for lett.					
iii) Jeg fikk vist en stor del av min kompetanse fra kapittel 6 på kortsvarsprøven.					
iv) Jeg foretrekker kortsvarsoppgaver fremfor flervalgsoppgaver som underveisvurdering (vanlige prøver i løpet av skoleåret).					
v) Jeg foretrekker kortsvarsoppgaver fremfor flervalgsoppgaver som sluttvurdering (tentamen og eksamen).					
vi) Jeg synes det er bra at det er kortsvarsspørsmål på eksamen i Biologi 2.					

b. Utfyllende spørsmål om kortsvarsprøven

i. Hvordan forberedte du deg til denne kortsvarsprøven?

ii. Fikk du vist din kompetanse fra kapittel 6 gjennom kortsvarsprøven? Begrunn svaret.

iii. **Hva slags fordeler har kortsvarsprøve i Biologi 2 for deg?**

iv. **Hva slags ulemper har kortsvarsprøver i Biologi 2 for deg?**

4. Felles spørsmål om flervalgsprøven og kortsvarsprøven

a. **Hva likte du best av flervalgsoppgaver eller kortsvarsoppgaver, i forhold til de to ulike metodene å bruke som spørsmål? Ikke ta hensyn til vanskelighetsgrad på oppgavene. Begrunn svaret!**

b. **I forhold til eksamen i Biologi 2, i hvilken av flervalgsoppgaver eller kortsvarsoppgaver føler du at du får vist din kompetanse <u>BEST</u>? Begrunn svaret!**

Tusen takk for hjelpen!

Vedlegg 4: Intervjuspørsmål til elever

I dette vedlegget er det en oversikt over de ulike spørsmålene som ble stilt til de fire elevene som ble intervjuet.

1. Hvordan forbereder du deg til en flervalgsprøve?
2. Hvordan forbereder du deg til en kortsvarsprøve?
3. Hvordan var vanskelighetsgraden på flervalgsprøven?
4. Hvordan var vanskelighetsgraden på kortsvarsprøven?
5. Gjettet du?
    a. Var det vilkårlig avkrysning, eller kunne du eliminere noen svaralternativer?
6. Fikk du uttrykt din kompetanse på flervalgsprøven?
7. Fikk du uttrykt din kompetanse på kortsvarsprøven?
8. Hva slags fordeler vil en slik oppkjøring gi deg til eksamen?

Vedlegg 5: Intervjuspørsmål til lærer

I dette vedlegget er det en oversikt over de ulike forhåndsbestemte spørsmålene som ble stilt til læreren under intervjuet.

1. Hva synes du om bruk av flervalgsoppgaver som vurderingsform?
    a. Vurdering av læring
    b. Vurdering for læring
2. Hva slags nyttig verdi ser du ved bruk av flervalgsoppgaver på it's learning?
3. Hvordan er den tekniske bruken av Test 2.0?
4. Hva slags tanker har du om hvordan man kan måle kunnskapsnivå ved bruk av flervalgsoppgaver?
    a. I forhold til Blooms taksonomi?
    b. I forhold til læreplanmål/kompetansemål?
5. Hva slags fordeler er det ved bruk av flervalgsoppgaver som vurderingsform?
6. Hva slags ulemper er det ved bruk av flervalgsoppgaver som vurderingsform?
7. Hvordan vil du benytte denne vurderingsformen videre i undervisningen?

Vedlegg 6: Flervalgsprøven

I dette vedlegget er det en oversikt over de 24 flervalgsoppgavene som ble gitt på flervalgprøven som elevene hadde tirsdag 13.01.09. Svaralternativet i kursiv viser nøkkelen i flervalgsoppgavene.

1. **Den fulle, engelske betegnelsen på DNA er**
   1) *DeoxyriboNucleicAcid*
   2) DeoxyNucleicAcid
   3) DeoxyriboNutrientAcid
   4) DeriboNucleicAcid
   5) DiNitroAdenin

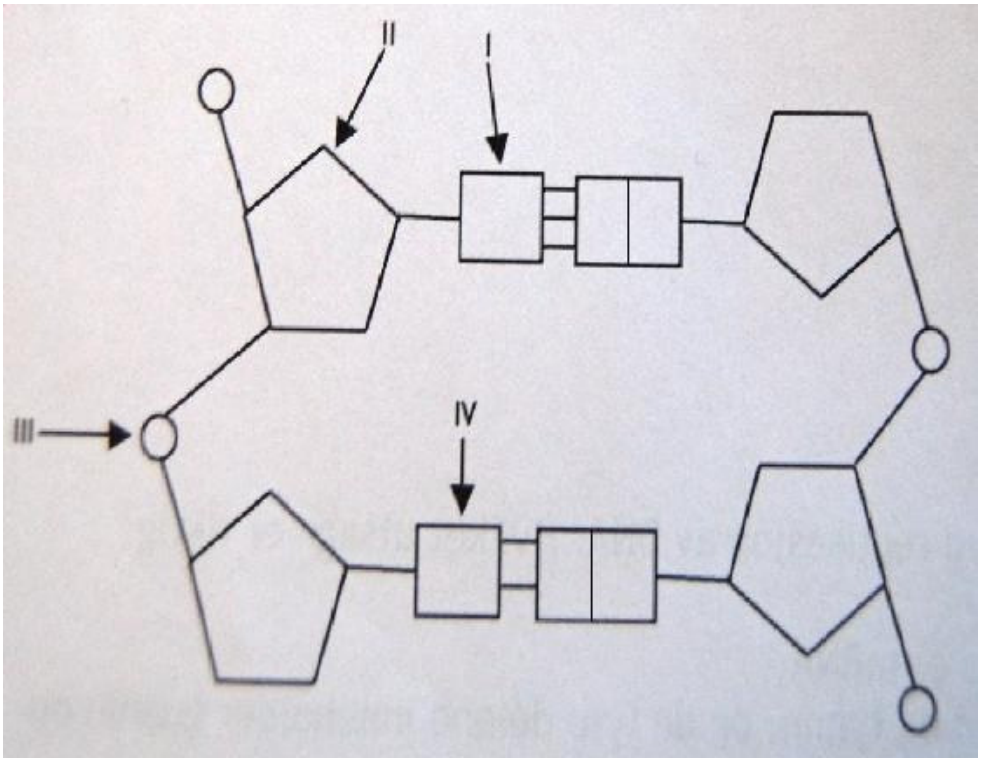
2. **Hva er komponentene i et DNA-nukleotid?**
   1) Par av komplementære baser som holdes sammen av hydrogenbindinger
   2) En dobbelhelix som holdes sammen av hydrogenbindinger
   3) En dobbelhelix som holdes sammen av kovalente bindinger
   4) Sukkeret deoxyribose og fosfat
   5) *Sukkeret deoxyribose, en base og fosfat*

3. **Hvilke av disse nitrogenbasene er <u>IKKE</u> med i oppbygningen til DNA?**
   1) Adenin
   2) Guanin
   3) Tymin
   4) Cytosin
   5) *Uracil*

4. **Hvilke av disse nitrogenbasene er komplementære i DNA?**
   1) Guanin og tymin
   2) *Adenin og tymin*
   3) Tymin og uracil
   4) Adenin og cytosin
   5) Guanin og adenin

5. **Nitrogenbasene i et DNA-molekyl holdes sammen av**
   1) *Hydrogenbindinger*
   2) Ionebindinger
   3) Kovalente bindinger
   4) Både hydrogenbindinger og kovalente bindinger
   5) Det er ingen bindinger som holder dem sammen

6. **Figuren viser en kort del av DNA. Dersom enheten som er merket I er cytosin, hva er da enhetene II, III og IV?**



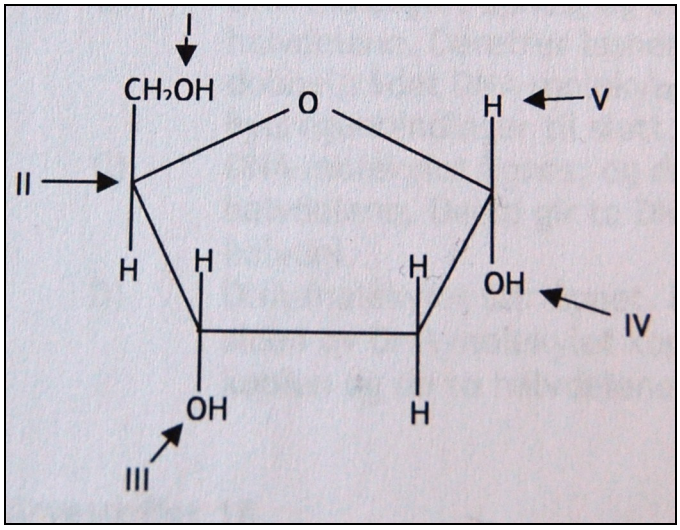
1)	II: Guanin	III: Deoksyribose	IV: Cytosin
2)	II: Guanin	III: Fosfat	IV: Cytosin
3)	II: Fosfat	III: Deoksyribose	IV: Tymin
4)	*II: Deoksyribose*	*III: Fosfat*	*IV: Tymin*
5)	II: Deoksyribose	III: Fosfat	IV: Adenin

7. **Hvilke av disse basesekvensene kan ved baseparing kobles til den ”klebrige enden” på DNA-biten på figuren?**

1) ACTG
2) *ACGT*
3) ACGA
4) TGCA
5) ACGU

8. **Ledetråden i DNA dannes …**
1) Stykkevis i 5’-3’-retning
2) *Sammenhengende i 5’-3’-retning*
3) Sammenhengende i 3’-5’-retning
4) Stykkevis i 3’-5’-retning
5) Begge retninger er ledetråd

9. Hvilke deler av deoksyribose bindes det fosfat til i DNA?



1) II og V
2) III og IV
3) I og IV
4) II og IV
5) *I og III*

10. Antall hydrogenbindinger mellom adenin og tymin i DNA er

1) 0
2) 1
3) *2*
4) 3
5) Det varierer mellom 2 og 3

11. Hva skjer i replikasjonen?

1) Ribosomene leser av mRNA og danner proteiner
2) Det dannes RNA fra DNA
3) *DNA blir kopiert*
4) Kromatider blir kopiert, og dannes søsterkromatider
5) Det blir kopiert nye nukleotidebaser i cytosol

12. To tettpakkede søsterkromatider er festet sammen i et område som kalles …

1) *Sentromeren*
2) Genomet
3) Nukleosomet
4) Spindelen
5) Sentrum

13. Hva er mitose?

1) En bakterie reproduserer seg ved en todeling av cellen
2) Arvestoffet i kjernen fordeles på to datterceller som får halvparten så mange kromosomer som morcellen
3) *Arvestoffet i kjernen kopieres og fordeles i to datterceller*
4) Det blir dannet haploide celler
5) Arvestoffet i kjernen fordeles på fire datterceller som får halvparten så mange kromosomer som morcellen

14. Hva skjer i mitosens anafase?

1) Kopierte kromosomer blir pakket
2) Spindelen trekker kromosomene slik at de legger seg i et plan midt i cellen
3) At det blir dannet en ny membran rundt de to nye kjernene
4) *Kromosomene blir delt ved at identiske søsterkromatider trekkes mot hver sin pol*
5) At cellen deler seg i to

15. Hva er forskjellen på mitose og meiose?

1) Mitose gir opphav til haploide celler, mens meiose gir opphav til diploide celler
2) Det er to kjernedelinger i mitosen
3) Det forekommer ikke overkrysning i meiosen
4) Mitosen produserer kjønnsceller
5) *Mitose gir opphav til diploide celler, mens meiose gir opphav til haploide celler*

16. Hva er <u>DEN MEST KORREKTE</u> påstanden om hovedfunksjonen til DNA?

1) Utføre DNA-replikasjon
2) *Lagre og overbringe genetisk informasjon*
3) Kontrollere kjemiske prosesser i cellene
4) Forhindre mutasjoner
5) Utføre meiose

17. Hvor skjer transkripsjon og translasjon i eukaryote celler?

1)	Transkripsjon: ribosomer	Translasjon: ribosomer
2)	Transkripsjon: ribosomer	Translasjon: kjernen
3)	*Transkripsjon: kjernen*	*Translasjon: ribosomer*
4)	Transkripsjon: kjernen	Translasjon: kjernen
5)	Transkripsjon: kjernen	Translasjon: Mitrokondriet

18. Hva kalles prosessen der RNA dannes i cellekjernen?

1) Translokasjon
2) Translasjon
3) Transduksjon
4) *Transkripsjon*
5) Replikasjon

19. Ut fra DNA-tråden, hva blir mRNA-tråden?

DNA-tråden: ATGACT

1) UACTGA
2) UAGUGA
3) TAGTGA
4) *UACUGA*
5) TACTGA

20. Hva er translasjon?

1) *Ribosomene leser av mRNA og danner proteiner*
2) Polypeptider foldes til proteiner
3) DNA oversettes til RNA
4) DNA kopieres
5) Det dannes søsterkromatider i kromosomene

21. Den genetiske koden i DNA blir bestemt av

1) Antallet DNA-molekyler
2) *Rekkefølgen av nitrogenbasene*
3) Rekkefølgen av sukker og fosfat
4) Antall kromosomer
5) Rekkefølgen av aminosyrene

22. Antall baser i et kodon er

1) 1
2) 2
3) *3*
4) 4
5) 5

23. Funksjonen til tRNA er å

1) Syntetisere DNA
2) Syntetisere mRNA
3) Danne ribosomer
4) *Overføre aminosyrer til ribosomene*
5) Lage polypeptider

24. Ut i fra mRNA-tråden og tabellen, hva er aminosyresekvensen i korrekt rekkefølge?

mRNA-tåd: AUGAAUGCGCGCUAUUAG

Tabellen:

1. BASE	2. BASE U	2. BASE C	2. BASE A	2. BASE G	3. BASE
U	Fenylalanin	Serin	Tyrosin	Cystein	U
	Fenylalanin	Serin	Tyrosin	Cystein	C
	Leucin	Serin	Stopp	Stopp	A
	Leucin	Serin	Stopp	Tryptofan	G
C	Leucin	Prolin	Histidin	Arginin	U
	Leucin	Prolin	Histidin	Arginin	C
	Leucin	Prolin	Glutamin	Arginin	A
	Leucin	Prolin	Glutamin	Arginin	G
A	Isoleucin	Treonin	Asparagin	Serin	U
	Isoleucin	Treonin	Asparagin	Serin	C
	Metionin	Treonin	Lysin	Arginin	A
	Metionin	Treonin	Lysin	Arginin	G
G	Valin	Alanin	Asparagin-	Glycin	U
	Valin	Alanin	syre	Glycin	C
	Valin	Alanin	Glutamin-	Glycin	A
	Valin	Alanin	syre	Glycin	G

1)	Isoleucin	Asparagin	Alanin	Arginin	Serin	Stopp
2)	Metionin	Lysin	Alanin	Tryptofan	Tyrosin	Stopp
3)	Metionin	Asparagin	Valin	Arginin	Tyrosin	Tyrosin
4)	*Metionin*	*Asparagin*	*Alanin*	*Arginin*	*Tyrosin*	*Stopp*
5)	Metionin	Asparagin	Valin	Tryptofan	Serin	Stopp

Vedlegg 7: Resultater fra flervalgsprøven uten og med korrigering for gjetting

I dette vedlegget er det en oversikt som viser resultatene fra flervalgprøven til elevene. Vedlegget viser også en oversikt på hva resultatet hadde blitt, dersom det var blitt benyttet korrigering for gjetting. Disse metodene er nevnt i kapittel 2.6.

Ved bruk av metode 1, benyttes formel 2 som sier: Skår = Rett – Galt/(n-1).

Ved bruk av metode 2, blir det gitt 3 poeng for hvert korrekte svar, -2 poeng for hvert ukorrekte svar og 0 poeng for hver ubesvarte oppgave.

Ved bruk av metode 3, blir det gitt 2 poeng for hvert korrekte svar, 0 poeng for hvert ukorrekte svar og ½ poeng for hver ubesvarte oppgave.

	Resultat uten korrigering	**Resultat med korrigering ved metode 1**	**Resultat med korrigering ved metode 2**	**Resultat med korrigering ved metode 3**
TOTAL POENG	**24 poeng**	**24 poeng**	**72 poeng**	**48 poeng**
Elev A	22 poeng	21,5 poeng	62 poeng	44 poeng
Elev B	22 poeng	21, 5 poeng	62 poeng	44 poeng
Elev C	21 poeng	20,25 poeng	57 poeng	42 poeng
Elev D	21 poeng	20,25 poeng	57 poeng	42 poeng
Elev E	21 poeng	20,25 poeng	57 poeng	42 poeng
Elev F	21 poeng	20,25 poeng	57 poeng	42 poeng
Elev G	19 poeng	17, 75 poeng	47 poeng	38 poeng
Elev H	19 poeng	17, 75 poeng	47 poeng	38 poeng
Elev I	19 poeng	17, 75 poeng	47 poeng	38 poeng
Elev J	19 poeng	17, 75 poeng	47 poeng	38 poeng
Elev K	18 poeng	16,5 poeng	42 poeng	36 poeng
Elev L	17 poeng	15, 25 poeng	37 poeng	34 poeng
Elev M	17 poeng	15, 25 poeng	37 poeng	34 poeng
Elev N	13 poeng	10, 25 poeng	17 poeng	26 poeng
Elev O	12 poeng	9 poeng	12 poeng	24 poeng
Elev P	12 poeng	9 poeng	12 poeng	24 poeng
Elev Q	11 poeng	7,75 poeng	7 poeng	22 poeng
Elev R	ikke møtt			

Vedlegg 8: Kortsvarsprøven

I dette vedlegget er det en oversikt over de 5 kortsvarsoppgavene som ble gitt på kortsvarsprøven som elevene hadde fredag 23.01.09.

På denne prøven skal du svare så godt du kan på alle spørsmålene, og du må besvare alle spørsmålene. Prøven har en varighet på 45 minutter, og du skal bruke maks en A4 side per oppgave.

Tegn og forklar

1. Hvordan DNA er bygget opp?
2. Hvordan transkripsjon foregår?
3. Hvordan translasjon foregår?
4. Likheter og ulikheter mellom mitose og meiose.
5. Genregulering i eukaryote celler.

Vedlegg 9: Resultater fra flervalgsprøven, kortsvarsprøven og total karakter.

I dette vedlegget er det en oversikt fra resultater som elevene fikk på flervalgsprøven, kortsvarsprøven og total karakter. Hovedvekten for vurderingen var resultatene fra kortsvarsprøven, og resultatene fra flervalgsprøven ble benytter for justering av karakteren i forhold til resultatene fra kortsvarsprøven. Dersom elevene kun møtte opp til flervalgsprøven, var det ikke grunnlag for å sette karakter, siden hovedvekten for total karakter var resultatene fra kortsvarsprøven. Dersom elevene kun møtte opp til kortsvarsprøven, ble det gitt total karakter.

	Poengsum fra flervalgsprøven	**Karakter fra kortsvarsprøven**	**Total karakter**
Elev A	22 poeng	5	5 +
Elev B	22 poeng	ikke møtt	ikke karakter
Elev C	21 poeng	5 -	5
Elev D	21 poeng	4	4/5
Elev E	21 poeng	4 -	4/5
Elev F	21 poeng	2	2/3
Elev G	19 poeng	6 -	6/5
Elev H	19 poeng	5	5
Elev I	19 poeng	4 -	4/5
Elev J	19 poeng	ikke møtt	ikke karakter
Elev K	18 poeng	4 -	4 +
Elev L	17 poeng	4	4 +
Elev M	17 poeng	ikke møtt	ikke karakter
Elev N	13 poeng	ikke møtt	ikke karakter
Elev O	12 poeng	2/1	2
Elev P	12 poeng	2/1	2
Elev Q	11 poeng	4 -	4 -
Elev R	Ikke møtt	2	2

Vedlegg 10: Postintervjuspørsmål til lærer

I dette vedlegget er det en oversikt over de ulike forhåndsbestemte spørsmålene som ble stilt til læreren under postintervjuet.

1. Hva har du gjort etter datainnsamlingen i forhold til:
    a. Bruk av flervalgsoppgave som vurderingsform?
    b. Bruk av Test 2.0 ved underveisvurdering?
2. I hvilken grad har du benyttet Test 2.0 som et verktøy for å lage (flervalgs)oppgaver?
3. I hvilken grad har du forberedt elevene til skriftlig eksamen i Biologi 2?
    a. I forhold til metodebruk?
    b. I forhold til bruk av Test 2.0?
4. Er din kompetanse økt for å lage flervalgsoppgaver for å forberede elevene til skriftlig eksamen?