

Bruk av parprøve til vurdering i matematikk

*En casestudie av et parprøveforløp fra
oppgavedesign til retteskjema*

Jenny Engh



Masteroppgave
Lektorprogrammet
30 studiepoeng

Institutt for lærerutdanning og skoleforskning
Det utdanningsvitenskapelige fakultet

UNIVERSITETET I OSLO

Juni 2022

Bruk av parprøve til vurdering i matematikk

En casestudie av et parprøveforløp fra oppgavedesign til retteskjema

Jenny Engh

Masteroppgave

Lektorprogrammet

30 studiepoeng



Institutt for lærerutdanning og skoleforskning

Det utdanningsvitenskapelige fakultet

UNIVERSITETET I OSLO

Juni 2022

© Jenny Engh

2022

Bruk av parprøve til vurdering i matematikk: En casestudie av et parprøveforløp fra oppgavedesign til retteskjema

Jenny Engh

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

IV

Sammendrag

Problemstillingen for denne oppgaven er «*På hvilke måter og i hvilken grad utnytter en lærer parprøvens potensiale til å vurdere elevenes matematiske kompetanse?*». Studien er designet som en casestudie, og går i dybden på et parprøveforløp i en 2P-klasse, gjennom intervju, observasjon og oppgaveanalyse. Hovedfokuset er rettet mot lærerens intensjon med vurderingen og hvilke grep han tar i prøvedesignet, gjennomføringen og etterarbeidet for å ivareta parprøvens potensiale til å vurdere elevenes kompetanse. Oppgaven tar også for seg hvordan lærerens grep gjenspeiles i praksis, noe som gir grunnlag for å diskutere muligheter og begrensninger ved prøveformen. Datamaterialet består av transkripsjon fra intervju med læreren, lydopptak fra observasjon av tre elevpar som gjennomfører parprøven og oppgavesettet. Datamaterialet fra intervjuet er blitt analysert ved tematisk analyse (Braun & Clarke, 2017). I tillegg er alt datamaterialet analysert i lys av et rammeverk for matematisk kompetanse, MCRF+pd, designet for denne studien, basert på *Mathematical Competancy Research Framework* (Lithner et al., 2010) og kompetansen produktiv disposisjon inspirert av *Adding it up* (Kilpatrick et al., 2001). Funnene indikerer at slik parprøven er designet og gjennomført, har den potensiale til å aktivere kompetanse hos elevene, men at tilsvarende kompetanse ikke blir vist i like stor grad i elevenes besvarelser. Læreren vektlegger verdien av å se vurderingen som en læringssituasjon, samtidig er det et behov for at parprøven gir elevene mulighet til å vise sin kompetanse dersom parprøven skal utgjøre en vurdering for læring (Utdanningsdirektoratet, 2019). Læreren beskriver grep for å sørge for at elevene får vist ulike kompetanser, som å bruke egnede oppgaver og tilpasse nivået på oppgavene. Han legger også vekt på observasjon og samtale med elevene som et viktig grunnlag for den uformelle vurderingen av elevene. Studien impliserer at det kan være nyttig å se helhetlig på vurdering av matematisk kompetanse, og undersøke hvordan man kan utvikle en vurderingspraksis som benytter ulike vurderingsformer, og som i noe grad er frigjort fra kravet om å oversette informasjonen om elevenes kompetanse til en karakter. Dette vil være i tråd med forskriften (Forskrift til opplæringslova, 2006), som sier at elevene skal vurderes på flere og varierte måter, samt at en slik praksis trolig vil sørge for at elevene får aktivert og målt de essensielle kompetansene som utgjør matematisk kompetanse.

Forord

Her kommer forordet. Jeg har forordet for mine føtter. Forordet sier noe om meg og hvordan dette har gått. Jeg har sett frem til å skrive forord. En slags avslutning på det hele.

Innledningsvis.

Jeg vil benytte forordet til å takke for dette semesteret hvor jeg har fått skrive master. Det er ikke alle som får det.

Jeg vil benytte forordet til å takke min veileder Alexander Jonas Viktor Selling. Takk for at du har lært meg å skrive masteroppgave. Takk for gode innspill, raske svar, fleksibilitet og tilgjengelighet helt til siste slutt.

Jeg vil benytte forordet til å takke læreren og elevene som deltok i studien min. Takk for at jeg fikk et innblikk i deres praksis så jeg kunne realisere dette prosjektet. Takk for alt dere har lært meg. Jeg tar vare på det.

Jeg vil benytte forordet til å takke Øyvind Engh. Selvfølgelig. Takk for at du har lest korrektur. Jeg tar eget ansvar fra nå. Takk for at du alltid gjør deg fortjent til en plass i rulleteksten.

Oppgaven er ferdig. Vi er i havn og på land – heldigvis.

Vi. Takk.

Oslo, juni 2022

Jenny Engh

Innholdsfortegnelse

1	Innledning.....	1
1.1	Problemstilling og forskningsspørsmål	3
1.2	Avgrensning og utforming.....	3
2	Teori og tidligere forskning.....	6
2.1	Kompetanse	6
2.1.1	Kompetanserammeverk.....	9
2.2	Vurdering.....	11
2.2.1	Parprøven	13
2.2.2	Andre alternative vurderingsformer	14
2.3	Oppgaver	18
2.3.1	Klasseromsnormer.....	18
2.3.2	Mathematical task knowledge for teaching.....	19
2.3.3	Oppgavenivå.....	20
2.3.4	Perspektiver på oppgavers potensial	20
3	Metode.....	21
3.1	Valg av metode.....	21
3.1	Utvalg	23
3.1.1	Lærer	23
3.1.2	Tre par	25
3.1.3	Oppgaver	25
3.2	Datainnsamling - metoder og gjennomføring.....	26
3.2.1	Intervju	26
3.2.2	Observasjon	28
3.3	Datamaterialet og analysemetode	29
3.3.1	Rammeverk	30
3.3.2	Intervju	35
3.3.3	Observasjon	36
3.3.4	Oppgavene.....	37
3.4	Etiske betraktninger	38
3.5	Validitet, Reliabilitet og generaliserbarhet	40
3.5.1	Validitet og reliabilitet	40

3.5.2	Generaliserbarhet	42
4	Resultat.....	43
4.1	Intervju.....	43
4.1.1	Muligheter: hvorfor bruke parprøver og hvilke kompetanser kan parprøver måle? 43	
4.1.2	Hvilke grep tar Jan i prøvedesign og gjennomføring.....	47
4.1.3	Hvilke grep tar Jan i retteprosessen.....	49
4.2	Oppgavene	51
4.2.1	Analyse av oppgavesettet ved hjelp av rammeverket	51
4.2.2	Observasjon.....	53
5	Diskusjon og konklusjon.....	59
5.1	Diskusjon.....	59
5.1.1	Vurdering som læring.....	59
5.1.2	De essensielle komponentene i matematisk kompetanse	60
5.1.3	Oppgavedesign	61
5.1.4	Normer	63
5.1.5	Observasjon.....	64
5.1.6	Parsammensetning.....	65
5.1.7	Retteskjema	66
5.1.8	Læreren som forutsetning.....	66
5.1.9	Kompetanseprofiler	67
5.1.10	Helhetlig syn på vurdering	69
5.2	Konklusjon.....	69
5.2.1	Implikasjoner.....	71
	Vedlegg 1: Parprøven.....	77
	Vedlegg 2: Retteskjema	80
	Vedlegg 3: Informasjonsskriv og samtykkeskjema, lærer	81
	Vedlegg 4: Informasjonsskriv og samtykkeskjema, elev	85
	Vedlegg 5: Intervjuguide.....	89
	Figur 1 Kodeskjema for den tematiske analysen av intervju materialet	36

1 Innledning

Vurdering i matematikk bør speile den matematikken som er verdsatt og ansett som viktig å lære (Suurtamm et al., 2016). Dette innebærer at elevene skal vurderes i fagkunnskap og ferdigheter, men også i de kognitive prosessene som er nødvendige for å mestre matematikk – det vi samlet kaller matematisk kompetanse (Kilpatrick, 2020). I forskriften til opplæringsloven står det at «[e]levene skal ha mulighet til å vise sin kompetanse på flere og varierte måter» (Forskrift til opplæringslova, 2006). I tillegg til at ulike vurderingsformer er elevenes rett, finnes det mange argumenter som underbygger å variere vurderingsformene. Suurtamm et al. (2016) er blant de som oppfordrer lærere til å ta i bruk et bredt spekter av vurderingsformer. I takt med at læreplanene gjennom fagfornyelsen 2020 (Utdanningsdirektoratet, u.å.) dreies i retning av kompetanse- og dybdelæring må lærerne endre sin undervisnings- og vurderingspraksis slik at elevene får mulighet til å utvikle ønsket kompetanse. Matematikkundervisningen i Norge har vært preget av prøver som tester prosedyrekunnskap (Pettersen & Nortvedt, 2017), men disse prøvene er kanskje ikke lenger tilstrekkelig til å vurdere elevenes læring i matematikk – en prøve som kun krever prosedyrekunnskap gir ikke valid informasjon om elevenes evne til å se sammenhenger, noe som er en viktig del av dybdelæringen (Kunnskapsdepartementet, 2020). Niss (2007) uttrykker et behov for flere alternative vurderingsformer og begrunner det med at de vurderingsformene som brukes er så snevre at de kun egner seg til å vurdere et begrenset utvalg ferdigheter. Et annet aspekt som motiverer bruk av varierte vurderingsformer, er at elevene tilegner seg kunnskap på forskjellige måter og vurderingssituasjonene i matematikk må dermed ta høyde for elevenes ulike lærestiler slik at alle får mulighet til å vise kompetanse (Suurtamm et al., 2016; Watt, 2005). For det tredje bør det være samsvar mellom undervisning og vurdering (Suurtamm et al, 2016). Dersom elevene jobber mye med læringspartner i timene, noe som er blitt svært vanlig i matematikkundervisningen (Liljedahl, 2014), kan det argumenteres for at noen av vurderingssituasjonene bør gjennomføres med en partner.

Fagfornyelsen (Utdanningsdirektoratet, u.å.) og utspring av nye undervisningsformer krever altså at også vurderingsformene utvikles, og det er oppstått et behov for alternativer til den skriftlige matematikkprøven. De alternative vurderingsformene bør oppleves som meningsfulle, og aktivitetene bør legges til rette for at elevene får brukt forkunnskaper i nye situasjoner, de bør tillate matematisk kommunikasjon og oppfordre elevene til å ta sjanser

(Suurtamm, 2004). I Forskrift til opplæringslova (2006) under §3-3 *Vurdering i fag* står det at «[f]ormålet med vurdering i fag er å fremje læring og bidra til lærelyst undervegs, og å gi informasjon om kompetanse undervegs og ved avslutninga av opplæringa i faget».

Matematikk skiller seg fra andre fag ved at mange elever har en oppfatning om at oppgavene har én riktig løsning og at risikoen for å gjøre feil er stor; dette fører til at mange kvier seg for å ta sjanser når de jobber med matematikkoppgaver (Chinn, 2009). Data fra PISA 2012 viser at matematikkangst er negativt korrelert med matematikkprestasjoner (Foley et al., 2017). Å kunne utvikle gode vurderingssituasjoner, hvor elevene opplever mestring og vurdering *som* læring vil trolig være gunstig for elevenes motivasjon og holdninger til faget.

Det finnes mye forskning på hvordan gruppearbeid kan bidra til bedret prestasjon og motivasjon (f.eks. Leikin & Zaslavsky, 1999; Wiedmann et al., 2012). Jeg synes det er relevant å undersøke gruppevurderinger i matematikk, da jeg mener dette er en form som gir elevene gode læringsmuligheter i vurderingssituasjonen fordi de må bruke språket muntlig. Bjørk & Theodorsen (2019) gjennomførte en aksjonsstudie på egne 1P-elever, hvor elevene samarbeidet i par på alle matematikkprøver med unntak av to terminprøver gjennom et helt år. De fant at elevene følte seg mindre stresset, ble mer motivert og hadde bedre utholdenhet både når de jobbet på skolen og hjemme. Studien impliserte videre at lærerne hadde et behov for en måte å vurdere elevene ved hjelp av parprøve, som ga sikrere informasjon om hver enkelt elevs kompetanse. Både i Australia og Sverige er det gjort forskning på alternative vurderingsformer i matematikk, med liknende resultater som hos Bjørk & Theodorsen (2019). Watt (2005) fant at lærerne føler seg usikre når de skal vurdere alternative vurderingsformer, og at de mener vurderingen blir subjektiv. Frykedal & Chiriac (2011) fant at lærerne savner et profesjonelt språk med terminologi og modeller for vurdering av faglig kompetanse i gruppevurderinger.

Forsell et al. (2020) har forsøkt å få oversikt over forskningen på *vurdering* av gruppearbeid gjennom en «litterature review» av 83 artikler. Da de så på hva som blir målt i gruppevurderinger fant de at *innsats* og *evne til å jobbe som en gruppe* var sterkest representert, noe som er problematisk i matematikk da det ikke følger av læreplanen at innsats skal vurderes (Forskrift til opplæringslova, 2006). Dette bidrar til at gruppearbeidet blir brukt som et mål i seg selv, og ikke som et middel til å aktivere og måle kompetanse (Forsell et al., 2020; Frykedal & Chiriac, 2011). Dette tyder på at det må gjøres noe med måten gruppevurdering brukes dersom den skal kunne benyttes til å vurdere kompetanse i fag. Det er

svært viktig for vurderingens validitet at den måler det vi tror at den måler. I tillegg er det viktig slik at elevene kan få en konstruktiv tilbakemelding. Dersom de alternative vurderingsformene ikke er valide må de etterprøves med tradisjonelle vurderingsformer, noe som er problematisk med tanke på at de alternative vurderingsformene allerede er tidkrevende (Black & Wiliam, 2018; Suurtamm, 2004).

1.1 Problemstilling og forskningsspørsmål

På grunnlag av behovet for et bredt utvalg vurderingsformer og mulighetene gruppevurderinger i matematikk gir til vurdering *som* læring mener jeg det er interessant å undersøke bruk av parprøver som vurderingsform i matematikk. Det finnes allerede noe forskning på elevenes lærings- og trivselsutbytte; jeg vil ta tak i usikkerheten lærerne kan føle på i vurderingssituasjonen, og målet for denne oppgaven er å komme nærmere noen felles prosedyrer for bruk av parprøver til måling av elevenes kompetanse. Den overordnede problemstillingen jeg skal forsøke å besvare lyder

På hvilke måter og i hvilken grad utnytter en lærer parprøvens potensiale til å vurdere elevenes matematiske kompetanse?

Da det finnes lite forskning på «parprøvens potensiale» har jeg valgt å diskutere problemstillingen med utgangspunkt i lærerens oppfatning av parprøvens potensiale, og en analyse av oppgavesettet med tanke på hva som kreves av elevene for å løse oppgavene.

Problemstillingen skal besvares gjennom følgende forskningsspørsmål

1. hva mener læreren at er parprøvens potensiale?
2. hvilke grep tar læreren i prøvedesignet, under gjennomføringen av prøven og i etterarbeidet for å utnytte parprøvens potensiale?
3. hvilke kompetansekrav ligger i oppgavene i parprøven?
4. i hvilken grad gjenspeiles lærerens grep i prøvedesignet og under gjennomføringen av prøven?

1.2 Avgrensning og utforming

Oppgaven tar for seg et parprøveforløp med fokus rettet mot lærerens hensikt og valg. Denne delen av oppgaven er ivaretatt av forskningsspørsmål 1 og 2. Da målet er å bidra til å komme nærmere noen felles prosedyrer for bruk av parprøver til vurdering av elevenes kompetanse, er det også relevant å undersøke hvordan grepene gjenspeiles i prøvedesignet og under

gjennomføringen av prøven; dette for å kunne si noe om hvorvidt grepene fungerte etter hensikten og belyse eventuelle begrensninger. Denne delen av oppgaven er ivaretatt av forskningsspørsmål 3 og 4.

Oppgaven er utformet som en casestudie (Cohen et al., 2018). I forbindelse med dette prosjektet er det blitt gjennomført en parprøve i en 2P-klasse. Datamaterialet omfatter data fra intervju med læreren som har laget prøven, lydopptak fra observasjon av tre elevpar som løser oppgavesettet, og oppgavesettet. Datamaterialet fra intervjuet er blitt analysert ved tematisk analyse (Braun & Clarke, 2017). I tillegg er alt datamaterialet analysert i lys av et rammeverk for matematisk kompetanse, MCRF+pd, designet for denne studien, basert på *Mathematical Competency Research Framework* (Lithner et al., 2010) og kompetansen *productiv disposition* inspirert av *Adding it up* (Kilpatrick et al., 2001). Funnene danner grunnlaget for å besvare problemstillingen.

I kapittel 2 blir relevant teori og tidligere forskning presentert. Hovedtemaene er kompetanse, vurdering og oppgaver. Under *kompetanse* vil jeg presentere ulike forståelser av begrepet, og de rammeverkene for matematisk kompetanse som ligger til grunn for analysen i denne oppgaven. Under *vurdering* vil jeg presentere formålet med vurdering i matematikk og prinsippene for vurdering av, for og som læring. Parprøven innehar et element av læring i vurderingssituasjonen. Derfor er det relevant å inkludere tidligere forskning på gruppearbeid, uten vurderingsaspektet. I tillegg vil jeg presentere tidligere forskning som omhandler par- og gruppevurdering. Under *oppgaver* vil jeg presentere perspektiver på oppgavers potensiale og betydningen av hvordan og i hvilken kontekst de blir implementert; samt perspektiver på hva som kreves av læreren for at en aktivitet med oppgaver skal fungere etter hensikten.

I kapittel 3 blir forskningsdesignet presentert. Jeg begrunner valg av metoder og utvalg, og presenterer casen. Deretter beskrives datainnsamling og dataanalysen. Kapittelet avsluttes med etiske betraktninger og mine refleksjoner rundt oppgavens kvalitet i lys av begrepene validitet, reliabilitet og generaliserbarhet.

I kapittel 4 blir funn fra analysen presentert. Kapittelet er delt i to. Første del omhandler intervjuet, og funn fra intervjumaterialet presenteres tematisert etter forskningsspørsmål 1 og 2. Andre del omhandler oppgavene og gjennomføringen av prøven. I denne delen vil først funn fra analysen av oppgavesettet presenteres. Deretter ser jeg nærmere på tre av oppgavene

fra oppgavesettet, gjennom funn fra oppgaveanalysen, analysen av intervjumaterialet og analysen av observasjonsmaterialet.

I kapittel 5 blir funnene diskutert i lys av teorien og tidligere forskning, presentert i kapittel 2. Diskusjonen skal bidra til å besvare forskningsspørsmålene og problemstillingen. Kapitlet avsluttes med noen konkluderende betraktninger og forslag til implikasjoner for parprøvepraksis og videre forskning.

2 Teori og tidlige forskning

I dette kapitlet vil jeg gjøre rede for relevant teori og tidligere forskning knyttet til temaet for oppgaven. Jeg vil først ta for meg de mest sentrale begrepene i denne oppgaven; *kompetanse* og *vurdering*. Dette er begreper som er hyppig brukt i ulike kontekster både i praksis på skolene og i utdanningsforskning, og det blir derfor viktig å definere hva begrepene skal bety i denne oppgaven. Begrepene er nært knyttet til hverandre; dersom man skal gjøre en vurdering av elevers læring, har man behov for terminologi for å beskrive og gjenkjenne læring, og et rammeverk for matematisk kompetanse kan være et svar på dette behovet (Kilpatrick, 2020; Niss & Højgaard, 2019). Mot slutten av kapitlet vil jeg presentere teori og tidligere forskning som omhandler *oppgaver*. Oppgaver er en sentral del av matematikkopplæringen og brukes både til å engasjere elever i matematiske aktiviteter slik at de tilegner seg kompetanse og for å vurdere elevenes kompetanse (Pettersen, 2019).

2.1 Kompetanse

Kompetansebegrepet er ifølge Kilpatrick (2020) et av de mest flyktige konseptene i utdanningslitteraturen. I kunnskapsløftet 2020 er kompetanse definert som «å kunne tilegne seg og anvende kunnskaper og ferdigheter til å mestre utfordringer og løse oppgaver i kjente og ukjente sammenhenger og situasjoner. Kompetanse innebærer forståelse og evne til refleksjon og kritisk tenkning» (Kunnskapsdepartementet, 2020). Denne definisjonen gir en generell forståelse av begrepet, men ofte vil det være behov for en mer spisset og anvendbar definisjon om man skal bruke den aktivt i praksis i skolen eller til forskning. Derfor er det utviklet ulike rammeverk for matematisk kompetanse som forsøker å møte dette behovet. Før jeg går nærmere inn på noen av rammeverkene, vil jeg si litt mer om hvordan begrepet i seg selv kan forstås.

Blömeke et al. (2015) definerer kompetanse som «de latente kognitive og affektive trekkene som underbygger domenespesifikk atferd i ulike situasjoner» (s. 2, min oversettelse). De beskriver videre hvordan det ut fra definisjonen har vokst frem to ulike oppfatninger av hvordan kompetanse skal defineres og problematiserer denne dikotomien. Ifølge Blömeke et al. (2015) har vi på den ene siden en oppfatning av kompetanse som en sum av kognitive og affektive trekk som fører til en gitt atferd. Her er det høy grad av korrelasjon mellom kompetanse og atferd, og selv om de kognitive og affektive trekkene ligger latent og ikke er

mulig å observere, kan man trekke valide slutninger om aktørens kompetanse basert på observerbare prestasjoner. Den andre siden av dikotomien, betrakter kompetanse som ulike tilstander av kognitive evner og affektive trekk, som endrer seg gjennom handling. Dette innebærer at to personer som gjør samme prestasjon kan inneha ulike «kompetanseprofiler». Dette siste perspektivet stammer fra utdanningsforskningen hvor hensikten er å fremme *utvikling* av kompetanse på individ-, klasserom- og systemnivå. Vesentlig for dette perspektivet er at en aktørs underliggende trekk, som ligger til grunn for vedkommendes atferd, kan *påvirkes* av ytre stimuli, slik at det finnes en relasjon mellom undervisning og utvikling av kompetanse. Blömeke et al. (2015) sin problematisering av dikotomien bygger på oppfatningen om at kompetanse i realiteten ligger et sted mellom observerbar atferd og uobserverbare latente kognitive, konative og affektive trekk. De foreslår å betrakte kompetanse som en prosess, og de beskriver det som et kontinuum med mange steg. Diskusjonen er interessant i lys av vurdering fordi læreplanen i Norge bygger på kompetansebegrepet, og elevene skal vurderes ut fra oppnådd kompetanse. Sammenhengen mellom kompetanse og atferd er relevant for hvordan vi kan måle elevenes kompetanse ut fra prestasjon.

Psykologen Franz Emanuel Weinert beskriver kompetanse som et system av kunnskap, erfaringer og evalueringer (Weinert, 2001), som minner om Blömeke et al. (2015) sitt bilde av kompetanse som et kontinuum. Han skriver at kompetansebegrepet refererer til en kombinasjon av komponenter bestående av «kognitive, motivasjonelle, moralske og sosiale ferdigheter» (Weinert, 2001, s. 2433) som gjør en person i stand til å forstå og prestere i ulike situasjoner. I følge Weinert (2001) er konseptet uklart og mulighetene for å fjerne eller erstatte komponenter er lite utforsket. Det gjør at det er vanskelig å trekke linjer mellom ytre krav – for eksempel en oppgave – og et individs kompetanse. På en annen side trekker Weinert (2001) også frem at en del disipliner har en forståelse av kompetanse som et ganske spesifikt arsenal av evner som er nødvendige for å nå et mål – liknende hva Blömeke et al. (2015) beskriver.

På engelsk skiller det mellom begrepene «competence» og «competency». Det førstnevnte favner bredt, mens «competency» betegner komponentene som utgjør «competence». Disse komponentene kan være kognitive, konative, affektive eller motivasjonelle. På norsk bruker vi begrepet kompetanse til å beskrive begge disse nivåene. Vi sier for eksempel at vi har

rammeverk for matematisk kompetanse (competence), og at rammeverket er bygget opp av delkompetanser som for eksempel kommunikasjonskompetanse (competency).

Weinert (2001) skriver at jo mer generelt et kompetansekonstrukt er, desto mindre bidrag gir konstruktet til å løse utfordringer innenfor et felt. Det er blant annet behov for konstrukt som beskriver fagspesifikk kompetanse, og ikke bare generell og overordnet kompetanse, som for eksempel å beherske morsmålet sitt. Matematisk kompetanse er, som nevnt, forsøkt beskrevet og ordnet i rammeverk. Disse har noe ulik oppbygning og vektlegger ulike elementer; noen inkluderer både spesifikk matematisk fagkunnskap - som algebra og geometri – samt de kognitive prosessene for å jobbe med matematikk, mens andre beskriver bare prosessene (Kilpatrick, 2020). I følge Niss & Højgaard (2019) strekker kompetansebegrepet seg gjennom og forbi spesifikke matematiske fagområder og kan derfor ikke defineres ved hjelp av disse områdene. Videre forklarer de at matematisk kompetanse likevel er nært knyttet til de matematiske fagområdene fordi matematisk kompetanse blir aktivert i situasjoner som involverer matematisk fagkunnskap (Niss & Højgaard, 2019). Niss et al. (2016) skriver at det er lite hensiktsmessig å diskutere om en definisjon av kompetanse er riktig eller gal, men at det som er vesentlig er hvor godt konstruktet egner seg til det formålet vi ønsker under de gjeldende forholdene.

Når måling av elevenes kompetanse i en parprøvesetting i norsk skole skal diskuteres, er det nærliggende å gå til rammeverkene som ligger til grunn for de nasjonale læreplanene i Norge; deriblant er ideene i KOM-prosjektet (Niss & Højgaard Jensen, 2002) og *Adding it up* (Kilpatrick et al., 2001) bærende (Boesen et al., 2014; Valenta et al., 2015). Disse rammeverkene er i stor grad rettet mot å formulere mål for utvikling på policy-nivå. I denne oppgaven er behovet et kompetanserammeverk til bruk i analyse. Lester (2009) anbefaler å sammenlikne ulike teoretiske perspektiver og undersøke hvordan de belyser fenomenet man studerer, og så utvikle rammeverk basert på flere teoretiske perspektiver og kilder. Denne metoden kalles *bricoleurs* og ifølge (Lester, 2009) vil den både styrke forståelsen av fenomenet og øke nytten av forskningen. KOM-prosjektet (Niss & Højgaard Jensen, 2002) og *Adding it up* (Kilpatrick et al., 2001) er som nevnt relevante fordi ideene fra disse rammeverkene ligger til grunn i norsk skole og beskriver godt det vi i Norge – og mange andre land – anser som viktig i matematikkundervisningen (Boesen et al., 2014; Valenta et al., 2015). I tillegg vil jeg introdusere rammeverket, *Mathematical Competency Research Framework* (MCRF) utviklet av (Lithner et al., 2010). MCRF er basert på KOM-prosjektet

(Niss & Højgaard Jensen, 2002), *Adding it up* (Kilpatrick et al., 2001) og NCTM Standards (NCTM, 2000), og laget til bruk i empirisk forskning (Lithner et al., 2010). MCRF danner også grunnlaget for rammeverket jeg benytter i denne oppgavens analyse.

2.1.1 Kompetanserammeverk

På 1990-tallet var den voksende trenden å fremheve relasjonen mellom matematisk innhold og matematisk praksis, og prinsippet «knowing mathematics is ‘doing’ mathematics» ble lagt til grunn (Boesen et al., 2014). Dette brøt med det som da var regnet som tradisjonell undervisning av matematikk – ofte assosiert med innlæring av prosedyrekunnskap – og åpnet opp for et rikere bilde av hva det vil si å mestre matematikk. I lys av dette ble de kognitive prosessene trukket frem og vektlagt. Det ser man blant annet i NCTMs Principles and Standards for School Mathematics fra 2000 (NCTM, 2000), hvor fem av standardene omhandler matematiske fagområder (Number and Operations, Algebra, Geometry, Measurement, Data Analysis and Probability), mens de resterende fem standardene omhandler prosess (Problem Solving, Reasoning and Proof, Communication, Connections, and Representation).

Adding it up

Prinsippet «knowing mathematics is ‘doing’ mathematics» er også grunnlaget for et rammeverk utviklet av Kilpatrick et al. (2001) presentert i *Adding it up: Helping Children Learn Mathematics*, hvor det er beskrevet hvilke kognitive endringer hos elevene de ønsker å fremme gjennom sitt rammeverk. Rammeverket består av de fem komponentene; conceptual understanding, procedural fluency, strategic competence, adaptive reasoning og productive disposition; som sammen utgjør det de kaller «mathematical proficiency». Proficiency tilsvarer kompetansebegrepet (Kilpatrick, 2020; Kilpatrick et al., 2001). I dette rammeverket legges det særlig vekt på at komponentene skal betraktes som sammenvevde og i samspill med hverandre; det kreves evne til å se sammenhenger for å oppnå dyp forståelse, og disse sammenhengene er også avgjørende for om elevene klarer å bruke kunnskapen sin til å løse problemer (Kilpatrick et al., 2001). Komponenten productive disposition - heretter kalt produktiv disposisjon - dekker det affektive aspektet ved matematisk kompetanse, og denne komponenten gjør at dette rammeverket skiller seg fra rammeverk som ikke inkluderer det affektive (Kilpatrick, 2020).

KOM-prosjektet

I Danmark ble det på samme tid opprettet et prosjekt kalt KOM-prosjektet, hvor ideen var at det skulle utvikles et grunnlag for matematikkopplæringen som var basert på matematisk kompetanse, heller enn en liste med matematiske fagområder. Resultatet ble et rammeverk for matematisk kompetanse bestående av åtte komponenter; tankegangskompetanse, problemløsningskompetanse, modelleringskompetanse, resonnementskompetanse, representasjonskompetanse, symbol- og formalismekompetanse, kommunikasjonskompetanse, hjelpemiddelkompetanse (Niss & Højgaard Jensen, 2002). I likhet med rammeverket i Kilpatrick et al. (2001), er også KOM-prosjektets komponenter regnet som sammenvevde og i samspill (Niss & Højgaard Jensen, 2002). De fire første komponentene omhandler å stille og besvare spørsmål om og med matematikk, mens de fire neste komponentene omhandler anvendelse av matematisk språk og verktøy (Niss, 2003).

MCRF

NCTM, *Adding it up* og KOM-prosjektet er i stor grad ment å virke på policy-nivå. Lithner et al. (2010) har utviklet et rammeverk, *Mathematical Competency Research Framework* (MCRF), basert på disse tre, men med formål å kunne brukes til empirisk forskning. Dersom et rammeverk skal være et verktøy for kategorisering av data er det hensiktsmessig at de ulike komponentene er avgrenset, og overlapper i så liten grad som mulig, slik at et fenomen ikke havner i flere kategorier (Lithner et al., 2010). MCRF er utviklet i tråd med dette prinsippet, og består av de seks komponentene (1) problemløsningssevne, (2) resonneringsevne, (3) prosedyresevne (4) representasjonsevne, (5) connectionsevne og (6) kommunikasjonsevne. Selv om det er et poeng at komponentene er tydelig avgrenset, betyr likevel ikke det at komponentene er uavhengige, og utviklingen av kompetanse er fortsatt et samspill mellom komponentene (Lithner et al., 2010). Boesen et al. (2014) har benyttet MCRF til analyse av datamateriale fra intervju og observasjon. MCRF er også blitt brukt til oppgaveanalyse (Boesen et al., 2018) og til analyse av datamateriale fra podcast (Bergqvist, 2013).

Produktiv disposisjon

Bjørk & Theodorsen (2019) fant i sin studie, som jeg kommer tilbake til i 2.2.1, at parprøver har en positiv effekt på elevenes motivasjon og utholdenhet. Disse aspektene dekkes av kompetansekompenten produktiv disposisjon, og det er derfor relevant å vie noe ekstra

plass til denne komponenten. Som nevnt finnes produktiv disposisjon i *Adding it up* (Kilpatrick, 2001) og dekker det affektive aspektet ved matematisk kompetanse. Produktiv disposisjon innebærer evnen til å se matematikk som noe meningsfullt og nyttig, evnen til å se seg selv som en som kan lære og bruke matematikk effektivt, og evnen til å forstå at hardt arbeid kan utvikle og øke ens egen matematiske kompetanse (Kilpatrick et al., 2001). Komponenter finnes ikke i KOM-prosjektets rammeverk eller MCRF, men den tyske psykologen Franz Emanuel Weinert definerer kompetanse generelt, altså ikke fagspesifikt, og i tillegg til kognitive evner og ferdigheter inkluderer han «the associated motivational, volitional and social readiness» (Niss et al., 2016, s. 622). Dette handler nettopp om disposisjonelle og affektive trekk som for eksempel holdninger, følelser og vilje (Niss et al., 2016). I NCTM-standarden fra 1989 finner vi også mål som involverer disposisjonelle og affektive trekk; (1) elevene skal lære å verdsette matematikk og (2) elevene skal lære å bli selvsikre og stole på sin evne til å gjøre matematikk (Niss et al., 2016).

2.2 Vurdering

Vurdering i matematikk handler ifølge Niss (1993) om å bedømme elevenes kompetanse og prestasjoner og sier altså noe om utbyttet av undervisningen. Vurdering foregår på ulike nivåer, og vi skiller ofte mellom storskalavurdering og klasseromsvurdering (Suurtamm et al., 2016). I et storskalaperspektiv gjøres vurderinger blant annet for å sikre kvalitet i undervisningen ved å sammenlikne resultater fra ulike skoler eller ulike land (Suurtamm et al., 2016). I Norge er eksamen, PISA og TIMSS eksempler på denne type vurderinger. I et klasseromsperspektiv er hovedformålet med vurderingene å samle informasjon om elevenes læring blant annet for å gi elevene et bilde av hvordan de ligger an. Klasseromsvurderingene i norsk skole er inndelt i sluttvurdering og underveisvurdering, og elevene har krav på begge disse. Sluttvurderingen skal gi informasjon om elevens kompetanse etter opplæringen i faget, det vil si ved terminslutt (Forskrift til opplæringslova, 2006). Underveisvurdering er, som begrepet indikerer, vurderinger som gjøres underveis i opplæringen. I Forskrift til opplæringslova (2006) § 3-3. Vurdering i fag, står det at underveisvurderingen skal «fremje læring, tilpasse opplæringa og auke kompetansen i fag».

I følge Suurtamm et al. (2016) bør det sentrale formålet med både storskalavurdering og klasseromsvurdering være å støtte og fremme elevenes læring. I et storskalaperspektiv handler dette om å bidra til endringer på systemnivå for å øke elevens muligheter for læring, mens

underveisvurderingen i klasserommet foregår på individnivå og skal fremme læring for hver enkelt elev. Black & Wiliam (1998) skriver at formativ vurdering øker læringsutbyttet for elevene. Formativ vurdering er vurdering som gir eleven informasjon om egen kompetanse, læringsprosesser og målene for opplæringen, på en måte som gjør at eleven blir i stand til å oppnå progresjon (Suurtamm et al., 2016). Motsatsen er *summativ* vurdering, hvor formålet er å gjøre opp status, ofte i form av en karakter eller annen plassering på skala. Vurdering i seg selv er ikke formativ eller summativ; det er hvordan vi velger å bruke den innhentede informasjonen som avgjør om vurderingen er formativ eller summativ (Black & Wiliam, 2018). I 2010 ble det av Utdanningsdirektoratet startet et arbeid med å sørge for mer «læringsfremmende vurderingspraksis og vurderingskultur» gjennom satsningen «Vurdering for Læring» (VfL) (Utdanningsdirektoratet, 2019). VfL er i denne sammenheng synonymt med formativ vurdering.

Et tredje aspekt ved vurdering er vurdering *som læring* (VsL). Begrepet er sentralt blant annet i Mok (2013) sitt rammeverk *Self-directed learning oriented assessment* (SLOA) som er ment å støtte utvikling av en lærings- og vurderingspraksis som i stor grad styres av den som skal lære. Hun forklarer VsL som en metakognitiv vurderingsprosess hvor eleven aktivt setter egne læringsmål, velger læringsstrategier og overvåker sin egen læring i vurderingssituasjonen, og som resultat av dette tilegner seg ny forståelse, ser sammenheng mellom ny informasjon og forkunnskaper og utvikler nye strategier og mål. Mok (2013) beskriver VsL som det “ultimate mål” (s. 8) for VfL. Yan & Yang (2022) beskrivelse av en VsL-oppgave er i tråd med Mok (2013) sin definisjon; oppgaven må utgjøre en læringssituasjon som aktiverer mer enn å reprodusere forkunnskaper og sørge for at eleven får mulighet til å utvikle metakognisjon og selvregulering, slik at de kan oppdage hva de trenger for progresjon og hvordan de kan møte disse behovene. Fordi elevene er aktivt deltakende i selve vurderingsarbeidet, er ikke lenger læreren den eneste kilden til tilbakemelding og elevene utgjør en viktig ressurs for seg selv og andre elever. Yan & Yang (2022) skriver videre at denne måten å se vurdering på vil hjelpe på noen av utfordringene som følger med innføringen av nye vurderingsformer, deriblant den økte arbeidsmengden for lærerne. Black & Wiliam (2018) ser også det gunstige i å kombinere læring og vurdering; de trekker frem at det for eksempel er tidkrevende å vurdere elevene gjennom autentiske oppgaver og at det kan være vanskelig å legitimere denne tidsbruken om elevene ikke lærer noe underveis. Samtidig ytrer de en skepsis og et behov for å begrense formativ vurdering til å omhandle nettopp *vurdering*. Black & Wiliam (2018) mener at faren ved å akseptere en vid definisjon av

formativ vurdering, som inkluderer vurdering som læring, er at begrepet blir en “teori om alt”(s.552) uten tyngde.

Vurdering i matematikk bør speile den matematikken som er verdsatt og ansett som viktig å lære (Suurtamm et al., 2016). Dette innebærer at vurderingen må dekke matematisk kompetanse i tillegg til matematiske fagkunnskaper. I følge Suurtamm et al. (2016) er vurderingspraksisen mest effektiv når den samsvarer med undervisningspraksisen. De oppfordrer også til å bruke ulike vurderingsformer og gi elevene varierte muligheter til å vise sin kompetanse. Dette er i tråd med Forskrift til opplæringslova (2006) som sier at elevene skal ha fått mulighet til å vise sin kompetanse på flere og varierte måter. I 1993 skrev Niss (1993) at formål, teori og praksis i matematikkundervisningen hadde utviklet seg de seneste tiårene, mens han påpekte mangel på tilsvarende utvikling når det gjaldt vurdering. I 2007, drøyt fem år etter at kompetanserammeverkene til Kilpatrick og KOM-prosjektet ble publisert og fikk sin innvirkning på matematikkopplæringen, trekker han frem de samme utfordringene (Niss, 2007). Niss (2007) påpeker at det mangler alternative vurderingsformer, og at de vurderingsformene som brukes er så snevre at de kun egner seg til å vurdere et begrenset utvalg ferdigheter. Det har skjedd en utvikling i forskningen på vurdering, men han mener lite fremgang er gjort innen forskning på vurderingsformer som dekker det han kaller de essensielle komponentene i matematisk kompetanse; deriblant kommunikasjon, bygge argumenter og formelle bevis, beherske ulike representasjoner og veksle mellom ulike representasjoner.

2.2.1 Parprøven

Denne oppgaven skal undersøke parprøver som vurderingsform i matematikk. Jeg ønsker å definere en parprøve slik, «en vurderingssituasjon hvor to elever løser en oppgave sammen og leverer et felles skriftlig produkt», tilsvarende beskrivelsen av parprøve i Bjørk & Theodorsen (2019). Vurderingen skal tilfredsstillende det sentrale formålet: støtte og fremme læring i matematikk. For å oppnå formålet må parprøven være et middel for elevene for å vise matematisk kompetanse. Dersom prøven skal være valid må lærer vite hvilke kompetanser oppgavene krever og i tillegg kunne gjenkjenne tegn på oppnådd kompetanse i besvarelsen. Skal vurderingen være formativ må informasjonen som hentes inn også brukes til å hjelpe eleven til videre læring. Jeg vil anse parprøver som en alternativ vurderingsform i matematikk fordi den skiller seg fra den tradisjonelle, individuelle, skriftlige matematikkprøven. Det

sosiale aspektet gjør at elevene får mulighet til å kommunisere, noe jeg mener er en god forutsetning for vurdering som læring.

Tidligere forskning på parprøver

Bjørk & Theodorsen (2019) gjennomførte et parprøveprosjekt ved en videregående skole i Norge. Gjennom hele skoleåret var alle prøver for 1P, med unntak av to terminprøver og en kartleggingsprøve, gjennomført som parprøver. Målet var å «fremme vurdering for læring, øke elevenes motivasjon og minske stress» (Bjørk & Theodorsen, 2019). Prosjektet var også motivert av ønsket om å «skape en sosiokulturell vurderingspraksis, som er mer i tråd med strømningene i skolen ellers». Prosjektet var en lærerdrevet aksjonsstudie. Det ble gjennomført to spørreundersøkelser for elevene som deltok ($n_1 = 117$ og $n_2 = 102$) og gruppeintervju med lærerne som deltok. De to lærerne ved skolen som ledet studien og som skrev rapport i etterkant, deltok også med sine egne klasser. Parene var satt sammen av elever på omtrent likt faglig nivå, og elevene fikk ny partner til hver prøve i den grad det lot seg gjøre. Gjennom prosjektet fant de at elevene var mer motiverte til å jobbe i timene og hjemme, at de ikke ga opp like lett som før, og at parprøver fører til mindre stress (Bjørk & Theodorsen, 2019). Når det gjelder vurdering for læring oppga elevene at de lærer noe nytt under parprøven, at samarbeidspartneren hjelper dem å forstå, og at de opplever et positivt læringsutbytte – dette gjelder også elevene på lavt faglig nivå, som i liten grad mestrer å delta i den matematiske samtalen. At parprøven kan gi økt motivasjon og utholdenhet bidrar også til at vurderingsformen blir en vurdering *for* læring. En utfordring de fant – som er særlig relevant for denne oppgaven – var at lærerne opplevde usikkerhet ved om karakterene elevene fikk på parprøvene reflekterte elevenes kompetanse (Bjørk & Theodorsen, 2019). Usikkerheten skyldtes blant annet at elevene fikk lavere karakter på terminprøven enn på underveisvurderingene, og differansen mellom karakterene var større enn lærerne til vanlig opplevde.

2.2.2 Andre alternative vurderingsformer

Alternative vurderingsformer er også undersøkt av Watt (2005). I 2005 spurte hun, gjennom en spørreundersøkelse, 60 australske lærere om deres holdninger til alternative vurderingsformer i matematikk (muntlig, praktisk, observasjon, elevjournaler, egenvurdering, foreldrevurdering). De 60 lærerne representerte åtte offentlige skoler og tre privatskoler,

deriblant to jenteskoler og en gutteskole; alle i Sydney. Spørreskjemaet besto av tre spørsmål med svaralternativer (eks: på en skala fra 1-5, hvor godt egner tradisjonelle matematikkprøver seg til å vurdere elevenes kompetanse?), og fem spørsmål hvor lærerne kunne fortelle fritt (eks: fra denne listen med alternative vurderingsformer, forklar hvorfor/hvorfor ikke du benytter deg av disse). Watt (2005) fant at lærerne syntes skriftlige, individuelle matematikkvurderinger egnet seg godt for elever i aldersgruppen 11-12 år, mindre godt for aldersgruppen 9-10 år, mens de syntes denne typen prøver egnet seg dårligst for aldersgruppen 7-8 år. Hun påpeker at denne praksisen kan komme på bekostning av det økende behovet for å vurdere matematisk kompetanse på høyt nivå etter hvert som elevene når høyere trinn - slik også Niss (2007) og Suurtamm et al. (2016) peker på. Watt (2005) fant at hovedårsaken til at lærerne er restriktive med bruken av alternative vurderingsformer er at de blir usikre og mener at vurderingen blir for subjektiv.

Suurtamm (2004) undersøker i *Developing Authentic Assessment* fem kanadiske læreres opplevelser med bruk av “autentisk vurdering” på fire ulike ungdomsskoler. Autentisk vurdering, som tilsvarende begrepet alternativ vurdering, er her ment å beskrive vurdering som engasjerer elevene i oppgaver som er meningsfulle og som har fellestrekk med læringssituasjoner. Dette innebærer at aktivitetene skal oppfordre elevene til å “ta sjanser, tillate matematisk kommunikasjon og sørge for at elevene tar i bruk forkunnskaper i ukjente situasjoner” (Suurtamm, 2004, s. 499, min oversettelse). Vurderingspraksisen innebærer også å bruke ulike vurderingsformer for å dekke alle aspekter ved matematikken. Målet med studien til Suurtamm (2004) var å beskrive hvordan de fem lærerne gjennomfører autentisk vurdering og avdekke teoretiske og praktiske utfordringer. Det ble gjennom ett år gjort individuelle intervjuer, gruppeintervjuer og observasjoner, samt at vurderingsmateriell og noen elevbesvarelser var inkludert i datamaterialet. I tillegg førte lærerne loggbøker hvor de beskrev sin praksis og refleksjoner rundt gjennomføringen. Tre hovedutfordringer ble trukket frem. For det første føler lærerne seg dratt mellom kravet om å dekke pensum og ønsket om å skape læringssituasjoner som gir elevene mulighet til å oppnå dypere forståelse, jobbe med sammenhenger og reflektere over eget arbeid (Suurtamm, 2004). Lærerne trekker frem eksamen som en medvirkende årsak til at de føler seg bundet; elevene må forberedes til eksamen som gjennomføres tradisjonelt tilsvarende såkalte “paper and pencil”-tester, samtidig som undervisningspraksisen skal endres i retning mer problemløsning, mer samarbeidslæring og fokus på kommunikasjon. Andre utfordring dreide seg om at lærerne følte seg isolerte i arbeidet med autentisk vurdering, og at mangelen på et fellesskap førte til mindre konsistent

vurderingspraksis blant lærerne på skolen (Suurtamm, 2004). Lærerne rapporterer dette som problematisk fordi store forskjeller i vurderingspraksis kan skape usikkerhet rundt troverdigheten til vurderingene. Tredje utfordring dreide seg om karaktersetting. Autentisk vurdering består av mange uformelle vurderingssituasjoner, samt at tilbakemeldinger blir gitt i ulik form, og lærerne opplevde det som vanskelig å overføre dette og få det til å samsvare med tradisjonell måte å rapportere elevenes måloppnåelse (Suurtamm, 2004). Flere av dem var enige om at elevene og foreldrene fikk mer ut av en utfyllende rapport om elevens prestasjoner, samtidig som det var krav om sluttvurdering med karakter. På tross av utfordringene ønsket lærerne å fortsette å bruke autentiske vurderingsformer fordi de hadde en klar oppfatning om at det ville komme elevene til gode. De ønsket også bedre muligheter og flere anledninger for lærere til å snakke sammen om praksis, lære nye vurderingsformer og øve på å gjennomføre dem. Flere av lærerne foreslo en ny læreplan med mer rom til vurderingsformer som støtter den matematikken de anså som viktig (Suurtamm, 2004). Lærerne i Watt (2005) sin studie ytret et tilsvarende ønske.

Kommunikasjon og muligheten til å samarbeide er viktige aspekter ved parprøven. De samme aspektene finner vi ved gruppevurderinger og det er derfor relevant å inkludere noen funn fra forskning på gruppearbeid. Frykedal & Chiriac (2011) var interessert i vurdering av gruppearbeid og gjorde fokusgruppeintervjuer med 11 svenske lærere for å få mer kunnskap om hvordan lærere bruker gruppearbeid som vurderingssituasjon. Lærerne underviste elever i alderen 11-16 og datainnsamlingen ble gjort våren 2009. De fant at lærerne savner et profesjonelt språk med terminologi og modeller for vurdering av faglig kompetanse i gruppevurderinger. Dette samsvarer med det Watt (2005) fant - lærerne opplever vurderingspraksisen som usikker og subjektiv.

Forsell et al. (2020) har forsøkt å få oversikt over forskningen på vurdering av gruppearbeid gjennom en «litterature review» av 83 artikler. De så på hva forskningen sier om (a) hensikten med gruppearbeid, (b) hva som blir vurdert når man vurderer gruppearbeid, (c) metoder for vurdering av gruppearbeid (d) effekter og konsekvenser av vurdering av gruppearbeid, og (e) kvaliteten på vurderingen av gruppearbeid. De fant blant annet at forskningen i stor grad omhandler sosiale ferdigheter og gruppeprosess. Da de så på *hva* som blir vurdert i vurdering av gruppearbeid, i lys av kategoriene «contribution», «knowledge» og «product», fant de at contribution var tydelig representert i alle studiene, mens knowledge kun dukket opp noen få ganger. Dette innebærer blant annet at det ikke er prosessen som blir vurdert, men elevenes

bidrag til prosessen (Forsell et al., 2020). Dette er problematisk i matematikk da det ikke følger av læreplanen at innsats skal vurderes (Forskrift til opplæringslova, 2006). Forsell et al. (2020) finner generelt at de fleste får en gruppekarakter som er høyere enn det nivået de normalt ligger på. Dette samsvarer med funnene til Bjørk & Theodorsen (2019).

Gruppesammensetning

Gruppesammensetning påvirker elevenes mulighet til å prestere, både når de jobber i par og i større grupper. Wiedmann et al. (2012) skriver at for best læringsutbytte bør det være minst én elev på «høyere faglig nivå». Høyt presterende elever får best utbytte av å være på gruppe med andre høyt presterende. Lavt presterende og middels presterende har best utbytte av å være på blandede grupper. Lavt presterende kan nå sin proksimale utviklingszone i samhandling med den «kompetente andre», mens den middels presterende får utbytte av å forklare (Wiedmann et al., 2012). Leikin & Zaslavsky (1999) konstaterer at lavt presterende elever kan bli mer motivert og få økt tro på at de klarer å løse oppgaven dersom de får jobbe med høyere presterende. Elevers selvtillit spiller inn på hvor aktive de er i interaksjonen når de løser matematiske problemer, og et parsamarbeid mellom to lavt presterende elever kan fort stagnere (Samuelsson & Frykedal, 2014). Likevel er det viktig at ikke nivåforskjellen på gruppe medlemmene er for stor da de lavt presterende melder seg ut og de høyere presterende ikke får noe ut av å forklare på det nivået de lavt presterende har behov for.

Liljedahl (2014) studerer effekten av at elevene er satt sammen i *synlig* tilfeldige grupper i matematikkundervisningen. Studien ble gjennomført i en 10. klasse vest i Canada. Klasse- og skolemiljøet var preget av at elevene tilhørte en av to omtrent like store etniske grupper, som læreren i studien og Liljedahl opplevde at eksisterte side om side, nesten uten kontakt eller samhandling. Datamaterialet består av notater fra observasjoner og samtaler med elevene, samt materiale fra intervju med lærer og et utvalg elever. Liljedahl (2014) fant blant annet at de sosiale barrierene forsvant og at elevene aksepterte å jobbe med hvem som helst fra klassen. Avgjørende for dette var at de oppdaget at gruppene var tilfeldig sammensatt. Elevene ble også mer engasjerte i matematikk, mindre avhengige av læreren og stolte mer på seg selv og medelevenes kompetanse. Liljedahl (2014) skriver at når læreren setter sammen parene er det ofte for å oppnå et mål, og at dette målet ofte er et annet enn elevene sitt mål. Det kan gjøre at elevene forsøker å inngå avtaler eller gjøre endringer på gruppene så også deres interesser blir ivaretatt. Med tilfeldige sammensatte grupper - som elevene vet at er

tilfeldig - kan man oppnå at det ikke er rom for forhandlinger fordi ingens interesser er blitt tatt hensyn til, samtidig som all undring og mistenksomhet rundt gruppesammensetningen blir borte og fokuset rettes mot matematikken.

2.3 Oppgaver

Parprøven er en skriftlig prøve bestående av oppgaver, hvor konteksten oppgavene skal løses i skiller seg fra en tradisjonell individuell matematikkprøve. I dette delkapittelet blir det presentert noen perspektiver på oppgaver, refleksjoner rundt hvordan man avgjør en oppgaves potensiale, hva kontekst har å si, og hva som kreves av læreren for at en oppgave skal fungere etter hensikten.

Matematikkoppgaver blir brukt til ulike formål og vi skiller ofte mellom oppgaver som brukes i undervisning og som har til hensikt å fremme læring, og oppgaver som brukes i vurderingssituasjoner og som har til hensikt å gi informasjon om elevens læring (Pettersen, 2019). Gjennom arbeid med oppgaver får elever mulighet til å utvikle matematisk kompetanse fordi de engasjerer seg i matematisk tenkning (Pettersen, 2019). Konteksten oppgaven skal løses i vil kunne endre i hvilken grad oppgaven er kognitivt utfordrende. Konteksten kan blant annet innebære hva eleven kan fra før, elevens motivasjon og interesser, lærerens organisering av klassen og tilgang på instrumentell og emosjonell støtte.

Det er altså viktig at lærer velger oppgaver som gir elevene mulighet til å nå høyere ordens tenkning, men også implementerer oppgavene på en måte som gjør at elevene opprettholder nivået på tenkningen. For å lage gode oppgaver som stimulerer til utvikling av kompetanse og vurdering av kompetanse må kompetansebegrepet integreres i lærerens praksis. En vanlig mekanisme er at lærere plukker opp de store linjene i en reform og tolker disse med hensyn på egne holdninger i stedet for å bruke reformen til å endre holdninger og undervisning (Boesen et al., 2014).

2.3.1 Klasseromsnormer

Yackel & Cobb (1996) skriver om hvordan sosiomatematiske normer påvirker elevenes mulighet til å lære. Sosiomatematiske normer er normer som skiller seg fra sosiale normer ved at de er spesifikke for den matematiske konteksten (Yackel & Cobb, 1996). Normene påvirker hvordan elevene jobber med oppgaver, for eksempel er det en sosial norm at du skal

begrunne svaret ditt, mens de sosiomatematiske normene sier noe om hva som er tilstrekkelig matematisk argumentasjon. Hva som er normativt i et klasserom avgjøres av læreren og elevenes holdninger, forventninger og mål (Yackel & Rasmussen, 2002) og kan være vanskelige å endre. Peter Liljedahl har forsket på etablering og utvikling av det han kaller «thinking classrooms» (eks: Liljedahl, 2016; Liljedahl, 2018). Et tenkende klasserom er et klasserom hvor elevene tenker individuelt og sammen, og konstruerer kunnskap og forståelse gjennom aktivitet og diskusjon – i motsetning til et ikke-tenkende klasserom hvor elevene gir opp i møte med matematiske utfordringer (Liljedahl, 2016). En forutsetning er at de sosiale og de sosiomatematiske normene tillater tenkning. Liljedahl (2016) fant at de dominerende, eksisterende klasseromsnormene virker hemmende på tenkende klasserom, og at det er vanskelig å bryte normer som allerede er etablert. Lærere har derfor behov for verktøy som bidrar til å trenge gjennom de gjeldende klasseromsnormene før et tenkende klasserom kan oppstå og opprettholdes (Liljedahl, 2018). Liljedahl (2018) trekker blant annet frem oppgavetyper, implementering av oppgavene, gruppesammensetning, fysisk organisering av klasserommet, og faglig nivå på aktivitetene, som aspekter lærer kan påvirke for å bidra til at elevene tenker mens de løser oppgaver.

2.3.2 Mathematical task knowledge for teaching

For å velge eller designe egnede oppgaver til et formål må læreren vite hva som gjør at en oppgave stimulerer til læring. Chapman (2013) mener også at konteksten og implementeringen av oppgaven i undervisningen er avgjørende for oppgavens potensiale. Dette krever det han kaller “mathematical task knowledge for teaching” som innebærer kunnskapen lærere trenger for å «(a) velge og utvikle oppgaver som fremmer elevenes konseptuelle forståelse av matematikk, som støtter utvikling av deres matematiske tenkning, og som fanger deres interesse og nysgjerrighet, og (b) optimere læringspotensialet i slike oppgaver» (Chapman, 2013, s. 1, min oversettelse). De to punktene er utdypet i seks underpunkter som konkretiserer ytterligere hva læreren må kunne; læreren må (1) vite hva som kjennetegner en god oppgave (inneholder viktig matematikk, kan løses på flere måter, krever høyere ordens tenkning, osv.), (2) ha evne til å identifisere, designe og velge gode oppgaver tilpasset behovet elevene har, (3) ha kunnskap om oppgavens nivå, hva krever den av elevens kognitive ferdigheter, (4) ha kunnskap om elevenes forkunnskaper, lærestiler og interesser, (5) forstå at lærerens valg og implementering av oppgaver har noe å si for elevenes læring (6) vite hvordan arbeidet med oppgavene bør organiseres, hvilke støttestrukturer som

skal gjøres tilgjengelig osv. uten å redusere den stimulerende utfordringen for elevene i for stor grad (Chapman, 2013).

2.3.3 Oppgavenivå

Samuelsson & Frykedal (2014) har studert oppgavens påvirkning på gruppearbeid. Utvalget besto av tre grupper på fire elever og én gruppe på tre elever. De skriver at gruppearbeid kan sees som et mål i seg selv (da vurderes gruppeprosess, dynamikk, osv.) eller som et middel til å oppnå kunnskap. Dersom elevene klarer å jobbe *som* en gruppe (collaboration) kan de oppnå kompetanse. Når det gjelder oppgavens påvirkning har de funnet ut at dersom oppgavene er for vanskelige eller for lette jobber elevene *i* gruppe (cooperation) og har mulighet til å trene blant annet sosiale ferdigheter, mens om oppgavene er innenfor hva elevene kan løse med litt arbeid jobber de som en gruppe og er følgelig i bedre posisjon til å tilegne seg matematisk kompetanse (Samuelsson & Frykedal, 2014).

2.3.4 Perspektiver på oppgavers potensial

Skal man vurdere en oppgaves potensial, er én mulighet å betrakte oppgavens potensial til å aktivere ulike kognitive prosesser hos elevene, uavhengig av konteksten oppgavene gjennomføres i. Fokuset blir da å studere selve oppgaven og vurdere hva den krever av elevene uten at man tar hensyn til undervisningssituasjonen (Pettersen, 2019). Dette kan man gjøre ved hjelp av et kompetanserammeverk, som for eksempel MCRF (Boesen et al., 2018). Analyserer man en parprøve på denne måten, vil man i noen grad gå glipp av parprøvens særtrekk; det sosiale aspektet. Fra et sosiokulturelt perspektiv på oppgaveanalyse, studerer man også konteksten oppgavene skal løses i for å avgjøre oppgavens potensial (Pettersen, 2019). Fra et slikt perspektiv er den praktiske gjennomføringen, og eventuell interaksjon med andre elever eller lærer, relevant å inkludere når man studerer hvordan matematikkoppgaver virker. TIMSS 1999 Video Study er et eksempel på en studie som tar hensyn til begge disse perspektivene ved både å analysere oppgavene og finne hvilket potensial som ligger i dem i sin skrevne form, og i tillegg analysere videodata fra implementeringen av oppgavene (Pettersen, 2019).

3 Metode

I dette kapittelet skal jeg gjøre rede for mitt forskningsdesign og forklare hvordan datainnsamlingen og analysemetoden kan bidra til å besvare forskningsspørsmålene (1) hva mener læreren at er parprøvens potensiale? (2) hvilke grep tar læreren i prøvedesignet, under gjennomføringen av prøven og i etterarbeidet for å utnytte parprøvens potensiale? (3) hvilke kompetansekrav ligger i oppgavene i parprøven? (4) i hvilken grad gjenspeiles lærerens grep i prøvedesignet og under gjennomføringen av prøven? Jeg vil minne om at målet for oppgaven er å komme nærmere noen felles prosedyrer for bruk av parprøver til måling av elevenes kompetanse. Derfor er det interessant å innhente intervjudata som sier noe om en lærers erfaringer med parprøver og lærerens tanker rundt hvilket potensial som ligger i parprøven, hva som er en god parprøve, og hvilke utfordringer som kan oppstå ved bruk av parprøver. I tillegg til lærerens refleksjoner er det hensiktsmessig å også observere en vurderingssituasjon hvor parprøven blir brukt, for å få et rikere bilde av konteksten og kunne si noe om hvordan lærerens visjon gjenspeiles i praksis. Å analysere oppgavesettet i lys av kompetansekrav vil gi ytterligere forståelse for hvordan parprøven kan brukes til vurdering av kompetanse, og for hva lærerens grep har å si.

3.1 Valg av metode

Forskningsspørsmålene ligger til grunn for forskningsdesignet. Jeg skal studere hvordan én lærer med erfaring med parprøver utøver sin vurderingspraksis for å få dypere innsikt i hvordan denne læreren bruker parprøver til å vurdere elevenes kompetanse. Studien min er derfor designet som en kvalitativ casestudie. Cohen et al. (2018) skriver at en fordel med casestudier er at de bidrar til «en verden i bevegelse» ved å gi et detaljert bilde av reelt et tilfelle her og nå, som videre kan tolkes og brukes direkte til å informere og utvikle individer, institusjoner eller utdanningspolitikken. Jeg vil forsøke å gi et beskrivende bilde av hvordan lærerens praksis er; denne type studier utgjør en sentral del av utdanningsforskningen fordi det er nødvendig å kartlegge hva som kjennetegner et fenomen før man kan forklare hvorfor noe skjer (Blikstad-Balas & Dalland, 2021). Fordi hensikten er å gå i dybden på et konkret tilfelle er en kvalitativ tilnærming hensiktsmessig (Patton, 2014).

En casestudie gir meg altså mulighet til å undersøke et spesifikt tilfelle i en gitt kontekst. En case kan være kompleks med mange variabler og det brukes ofte ulike datainnsamlingsmetoder i samme studie for å gi et helhetlig bilde (Cohen et al., 2018). Selv om mange beskriver casestudien som holistisk, det vil si basert på en idé om helhet og at alt henger sammen, argumenterer Verschuren (2003) for at en bør konsentrere seg om de aspektene som er relevant for det man vil studere. Casestudie har ingen entydig definisjon (Patton, 2014), og jeg vil derfor være påpasselig med å presentere mine avgrensninger i det følgende. Cohen et al. (2018) skriver at tid og sted kan være med på å avgrense casen; det samme kan karakteristikkene for institusjonen, gruppen eller individet man ønsker å studere. I denne oppgaven er casen et parprøveforløp sentrert rundt læreren og hans mulighet til å aktivere og måle kompetanse hos elevene. Casen er altså begrenset til å omfatte én lærer og gjennomføringen av en parprøve i én av hans 2P-klasser. Jeg vil undersøke lærerens *design* av parprøven ved å intervju læreren om hensikten med designet og eventuelle hensyn som er tatt. Jeg vil også analysere oppgavene i prøvesettet (se vedlegg 1) med hensyn på kompetansekrav. Gjennom intervju vil jeg også innhente data som omfatter lærerens refleksjoner rundt *gjennomføringen* av parprøven og *etterarbeidet*. Fordi jeg ønsker at denne studien skal bidra til å gjøre det enklere å ta i bruk parprøver er det interessant å inkludere data som kan si noe om hvordan gjennomføringen av parprøven foregår, hvilke utfordringer som oppstår, og hvordan de ulike hensynene læreren beskriver utspiller seg i praksis. Derfor har jeg valgt å inkludere observasjon av elevene under gjennomføringen av prøven i mitt datamateriale. I **tabell 1** har jeg presentert en oversikt over datamaterialet, hvilke forskningsspørsmål de skal bidra til å besvare, og analysemetodene som skal benyttes.

Tabell 1 Oversikt over datamaterialet, hvilke forskningsspørsmål materialet skal bidra til å besvare og analysemetodene.

	Intervju	Oppgavene	Observasjon
Objekt	Læreren Jan	Oppgavesett designet av Jan til bruk som parprøve i 2P	Tre par, A, B og C fra Jans 2P-klasse
Forsknings-spørsmål	(1) hva mener læreren at er parprøvens potensiale? (2) hvilke grep tar læreren i prøvedesign, under gjennomføringen av prøven og i etterarbeid for å utnytte parprøvens potensial (4) i hvilken grad gjenspeiles lærerens grep i prøvedesignet og under gjennomføringen av prøven?	(3) hvilke kompetansekrav ligger i oppgavene i prøven? (4) i hvilken grad gjenspeiles lærerens grep i prøvedesignet og under gjennomføringen av prøven?	(4) i hvilken grad gjenspeiles lærerens grep i prøvedesignet og under gjennomføringen av prøven?
Formål	Få innsikt i <ul style="list-style-type: none"> - hvilket potensiale Jan mener ligger i parprøver - hvilke kompetanser Jan mener parprøver egner seg til å måle - hvilke grep og hensyn Jan mener han har tatt i prøvedesignet, gjennomføringen av prøven og etterarbeidet, og hensikten med disse 	Avgjøre hvilke kompetanser som kreves for å besvare oppgavene i prøven. Gi grunnlag for betraktninger rundt effekten av de grep og hensyn Jan beskriver at han har tatt i design, gjennomføring og etterarbeid	Få innsikt i hvordan parprøven fungerer i praksis. Gi grunnlag for betraktninger rundt effekten av de grep og hensyn Jan beskriver at han har tatt i design, gjennomføring og etterarbeid
Analyse-metode	Tematisk analyse Analyse v.h.a. MCRF+pd	Analyse v.h.a. MCRF+pd	Analyse v.h.a. MCRF+pd

3.1 Utvalg

3.1.1 Lærer

Det er foretatt et strategisk utvalg. Ifølge Cohen et al. (2018) kan et strategisk utvalg være hensiktsmessig når man har behov for å rette søkelyset mot en spesifikk case. Formålet med studien innebærer å presentere et enkelttilfelle og derfor er det tilstrekkelig med én lærer i utvalget. Utvalgskriteriene var (a) at læreren skulle ha gjennomført minst én parprøve tidligere, (b) at læreren hadde planlagt å gjennomføre minst én parprøve med egen klasse i løpet av høsten 2021, og (c) at hen lot meg observere gjennomføringen av parprøven forutsatt at elevene samtykket til det. Punkt (c) innebar at læreren stilte som portvakt og bidro til at jeg

kom i kontakt med klassen. En strategisk utvelgelse etter disse kriteriene gjør at jeg får et utvalg som egner seg til oppgavens formål, men utvalget er ikke representativt for alle lærere som bruker parprøver. I utvalgsprosessen ble det gjort avgrensninger basert på begrensede ressurser i prosjektet - blant annet er søket avgrenset til Østlandet. Å avgrense utvalgsprosessen basert på slike ytre faktorer gjør at utvalget også er et bekvemmelighetsutvalg (Blikstad-Balas & Dalland, 2021). Blikstad-Balas & Dalland (2021) skriver at en utfordring med å velge det som er bekvemmelig er at man risikerer å få et utvalg som man ikke har så mye informasjon om. I denne studien blir denne utfordringen i noe grad kompensert for ved at utvalget også er strategisk.

Jeg tok kontakt med 30 avdelingsledere for realfag ved videregående skoler på Østlandet for å kartlegge hvor det benyttes parprøver. Avdelingslederne fungerte videre som portvakter, og av de 15 som svarte var det syv som videresendte min forespørsel til kollegiet slik at interesserte kunne kontakte meg, og én som tipset om to matematikklærere ved sin skole. De resterende syv, svarte at ingen ved deres skole benyttet parprøver. Jeg kontaktet de to jeg ble tipset om hvorav én av dem svarte positivt, samt at to til fra to andre skoler meldte sin interesse. Til slutt satt jeg igjen med tre kandidater som møtte alle kravene til den strategiske utvelgelsen. Det ble så gjort et bekvemmelighetsutvalg og læreren jeg endte opp med, jeg kaller ham Jan, hadde det mest gunstige tidspunktet for gjennomføring av sin parprøve med tanke på fremdriften til min studie. I tillegg uttrykte han stort engasjement for prosjektet mitt og et ønske om å bidra til forskning på parprøver; jeg opplevde han som en som var villig til å dele informasjon åpent og ærlig, noe som er en viktig forutsetning for kvalitet i prosjektet (Svenkerud, 2021). Samtidig er det viktig å ta høyde for at informantens engasjement for fenomenet som studeres vil kunne påvirke resultatene i studien (Blikstad-Balas & Dalland, 2021); for eksempel vil en som er positiv til bruk av parprøver antakelig legge større vekt på fordelene enn ulempene i sine uttalelser.

Jan underviser i matematikk ved en skole på Østlandet. Han har tidligere benyttet parprøver på denne skolen og på sin forrige arbeidsplass - primært i 1P og 1T. Koronapandemien og restriksjoner som innebar krav om avstand mellom elevene, samt mye digital undervisning, gjorde at Jan valgte ikke å gjennomføre parprøver gjennom hele 2020 og våren 2021. Fra høsten 2021 har han tatt opp bruken igjen, og planlegger å gjennomføre parprøver jevnlig som en del av sin vurderingspraksis. Denne studien er gjennomført i Jans 2P-klasse.

3.1.2 Tre par

Tre par ble observert under gjennomføringen av prøven. For å velge par ble det gjort et bekvemlighetsutvalg blant elevene i Jans 2P-klasse. Elevene var allerede satt sammen i par etter hvilket faglig nivå de lå på, og for å kunne delta i studien måtte begge i paret ha signert samtykkeskjemaet. Det var fem par som møtte kravet og av dem valgte vi to par på middels nivå og et par på lavt nivå. Utvelgelsen var blant annet basert på Jans opplevelse av disse elevenes muntlige aktivitet. Da prøvedagen kom gjorde fravær at det paret som lå på lavt nivå utgikk, og ble erstattet av et nytt par som ble satt sammen for anledningen av to elever på lavt nivå som begge hadde samtykket. Jeg kaller parene, A (middels nivå) , B (middels nivå) og C (lavt nivå).

3.1.3 Oppgaver

Temaet for prøven var «likninger og ulikheter». Hele prøven skal analyseres med tanke på kompetansekrav. Jeg har valgt tre oppgaver, se **tabell 2**, som i tillegg til å analyseres ved hjelp av rammeverket MCRF+pd, brukes som artefakt i intervjuet når Jan skal fortelle om hensikten med oppgavene og ulike grep han har tatt. Disse tre oppgavene skal også sees i lys av observasjoner fra gjennomføringen av prøven. Dette danner grunnlag for å vurdere om Jans intensjoner og grep gjenspeiles i kompetansekravet i prøven og det som utspiller seg i praksis når elevene jobber med oppgavene.

Oppgavene er valgt både fordi de har formuleringer og/eller en vanskegrad som gjør at jeg er interessert i å høre hensikten med designet og se hvordan oppgavene virker i praksis.

Tabell 2 Oversikt over oppgavene som skal analyseres ved hjelp av MCRF+pd. Alle oppgavene er hentet fra parprøven.

Oppgave 1

c) Løs likningen $8^{3x} = 64$

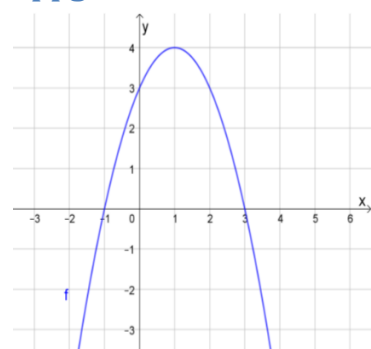
Oppgave 3

Sigurd skal løse en likning og starter slik:

$$x = \frac{-7 \pm \sqrt{7^2 - 4 * 1 * 10}}{2 * 1}$$

- a) Hvilken likning er det Sigurd prøver å løse? Skriv opp likningen.
- b) Fullfør løsningen av likningen.

Oppgave 4



I koordinatsystemet ovenfor har vi tegnet grafen til en annengradsfunksjon f gitt ved

$$f(x) = -x^2 + 2x + 3$$

- a) Hva er løsningen på likningen $-x^2 + 2x + 3 = 0$?

- b) Løs ulikheten $f(x) > 3$

3.2 Datainnsamling - metoder og gjennomføring

3.2.1 Intervju

En styrke ved casestudier er at man har mulighet til å studere et reelt tilfelle i detalj og avdekke unike aspekter som man ikke kunne oppdaget ved hjelp av for eksempel storskalaundersøkelser (Cohen et al., 2018). Fordi man er interessert i å gå i dybden er det kvalitative intervjuet en hyppig brukt metode. Også i denne studien er kvalitativt intervju egnet; to av forskningsspørsmålene dreier seg om lærerens ønsker og hensikter, og hvilke grep/hensyn læreren har tatt. Disse aspektene er ikke observerbare, og spørsmålene kan derfor vanskelig besvares ved hjelp av tekstanalyse eller observasjon. Det kvalitative intervjuet er derimot en svært sensitiv metode som kan gi innsikt i informantens opplevelser og erfaringer fra informantens perspektiv, og dermed et innblikk i informantens opplevde verden (Brinkmann & Kvale, 2018). Samtidig er det viktig å være bevisst forskerens påvirkning i intervjusituasjonen fordi intervjuet er en interaksjon mellom forsker og informant (Brinkmann & Kvale, 2018). I denne oppgaven skulle intervjuet gi innsikt i hvilke hensyn og valg Jan hadde tatt i designet av prøven og i retteprosessen. I tillegg ville jeg vite hva han mener

parprøven egner seg til å måle, og hvilket potensial han mener ligger i bruk av parprøver til å måle elevers kompetanse. Jan er en sentral del av denne casen og for å få et detaljert og riktig bilde var det viktig for meg å følge opp de temaene som Jan syntes var relevante. På grunnlag av dette valgte jeg å legge opp til et semistrukturert intervju; det ga meg mulighet til å styre samtalen retning med forberedte spørsmål, samtidig som det var rom for å gå i dybden på interessante og noen ganger uventede momenter (Gleiss & Sæther, 2021).

Pilotering og bearbeidelse av intervjuguide

Intervjuguiden ble utformet med utgangspunkt i forskningsspørsmålene. Jeg ønsket at dataene skulle romme Jans erfaringer, hensikter og opplevelser, og bidra til å gi et bilde av hans praksis. Jeg la derfor opp til åpne spørsmål med anledning for Jan til å forklare og beskrive. I forkant av datainnsamlingen hadde jeg en samtale med en lærer fra en videregående skole i Oslo. Gjennom samtalen fikk jeg innspill som ville gjøre prosjektet mer praksisnært, det dreide seg særlig om det formative aspektet ved vurdering og hvordan *begge* elevene i paret kan få noe ut av vurderingssituasjonen. Etter samtalen gjorde jeg endringer i intervjuguiden og inkluderte temaene vi hadde diskutert på en mer konkret måte.

Som forberedelse til intervjuet gjennomførte jeg to pilotintervjuer. Det første var primært for å øve på å stille oppfølgingsspørsmål da jeg hadde erfart fra tidligere intervjuer at det var nødvendig for meg. Intervjuobjektet var en medstudent som hadde laget, gjennomført og rettet en prøve for sin praksisklasse. Jeg tok lydopptak av intervjuet slik at jeg kunne evaluere meg selv i etterkant. Det andre pilotintervjuet gjorde jeg for å sørge for at formuleringene mine var enkle og forståelige, og for å sørge for en god struktur i intervjuguiden.

Intervjuobjektet var en konsulent som jobber med brukermedvirkning og veiledning av forskere som skal gjennomføre datainnsamling. Etter dette pilotintervjuet gjorde jeg noen endringer i rekkefølgen på spørsmålene og i formuleringene; den endelige intervjuguiden ligger vedlagt (se vedlegg 5).

Gjennomføring av intervju

Intervjuet med Jan ble gjennomført i starten av februar 2022, i et grupperom på Jans arbeidsplass. Ideelt sett burde intervjuet funnet sted ganske raskt etter gjennomføringen av parprøvene slik at vi hadde erfaringene friskt i minnet (Svenkerud, 2021). Vi valgte å utsette intervjuet til februar for å få tilstrekkelig tid til forberedelser; trolig vil gevinsten av

forberedelsene, pilotintervjuer inkludert, veie opp for ulempene ved lang tid mellom observasjon og intervju. I dagene før intervjuet gikk jeg gjennom opptakene og notatene fra observasjonene. Cohen et al. (2018) skriver at det kan være hensiktsmessig å gjøre intervjuet mot slutten av datainnsamlingsprosessen slik at man får mulighet til å skape seg et eget bilde av informantens verden og dermed kan få mer ut av diskusjonen. Jan fikk tilsendt en oppsummering av intervjuguiden slik at han fikk litt tid til å hente frem de refleksjonene han hadde gjort seg mens han designet, gjennomførte og rettet prøven. For å stimulere en informants hukommelse kan man bruke det som kalles “stimulated recall”. Metoden går ut på at man viser informanten opptak av deres egen atferd for at de lettere skal kunne huske situasjonen og refleksjonene de gjorde underveis i situasjonen (Dempsey, 2010). I dette tilfellet ønsket jeg at Jan skulle snakke om prosessen rundt designet av prøven, og jeg anså derfor prøven som egnet artefakt for å stimulere Jans hukommelse. Prøven ga også Jan mulighet til å henvise til konkrete elementer som formuleringer, grafiske fremstillinger og symboler.

3.2.2 Observasjon

Observasjon kan gi rike autentiske data og mye informasjon om en kontekst, og er den mest brukte datainnsamlingsmetoden i casestudier (Cohen et al., 2018). For å gi et helhetlig bilde av parprøveforløpet er det relevant å samle inn data som sier noe om det som gjør parprøven til en særegen vurderingssituasjon - nemlig arbeidet i par. Observasjon gir innsikt i konteksten (Gleiss & Sæther, 2021) og gjennom observasjon og analyse av samtalen elevene har under prøven får jeg mulighet til å få innsikt i hvilke kompetanser prøven aktiverer hos elevene. Et av forskningsspørsmålene dreier seg om prøvens kompetansekrav, og en vil fra et sosiokulturelt perspektiv mene at det er nødvendig å se oppgavene i sammenheng med den konteksten de skal løses i dersom man skal kunne si noe om hva oppgavene krever av elevene (Pettersen, 2019). I tillegg kan observasjonsdataene sees i sammenheng med Jans uttalelser om prøvedesignet. Hva som faktisk skjer i en situasjon kan være noe annet enn hva informanten ønsker at skal skje eller hva informanten beskriver at pleier å skje, og ved hjelp av observasjon kan man fange opp denne forskjellen (Cohen et al., 2018). Slike data er relevante fordi målet for denne oppgaven er å si noe om hva parprøven kan brukes til, og ikke bare hva det er ønsket at den kan brukes til.

Gjennomføring av observasjon

Observasjonen av parprøvene ble gjennomført i midten av november 2021. Klassen ble delt i to; de som ikke ønsket å delta i prosjektet ble med Jan til et grupperom, mens de som hadde samtykket gjennomførte prøven i klasserommet. Det ble tatt lydopptak av de tre utvalgte parene. Lydopptak sørger for nøyaktig dokumentasjon av hva som blir sagt i parene og gir mulighet for å spole frem og tilbake – noe som kan lette analysearbeidet (Gleiss & Sæther, 2021). I tillegg minker faren for observatøreffekt fordi man ikke er avhengig av hva forskeren får med seg i situasjonen (Cohen et al., 2018). Jeg var også fysisk til stede i situasjonen og gjennomførte det Cohen et al. (2018) kaller semistrukturert observasjon. Dette innebar at jeg hadde noen aspekter jeg ønsket å se etter, samtidig som jeg ville være åpen for nye aspekter ved situasjonen som jeg ikke var forberedt på (Cohen et al. 2018). Aspektene jeg så etter var hovedsakelig det som kan være vanskelig å fange opp på lydopptakene; kroppsspråk, tegn til engasjement, dynamikk mellom parene og hva som skjer når samtalen stopper. Disse aspektene er særlig viktige for å si noe om elevenes produktive disposisjon. Observasjonene noterte jeg i en bok organisert slik at hver gruppe hadde sin del av siden. Fordi klassen var delt og Jan var på grupperommet, ledet jeg aktiviteten i klasserommet, det vil si at jeg fungerte som lærer mens elevene gjennomførte prøven. Jeg hadde altså en rolle som fullt deltakende observatør (Dalland et al., 2021). I en slik rolle kommer man tett på situasjonene fordi man deltar i aktiviteten, samtidig kan det gjøre at man må prioritere oppgaven som lærer over forskerformålet i visse situasjoner. Betraktninger rundt min påvirkning av studien, og andre aspekter knyttet til kvalitet i forskningen følger i delkapittel 3.5.

3.3 Datamaterialet og analysemetode

Datamaterialet fra intervju, observasjon og oppgavetekst gir grunnlag for analysearbeidet. Mye av analysen er gjort ved hjelp av et rammeverk for kompetanse, *MCRF + produktiv disposisjon (MCRF+pd)*, og jeg vil først presentere dette rammeverket og hvordan det er operasjonalisert. Deretter følger presentasjon av bearbeidelsen av datamaterialet, og gjennomføringen av analyse for de tre typene datamateriale.

3.3.1 Rammeverk

I arbeidet med å utforme et rammeverk til analyse av datamaterialet mitt har jeg tatt utgangspunkt i MCRF (Lithner et al., 2010). Jeg valgte MCRF som utgangspunkt fordi det er utviklet til bruk i empirisk forskning. Behovet var et rammeverk som kunne brukes til oppgaveanalyse og analyse av datamateriale fra intervju og observasjon; MCRF er blitt benyttet på alle disse typene datamateriale (se for eksempel: Boesen et al., 2014; Boesen et al., 2018). For å kunne gjøre analyse, for eksempel av samtalen elevene har når de løser oppgaver i par, er det ikke tilstrekkelig med definisjonen av kompetansene, det er nødvendig med noen kjennetegn på hva det vil si å mestre de ulike kompetansene. Lithner et al. (2010) introduserer tre kompetanserelaterte aktiviteter; (I) tolke, (II) anvende og (III) evaluere; som for hver av de seks hovedkategoriene skal beskrive hva det vil si å mestre den respektive kompetansen. Underkategoriene gjør rammeverket mer anvendelig, og i **tabell 3-8** er kompetansekategoriene presentert med underkategorier, som beskrevet i Lithner et al. (2010). Komponentene i rammeverket betegnes med «-evne», som er en oversettelse av «ability». Hverken evne eller kompetanse er strengt definert, og overlapper i stor grad (Kilpatrick, 2020). Jeg har valgt å beholde navnene på komponentene fra Lithner et al. (2010) fordi definisjonene av komponentene er så like definisjonene fra det opprinnelige rammeverket, MCRF, men jeg gjør oppmerksom på at komponentene også kunne vært betegnet med -kompetanse som suffiks.

Tabell 3 Presentasjon av kompetansekategorien "Problemløsningsevne", som utgjør en del av rammeverket MCRF+pd

Problemløsningsevne		
Evnen til å ta del i arbeidet med å løse et problem; det vil si en oppgave hvor løsningsmetoden ikke er kjent.		
Kommentar: en oppgave kan være en rutineoppgave for én elev, og et problem for en annen – det kommer an på elevenes faglige nivå. Dette aspektet vil være ulikt mellom parene og trolig innad i parene.		
Tolke	Anvende	Evaluere
Forstå og gjenkjenne komponentene til et problem og forstå metodene, verktøyene og målene ved problemløsning	Bruke matematikk til å løse ulike problemer, bruke og tilpasse ulike passende problemløsningsstrategier og -metoder. Sette opp og spesifisere ulike typer problemer.	Vurdere troverdigheten til en løsning, samt følge og reflektere over løsningsprosessen.

Tabell 4 Presentasjon av kompetansekategorien "Resonneringsevne", som utgjør en del av rammeverket MCRF+pd

Resonneringsevne		
<p>Evnen til å begrunne og gjøre rede for valg og konklusjoner ved hjelp av matematiske argumenter. Matematiske argumenter er argumenter som er forankret i egenskapene til de matematiske komponentene resonnementet omfatter, og som sier hvorfor en konklusjon er troverdig og plausibel.</p>		
Tolke	Anvende	Evaluerer
<p>Forstå og tolke egne og andres resonnement.</p>	<p>Velge og bruke uformelle og formelle argumenter som støtter valg og konklusjoner. Bruke resonnering til å konstruere tolkninger av informasjon.</p>	<p>Vurdere egne og andres resonnement og reflektere rundt resonneringens rolle og kvaliteter ved valid resonnering. Metakunnskap om resonnering og kunne skille ulike resonnementstyper; f.eks. kunne skille bevis og andre typer resonnement</p>

Tabell 5 Presentasjon av kompetansekategorien "Prosedyresevne", som utgjør en del av rammeverket MCRF+pd

Prosedyresevne		
<p>Evne til å bruke prosedyrer for å løse oppgaver. En prosedyre er sekvenser av matematiske handlinger som utgjør en akseptabel måte å løse en oppgave på.</p> <p><u>Kommentar:</u> I noen tekster brukes begrepet algoritme om prosedyrer i matematikk (eks: Lithner et al., 2010)</p>		
Tolke	Anvende	Evaluerer
<p>Forstå og tolke egne og andres prosedyrer.</p>	<p>Velge og bruke prosedyrer til å nå konklusjoner, og gjøre dette på en hensiktsmessig og effektiv måte slik at fordelene ved bruk av prosedyren utnyttes.</p>	<p>Vurdere utførelsen og utfallet av egne og andres prosedyrer, samt generelle refleksjoner rundt prosedyrens rolle og funksjon.</p>

Tabell 6 Presentasjon av kompetansekategori "Representasjonsevne", som utgjør en del av rammeverket MCRF+pd

Representasjonsevne		
Evne til å håndtere representasjoner.		
Kommentar: Representasjoner er her konkrete fremstillinger av abstrakte matematiske enheter (Lithner et al., 2010).		
Tolke	Anvende	Evaluerer
Forstå og tolke egne og andres representasjoner	Velge og bruke representasjoner til å organisere, formidle, løse problemer, modellere og tolke fysiske, sosiale og matematiske fenomen, og kommunisere matematiske ideer. Velge og bruke tolkninger av representasjoner	Vurdere egne og andres representasjoner, og generelle refleksjoner rundt representasjonens rolle og funksjon

Tabell 7 Presentasjon av kompetansekategori "Connectionsevne", som utgjør en del av rammeverket MCRF+pd

Connectionsevne		
Evne til å sammenkoble matematiske enheter eller representasjoner av matematiske enheter. Dette innebærer sammenhengen mellom representasjoner av ulike enheter, mellom ulike representasjoner av samme enhet, mellom ulike deler av en representasjon, mellom ulike enheter og mellom ulike deler av samme enhet.		
Tolke	Anvende	Evaluerer
Forstå og tolke egne og andres sammenkoblinger. Forstå hvordan matematiske ideer bygger på hverandre og utgjør en helhet.	Velge og bruke sammenhenger til å organisere, løse problemer, og modellere og tolke fysiske, sosiale og matematiske fenomen.	Vurdere egne og andres sammenkoblinger, samt generelle refleksjoner rundt sammenhengenes rolle og funksjon.

Tabell 8 Presentasjon av kompetansekategorien "Kommunikasjonsevne", som utgjør en del av rammeverket MCRF+pd

Kommunikasjonsevne		
Evne til å delta i prosesser hvor informasjon er utvekslet mellom individer gjennom et felles system av symboler, tegn eller atferd.		
<u>Kommentar:</u> Dette krever en sender og en mottaker. Under en parprøve er elevene selv både sendere og mottakere i dialogen med hverandre, de er mottakere av informasjon fra oppgaveteksten, og i tillegg er paret sender av produktet de leverer, og lærer mottaker.		
Tolke	Anvende	Evaluerer
Forstå og tolke informasjon fra en sender, og kunne tolke skriftlige, muntlige og visuelle matematiske påstander.	Konstruere og formulere informasjon til en mottaker. Dette innebærer å kunne veksle adekvat mellom ulike nivåer av teknisk og teoretisk presisjon.	Vurdere egen og andres kommunikasjon, samt gjøre generelle refleksjoner rundt kommunikasjons rolle og funksjon.

Parprøver har et sosialt aspekt som individuelle prøver ikke har. Elevene må forholde seg til en partner, og hvordan de opplever, utnytter og påvirker samspeillet vil ha noe å si for hva de kan oppnå i vurderingssituasjonen. Dette gjør det relevant å trekke inn en komponent i rammeverket som belyser denne kompetansen. Jeg har valgt å kalle denne komponenten *produktiv disposisjon*. Komponenten omhandler det affektive aspektet ved matematisk kompetanse, og er basert på NCTM-standarden fra 1989 (Niss et al., 2016) og *Adding it up* (Kilpatrick et al., 2001). Fordi komponenten skiller seg fra kategoriene i MCRF har jeg valgt å behandle denne kategorien for seg, og analysen er gjort i to steg; først i lys av MCRF, og deretter i lys av produktiv disposisjon.

Kategoriene er operasjonalisert basert på datamaterialets karakter. Først gikk jeg gjennom transkripsjonen fra intervjuet og markerte uttalelser som jeg umiddelbart vurderte at indikerte de ulike kompetansene i MCRF. For eksempel *begrunne*, *bruke formel* og *snakke matematikk*, som kan knyttes til henholdsvis resonneringsevne, prosedyreevne og kommunikasjonsevne. Deretter lagde jeg en liste med indikatorord for hver kompetansekategori, presentert i **tabell 9**. Ordene er basert på første gjennomgang av transkripsjonen og på Lithner et al. (2010) sin definisjon av hver kategori.

Tabell 9 Oversikt over indikatorordene for seks av kompetansekategoriene i MCRF+pd

Kompetansekategori	Indikatorord
Problemløsningsevne	finn en måte å, prøv deg frem, undersøk
Resonneringsevne	begrunnelse, argument, resonnement, forklare, bevise, motbevise
Prosedyreevne	utfør, regn ut, løs, algoritme, fullfør
Representasjonsevne	fremstille, presentere, graf, symbol
Connectionsevne	oversett, omgjør, sammenheng, relasjon, sammenlikne
Kommunikasjonsevne	formuler, definer, forklar, diskuter

Lithner et al. (2010) har laget en liste med verb som indikerer de ulike underkategoriene. Jeg har laget en liste med indikatorord for hver underkategori basert på Lithner et al. (2010) og tilpasset etter inntrykket fra første gjennomlesing av transkripsjonen. Listen er presentert i **tabell 10**.

Tabell 10 Oversikt over indikatorordene for de tre kompetanserelaterte aktivitetene; tolke, anvende og evaluere

Kompetanseaspekt	Indikatorord
Tolke	forstå, tolke, identifisere, gjenkjenne
Anvende	bruke, velge, gjøre, tilpasse, lage, sette opp, spesifisere
Evaluere	vurdere, reflektere, evaluere, skille, sammenlikne

Til kategorien produktiv disposisjon lagde jeg en liste med kjennetegn basert på NCTM-standarden fra 1989 (Niss et al., 2016) og *Adding it up* (Kilpatrick et al., 2001), se **tabell 11**. Basert på første gjennomlesing av transkripsjonen lagde jeg også en liste med indikatorord til produktiv disposisjon, også presentert i **tabell 11**.

Tabell 11 Presentasjon av kjennetegn og indikatorord for kompetansekategorien "Produktiv disposisjon"

Produktiv disposisjon	
Kjennetegn	<ul style="list-style-type: none"> - Eleven forstår at matematikk bør gi mening, og ved å jobbe kan man oppnå dette - Eleven forstår at matematikk kan læres og at innsats vil gi resultater - Eleven har selvtillit når det gjelder egen kunnskap og evne - Eleven viser utholdenhet og prøver igjen dersom noe ser feil ut - Eleven har en konstruktiv og oppmuntrende holdning til medelevens deltakelse i arbeidet - Eleven har fokus på hvordan oppgaven kan løses, og ikke på hvem som løser den eller hvem sin feil det ev. er om oppgaven ikke løses.
Indikatorord	utholdenhet, samarbeid, motivasjon, prøve igjen, selvtillit

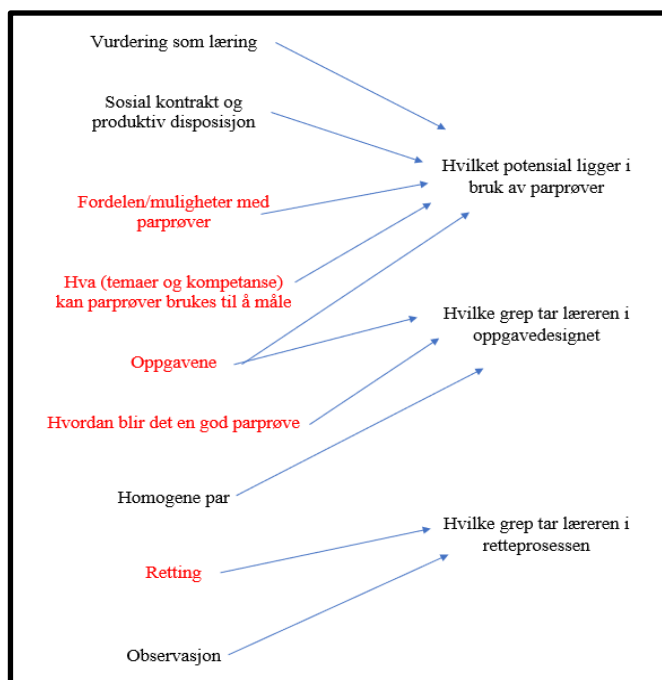
3.3.2 Intervju

Da intervjuet var gjennomført transkriberte jeg opptakene i sin helhet. Jeg brukte f4transkript (UiO, 2017) som verktøy for transkriberingen. Transkripsjonen fra intervjuet ble analysert i lys av rammeverket MCRF + pd, i tillegg til at det ble gjort en tematisk analyse. Jeg vil først beskrive den tematiske analysen, og deretter analysen som ble gjort ved hjelp av rammeverket.

Tematisk analyse

Tematisk analyse er en metode for å “identifisere, analysere og rapportere mønstre (temaer) i datamaterialet” (Braun & Clarke, 2006, min oversettelse). Et tema omtaler noe av vesentlig relevans for forskningsspørsmålene, og ved hjelp av den tematiske analysen kan man organisere og gi mening til datamaterialet (Braun & Clarke, 2006). Intervjuet i denne studien var semistrukturert, og dermed var dataene preget av mine spørsmål, men også i stor grad av Jans assosiasjoner og temaer han syntes var relevante. Det gjorde at datamaterialet fremsto ustrukturert og det var behov for å identifisere temaene i materialet og systematisere dem. Tematisk analyse egner seg til dette, i tillegg egner metoden seg på små utvalg og casestudier (Braun & Clarke, 2017) – noe som underbygget valget. Jeg gjorde den tematiske analysen med utgangspunkt i forskningsspørsmålene og intervjuguiden. I **figur 1** er kodeskjemaet

presentert. Jeg brukte en abduktiv tilnærming (Gleiss & Sæther, 2021) hvor jeg hadde laget noen kategorier på forhånd (skrevet i rødt i figur 1), mens de andre (skrevet i svart i figur 1) oppsto fordi det var temaer som ble lagt vekt på i lærerens uttalelser. Dette gjorde det mulig å oppdage temaer som Jan introduserte. Kategoriene samlet jeg så under tre hovedtemaer.



Figur 1 Kodeskjema for den tematiske analysen av intervjumaterialet

Analyse ved hjelp av rammeverket MCRF + pd

Som nevnt brukte jeg første gjennomlesning av transkripsjonen til å bestemme indikatorord. Den endelige kategoriseringen av datamaterialet fra intervjuet gjorde jeg ved hjelp av indikatorordene, og markerte i transkripsjonen hver gang en indikasjon på kompetanse dukket opp. Deretter gikk jeg til resultatet av den tematiske analysen og fant under hvilke temaer de ulike indikasjonene på kompetanse dukket opp. Dette gjorde jeg for å systematisere funnene slik at de kunne brukes til å besvare forskningsspørsmålene.

3.3.3 Observasjon

Fra lydopptakene var det kun relevant å transkribere de delene hvor elevene jobbet med oppgave 1c, oppgave 3 eller oppgave 4. Transkripsjonen ble gjort i f4transkript (UiO, 2017). Jeg brukte også notatene fra mine observasjoner under gjennomføringen.

Transkripsjonen ble analysert ved hjelp av rammeverket, MCRF + pd. Jeg brukte indikatorordene for hver kompetansekategori, og markerte indikasjon på kompetanse. Videre så jeg nærmere på funnene og vurderte om det var indikasjon på kompetanse *vist* eller kompetanse *aktivert*. Jeg markerte for kompetanse *vist* dersom det kom frem av opptaket, eller fra mine observasjonsnotater, at elevene viste kompetansen i besvarelsen, ved å for eksempel skrive ned utregning eller begrunnelse. Kompetanse *aktivert* betyr at elevene viser kompetanse i samtalen de har mens de løser oppgaven.

Tabell 12 Eksempel på fremstilling av kompetanse aktivert og kompetanse vist

PAR	AKTIVERT	VIST
A	kommunikasjon resonnement prosedyre connection produktiv disposisjon	ingen

3.3.4 Oppgavene

For å besvare forskningsspørsmålet om hvilke kompetansekrav som ligger i oppgavene, gjorde jeg en oppgaveanalyse ved hjelp av rammeverket, MCRF + pd. Ved å analysere oppgavene ved hjelp av et kompetanserammeverk kan jeg avdekke hvilke kompetanser som kreves, og resultatet av analysen blir en beskrivelse av prøvens potensiale til å måle elevers kompetanse uavhengig av konteksten den skal løses i. For å analysere oppgavene tok jeg utgangspunkt i analysen til (Boesen et al., 2018) og fulgte deres prosedyre. Jeg tok for meg én og én oppgave og analyserte kategori for kategori. Jeg begynte med å se etter indikatorord for kategori nr. 1, problemløsningsevne. Dersom jeg fant at oppgaven kunne plasseres i den kategorien gikk jeg videre og avgjorde om det var indikasjoner på tolknings-, anvendelses- og/eller evalueringsaspektet ved kategorien. Deretter gikk jeg videre til neste kompetansekategori.

3.4 Ethiske betraktninger

Ansvar for at er gjennomført på en måte som er etisk forsvarlig tar jeg på alvor. Dette prosjektet har fått godkjenning fra Norsk samfunnsvitenskapelig datatjeneste (NSD). All databehandling er skjedd i tråd med NSD og UiO sine retningslinjer.

Et viktig forskningsetisk prinsipp er at alle deltakere skal ha gitt informert og fritt samtykke (Befring, 2015). At samtykket er informert innebærer at deltakerne skal ha fått ærlig informasjon formidlet på en måte som deltakerne forstår. Informasjonen skal omfatte hva det vil si å delta, inkludert hva som er forventet av dem, tidsaspekt, og hvem som er ansvarlig for forskningen (Everett & Furseth, 2012). Jan fikk tilsendt samtykkeerklæringen (vedlegg 3) etter at jeg ved flere anledninger hadde lagt frem hva det ville innebære for ham å delta. Elevene fikk et annet samtykkeskjema (vedlegg 4) enn Jan fordi deres deltakelse innebar noe annet hans deltakelse. I tillegg tilpasset jeg språket slik at de lett kunne få med seg innholdet. Jeg besøkte skolen én uke før parprøven. Da introduserte jeg prosjektet for elevene. Vi gikk gjennom samtykkeerklæringen og de fikk mulighet til å stille spørsmål til både Jan og meg. Kravet om informert og fritt samtykke innebærer også at det skal være en reell mulighet for å takke nei eller trekke seg fra prosjektet, samt at deltakerne skal gjøres klar over at de kan trekke seg på et hvilket som helst tidspunkt uten at det vil få noen konsekvenser for dem (Befring, 2015). For å ivareta dette var det viktig å presisere at selv om parprøven utgjorde en reell vurderingssituasjon for elevene, ville ikke deltakelse ha noen innvirkning på vurderingen av dem. Jeg understreket dette ved å forsikre klassen om at Jan ikke ville få tilgang til lydopptakene eller mine notater fra observasjonene, og at intervjuet med Jan ville foregå etter at elevene hadde fått sine prøver vurdert. Jeg viste i plenum hvor i skjemaet de kunne finne min kontaktinfo og poengterte at de når som helst kunne ta kontakt for å stille spørsmål eller trekke seg. Noen av elevene ønsket ikke å delta fordi det skulle tas lydopptak og for enkelte ble det viktig å presisere at de ikke ønsket å delta. Jeg besluttet da å la elevene enten samtykke, eller aktivt skrive på skjemaet at de ikke ønsket å delta. I utgangspunktet tenkte jeg at det ville bli større terskel for å si nei dersom de aktivt måtte skrive det, i stedet for å bare la være å svare på skjemaet, men i denne klassen opplevde jeg dette som et grep som ga trygghet til elevene. Jeg sørget for å presentere deltakelse og ikke-deltakelse som to likeverdige alternativer. Omtrent halvparten av elevene samtykket og da prøven skulle gjennomføres plasserte vi deltakerne på et eget rom slik at det ikke var noen fare for at ikke-deltakere skulle komme med på opptaket.

Et annet forskningsetisk prinsipp er at deltakelse i prosjektet ikke skal utgjøre noen ulempe for de som deltar (Everett & Furseth, 2012). Det viktigste hensynet jeg tok i denne forbindelse dreide seg om å balansere min rolle som forsker og min rolle som deltakende observatør. Som forsker var jeg i hovedsak interessert i samtalen mellom elevene i parene, og hvilke kompetanser som ble aktivert, og jeg ville i så liten grad som mulig påvirke denne. Samtidig hadde vi plassert de som hadde samtykket på et eget rom, og det var jeg som skulle lede vurderingssituasjonen deres. Det vil si at jeg var deltakende observatør, og at jeg i rollen som lærer i en reell vurderingssituasjon måtte komme med noen faglige innspill og opptre motiverende, for å unngå at de som deltok i prosjektet fikk en ulempe av å ikke ha Jan til stede. Dersom det ikke hadde vært en reell vurderingssituasjon ville jeg ikke bidratt til elevenes arbeid, slik at resultatene mine utelukkende beskrev hva som skjedde i paret. På tross av hensyn jeg tok vil det at jeg var til stede under prøven, at elevene hadde en lydopptaker på pulten, at Jan ikke var til stede, og at det ble gjort organisatoriske grep – klassen ble delt i to – påvirke elevene i vurderingssituasjonen. Jeg informerte om det praktiske rundt gjennomføringen, og oppfordret elevene til å stille meg spørsmål til prosjektet og gjennomføringen på forhånd slik at minst mulig skulle være uklart underveis. Likevel er det sannsynlig at disse elementene har virket forstyrrende på noen av elevene.

Et tredje forskningsetisk prinsipp er kravet om konfidensiell og anonym deltakelse (Befring, 2015). Dette innebærer at alle personopplysninger skal behandles konfidensielt, og at informanter skal bli anonymisert i datamaterialet. Dette er gjort i tråd med retningslinjer fra NSD og all databehandling er gjort i tråd med lagringsguiden til UiO. For å ta opp lyden under intervjuet brukte jeg UiOs mobilapp Nettskjema-diktafon. Opptakene blir umiddelbart kryptert og sendt til Nettskjema; ingenting lagres på telefonen. Opptakene fra gjennomføringen av parprøvene ble gjort med opptaksutstyr lånt fra UiO. Materialet ble lagret på UiO-driftet desktop samme dag som opptakene ble tatt, og deretter ble materialet slettet fra minnekortene, og minnekortene ble formatert. Intervjudataene ble transkribert i sin helhet ved hjelp av f4transkript. Alle navn er erstattet med fiktive navn, og informasjon om skolen er utelatt. All personidentifiserende informasjon er fjernet. I den forbindelse har jeg utelatt det meste av informasjonen om lærerens bakgrunn, fordi antallet lærere som bruker parprøver ikke er så stort og det derfor ikke skal mye informasjon til før det blir personidentifiserende. Jeg har også vært opptatt av å ivareta informantenes integritet og vært påpasselig med å gjengi dem på en korrekt måte.

3.5 Validitet, Reliabilitet og generaliserbarhet

3.5.1 Validitet og reliabilitet

For at prosjektet skal utgjøre et reelt bidrag til forskningsfeltet er det en rekke hensyn man må ta slik at forskningen blir valid, også kalt troverdig. Troverdighet i forskningen handler om at man kan trekke pålitelige slutninger fordi dataene man har er relevante og gyldige for det man ønsker å undersøke (Everett & Furseth, 2012). Det finnes ulike måter å styrke validiteten i et prosjekt og triangulering er en av dem. I denne oppgaven er det benyttet metodetriangulering, som vil si at flere metoder tatt i bruk for å gi data om samme fenomen (Johnson, 2017). De ulike metodene har styrker og svakheter og ved å benytte flere metoder som utfyller hverandre vil man få et mer troverdig bilde av det man studerer. I tillegg vil man kunne sammenlikne funnene, og i hvilken grad de konvergerer kan styrke eller svekke funnenes gyldighet. I dette tilfellet var det særlig verdifullt å innhente data fra gjennomføringen av prøven slik at man fikk innsyn i hvordan prøvene fungerte i praksis. I et intervju er det stor sjanse for at informantens beskrivelser av undervisning og vurdering er farget av informantens forestilling om hva som er ansett som god og dårlig praksis. Dette kan skje ubevisst og er noe man bør ta høyde for i kvalitative intervju. I intervjuet i forbindelse med denne oppgaven snakket vi for eksempel om ulike grep og hensyn læreren hadde tatt i design, gjennomføring og etterarbeid, for å lage en vellykket vurderingssituasjon. Med flere datainnsamlingsmetoder kan situasjonen belyses fra ulike perspektiver, og et rikere bilde av vurderingssituasjonen vil gi et bedre grunnlag for å si noe om hvor godt lærerens grep fungerer.

Jeg fungerte som fullt deltakende observatør under gjennomføringen av prøven. Det vil si at elevene var svært klar over min tilstedeværelse, noe som potensielt kan svekke validiteten på grunn av det som kalles observatøreffekten. Observatøreffekten vil si at man som forsker påvirker deltakernes atferd ved å være til stede (Dalland et al., 2021). En måte å minske observatøreffekten er å sørge for å presentere seg selv og prosjektet nøye slik at alle forstår hvorfor forskeren er til stede (Dalland et al., 2021). For meg ble det viktig å presisere at jeg i hovedsak var der for å studere Jan sitt arbeid, og at jeg ikke skulle vurdere elevenes ferdigheter. Jeg presiserte også at jeg ikke hadde noen erfaringer med parprøver og dermed ingen tanker om hva som er riktig måte å gjøre det på - alle typer samtaler i parene ville være interessant for min studie. Forskning viser at elever og lærere ofte blir fort vant til å bli

observert og at de oppfører seg ganske likt som i en normalsituasjon (Dalland et al., 2021), det var også min erfaring. En fordel med å være fullt deltakende observatør i en klasse er at elevene, som for eksempel er vant med å av og til ha vikar, raskt aksepterer meg som en naturlig del av konteksten. Elevene henvendte seg til meg med faglige spørsmål som om jeg var læreren deres, og jeg opplevde at de i stor grad glemte at de ble observert. Min deltakelse i samtalene er fanget opp av diktafonene og i arbeidet med opptakene vil jeg kunne avdekke hvilke elevresonnement jeg har påvirket.

Cohen et al. (2018, fra Nisbet & Watt, 1984) trekker frem faren for forskerbias ved casestudier. Forskerbias er en trussel for validiteten, og innebærer at forskerens forkunnskaper, ønsker og hensikter kan påvirke datainnsamlingen slik at resultatene blir i tråd med det forskeren har forespeilet seg (Johnson, 2017). Faren er spesielt at en ubevisst blir selektiv i datainnsamlingsprosessen, oppdager det som støtter egne teorier og kanskje går glipp av funn som taler imot. Det å være reflektiv er en måte å redusere effekten av forskerbias på. Refleksivitet handler om at forskeren undersøker sine egne fordommer og hensikter, samt overvåker prosessen og hvordan forståelsen av fenomenet utvikler seg (Johnson, 2017). Jeg var i utgangspunktet positiv til bruk av parprøver og denne holdningen var jeg bevisst på underveis; jeg sørget for å lese forskning som utfordret mine tanker på forhånd slik at jeg skulle være åpen for flere perspektiver og ha evne å stille kritiske spørsmål. Jeg var også bevisst mitt formål - å komme nærmere prosedyrer for bruk av parprøver - og visste at både funn som indikerte fordeler og funn som indikerte ulemper ved bruk av parprøver ville støtte dette formålet. I analysearbeidet diskuterte jeg med veilederen min, særlig der vi hadde ulik oppfatning av hva funnene impliserte. Å evaluere og diskutere med en som kjenner fagfeltet, men ikke deltar i gjennomføringen av studien gjør at man kan oppdage forskerbias og bidrar til å heve kvaliteten på studien (Creswell & Miller, 2000). Dette styrker forskningens intrareliabilitet.

Hva en casestudie er avgjøres av hvilket fenomen som skal studeres. Fordi studien er kontekstavhengig er det viktig å gi rike beskrivelser av casen, slik at leseren får forståelse for under hvilke forhold studien er gjort; det vil si hva som har gitt grunnlaget for funnene. Rike og tykke beskrivelser er med på å styrke studiens validitet fordi det bidrar til at leseren kan vurdere om grunnlaget for funnene og slutningene som er trukket er troverdige (Creswell & Miller, 2000). Rike beskrivelser gjør også at leseren selv kan avgjøre forskningens

implikasjoner, og for hvilke andre situasjoner resultatene kan være gyldige – dette handler om å være transparent (Gleiss & Sæther, 2021)

3.5.2 Generaliserbarhet

Denne studien er kvalitativ og svært kontekstbasert. Hensikten er å gi et detaljert bilde av et enkelttilfelle, og slik bidra til utvikling innen vurderings- og parprøvepraksis. Fordi utvalget er lite og casen ikke er representativ for matematikklærere som bruker parprøver, er ikke studien statistisk generaliserbar til en større populasjon (Gleiss & Sæther, 2021) – det er heller ikke formålet. I stedet kan det gjøres en analytisk generalisering av resultatene, som innebærer at funnene kan gi implikasjoner for liknende situasjoner (Gleiss & Sæther, 2021), for eksempel parprøveforløp i andre klasser, eller vurderingspraksis med andre alternative vurderingsformer. Ved hjelp av rike beskrivelser og muligheten til å studere prosessene i dybden, kan studien tilby viktig forståelse som kan bidra til å utvikle feltet.

4 Resultat

I dette kapitlet skal jeg presentere funnene fra analysen av datamaterialet. Jeg har gjort en tematisk analyse av transkripsjonen fra intervjuet, og jeg har brukt MCRF + pd, for å kategorisere indikasjoner på kompetanse i Jans utsagn. I del 1 av kapitlet, vil jeg fokusere på analysen av intervjuet, og denne delen er tredelt; først vil jeg presentere (1) hvilke muligheter Jan mener finnes i parprøver. Dette innebærer hvorfor Jan velger å bruke parprøver som vurderingsform, og hvilke kompetanser han mener parprøver egner seg til å måle. Så vil jeg presentere (2) hvilke grep Jan mener han tar i *prøvedesignet og gjennomføringen* av prøven for å utnytte det han mener er prøvens potensiale. Deretter vil jeg presentere (3) hvilke grep Jan mener han tar i *retteprosessen* for å utnytte det han mener er prøvens potensiale.

For å gi et helhetlig bilde av casen og besvare problemstillingen, består datamaterialet også av oppgavesettet klassen til Jan løste under parprøven, og opptak fra tre par under gjennomføringen av parprøven. I del 2 av kapitlet presenterer jeg en analyse av oppgavesettet ved hjelp av MCRF+pd. Avslutningsvis vil jeg presentere analysen av observasjonsmaterialet og se funnene i lys av funn fra intervjuet og funn fra oppgaveanalysen

4.1 Intervju

4.1.1 Muligheter: hvorfor bruke parprøver og hvilke kompetanser kan parprøver måle?

Jan mener det kan være nyttig å ha parprøver i de temaene elevene tradisjonelt synes er vanskelige, fordi det er da elevene har mest behov for å prate sammen. Det var årsaken til at han valgte å ha parprøver i temaet *likninger og ulikheter*.

Jan: Det var ikke tilfeldig at det ble parprøver i det kapitlet [...]. Det mest krevende, sånn rent sånn matematiske temaet i 2P. Mest algebra, mest [...] forståelse, sammenheng graf og likning, og sånn - egentlig ganske abstrakt for en del.

Jan forklarer at han gjennom parprøvene kan gå inn og “prøve en ganske dyp forståelse av hva temaet handler om - nettopp fordi de er to og kan sparre litt i den situasjonen”. Å “sparre” kan i denne sammenhengen tolkes som en faglig samtale hvor man vurderer partnerens

raisonnement og påstander, og så korrigerer for feil, bygger videre på det man oppfatter som rimelig, eller kommer med andre raisonnement og påstander. Dette er aktivering av kommunikasjonsevne og resonneringsevne. Utsagnet om sparring, tyder på at disse to kompetansekategoriene er *gunstige* for elevene å mestre, og at parprøven gir elevene mulighet til å utnytte disse kompetansene for å oppnå dypere forståelse.

Jan trekker også frem at parprøven, uansett tema, bidrar til å «utvide eller øve den ferdigheten det er å snakke matematikk» og at når du løser en parprøve, blir du nødt til å «fremlegge et raisonnement, argumentere for din løsning». Disse utsagnene indikerer også henholdsvis kommunikasjonsevne og resonneringsevne – i tråd med avsnittet over. Forskjellen på disse uttalelsene og uttalelsene i forrige avsnitt er at vi nå ser utvikling av disse kompetansene som et mål i seg selv og ikke bare som gunstige for å oppnå dypere forståelse.

Jan nevner blant annet sammenhengen mellom graf og likning når han snakker om hvorfor *likninger og ulikheter* er et tema hvor parprøve er egnet vurderingsform. At elevene skal jobbe med *sammenhengen* mellom graf og likning indikerer at connectionsevne blir aktivert. Connectionsevne blir også indikert, når Jan videre trekker frem modellering som et tema der parprøver er en egnet vurderingsform. I følge Lithner et al (2010) er modellering å lage en matematisk modell som representerer et fenomen fra virkeligheten, behandle situasjonen matematisk, og så tolke resultatet i lys av virkeligheten. Denne syklusen innebærer å *anvende og tolke* connections mellom modellen og virkeligheten. I tillegg blir også representasjonsevnen aktivert, fordi man lager en representasjon når man lager modellen og man anvender og tolker en representasjon når man behandler modellen (Lithner et al. 2010).

Jan mener modellering, sammen med statistikk, er et aktuelt tema for parprøve blant annet fordi det er rom for å lage åpne diskusjonsoppgaver, hvor prøven nærmest kan bestå av kun én oppgave; for eksempel «fremstill dataen[e] og diskuter hva det er dataene viser deg». Han forklarer videre at denne type oppgaver krever at elevene anvender noe de har lært på en ny situasjon, hvilket er i tråd med slik han oppfatter de nye læreplanene. Etter Jans inntrykk er det i de nye læreplanene lagt mer opp til åpne oppgaver uten deloppgaver, og så må elevene selv inn og definere hvordan de vil gå frem. Han forklarer at de må «gjøre forutsetninger - si hvilke forutsetninger [de] gjør, og sette opp uttrykket». Oppgaver der elevene må sette opp og spesifisere problemer aktiverer anvendelsesaspektet ved problemløsningsevnen. Han sier videre at denne type åpne oppgaver “gir stor frihet i oppgaveløsning, men så krever [det] mer av elevene”. Noen av oppgavene er nesten på firenivå allerede fra start, mener Jan. Da kan

elevene få utbytte av muligheten til å snakke sammen. Stor frihet i oppgaveløsning tolker jeg som en indikasjon på at elevene må gjøre vurderinger, ta valg og evaluere valgene. Det krever ofte et metaperspektiv på prosessen og dette kan tolkes som en indikasjon på evalueringsaspektet ved kompetansene.

Som nevnt mener Jan elevene får mulighet til å aktivere dyp forståelse gjennom diskusjonen i parprøven, og han trekker frem dette som en viktig forskjell på individuelle prøver og parprøver. I stedet for å teste det han omtaler som det “mer mekaniske”, som indikerer prosedyreevne, får man mulighet til å trekke inn det mer prinsipielle; Jan mener parprøven gir mulighet til å stille spørsmål om hva noe er, for eksempel «hva er det som forteller deg at det her er en ulikhet?», i stedet for å be elevene løse en ulikhet. Dette aktiverer tolkningsaspektet og vurderingsaspektet ved representasjonsevnen, mens oppgaven “løs ulikheten” ofte kun aktiverer anvendelsesaspektet ved prosedyreevnen.

Jan kommer stadig tilbake til det han kaller “den sosiale kontrakten” mellom elevene. Den sosiale kontrakten innebærer at elevene føler på et samhold i paret, og et ansvar i relasjonen som fører til bedre utholdenhet. Han beskriver at på individuelle prøver er det mange elever som gir opp med en gang de møter en oppgave de ikke umiddelbart vet hvordan de skal løse, mens når de jobber i par er sjansen større for at de gjør noen forsøk. Han forteller at elevene bygger hverandre opp ved å si ting som «vi får det til» eller «vi prøver». Dette tyder på at elevene anser paret som i stand til å mestre matematikk og at forsøk vil føre frem. Det indikerer kompetansen produktiv disposisjon. Dette er et viktig aspekt ved parprøven, og ifølge Jan gir det elevene mulighet til å mestre flere oppgaver fordi de bruker mer tid på hver oppgave. Jans opplevelse er at særlig elevene på lavt faglig nivå har godt utbytte av parprøver - nettopp fordi de gjennom den sosiale kontrakten gjør flere forsøk på å løse oppgavene, mestrer mer og slik hever seg med tanke på resultat. Jan beskriver også at når elevene på lavt nivå viser kompetanse på parprøver, blir det lettere å få et inntrykk av deres kompetanse generelt, noe som gjør det enklere å sette en standpunktkarakter.

Som nevnt snakker Jan om at parprøven egner seg til å vurdere temaet modellering og problemløsning gjennom åpne og krevende oppgaver. Jeg tolker «krevende oppgaver» som oppgaver som ligger på litt høyt nivå – i tråd med det Jan sier om at noen av oppgavene er på firernivå fra start. Krevende oppgaver kan også bety oppgavetyper eller oppgaveformuleringer som er ukjente for elevene. Slike oppgaver krever at elevene er utholdende, forstår at innsats vil gi resultater selv om de ikke gjenkjenner problemet med en gang, og at de tør å prøve på

nytt dersom de gjør noe feil. Basert på denne tolkningen, mener jeg at «krevende oppgaver» indikerer produktiv disposisjon.

Jan mener at parprøver er en mulighet for å “snu vurderingen om til lærings situasjoner”. I vurderingssituasjoner er elevene mottakelige for læring, og både lærer og elev får godt utbytte av å utnytte vurderingssituasjonen til læring. Jan mener at elevene aldri jobber så godt som under en parprøve. Har man parprøver jevnlig med elevene kan det også bidra til at de lærer å bruke samtalen i andre lærings situasjoner, noe som vil kunne øke læringsutbyttet på sikt. Han nevner spesifikt at etter to til tre parprøver kan elevene få tanker som «vi kan jo egentlig prate sammen også i timen», som indikerer en refleksjon rundt egen læring og arbeidsprosess. Dette tyder på at Jan i tillegg til å samle informasjon om elevenes kompetanse som kan brukes formativt for å støtte elevenes læring - det vi kaller vurdering *for* læring - ønsker at parprøven også skal fungere som vurdering *som* læring.

Funnene som omhandler Jans refleksjoner rundt potensialet i parprøven er oppsummert i **tabell 13**. Vi ser at ifølge Jan, ligger det muligheter i parprøven til å vurdere elevenes kommunikasjons-, resonnerings-, connection-, representasjons- og problemløsningsevne. Vi ser også at alle de tre kompetanserelaterte aktivitetene er indikert. Vesentlig er at flere av utsagnene dreier seg om å utnytte muligheten til å kommunisere og samarbeide, mens få av utsagnene dreier seg om å *måle* kompetanse.

Tabell 13 Oppsummering av hva Jan mener er potensialet i parprøven, og oversikt over hvilke kompetansekategorier som blir indikert

Parprøven gir mulighet til å	Kompetansekategori
«prøve dyp forståelse» fordi elevene kan «sparre»	resonnering: tolke, anvende, evaluere kommunikasjon: tolke, anvende, evaluere
øve på å snakke matematikk og argumentere	resonnering: tolke, anvende, evaluere kommunikasjon: tolke, anvende, evaluere
vurdere modellering	connection: tolke, anvende representasjon: tolke, anvende
å vurdere med åpne oppgaver	problemløsning: tolke, anvende, evaluere generelt: evaluere
teste elevenes forståelse for det prinsipielle (spørsmål om hva noe er)	representasjon: tolke, evaluere
utnytte den sosiale kontrakten til å aktivere flere kompetanser	produktiv disposisjon
teste elevene i krevende oppgaver	Produktiv disposisjon
vurdere <i>for</i> og <i>som</i> læring	

4.1.2 Grep Jan tar i prøvedesign og gjennomføring

I dette delkapittelet vil jeg presentere funnene som omhandler hvilke grep Jan tar i prøvedesign og gjennomføring for å utnytte det han mener er parprøvens potensiale. Jeg vil beskrive grepene og Jans refleksjoner rundt hensikten og effekten av dem. I tillegg vil jeg se uttalelsene i lys av kompetanserammeverket slik at det blir tydeligere hva grepene har å si for muligheten til å aktivere og måle elevenes kompetanse.

Jan mener det er forskjell på å designe en parprøve og en individuell prøve. På spørsmål om hvordan en parprøve blir god, svarer Jan følgende

Jan: På et nivå må den teste det helt grunnleggende, som man gjør i prøver. Oppgave 1a og 1b bør på en måte være greit angripelig for alle på nesten uansett hvilket nivå de er på. Men ganske raskt bør [...] oppgavene bevege seg ut i at du må bruke noe for å vurdere noe. At du får fremsatt påstander, for eksempel [...] «alle rektangler er kvadrat, bevis eller motbevis det»

Dette grepet dreier seg om valg av oppgavetyper. Jan mener slike oppgaver vil generere diskusjon gjøre at man får testet dypere forståelse. Ser vi utdraget i lys av MCRF + pd, kan vi finne indikasjoner på flere kompetanser. At «du må bruke noe for å vurdere noe» indikerer direkte både anvendelsesaspektet og vurderingsaspektet ved kompetansene.

Tolkningsaspektet blir indikert når han videre forklarer at man får fremsatt en påstand. Som eksempel nevner Jan en oppgave som handler om å bevise eller motbevise noe. Et bevis er et matematisk resonnement og generelt kan vi si at slike oppgaver krever resonneringsevne, ofte på alle tre nivåer: elevene må *tolke* påstanden eller resonnementet som er presentert i oppgaven, *evaluere* om påstanden eller resonnementet er rimelig, og *anvende* matematisk resonnement for å bevise eller motbevise påstanden. I dette konkrete eksempelet, «alle rektangler er kvadrat», er påstanden en sammenheng mellom to enheter. Oppgaven krever at man tolker, forstår og evaluerer denne sammenhengen – kjenner og sammenlikner rektangelets egenskaper og kvadratets egenskaper – det indikerer at tolkningsaspektet og vurderingsaspektet ved connectionevnen blir aktivert. I tillegg innebærer det å etablere den korrekte sammenhengen mellom rektangler og kvadrat, som vil være å aktivere anvendelsesaspektet. Oppgaven vil også trolig aktivere representasjonsevnen da en adekvat

måte å løse oppgaven på inkluderer *anvendelse* av representasjoner som formler eller figurer, for å kunne bevise/motbevise påstanden presist.

Jan mener valg av oppgavetyper er viktig for å utnytte parprøvens potensiale til å aktivere elevenes kompetanse, men at det også avhenger av oppgaveformuleringen.

Oppgaveformuleringen er viktig for at elevene skal få *vist* kompetanse. Et grep han tar i prøvedesignet er å passe på at det er rom for diskusjon ved å lage gode problemstillinger; for eksempel spørre *hvorfor* b er riktig og a og c er feil, i stedet for å spørre *hvilket* av alternativene som er riktig. «Der har du allerede en vri på spørsmålet som tester akkurat det du er ute etter, nemlig kan de begrunne hvorfor b er riktig», sier han. At elevene skal begrunne indikerer anvendelsesaspektet ved resonneringsevnen og ved å spørre hvorfor, krever oppgaven ikke bare at kompetansen aktiveres, men også at elevene viser kompetansen.

Jan mener nivået på oppgavene også påvirker muligheten til å få riktig informasjon om elevenes kompetanse. Han forteller at de fleste elevene får et bedre resultat på parprøver enn på individuelle prøver. Dette kompenserer han for ved å øke vanskelighetsgraden. Økt vanskelighetsgrad gjør også at elevene får behov for å diskutere og dermed kan det bidra til å aktivere kommunikasjons- og resonneringsevnene, samt produktiv disposisjon.

Jan har også tatt organisatoriske grep i gjennomføringen av prøven. Han mener læreren bør sette sammen parene; det er en forutsetning for at parprøvene skal fungere at det ikke er noen diskusjon rundt parsammensetningen. Jan forsøker så langt det går å sette sammen elever på omtrent samme faglige nivå fordi det av hans erfaring har fungert godt. Jan forklarer at dersom han blir nødt til å lage et par av to som ligger på ulikt faglig nivå, kan det dukke opp utfordringer i retteprosessen. Dette mener han er lærerens ansvar å korrigere for, ved å gjøre observasjoner mens prøven gjennomføres, snakke med elevene om resultatet og sørge for at begge elevene forstår hvilket nivå de ligger på.

Jan forteller at skeptikerne ofte peker på deling av svar mellom parene som en utfordring, og at det kan gi et feil bilde av elevenes kompetanse. Jan mener det er flere måter å bøte for denne utfordringen på, slik at man kan være ganske sikker på at elevene har kommet frem til løsningene i besvarelsen sin selv. Det ene tiltaket er å gjøre det klart for elevene at det er begrunnelsene som gir uttelling, og at riktig svar uten begrunnelse i verste fall gir null poeng. Det andre tiltaket er å lage flere versjoner av prøven, hvor oppgavetyperne er like, men

verdiene er ulike. Jan valgte å ha to versjoner av prøven denne gangen, men påpeker at man kan se an gruppen man har og vurdere ut ifra hvor sosiale de er i en prøvesituasjon.

Som nevnt mener Jan at parprøven er en anledning til både å oppnå læringsutbytte og gjøre en vurdering av elevenes kompetanse, men det krever at man går systematisk til verks. Etter Jans erfaring er parprøver noe som må gjennomføres flere ganger før man får det skikkelig til. Han anbefaler også å bruke tid med elevene, forberede dem på vurderingsformen, avklare hva som er forventet av dem og hvordan parprøven skal gjennomføres rent praktisk - “ikke forvent at de kan det, for det kan de ikke”, sier han. Tidligere har han formalisert dette ved å lage partnerkontrakter slik at elevene lærer å samarbeide. Et annet grep for at flere skal få utbytte av parprøven, er å la elevene jobbe i paret noen uker i forkant av prøven.

Funnene som omhandler hvilke grep Jan tar i prøvedesign og under gjennomføring av parprøven er oppsummert i **tabell 14**.

Tabell 14 Oppsummering av grepene Jan tar i prøvedesign og under gjennomføring for å utnytte parprøvens potensiale, og hensikten med grepene

Grep	Hensikt
Oppgavetyper	Aktivere flere kompetanser
Oppgaveformuleringer	Aktivere og flere ulike kompetanser
Oppgavenivå	Kompenserer for at elevene presterer bedre i par Gjør at elevene får behov for å diskutere
Parsammensetning	Dempe uro Unngå utfordringer i retteprosess
Flere prøveversjoner	Hindre deling av svar mellom parene
Forberede elevene på hva som forventes av dem og øve	Bidra til at elevene kan utnytte samarbeidet Bidra til at elevene begrunner svarene

4.1.3 Grep Jan tar i retteprosessen

I retteprosessen uttrykker Jan at det i stor grad er opp til læreren hvor mye man kan få ut av vurderingen: Han forklarer, «når du retter så retter du som en vanlig prøve. Så får de et resultat. Så er det opp til læreren å bruke det resultatet fornuftig». Han mener utfordringen ligger i å gi riktig vurdering og tilbakemelding til hver enkelt elev, og at han løser det ved å blant annet involvere elevene i rettingen. For å involvere elevene i vurderingsprosessen bruker Jan retteskjema. Dette er likt for alle prøver i hans matematikkparalleller. Parene får

sin egen besvarelse, et løsningsforslag og et tomt retteskjema (se vedlegg 2). I retteskjemaet skal elevene gi seg selv poeng, og eventuelt kommentarer. Jan forklarer at retteskjemaet tvinger elevene til å reflektere over sine egne svar sammenliknet med løsningsforslag. Når elevenes retteskjema så sammenliknes med lærerens, og man oppdager vesentlige forskjeller i poenggivning, gir det grunnlag for samtale om elevenes kompetanse. Elevene og læreren må da peke på hva de har funnet og hva som eventuelt mangler – hvilke bevis gjør at man kan gi så og så mange poeng. Jan mener innblikk i retteprosessen fører til at elevene forstår hva som forventes av dem, i tillegg til at det gir både elev og lærer informasjon om elevens kompetanse.

Jan forteller at retteskjemaet gir grunnlag for samtale med elevene, og han mener at samtaler med elevene om kompetanse og vurdering er viktig. Han forklarer at han opplever at matematikk er ansett som “greit” å vurdere fordi slutt karakteren sier seg selv basert på de individuelle prøvene. Med parprøver må man tenke litt annerledes, “nå må du inn og justere mer, med tanke på det inntrykket du sitter igjen med selv. Og det tvinger deg [til å] være enda mer bevisst på det her med hva observerer du i klasserommet” sier han. Han sammenlikner det med hvordan han jobber i språkfag, hvor han gjennom observasjon hele tiden bygger på det uformelle helhetsinntrykket og får en følelse av “hvor eleven er hen”. Han mener læreren må tydeliggjøre for elevene hva karakteren har å si, og sørge for at de forstår at standpunkt karakteren ikke bare er et snitt av resultatene de har oppnådd.

I utdraget over nevner Jan observasjon. Videre forteller han at han bruker observasjonen aktivt. Et formål er å korrigere; for eksempel dersom det i et par er veldig ulik arbeidsinnsats, vil dette kunne oppdages i observasjonen, og i samtale med paret kan han gi tilbakemelding om at resultatet ikke speiler begges kompetanse. Jan bruker også observasjonene til å få tilgang på de delene av prosessen som ikke kommer til uttrykk i den skriftlige besvarelsen. Han forklarer at det ikke er all prosess som havner på papiret og blir vurdert der og da, men hvis det er elever som han observerer at “mestrer samarbeidet godt, blir flinke til å prate og uttrykke seg, [og] argumenterer klart og godt” så vil det gi positivt utslag på helhetsvurderingen av eleven.

Jan ønsker å utnytte parprøvens potensiale til å være en læringssituasjon. Dette gjenspeiles når han snakker om retteprosessen og etterarbeidet; gjennom intervjuet nevner han flere ganger den sosiale kontrakten og andre indikatorer på produktiv disposisjon, og på spørsmål om han får tilgang på disse prosessene i besvarelsen svarer han følgende,

Jan: hvis de viser kunnskapen på papiret så er vi jo egentlig litt fornøyd med det [...], sånn i matematikkens vurderingsverden. Og så får vi heller [...] legge litt verdi i prosessen, [...] håpe at det fem-minuttet som de brukte på å diskutere seg frem til hva nullpunkt var, kanskje de har litt utbytte av det neste gang de skal finne et nullpunkt.

Han forklarer videre at i en ideell situasjon forlater elevene prøvesituasjonen med mer kunnskap enn de kom inn med. Ønsket er at det skal skje noen individuelle prosesser hos elevene og at disse kan komme til syne senere.

Funnene som omhandler hvilke grep Jan tar i etterarbeidet er oppsummert i **tabell 15**.

Tabell 15 Oppsummering av grepene Jan tar i etterarbeidet for å utnytte parprøvens potensiale, og hensikten med grepene

Grep	Hensikt
Retteskjema	Bidra til at elevene reflekterer rundt sin egen besvarelse og vurdering
Samtale med elevene	Bidra til at elevene og lærer forstår elevens kompetanse Bygge på det uformelle helhetsinntrykket
Bruke observasjonene	Bygge på det uformelle helhetsinntrykket Korrigere for ujevn innsats eller ujevnt nivå innad i paret
Stole på prosessen	Vurdere det som vises i besvarelsen, og samtidig la elevenes arbeid med prøven ha en verdi i seg selv som kan gi læringsutbytte på sikt

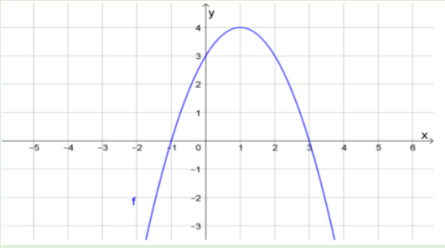
4.2 Oppgavene

4.2.1 Analyse av oppgavesettet ved hjelp av rammeverket

I dette delkapittelet skal jeg presentere resultatene fra analysen av oppgavesettet (vedlegg 1) som ble benyttet i denne studien.

Analysen av oppgave 1, oppgave 3 og oppgave 4, ved hjelp av MCRF+pd, er presentert inngående i **tabell 16**. Å lese en oppgavetekst og skrive ned et svar er indikasjoner på henholdsvis tolknings- og anvendelsesaspektet ved kommunikasjonsevnen. Dette er handlinger som jeg anser som implisitte i en skriftlig vurderingssituasjon og som vil gi utslag i alle oppgavene. Jeg har valgt ikke å inkludere disse i fremstillingen av mine funn; dette gjør jeg for at andre indikasjoner på kommunikasjonsevne skal bli synlige.

Tabell 16 Oppgave 1, oppgave 3 og oppgave 4 analysert ved hjelp av MCRF+pd

Oppgave	Kategori	Indikator/kommentar
Oppgave 1 c) Løs likningen $8^{3x} = 64$	Prosedyreevne: anvende Produktiv disposisjon	Det kreves at eleven gjenkjenner problemet og kan utføre prosedyren. Prosedyren er forventet kjent av denne 2P-klassen Oppgaven er på høyt nivå
Oppgave 3 Sigurd skal løse en likning og starter slik: $x = \frac{-7 \pm \sqrt{7^2 - 4 * 1 * 10}}{2 * 1}$ a) Hvilken likning er det Sigurd prøver å løse? Skriv opp likningen.	Prosedyreevne: tolke, anvende Representasjonsevne: tolke Produktiv disposisjon	Abc-formelen er forventet kjent for elevene. Oppgaveformuleringen er ukjent og krever at elevene tolker uttrykket de har fått og ser det i sammenheng med resten av prosedyren. Identifisere a, b og c i uttrykket Formuleringen/oppgavetypen er ukjent
Oppgave 3 b) Fullfør løsningen av likningen.	Prosedyreevne: anvende	Løsningsmetoden er forventet kjent og det gjenstår kun å fullføre utregningen
 Oppgave 4 I koordinatsystemet ovenfor har vi tegnet grafen til en annengradsfunksjon f gitt ved $f(x) = -x^2 + 2x + 3$ a) Hva er løsningen på likningen $-x^2 + 2x + 3 = 0$?	Prosedyreevne: anvende Representasjonsevne: tolke Connectionsevne: tolke, anvende	Kan løses med abc-formelen som er forventet kjent Må tolke uttrykket og grafen Sammenhengen mellom graf og annengradsfunksjon må tolkes. Anvendelsesaspektet aktiveres i det eleven bruker denne sammenhengen, og sammenhengen mellom likningen i oppgaven og annengradsfunksjonen til å finne løsningen
Oppgave 4 b) Løs ulikheten $f(x) > 3$	Prosedyreevne: anvende Representasjonsevne: tolke Connectionsevne: anvende Produktiv disposisjon	Kan løses ved utregning Må tolke ulikheten og forstå hva ulikhetstegnet innebærer Må kunne knytte ulikheten til grafen Oppgaven er på høyt nivå

I **tabell 17** er funnene fra analysen av hele parprøven presentert. Tabellen gir et bilde av hvilke kompetanser som er indikert i parprøven. Vi ser at det er hovedvekt av prosedyreevne, og ingen indikasjoner på problemløsningsevne.

Tabell 17 Funn fra analysen av parprøven. Analysen er gjort ved hjelp av MCRF+pd

	1a	1b	1c	2	3a	3b	4a	4b	5a	5b	6a	6b	7a	7b	sum
Problemløsning															0
Resonnering				X						X					2
Prosedyre	X	X	X		X	X	X	X	X			X	X	X	11
Representasjon					X		X	X			X				4
Connection							X	X			X				3
Kommunikasjon				X						X	X				3
Produktiv disposisjon			X		X			X			X				4

4.2.2 Observasjon

I dette delkapittelet vil jeg presentere funn fra observasjonen av oppgave 1c, oppgave 3 og oppgave 4. Jeg vil også presentere Jans refleksjoner rundt disse oppgavene. For hver av oppgavene har jeg vurdert om kompetanse er aktivert og/eller vist i besvarelsen. Resultatet av dette er presentert i **tabell 18-20**. Par C hoppet over oppgave 3 og 4 og det indikeres dermed ingen kompetanse for dette paret på disse oppgavene.

Oppgave 1c

Jan beskriver oppgave 1c) som en oppgave som tester om elevene gjenkjenner situasjonen og husker hvordan de skal gå frem for å løse den. I praksis, i en parprøvesetting, vil man få testet om minst én i paret har fått med seg hva man skal gjøre i en slik situasjon. Ifølge Jan er løsningsmetoden undervist og forventet kjent for elevene. Dette stemmer overens med min kategorisering av oppgaven som en vurdering av prosedyreevne. I tillegg mener jeg oppgaven

krever produktiv disposisjon fordi den har høy vanskelighetsgrad med tanke på at prøven gis i 2P.

Observasjonene underbygger også oppgaven som en aktivisering av prosedyreevne. I par A opplever elev A1 at oppgaven er kjent, mens A2 ikke vet hvordan den løses. A1 tar styringen og forklarer nøye hva han gjør og inkluderer A2 i utregningene. A1 aktiverer dermed anvendelsesaspektet ved kommunikasjons- og resonneringsevne fordi han er avhengig av å formidle det han tenker til A2 slik at A2 kan hjelpe han med utregningene. A2 aktiverer tolkningsaspektet ved disse kompetansene. Deler av prosedyren A1 husker er feil, men flere av forklaringene og resonnementene han har, deriblant forklaringen av sammenhengen mellom grunntall og eksponent, er korrekte. Løsningen de kommer frem til er feil, og da de ikke skriver ned utregningene som en del av besvarelsen får de ingen uttelling for at deler av prosedyren er riktig.

I par B er problemet kjent for begge elevene og de løser oppgaven, slik de har lært det, uten vanskeligheter.

I par C er det ingen av elevene som kjenner problemet. Det varierer også hvor godt den sosiale kontrakten fungerer. De har flere utsagn som «jeg tror vi egentlig skjønner det, vi må bare tvinge oss selv [...]». Dette er en indikasjon på produktiv disposisjon hvor de anser seg selv som i stand til å løse oppgaven. Samtidig bruker de mye tid på å snakke om at de er dårlige i matte, at de pleier å få dårlige karakterer, at de ikke pleier å øve fordi det ikke er noe vits når de aldri får til noe og at prøven er for vanskelig for dem. Dette er indikatorer på lite produktiv disposisjon. Likevel ender de med å fortsette å prøve, de fleste gangene de er blitt enige om å gi opp. Den sosiale kontrakten gjør at par C bruker mye tid på oppgaven, omtrent fem minutter, men de kommer ikke nærmere noen løsning og må til slutt gå videre til neste oppgave.

Tabell 18 Oversikt over kompetansekategorier indikert aktivert og indikert vist gjennom observasjon av elevenes arbeid med oppgave 1c

PAR	AKTIVERT	VIST
A	kommunikasjon resonnement prosedyre connection produktiv disposisjon	ingen
B	prosedyre	prosedyre
C	produktiv disposisjon resonnement	ingen

Oppgave 3

Ifølge Jan skal oppgave 3 først og fremst teste om elevene kan abc-formelen. I tillegg tester oppgaven om elevene kan “bevege seg begge veier”, det vil si sette inn i formel fra likning og motsatt. Jan forklarer at denne oppgaveformuleringen gjør at elevene må «bryte» med oppgaven og at det slik kan oppstå diskusjoner, som lærer kan observere, om hvor tallene i formelen kommer fra. Han sier også at dersom prøven skulle løses individuelt ville han gitt dem likningen og bedt dem sette inn i formelen – noe som er mer kjent for elevene.

Kategorisert ved hjelp av rammeverket er oppgave 3b) en ren prosedyreoppgave. Dette stemmer overens med Jans hensikt. 3a) tester elevenes kjennskap til abc-formelen, men er formulert på en måte som er ukjent for elevene slik at jeg vurderer den til å indikere produktiv disposisjon. På grunn av den ukjente formuleringen skapte oppgaven litt uro. Noen mente de ikke hadde lært å løse sånne oppgaver, og flere av parene søkte bekreftelse fra meg på om de hadde forstått oppgaven riktig.

Det er tydelig at klassen har jobbet med abc-formelen og alle parene gjenkjenner at uttrykket er fra formelen. I par A skjønner de raskt hva de skal gjøre og samarbeider godt ved å forklare for hverandre hva de tenker og begrunne forslagene sine. Det er tydelig at de kjenner formelen godt og at både 3a) og 3b) aktiverer deres prosedyreevne. Par A gjør noen feil i utregningene som krever at de gjør flere forsøk. Dette aktiverer vurderingsaspektet ved prosedyreevnen, og produktiv disposisjon.

I par B har B1 problemer med å forstå formuleringen av spørsmålet. B2 skjønner oppgaven og forsøker å forklare:

B2: det her er abc-formelen

B1: ja, si at det er en abc-formel og at [Sigurd] skal prøve å løse abc-formelen

B2: vi skal løse den her i b-oppgaven. Men a-oppgaven, vi skal lage en likning, det her er ikke likningen, det er oppgaven, det har vært en likning, de har putta inn liksom likningen her

B1: ja, det er akkurat det samme som det her

B2: det er formelen [...]

B1: ja det er likningen

B2: formel og likning er ikke det samme. Egentlig det skal være $6 - x + 6 = x$, noe sånt. Det skal være sånn greie.

B1: så hvordan får vi til det?

B2: det er det vi må finne ut

I utdraget over kommer par B inn på at formel og likning ikke er det samme. Som nevnt ser Jan parprøven som en anledning til å la elevene reflektere over hva noe er, i stedet for å bare be dem om å gjøre en utregning, og par B sin dialog kan regnes som et eksempel på dette. Senere kommer de tilbake til forskjellen på likning og formel og sammenhengen mellom dem. Par B har utfordringer med å løse oppgaven, men de viser tro på at de skal få det til ved at de oppfordrer hverandre til å forklare så de skal forstå bedre, og ved at de gjør mange forsøk – dette indikerer produktiv disposisjon. Par B løser oppgave 3b uten problemer.

Par C hoppet over denne oppgaven.

Tabell 19 Oversikt over kompetansekategorier indikert aktivert og indikert vist gjennom observasjon av elevenes arbeid med oppgave 3

PAR	AKTIVERT	VIST
A	prosedyre representasjon resonnering kommunikasjon produktiv disposisjon	prosedyre representasjon
B	prosedyre representasjon connection kommunikasjon produktiv disposisjon	prosedyre
C	ingen	ingen

Oppgave 4

Jan forteller at målet med oppgave 4a) er å teste elevenes forståelse for sammenhengen mellom algebra og graf, og at løsningen på abc-formelen også er nullpunktene på en graf. Oppgave 4b) tester forståelsen for hva en ulikhet er og sammenhengen med en grafisk fremstilling. Jan mener oppgaven ligger på høyt nivå fordi den ikke spør om $f(x) > 0$, men $f(x) > 3$, noe som er mer ukjent for elevene. Igjen beskriver Jan hvordan elevene kan dra nytte av den «sosiale kontrakten» i denne oppgaven. Jan opplever at de i større grad prøver å løse oppgavene, og sjeldnere gir opp. Han beskriver hvordan de gjennom samtale begynner å nøste i det han kaller «fragmenter av kunnskap» og at flere til slutt klarer å nå frem til riktig løsning.

I par A tar A1 styringen også på oppgave 4a). Han ser at likningen kommer fra annengradsfunksjonen og skjønner at de kan bruke grafen til å løse likningen. A2 hverken bekrefter eller avkrefter A1s forslag, men sier at hun ikke vet hvordan oppgaven skal løses. A1 beslutter at de skal løse likningen med abc-formelen for å sjekke om han har tenkt rett. De deltar begge i utregningene, men gjør noen feil, og når de ender opp med å skulle løse kvadratrotten av 14 bestemmer de seg for å bruke grafen og stole på at A1 har rett. På oppgave 4b) forsøker A1 å bruke grafen, men vet ikke helt hvordan. A2 sier hun kan løse oppgaven på PC, men at hun ikke vet hvordan det skal gjøres på papir. A1 kommer med noen forslag, men når A2 bare responderer med «vet ikke», fungerer ikke sparringen og de beslutter

å gå videre til neste oppgave. Par A viser produktiv disposisjon ved at de ønsker at det skal gi mening, og engasjerer seg i arbeidet selv om de ikke vet om det de gjør fører frem.

Par B løser oppgave 4a) med abc-formelen uten problemer. Ingen av dem forstår hvordan de skal gå frem på oppgave 4b). De begynner å snakke om hva ulikhetstegnet betyr, og bruker grafen til å prøve å finne en fremgangsmåte. De viser produktiv disposisjon ved at de gjør mange forsøk og oppmuntrer hverandre, noe som tyder på at de tenker de er i stand til å løse oppgaven dersom de prøver.

Tabell 20 Oversikt over kompetansekategorier indikert aktivert og indikert vist gjennom observasjon av elevenes arbeid med oppgave 4

PAR	AKTIVERT	VIST
A	resonnering representasjon connection	connection
B	prosedyre produktiv disposisjon problemløsning	prosedyre
C	ingen	ingen

5 Diskusjon og konklusjon

Målet for denne oppgaven er å reflektere rundt parprøvens funksjon som vurderingsform, for å komme nærmere en måte å bruke parprøven til å vurdere elevenes matematiske kompetanse. Oppgaven forsøker å besvare problemstillingen, «*På hvilke måter og i hvilken grad utnytter en lærer parprøvens potensiale til å vurdere elevenes matematiske kompetanse?*».

Undersøkelsene er gjort gjennom fire forskningsspørsmål, (1) hva mener læreren at er parprøvens potensiale? (2) hvilke grep tar læreren i prøvedesignet, under gjennomføringen av prøven og i etterarbeidet for å utnytte parprøvens potensiale? (3) hvilke kompetansekrav ligger i oppgavene i parprøven? (4) i hvilken grad gjenspeiles lærerens grep i prøvedesignet og under gjennomføringen av prøven? Spørsmålene er belyst gjennom tre ulike tilnærminger; intervju, observasjon og oppgaveanalyse; noe gir grunnlag for å gå i dybden på de ulike aspektene som diskuteres.

5.1 Diskusjon

5.1.1 Vurdering som læring

Det første forskningsspørsmålet dreier seg om hva Jan mener er potensialet i parprøven. Et viktig funn knyttet til dette er Jan sitt ønske om å bruke parprøven som en læringssituasjon. Å skape læringssituasjoner i vurderingssituasjonen, ved å sørge for at elevene får aktivert mer enn bare forkunnskaper og mulighet til å oppdage hva de trenger for å oppnå progresjon, omtales som vurdering *som læring* (Yan & Yang, 2022). Jan ønsker blant annet å legge til rette for diskusjon i vurderingssituasjonen slik at elevene kan tilegne seg kompetanse underveis. Både ved at de utvikler kompetanse i dialogen, og at de får mulighet til å erfare og reflektere rundt samtale som læringsmetode. Ifølge Jan innebærer arbeidet med parprøven noen viktige og lærerike prosesser som er verdifulle i seg selv uten at de nødvendigvis skal utgjøre vurderingsgrunnlag. At Jan ikke har vurdering som eneste mål for aktiviteten påvirker hvordan han vil definere en god parprøve, og hvilke grep han tar i design, gjennomføring og retteprosess. Det vil si at når Jan forteller om hvordan en god parprøve blir til, er noen av grepene han beskriver rettet mot å *aktivere* kompetanse hos elevene, heller enn å *måle* kompetansen deres. For å dra nytte av dette prosjektet til videre parprøvepraksis er det viktig å være oppmerksom på forskjellen mellom at parprøven aktiverer kompetanse og at parprøven

måler kompetanse. Sistnevnte gir elevene mulighet til å vise den kompetansen de har aktivert, som er vesentlig når formålet er vurdering.

5.1.2 De essensielle komponentene i matematisk kompetanse

Niss (2007) tar opp behovet for vurderingsformer som kan dekke det han i artikkelen omtaler som de essensielle komponentene i matematisk kompetanse; kommunikasjon, bygge på argumenter og formelle bevis, beherske ulike representasjoner og veksle mellom ulike representasjoner. Dette er indikatorer på henholdsvis kommunikasjonsevne, resonneringsevne, representasjonsevne og connectionevne (Lithner et al., 2010). Funn fra mine undersøkelser viser at det er disse kompetansene Jan legger vekt på når han snakker om hva parprøven egner seg til å aktivere og måle (tabell 13). Det vil si at parprøven ifølge Jan egner seg godt til å aktivere de essensielle komponentene i matematisk kompetanse.

Sammenlikner vi funnene som omhandler kompetanse fra intervjumaterialet (tabell 13), observasjonene (tabell 18-20) og analysen av oppgavesettet (tabell 16-17) finner vi varierende grad av samsvar. Jan nevner generelt at parprøver er egnet til å aktivere kommunikasjonsevne, resonneringsevne, representasjonsevne, connectionevne og problemløsningsevne. Dette gjenspeiles i oversikten over hvilke kompetanser som blir observert når elevene jobber med parprøven (tabell 18-20). Ser man derimot på funnene fra observasjonsmaterialet som sier noe om hva elevene *viser gjennom sin besvarelse*, er langt færre kompetanser representert. Vi ser i hovedsak prosedyreevne. Det virker altså som om parprøven aktiverer en rekke kompetanser – noe som er i tråd med hva Jan ønsker og i tråd med hva det er behov for at en alternativ vurderingsform skal gjøre (Suurtamm, 2004). Samtidig er det få av kompetansene som kommer til uttrykk i besvarelsene og dermed blir det formelle vurderingsgrunnlaget begrenset. En årsak til dette kan være at Jan også har til hensikt å lage en lærings situasjon for elevene, og at selve vurderingsaspektet – hvor læreren er avhengig av at elevene *viser* den kompetansen de har aktivert – ikke er blitt vektlagt i like stor grad i prøvedesignet. Dette er i tråd med Black & Wiliam (2018) som advarer mot å utvide begrepet formativ vurdering til å omfatte vurdering *som* læring, nettopp fordi et for vidt begrep kan gjøre at man mister fokuset på vurderingsaspektet.

5.1.3 Oppgavedesign

Kontekst

Ser vi på analysen av oppgavene, illustrert i **tabell 17**, er et funn at de fleste oppgavene måler prosedyreevnen, mens kun et fåtall måler resonneringsevnen, kommunikasjonsevnen og connectionevnen - på tross av at Jan trekker frem disse kompetansene som relevante å måle ved hjelp av parprøve. Pettersen (2019) skriver at implementeringen av oppgaven kan endre oppgavens karakter og i hvilken grad den er kognitivt utfordrende. Tidligere nevnte funn fra observasjonene sier at oppgavene i denne studien aktiverer flere kompetanser enn prosedyreevne, noe som kan tyde på at konteksten i seg selv påvirker prøvens potensiale til å aktivere kompetanse hos elevene på en positiv måte.

Oppgavetype

Funn fra denne studien indikerer at oppgave 1c er et eksempel på en oppgave som aktiverer flere kompetanser enn den måler. Jan forklarer at denne oppgaven er ment å teste om minst én i paret husker hvordan den skal løses, det vil si at den er designet for å måle anvendelsesaspektet ved prosedyreevnen, noe som samsvarer med oppgavens kategorisering ved hjelp av MCRF+pd. I utgangspunktet gir konteksten oppgaven skal løses i elevene mulighet til å snakke sammen, og bruke hverandre for å sammen løse oppgaven dersom de ikke husker prosedyren. Dette er en mulighet til å *aktivere* resonneringsevne, kommunikasjonsevne og produktiv disposisjon som ligger latent i parprøven fordi elevene jobber i par. Det er samtidig mulig at man kunne utnyttet parprøvesituasjonen ytterligere i denne oppgaven, slik at kompetansene som blir aktivert, også blir *målt*. Påstandsoppgaver og krav om begrunnelse, er oppgavetyper og formuleringer som Jan trekker frem som egnede i en parprøve for å øke elevenes utbytte av at de jobber i par, og for å få informasjon om elevenes kompetanse. Et alternativ kunne vært å legge til et krav om at elevene må vise utregning og forklare stegene. Et annet alternativ kunne vært å la elevene vurdere om en løsning er fornuftig, eller evaluere sin egen løsning. Da hadde lærer blant annet fått tilgang på elevenes resonnement og evalueringsaspektet ved prosedyreevnen.

I oppgave 4a får elevene presentert en funksjon, $f(x)$, grafen til funksjonen, og så skal de løse $f(x)=0$. Jan sier at hensikten med oppgaven er «å teste elevenes forståelse for sammenhengen mellom algebra og graf, og at løsningen på abc-formelen også er nullpunktene på en graf».

Dersom man kjenner sammenhengen mellom graf, funksjon og nullpunkt er det enklest å løse oppgaven grafisk ved å lese av punktene. Dersom elevene løser oppgaven slik, og begrunner svaret sitt, indikerer det connectionsevne og representasjonsevne, slik Jan ønsker. Oppgavens formulering tillater også at man finner nullpunktene ved hjelp av abc-formelen. Par B løser oppgave 4a ved utregning og viser dermed kun prosedyreevne. For å teste forståelse for sammenhengene, slik hensikten med oppgaven er, kunne man i stedet brukt formuleringer som, «bruk grafen til å løse $f(x)=0$ », «hvordan kan dere bruke grafen til å løse $f(x)=0$?» eller «hvordan kan dere finne nullpunktene til grafen ved regning?». Da tester oppgaven i større grad det den er ment å teste. En fordel med å lage oppgaver som kan løses på flere måter, slik som oppgave 4a, er at det gir mulighet til nivådifferensiering.

Oppgavenivå

Oppgave 1c er regnet som en oppgave på høyt nivå i 2P. Jan sier at han forsøker å legge prøven på et nivå som gjør at elevene må “bryte” litt med oppgavene. Det skal aktivere flere kompetanser, og analysen peker mot at det aktiverer både evalueringaspektet og tolkningsaspektet ved kompetansene i større grad enn oppgaver på lavere nivå; evalueringaspektet blir aktivert fordi vanskelige oppgaver krever at elevene reflekterer rundt sin egen fremgangsmåte og vurderer om de er på riktig vei, og tolkningsaspektet blir aktivert når elevene møter en helt eller delvis ukjent problemstilling som de må forstå. De må også tolke og evaluere hverandres faglige innspill. Et annet funn angående nivået på prøven, er at Jan øker vanskelighetsgraden på prøven for å kompensere for at elevene ofte gjør det bedre på parprøver enn på individuelle prøver. Tanken er at elevene får utbytte av den sosiale kontrakten – produktiv disposisjon – og muligheten til å kommunisere, og derfor mestrer oppgaver på høyere nivå. Ser vi på funn fra observasjonene fra oppgave 1c, finner vi at hos de parene som ikke opplever oppgaven som en triviell prosedyreoppgave - men derimot opplever at den ligger på et nivå over det de selv normalt ligger på - aktiveres flere kompetansekategorier. En utfordring ved å øke nivået for at elevene skal aktivere flere kompetanser eller for å kompensere for økte resultater, kan være at vurderingsformen da i noen grad favoriserer de som mestrer samarbeidet - de som besitter både kommunikasjonsevne og produktiv disposisjon. Samuelsson & Frykedal (2014) har funnet at oppgavens vanskelighetsgrad påvirker hvordan samarbeidet fungerer. Får elevene en oppgave som er på for høyt nivå vil de jobbe *i gruppe*, som vil si at de jobber individuelt med ulike deler av oppgaven, i stedet for å jobbe *som gruppe* og dra nytte av samarbeidet. Dette finner

vi eksempler på hos par C som ligger på lavt nivå og som gjør de fleste forsøkene på å løse oppgavene hver for seg. Vi ser det samme hos par A i møte med oppgave 4b.

Jan sier at det er viktig at parprøven også tester det «helt grunnleggende» og at «oppgave 1a og 1b bør [...] være greit angripelig for alle nesten uansett hvilket nivå de er på». Ser vi på oppgave 1a og 1b i prøvesettet (ganske enkle likninger) er det oppgaver som, i likhet med 1c, kun måler prosedyreevne og kun måler om minst én i paret husker prosedyren. Oppgavene er ikke på høyt nivå, og slike oppgaver sikrer at elever som er på lavt nivå og ikke mestrer samarbeidet, får muligheten til å vise kompetanse. Ulempen i en parprøvekontekst er at når oppgaven er på for lavt nivå, vil den i de aller fleste par mest sannsynlig kun aktivisere bare én elev, og kunne like gjerne blitt gitt på en individuell prøve. Samuelsson & Frykedal (2014) fant også i sin studie at dersom oppgaven er for lett vil dette føre til at elevene jobber hver for seg *i gruppe*. For å utnytte parprøvens mulighet til å aktivere flere kompetanser hos begge elevene, og samtidig holde et nivå som gjør oppgaven angripelig for alle, kunne en kanskje tatt i bruk en oppgavetype som Jan mener egner seg for parprøve – for eksempel en påstandsoppgave – og i tillegg lagt til noen støttestrukturer. Enten i form av formuleringer som får elevene til å diskutere eller argumentere, eller i form av øving på forhånd slik Jan mener at man bør. Ifølge Jan kan man ikke forvente at elevene vet hvordan de skal jobbe under en parprøve og elevene skal helst ha jobbet i par en stund i forkant av prøven slik at de blir vant med formen.

5.1.4 Normer

Å venne elevene til parprøveformen bidrar til en forventningsavklaring rundt vurderingssituasjonen. Jan sier at han blant annet vil gjøre det klart for elevene at det er begrunnelser som gir uttelling. Å snakke med elevene om hva som er et gyldig og tilstrekkelig svar handler om å etablere sosiomatematiske normer (Yackel & Cobb, 1996). Normene i klasserommet kan brukes til å få elevene til å jobbe på en måte som gjør at de får utbytte av at de er i par, ved at de for eksempel bruker argumenter på en spesifikk måte. Liljedahl (2018) bruker normer for å utvikle og opprettholde et tenkende klasserom. Oppgave 3a skaper noe uro i klassen fordi den er formulert på en måte som elevene ikke kjenner. Elevene søkte hjelp og bekreftelse fra meg i vurderingssituasjonen. I dette tilfellet kunne det vært hensiktsmessig om elevene var forberedt på at de kom til å møte ukjente oppgaver, og at de på forhånd hadde lært hvordan de skulle håndtere en slik situasjon. Sosiomatematiske normer kan også bidra til

at elevene viser den kompetansen de aktiverer ved at det for eksempel etableres at elevene skal skrive ned resonnementene sine rundt oppgaven selv om de ikke vet hvordan den skal løses. Denne klassen hadde ikke hatt parprøver tidligere, og hadde heller ikke vært gjennom alle forberedelsene Jan mener at de ideelt sett burde.

5.1.5 Observasjon

Et funn knyttet til parprøvens potensiale til å *måle* elevenes kompetanse, er at Jan trekker frem observasjon som en del av vurderingssituasjonen. Observasjonen inngår riktignok ikke i det formelle vurderingsgrunnlaget, men bidrar til et helhetsinntrykk som, slik det fremstår i intervjuet, kan påvirke Jans sluttvurdering av eleven. Han nevner blant annet at i oppgaver hvor paret skal bruke en prosedyre, kan han gjennom observasjoner plukke opp deler av prosessen og få et inntrykk av hvem i paret som husker prosedyren. Oftere å vurdere elevene uformelt er i tråd med praksisen til lærerne i Suurtamm (2004) sin forskning. En utfordring de møter er å oversette informasjonen de samler inn til en tallkarakter. Jan trekker også frem karakterer, og at en vanlig holdning er at ved individuelle prøver i matematikk kan man nærmest summere karakterene gjennom året og finne et snitt som blir elevens karakter, mens ved bruk av parprøver må lærer justere inntrykket ved hjelp av observasjoner. Dette betyr at mye avhenger av den enkelte lærers oppmerksomhet og evne til å tolke elevene i vurderingssituasjonen. Annen forskning på vurdering av gruppearbeid og andre alternative vurderingsformer trekker frem som en utfordring at vurderingen blir subjektiv (f.eks. Frykedal & Chiriac, 2011; Watt, 2005) – særlig på grunn av mangel på felles prosedyrer.

Et aspekt ved observasjon er lærerens kapasitet. Gjennom observasjon vil man få tilgang på kommunikasjonen mellom elevene, men en vesentlig begrensning er at det ikke vil være mulig for én lærer å observere 10 par på en måte som danner likt og valid vurderingsgrunnlag for alle. Fordi læreren ikke kan få med seg alt er det stor sjanse for at noen får vist mer kompetanse enn andre. En mulighet ved observasjon er at det tilbyr en fleksibilitet og en form for differensiering i vurderingssituasjonen fordi læreren kan velge seg par å observere. Det kan være en fordel dersom lærer trenger et rikere bilde av enkelte elevs kompetanse, eller dersom det er elever læreren vet har utfordringer med å uttrykke seg skriftlig. Jan forteller at han bruker observasjonene til å korrigere hvis han ser at én i paret bidrar i liten grad, og at slike observasjoner er en ressurs når vurderingen skal brukes formativt og det er viktig at begge i paret har forståelse for eget nivå.

5.1.6 Parsammensetning

Et organisatorisk grep Jan tar er ikke å involvere elevene når parene skal settes sammen. I tillegg forsøker han å så langt det går plassere elever på samme faglige nivå i samme par. Jan forklarer at dersom han blir nødt til å sette sammen to elever på ulikt nivå, korrigerer han for dette ved å snakke med elevene slik at de forstår hva karakteren betyr for hver enkelt. Dette tyder på at heterogene par først og fremst er en ulempe fordi det gjør den individuelle vurderingen av medlemmene i paret vanskeligere. Studier på gruppesammensetning viser at det ikke nødvendigvis er homogene par som gir best læringsutbytte. For eksempel viser både Leikin & Zaslavsky (1999) og Wiedmann et al. (2012) at lavtpresterende elever har best utbytte av å være på gruppe med elever på middels nivå slik at de kan dra nytte av den “kompetente andre” til å nå sin proksimale utviklingssone. Samtidig får de middels presterende utbytte av å forklare for den lavere presterende. Dette kan være verdt å merke seg, dersom man ønsker at parprøven også skal være en læringssituasjon. Samuelsson & Frykedal (2014) skriver at elevens selvtillit påvirker hvor aktive elevene er i samarbeidet, og at samarbeidet mellom to elever på lavt nivå fort kan stagnere når ingen har tro på at paret skal klare oppgaven. Jans erfaring er derimot at lavt presterende elever ofte presterer bedre på parprøver, og at det er grunnet den sosiale kontrakten. Observasjonene fra prøvesituasjonen gir delte indikasjoner; på den ene siden bruker paret på lavt nivå mye tid på å snakke om at de ikke mestrer og at de er “dumme”, men på den annen side gjør de flere ganger nye forsøk etter at de først har blitt enige om å gi opp. Dette kan tyde på at elevene har utbytte av å være i paret sammenliknet med å gjøre prøven individuelt, men at de trolig hadde prestert enda bedre om de hadde vært i par med en elev på litt høyere faglig nivå. Samuelsson & Frykedal (2014) bemerker at dersom differansen på elevenes nivå blir for stor vil den på lavest nivå ofte melde seg ut - og det er trolig dette Jan forsøker å unngå med homogene grupper.

Det at Jan setter sammen gruppene, er nødvendig for å få homogene par slik han ønsker. De to lavt presterende elevene i denne studien, par C, hadde ikke fått med seg at parene var homogene. Da de fant ut av dette skjønnte de at de begge var på lavt nivå og det så ut til å ha en demotiverende effekt på dem. Liljedahl (2014) undersøker i sin studie, effekten av å dele inn elevene i *synlig* tilfeldige grupper. Han fant blant annet at de sosiale barrierene forsvant og at elevene aksepterte å jobbe med hvem som helst fra klassen. Elevene ble også mer engasjerte i matematikk, mindre avhengige av læreren og stolte mer på seg selv og medelevenes kompetanse. Jan trekker frem at det er viktig at læreren bestemmer parene og at

det ikke er rom for diskusjon, dette antakelig for å få elevene til å bruke lite tid på å forhandle og raskt sette i gang med oppgavene. Liljedahl (2014) skriver at når læreren setter sammen parene er det ofte for å oppnå et mål, og at dette målet ofte er et annet enn elevene sitt mål. Med tilfeldige sammensatte grupper - som elevene vet at er tilfeldig - kan man oppnå at det ikke er rom for forhandlinger, samtidig som all undring og mistenksomhet rundt gruppesammensetningen blir borte og fokuset rettes mot matematikken.

5.1.7 Retteskjema

Jan bruker retteskjema når besvarelsene skal vurderes. Han legger vekt på samtale med elevene som en del av tilbakemeldingsarbeidet, og ønsker at elevene skal delta aktivt i vurderingen av seg selv. Han mener dette er viktig for elevenes forståelse av egen kompetanse og at samtalen bidrar til det uformelle vurderingsgrunnlaget. Jan har hatt behov for å systematisere vurderingen ved hjelp av retteskjema. Refleksjonene rundt vurdering og valget om å utvikle retteskjema er i tråd med teorien om formativ vurdering, og at det er hvordan man tar i bruk informasjonen om elevenes kompetanse som avgjør om vurderingen er formativ eller summativ (Black & Wiliam, 2018). En ulempe ved denne måten å vurdere på er at det er tidkrevende. Som Watt (2005) skriver, er behovet vurderingsformer som gir resultater som ikke må etterprøves, og om samtale med elevene er nødvendig for å få et riktig bilde av elevenes kompetanse kan det gjøre vurderingsformen for ressurskrevende. Samtidig er Jans hensikt at hele vurderingen skal gi læring og dersom elevene får sin kompetanse økt gjennom å gjøre parprøven og vurdere seg selv, kan tidsbruken i noe grad forsvares. Det er også verdt å minne om at Jan bruker retteskjema i alle sine matematikklærer, uavhengig av om det er parprøve. Han har brukt tid på å utvikle et system som han mener fungerer og ifølge han får elevene utbytte av det. Fordi han bruker det ved alle skriftlige prøver i matematikk er elevene godt trent, noe som antakelig også bidrar til at de forstår hvordan de skal bruke vurderingen til å støtte sin læring på en effektiv måte.

5.1.8 Læreren som forutsetning

Jan sier at det i stor grad er opp til læreren hvor mye man kan få ut av parprøven som vurderingsform. Han ønsker at prøven skal være en vurdering *for* og *som* læring, og legger vekt på observasjon, og tilpasning av blant annet oppgavetype og vanskelighetsgrad, for å nå målet. I tillegg gjør han organisatoriske grep for å bidra til at elevene får utbytte av

samarbeidet, og sørge for sikrere informasjon om elevenes kompetanse. Mye av det Jan tar opp er i tråd med Chapman (2013) sine punkter for hva en lærer må kunne for å lage gode oppgaver; kunnskap om hva en oppgave krever av elevene, elevenes forkunnskaper og hvordan klassedynamikken er i en vurderingssituasjon. Spesielt relevant er Chapman (2013) sine to siste punkter som handler om forståelse for at valg av oppgaver og implementering av oppgavene har noe å si for elevenes læring, og kunnskap om hvordan arbeidet med oppgavene bør organiseres. Lærer må gjøre liknende grep ved individuelle vurderinger, men en parprøve kan være ekstra krevende for læreren fordi praksisen ikke er like godt etablert. Jans refleksjoner rundt nivået på oppgavene og oppgavetyper tyder på at han er bevisst konteksten og ønsker å bruke egnede oppgaver og implementere dem på en måte som stimulerer elevene når de jobber i par. Han er engasjert i bruk av parprøver og jobber aktivt for å finne gode måter å bruke parprøven på. Han forteller at det er viktig at også læreren ser på vurderingen som en prosess, at formen kan utvikles og at man ikke nødvendigvis lykkes på første forsøk. En utfordring kan være mangel på fagfeller å utvikle vurderingsformen sammen med. Studien til Suurtamm (2004) finner at lærerne savner et fellesskap og at mangel på felles vurderingspraksis senker troverdigheten til de alternative vurderingsformene. Boesen et al. (2014) skriver at lærere også har en tendens til å ikke endre praksis fullt ut, men at de trekker inn elementer fra nye reformer og tilpasser dem til sin eksisterende praksis. Det kan også være utfordrende å introdusere nye metoder for elevene. 2P-klassen i dette prosjektet hadde aldri hatt parprøver før og flere stilte spørsmål om elementer som parsammensetning og oppgavetyper.

5.1.9 Kompetanseprofiler

Ser vi nærmere på hvordan parprøven *måler* kompetanse – det vil si elevenes mulighet til å vise kompetanse og ikke bare aktivere den – er det relevant å inkludere Jans utsagn om at “hvis de viser kunnskapen på papiret så er vi jo egentlig litt fornøyd med det”. Han forklarer videre at han stoler på at elevene får et læringsutbytte fra prosessene de må gjennom for å nå frem til det riktige svaret - selv om ikke alle deler av prosessen når frem til læreren og danner et direkte vurderingsgrunnlag. For eksempel mener Jan at kommunikasjonsevne er gunstig for elevene når de skal løse oppgavene. Funn fra observasjonen viser også at kommunikasjonsevnen blir aktivert når elevene jobber, men at den i liten grad kommer direkte til uttrykk i besvarelsene. Dette kan tolkes som et eksempel på at vurdering *som* læring er blitt vektlagt, og at det mangler noe i oppgavene som gjør at elevene får *vist* kompetanse.

Tidligere har jeg vært inne på at blant annet oppgavetype og oppgaveformulering kan bidra til at kompetansen kommer til syne i besvarelsen. Samtidig kan det være interessant å se dette i lys av definisjonen av kompetanse og Blömeke et al. (2015) sin diskusjon av kompetansebegrepet, kompetanseprofiler og korrelasjonen mellom kompetanseprofiler og atferd. Den ene delen av dikotomien dreide seg om at bak én atferd kan det ligge flere ulike kompetanseprofiler som er i endring gjennom atferden. Dette fremstår som forenelig med Jans utsagn om at elevene viser noe på papiret, og så skjer det i tillegg prosesser som lærer ikke har tilgang på. I noen par har kanskje resonneringsevnen vært aktivert i sparringen mellom elevene for å nå frem til svaret; mens i andre par har det kanskje ikke vært behov for sparring, og det har i større grad vært prosedyreevnen som har vært aktivert for å nå svaret. Fra dette perspektivet kan ikke læreren vite noe sikkert om kompetanseprofilene som har vært aktivert; dersom ikke oppgaven krever at det kommer eksplisitt til uttrykk i besvarelsen. Funn fra intervjumaterialet tyder samtidig på at Jan oppfatter noe grad av korrelasjon mellom atferd og kompetanseprofil – som er i tråd med det andre perspektivet i dikotomien. Når Jan øker nivået på oppgavene og hans intensjon er at elevene skal bruke kommunikasjon som et verktøy for å løse oppgavene, vil det si at det legges opp til at kompetanseprofilen som ligger til grunn for atferden, “eleven løser oppgaven korrekt”, inneholder kommunikasjonsevne. Dermed kan et korrekt svar på oppgaven indikere kommunikasjonsevne selv om denne kompetansen ikke kommer direkte til uttrykk i besvarelsen. En tanke om korrelasjon mellom besvarelse og kompetanse aktivert, kan utnyttes til å si noe om elevenes kompetanse ut ifra besvarelsen, men det krever at lærer har god kjennskap til elevenes forkunnskaper.

For å konkretisere poenget over, kan vi se på prosedyreevne og problemløsningsevne. Et *problem* er en oppgavetype som elevene ikke har løst tidligere (Lithner et al., 2010). Er oppgavetypen kjent for elevene, og de kun skal ta i bruk prosedyrer de kjenner fra før er ikke oppgaven et problem. For å avgjøre om en elev kun har vist prosedyreevne, eller om eleven også har vist problemløsningsevne må læreren vite hva som kan ansees som et problem for den aktuelle elevgruppen. For eksempel vil mange oppgaver som krever prosedyreevne av en 1T-elev, kreve problemløsningsevne av en 10. klassing – selv om besvarelsen de leverer er lik. Dette vil si at dersom læreren vet hva oppgavene krever av sine elever, kan læreren i sitt oppgavedesign sørge for at visse kompetanser blir aktivert.

5.1.10 Helhetlig syn på vurdering

Jan viser gjennom sine uttalelser at han er bevisst på sammenhengen mellom hva oppgavene krever og hvilke kompetanser som aktiveres. En slik bevissthet kan bidra til et riktigere bilde av elevenes kompetanse, men det krever mye av læreren. Også her kan det oppstå utfordringer med å oversette vurderingen til en tallkarakter, slik som lærerne i Suurtamm (2004) beskrev. Dette kan tyde på at det er behov for en større og helhetlig endring i vurderingspraksis. Jans praksis nå er å gjennomføre rundt én parprøve i halvåret. Suurtamm (2004) skriver at en alternativ vurderingspraksis bør bestå av flere ulike vurderingsformer slik at alle aspekter ved matematisk kompetanse kan bli vurdert. En av utfordringene lærerne i Suurtamm (2004) studie møtte, var å kombinere alternative vurderingsformer med tradisjonelle forventninger om rapportering av resultater. De opplevde at for å forberede elevene til eksamen var de nødt til å gå gjennom mange temaer på kort tid, noe som førte til fokus på prosedyreevne og effektivitet, snarere enn det lærerne anså som viktig matematikk. Skulle de bruke alternative vurderingsformer ble det dermed et behov for å dreie de mot det elevene ville bli testet i på eksamen. Lærerne ønsket en ny læreplan som ga mer rom for problemløsning; da ville det bli lettere å bruke vurderingsformene som de mente dekket flere aspekter ved matematisk kompetanse (Suurtamm, 2004). I Norge ble det gjennomført et kunnskapsløft i 2020, og Jan mener parprøver er en egnet vurderingsform for å ivareta den matematikken som vektlegges i de nye læreplanene.

Jan sier at han av og til er nødt til å forklare elevene hva karakteren innebærer. På en side kan det hjelpe elevene å forstå hvilket nivå de ligger på, men for at vurderingen skal virke formativt er det viktig at de forstår hva de kan gjøre for å utvikle seg (Black & Wiliam, 1998). Jan forteller at det også er nødvendig å presisere for elevene at sluttvurderingen ikke er et snitt av karakterer gjennom året. Kanskje kan redusert bruk av karakterer i undervisvurdering, gjøre at lærer føler seg mindre bundet og at det slik kan bidra til en bedre forståelse av elevenes kompetanse for både lærer og elev.

5.2 Konklusjon

Et av oppgavens hovedfunn er at læreren i denne studien anser parprøven som en anledning til å praktisere vurdering *som læring*. Dette ser ut til å påvirke hvordan læreren utnytter parprøvens potensiale til å vurdere elevenes kompetanse. Han legger vekt på at det ideelt sett

skjer noen prosesser hos elevene når de jobber med parprøven, som gir dem mulighet til å aktivere og utvikle kompetanse og erfare at samarbeid kan øke læringsutbyttet. Dette er prosesser som ikke nødvendigvis kommer til uttrykk i elevenes besvarelse, men som ifølge læreren kan være verdifulle i et mer langsiktig læringsløp. Vi finner indikasjoner på at elevene aktiverer et bredt spekter av kompetanser i arbeidet med prøven, noe som tyder på at formen egner seg til vurdering som læring, slik læreren ønsker. Særlig interessant er det å se at kommunikasjons-, resonnerings- og connectionevne også i noen tilfeller blir aktivert når oppgaven, analysert ved hjelp av MCRF + pd, kun krever prosedyreevne. Dette tyder på at parprøvekonteksten kan bidra til å løfte oppgavens potensiale til å aktivere kompetanse. Det kan også ha sammenheng med at læreren øker det faglige nivået på oppgavene og at elevene derfor har behov for å aktivere flere kompetanser for å håndtere prosedyren.

Et annet hovedfunn dreier seg om parprøvens potensiale til å måle kompetanse. Funnene indikerer at parprøven aktiverer ulike kompetanser, men at elevene ofte ikke viser tilsvarende kompetanse i besvarelsen. Det er nødvendig at elevene får vist sin kompetanse for at prøven skal utgjøre et vurderingsgrunnlag i lys av vurdering *for* læring – et viktig prinsipp i norsk vurderingspraksis (Utdanningsdirektoratet, 2019). Læreren trekker frem at lærerens valg av oppgavetype og oppgaveformuleringer kan bidra til at elevene får vist kompetanse. Samtidig viser denne studien at det ikke alltid er samsvar mellom oppgavetyper og -formuleringene læreren beskriver som egnede, og oppgavetyper og -formuleringene vi finner i parprøven. Dette kan være en av årsakene til at elevene ikke viser den kompetansen de aktiverer. Læreren legger også vekt på at elevene må lære hvordan de skal jobbe for å få utbytte av potensialet i parprøven, og læreren må avklare forventningene til elevene. Dette innebærer å etablere sosiomatematiske normer som blant annet sier noe om hva som utgjør en tilstrekkelig besvarelse. Dette var første gang elevene i studien hadde parprøve, og det er mulig de ville vist mer kompetanse dersom de hadde mer erfaring og normene var bedre etablert i klassen.

Et tredje hovedfunn er at læreren legger vekt på observasjoner fra gjennomføringen av prøven, og samtale med elevene, som en del av vurderingen. Observasjonene og samtalene utgjør en del av det uformelle vurderingsgrunnlaget som bidrar til et helhetsinntrykk av elevenes kompetanse. Læreren beskriver at han gjennom observasjonene får tilgang på informasjon om prosessene i parene og hvem som aktiverer hvilke kompetanser. En utfordring er lærerens kapasitet – det vil ikke være mulig for én lærer å observere 10 par og danne et rettferdig og valid vurderingsgrunnlag for alle. Derimot gir observasjonene mulighet

til å differensiere, og læreren har mulighet til å velge ut enkelte par å fokusere på dersom det er hensiktsmessig.

5.2.1 Implikasjoner

Læreren trekker frem en rekke grep som skal bidra til å aktivere og måle elevenes kompetanse. Ikke alle disse grepene er indikert i prøvedesignet og observasjonene. For videre studier kunne det vært nyttig å designe en parprøve med utgangpunkt i grepene læreren foreslår og se hvordan disse påvirker elevenes mulighet til å vise kompetanse. Tidligere forskning etterspør noen felles prosedyrer for bruk av parprøver (Bjørk & Theodorsen, 2019) og vurdering av gruppearbeid (Frykedal & Chiriak, 2011) til å innhente informasjon om elevenes kompetanse, og en slik studie vil mulig kunne bidra til dette.

Jeg opplever en vilje hos læreren til å gi elevene flere anledninger og varierte muligheter til å vise sin kompetanse, i tråd med Forskrift til opplæringslova (2006). Lærerens fokus på vurdering *som* læring og utbyttet elevene får på sikt tyder på en mer helhetlig forståelse av vurdering. Læreren mener blant annet at det vil være gunstig å inkludere flere aktiviteter som bidrar til det uformelle vurderingsgrunnlaget. For videre studier mener jeg det kunne vært interessant å se på systematisering av det uformelle vurderingsgrunnlaget; og om man kan utvikle en vurderingspraksis som benytter ulike vurderingsformer, baserer seg mer på et helhetsinntrykk og som i noe grad er frigjort fra kravet om å oversette informasjonen om elevenes kompetanse til en karakter, men som samtidig gir et riktig bilde av elevenes kompetanse.

Funn fra denne studien indikerer at elevene ikke får vist all kompetanse de har aktivert. Samtidig tyder lærerens refleksjoner rundt oppgavedesignet og gjennomføringen på at han er bevisst parprøvens potensiale til å aktivere kompetanse. Funnene indikerer også at oppgavene ofte aktiverer de samme kompetansene som læreren beskriver at parprøven egner seg til å vurdere. Jeg mener det kunne vært interessant å studere sammenhengen mellom kompetanse aktivert og besvarelse for et større utvalg elever. Dette vil kunne gi mer informasjon til læreren om de eventuelle prosessene elevene har i samarbeidet som vanskelig kommer eksplisitt til uttrykk i besvarelsen.

Litteraturliste

- Befring, E. (2015). Forskningsetikk. In E. Befring (Ed.), *Forskningsmetoder i utdanningsvitenskap* (pp. 28-35). Cappelen Damm akademisk.
- Bergqvist, T. (2013). Podcasting Mathematics. *International Journal for Technology in Mathematics Education*, 20(4), 147.
- Bjørk, T. B., & Theodorsen, J. H. (2019, 03.10.2019). Et forsøk med samarbeidsprøver i matematikk ga økt motivasjon og minsket stressnivået blant elevene. *Utdanningsnytt.no*. <https://www.uttanningsnytt.no/fagartikkel-karakterer-laererarbeid/et-forsok-med-samarbeidsprover-i-matematikk-ga-okt-motivasjon-og-minsket-stressnivaet-blant-elevne/171518>
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in education : principles, policy & practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in education : principles, policy & practice*, 25(6), 551-575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Blikstad-Balas, M., & Dalland, C. P. (2021). Forskningsdesign - hva må du tenke på når du skal planlegge et forskningsprosjekt? In E. Andersson-Bakken & C. P. Dalland (Eds.), *Metoder i klasseromsforskning: Forskningsdesign, datainnsamling og analyse* (pp. 21-45). Universitetsforlaget.
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond Dichotomies: Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3-13. <https://doi.org/10.1027/2151-2604/a000194>
- Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2014). Developing mathematical competence: From the intended to the enacted curriculum. *The Journal of mathematical behavior*, 33(1), 72-87. <https://doi.org/10.1016/j.jmathb.2013.10.001>
- Boesen, J., Lithner, J., & Palm, T. (2018). Assessing mathematical competencies: an analysis of Swedish national mathematics tests. *Scandinavian journal of educational research*, 62(1), 109-124. <https://doi.org/10.1080/00313831.2016.1212256>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2017). Thematic analysis. *The journal of positive psychology*, 12(3), 297-298. <https://doi.org/10.1080/17439760.2016.1262613>
- Brinkmann, S., & Kvale, S. (2018). *Doing Interviews* (Second ed., Vol. 2). 55 City Road: SAGE Publications Ltd. <https://doi.org/10.4135/9781529716665>

- Chapman, O. (2013). Mathematical-task knowledge for teaching. *Journal of mathematics teacher education*, 16(1), 1-6. <https://doi.org/10.1007/s10857-013-9234-7>
- Chinn, S. (2009). Mathematics anxiety in secondary students in England. *Dyslexia*, 15(1), 61-68. <https://doi.org/10.1002/dys.381>
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8 ed., Vol. 1). London: Routledge. <https://doi.org/10.4324/9781315456539>
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. 39(3), 124-130. Retrieved 2000-08-01, from <https://www.jstor.org/stable/1477543>
- Dalland, C. P., Bjørnstad, E., & Andersson-Bakken, E. (2021). Observasjon som metode i barnehage- og klasseromsforskning. In E. Andersson-Bakken & C. P. Dalland (Eds.), *Metoder i klasseromsforskning: Forskningsdesign, datainnsamling og analyse* (pp. 125-152). Universitetsforlaget.
- Dempsey, N. P. (2010). Stimulated Recall Interviews in Ethnography. *Qualitative sociology*, 33(3), 349-367. <https://doi.org/10.1007/s11133-010-9157-x>
- Everett, E. L., & Furseth, I. (2012). Lettere sagt enn gjort - å utforme et metodisk opplegg for oppgaven. In I. Furseth (Ed.), *Masteroppgaven : hvordan begynne - og fullføre* (2. utg. ed., pp. 127-144). Universitetsforl.
- Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The Math Anxiety-Performance Link: A Global Phenomenon. *Curr Dir Psychol Sci*, 26(1), 52-58. <https://doi.org/10.1177/0963721416672463>
- Forsell, J., Forslund Frykedal, K., & Hammar Chiriac, E. (2020). Group Work Assessment: Assessing Social Skills at Group Level. *Small group research*, 51(1), 87-124. <https://doi.org/10.1177/1046496419878269>
- Forskrift til opplæringslova, (2006). <https://lovdata.no/forskrift/2006-06-23-724>
- Frykedal, K. F., & Chiriac, E. H. (2011). Assessment of students' learning when working in groups. *Educational research (Windsor)*, 53(3), 331-345. <https://doi.org/10.1080/00131881.2011.598661>
- Gleiss, M. S., & Sæther, E. (2021). *Forskningsmetode for lærerstudenter: Å utvikle ny kunnskap i forskning og praksis*. Cappelen Damm.
- Johnson, B. R. (2017). Validity of Research Results in Quantitative, Qualitative and Mixed Research. In L. B. Christensen (Ed.), *Educational research : quantitative, qualitative, and mixed approaches* (Sixth edition ed., pp. 277-316). SAGE Publications.
- Kilpatrick, J. (2020). Competency Frameworks in Mathematics Education. In S. Lerman & S. Lerman (Eds.), *Encyclopedia of Mathematics Education* (Second Edition ed., pp. 85-87). Springer International Publishing AG. <https://doi.org/10.1007/978-3-030-15789-0>
- Kilpatrick, J., Findell, B., Swafford, J., Council, N. R., Education, D. o. B., Sciences, S., Education, C. f., Committee, M. L. S., Kilpatrick, J., Swafford, J., & Findell, B.

- (2001). The strands of mathematical proficiency. In K. Jeremy, F. Bradford, & S. Jane (Eds.), *Adding It Up: Helping Children Learn Mathematics* (pp. 115-157). National Academies Press. <https://doi.org/10.17226/9822>
- Kunnskapsdepartementet. (2020). *Kompetanse i fagene*. Retrieved from <https://www.udir.no/lk20/overordnet-del/prinsipper-for-laring-utvikling-og-danning/kompetanse-i-fagene/>
- Leikin, R., & Zaslavsky, O. (1999). COOPERATIVE LEARNING IN MATHEMATICS. *The Mathematics teacher*, 92(3), 240-246.
- Lester, F. K. (2009). On the Theoretical, Conceptual, and Philosophical Foundations for Research in Mathematics Education. In (pp. 67-85). Berlin, Heidelberg: Berlin, Heidelberg: Springer Berlin Heidelberg.
- Liljedahl, P. (2014). The Affordances of Using Visibly Random Groups in a Mathematics Classroom. In (pp. 127-144). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-04993-9_8
- Liljedahl, P. (2016). Building Thinking Classrooms: Conditions for Problem-Solving. In (pp. 361-386). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-28023-3_21
- Liljedahl, P. (2018). Building Thinking Classrooms. In (pp. 307-316). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-92390-1_29
- Lithner, J., Bergquist, E., Bergquist, T., Boesen, J., Palm, T., & Palmberg, B. (2010). Mathematical Competencies: a Research Framework. *Papers*.
- Mok, M. M. C. (2013). Assessment Reform in the Asia-Pacific Region: The Theory and Practice of Self-Directed Learning Oriented Assessment. In M. M. C. Mok (Ed.), *Self-directed Learning Oriented Assessments in the Asia-Pacific* (pp. 3-22). Springer Netherlands. https://doi.org/10.1007/978-94-007-4507-0_1
- NCTM. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.
- Niss, M., Bruder, R., Planas, N., Turner, R., & Villa-Ochoa, J. A. (2016). Survey team on : conceptualisations of the role of competencies, knowing and knowledge in mathematics education research.
- Niss, M. A. (1993). *Investigations into Assessment in Mathematics Education : An ICMI Study* (1st 1993. ed., Vol. 2). Springer Netherlands : Imprint: Springer.
- Niss, M. A. (2003). MATHEMATICAL COMPETENCIES AND THE LEARNING OF MATHEMATICS: THE DANISH KOM PROJECT.
- Niss, M. A. (2007). Reflections on the State of and Trends in Research on Mathematics Teaching and Learning: From Here to Utopia. In F. K. Lester Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 1293-1312). Information Age Publishing.

- Niss, M. A., & Højgaard Jensen, T. (2002). *Kompetencer og matematiklæring*.
<https://www.matematikkensenteret.no/sites/default/files/attachments/page/Kompetencer%20og%20matematikk%C3%A6ring.pdf>
- Niss, M. A., & Højgaard, T. (2019). Mathematical competencies revisited. *Educational studies in mathematics*, 102(1), 9-28. <https://doi.org/10.1007/s10649-019-09903-9>
- Patton, M. Q. (2014). Module 29: Data Collection Decisions. In M. Q. Patton (Ed.), *Qualitative research & evaluation methods* (4. ed., pp. 255-263). Sage Publications.
- Pettersen, A. (2019). *Towards competency-oriented mathematics education: An investigation of task demands and teachers' knowledge of task demands from a competency perspective* [Doktorgradsavhandling Universitetet i Oslo, Universitetet i Oslo].
- Pettersen, A., & Nortvedt, G. A. (2017). Identifying Competency Demands in Mathematical Tasks: Recognising What Matters. *International journal of science and mathematics education*, 16(5), 949-965. <https://doi.org/10.1007/s10763-017-9807-5>
- Samuelsson, J., & Frykedal, K. F. (2014). Task Influences on Students' Collaboration in Mathematics Teaching. *Anadolu Journal of Educational Sciences International*.
<https://doi.org/10.18039/ajesi.12753>
- Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., Silver, E., Ufer, S., & Vos, P. (2016). *Assessment in Mathematics Education : Large-Scale Assessment and Classroom Assessment* (1st ed. 2016. ed.). Springer International Publishing : Imprint: Springer.
- Suurtamm, C. A. (2004). Developing Authentic Assessment: Case Studies of Secondary School Mathematics Teachers' Experiences. *Canadian journal of science, mathematics and technology education*, 4(4), 497-513.
<https://doi.org/10.1080/14926150409556630>
- Svenkerud, S. W. (2021). Intervjuer i klasseromsforskning. In E. Andersson-Bakken & C. P. Dalland (Eds.), *Metoder i klasseromsforskning. Forskningsdesign, datainnsamling og analyse* (pp. 91-103). Universitetsforlaget.
- UiO. (2017, 25.03.2021). *f4transkript*. Universitetet i Oslo.
<https://www.uio.no/tjenester/it/forskning/datafangst-og-analyse/f4-transcribe/>
- Utdanningsdirektoratet. (2019). *Erfaringer fra nasjonal satsing på vurdering for læring (2010-2018)*. <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/erfaringer-fra-nasjonal-satsing-pa-vurdering-for-laring-2010-2018/1.innledning/#>
- Utdanningsdirektoratet. (u.å.). *Fagfornyelsen*. <https://www.udir.no/laring-og-trivsel/lareplanverket/fagfornyelsen/>
- Valenta, A., Nosrati, M., Åsenhus, R., & Wæge, K. (2015). *Skisse av den "ideelle læreplan i matematikk"*. N. s. f. m. i. opplæringen.
file:///C:/Users/jenny/OneDrive/Documents/uio/Master/tekster/Teori/Valenta%20(2015)%20Skisse%20av%20den%20ideelle%20%C3%A6replan%20i%20matematikk.pdf

- Verschuren, P. J. M. (2003). Case study as a research strategy: Some ambiguities and opportunities. *International journal of social research methodology*, 6(2), 121-139. <https://doi.org/10.1080/13645570110106154>
- Watt, M. G. H. (2005). Attitudes to the Use of Alternative Assessment Methods in Mathematics: A Study with Secondary Mathematics Teachers in Sydney, Australia. *Educational studies in mathematics*, 58(1), 21-44. <https://doi.org/10.1007/s10649-005-3228-z>
- Weinert, F. E. (2001). Competencies and Key Competencies: Educational Perspective. In (pp. 2433-2436). Elsevier Ltd. <https://doi.org/10.1016/B0-08-043076-7/02384-6>
- Wiedmann, M., Leach, R. C., Rummel, N., & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional science*, 40(4), 711-730. <https://doi.org/10.1007/s11251-012-9204-y>
- Yackel, E., & Cobb, P. (1996). Sociomathematical Norms, Argumentation, and Autonomy in Mathematics. *Journal for research in mathematics education*, 27(4), 458-477. <https://doi.org/10.2307/749877>
- Yackel, E., & Rasmussen, C. (2002). Beliefs and norms in the mathematics classroom. In G. C. Leder, E. Pehkonen, G. Törner, & SpringerLink (Eds.), *Beliefs: A Hidden Variable in Mathematics Education?* (Vol. 31). Springer Netherlands.
- Yan, Z., & Yang, L. (2022). Assessment-as-learning in the global assessment reforms. In Z. Yan & L. Yang (Eds.), *Assessment as learning : maximising opportunities for student learning and achievement* (pp. 1-7). Routledge, Taylor & Francis Group.

Vedlegg 1: Parprøven

Kapittelprøve 3 Likninger og ulikheter

Del 1

Oppgave 1

Løs likningene

a $2x + 3 = 5$

b $\frac{10}{x} = \frac{5}{7}$

c $8^{3x} = 64$

Oppgave 2

Lise har løst en oppgave og fått løsningen $x = 3$ og $y = 2$. Hvilken av oppgavene under er det Lise har løst? Husk å begrunne svaret!

$17x - 12 = x^2 - 2$

$25 + x^2 \leq 21$

$36x + 3y = 24x - y = 1$

Oppgave 3

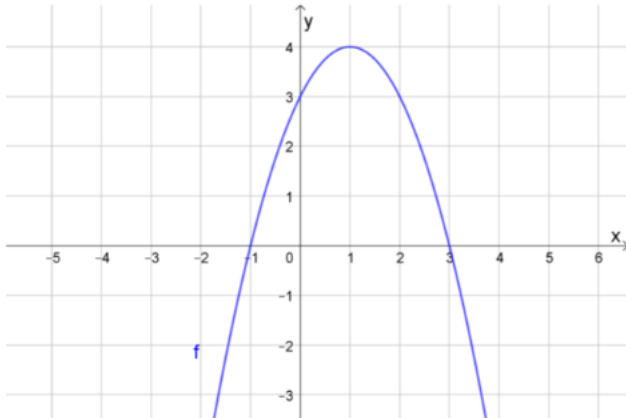
Sigurd skal løse en likning og starter slik:

$$x = \frac{-7 \pm \sqrt{7^2 - 4 \cdot 1 \cdot 10}}{2 \cdot 1}$$

aHvilken likning er det Sigurd prøver å løse? Skriv opp likningen.

bFullfør løsningen av likningen.

Oppgave 4



I koordinatsystemet ovenfor har vi tegnet grafen til en andregradsfunksjon f gitt ved $f(x) = -x^2 + 2x + 3$.

a) Hva er løsningen på likningen $-x^2 + 2x + 3 = 0$?

b) Løs ulikheten $f(x) > 3$.

Del 2

Oppgave 5

- Vis at $2x^2 - 6x + 8 = 0$ ikke har noen løsning.
- Samira påstår hun kan se om en andregradslikning har en, to eller ingen løsninger ved å kun regne en liten del av ABC-formelen. Er du enig eller uenig med Samira? Begrunn svaret!

Oppgave 6

Ni sitroner og sju epler koster til sammen 107, og fire sitroner og elleve epler koster til sammen 87. Hvor mye koster ett eple og en sitron?

- Sett opp likningssystem som beskriver situasjonen.
- Løs likningssystemet

Oppgave 7

Sylvi har løst ulikheten $2x + 3 < 9 - x$ og kommet frem til $x > 3$.

- a) Vis at Sylvi har fått feil svar
- b) Løs ulikheten grafisk

Vedlegg 2: Retteskjema

Retteskjema Prøve 2P 19.11.

Oppgave	Kommentar	Poeng
1a Lineær likning		/2
1b Kryssmultiplisering		/2
1c Potenslikning		/2
2 Matche løsning til oppgave		/2
3a Se andregradslikning ut fra abc-formelen		/2
3b Regne abc-formelen		/2
4a Løse likning fra graf		/2
4b Løse ulikhet fra graf		/2
5a Vise at en andregradslikning ikke har løsning		/2
5b Forklare når en andregradslikning har en, to eller ingen løsning		/2
6a Likningssystem fra tekst		/2
6b Løse likningssystem		/2
7a Vise at et svar er feil i ulikhet		/2
7b Løse ulikhet grafisk		/2
TOTALT		/28

Vedlegg 3: Informasjonsskriv og samtykkeskjema, lærer

Vil du delta i forskningsprosjektet

«*Vurdering av matematisk kompetanse gjennom parprøver*»?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å komme nærmere en felles terminologi og prosedyrer for bruk av parprøver til måling av elevenes kompetanse i matematikk. Parprøver er i dette prosjektet definert som vurderingssituasjoner hvor elevene jobber i par og leverer et felles skriftlig produkt. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Kunnskap om elevenes kompetanse er en viktig del av undervisningsvurderingen, der elevene både skal få forståelse for hva de mestrer, og få råd om hvordan de kan jobbe for å øke kompetansen sin. Det kan være vanskelig å vurdere hver enkelt elevs kompetanse i matematikk når elevene jobber i par og leverer et felles skriftlig produkt. Gjennom denne oppgaven vil jeg undersøke noen av mulighetene parprøver gir, og hvilke hensyn lærer må ta i oppgavedesign og retting for å få informasjon om hver enkelt elevs kompetanse i matematikk.

Oppgaven tar for seg parprøver som vurderingssituasjon. Med en læreplan i endring er det behov for å utvikle undervisningspraksisen slik at den støtter elevenes læring. Videre er det viktig at det er samsvar mellom undervisning og vurderingspraksis, og også mellom det som foregår på skolen og det elevene møter etter endt skolegang. Med et økende fokus på samarbeid på skolen og i arbeidslivet synes jeg det er interessant og relevant å undersøke hvordan samarbeid kan inkluderes i vurdering i skolen.

Den overordnede problemstillingen prosjektet skal forsøke å besvare er

Hvilke hensyn må lærere ta for å sikre informasjon om den matematiske kompetansen til hver enkelt elev i en prøvesituasjon hvor elevene jobber i par?

Videre ønsker jeg å finne ut:

1. Hvilke kompetanser er det ønsket at parprøvene skal måle og hvilke kompetanser kan parprøvene måle?
2. Hvilke hensyn tar lærer i oppgavedesignet for å sørge for informasjon om elevenes kompetanse gjennom en parprøve?

3. Hvilke hensyn tar lærer i retteprosessen for å sørge for informasjon om elevenes kompetanse gjennom en parprøve?

Problemstillingene skal besvares ved oppgaveanalyse av parprøver, observasjon av elever som jobber med parprøver og intervju av lærere.

Prosjektet er en masteroppgave og utgjør avslutningen på min (Jenny Engh) femårige utdanning ved Lektorprogrammet i realfag – matematikk og kjemi.

Hvem er ansvarlig for forskningsprosjektet?

Institutt for lærerutdanning og skoleforskning (ILS) ved Universitetet i Oslo (UiO) er ansvarlig for prosjektet. Masterstudent er Jenny Engh, og prosjektet er veiledet av stipendiat ved ILS Alexander Jonas Viktor Selling.

Hvorfor får du spørsmål om å delta?

Du får spørsmål om å delta fordi du underviser i matematikk på en videregående skole, og har minst én erfaring med å gjennomføre parprøver.

Hva innebærer det for deg å delta?

Hvis du velger å delta i prosjektet, innebærer det at du

- deler et oppgavesett fra en parprøve du selv har laget. Oppgavene skal i løpet av prosjektet løses av et utvalg elever, analyseres for å finne ut hvilke kompetanser som kreves for å løse dem og brukes som utgangspunkt i intervju
- deltar på to intervjuer. Det vil ta deg ca. 45 minutter per intervju. I intervjuene vil vi ta utgangspunkt i oppgavesettet du har laget. Du vil få spørsmål om designet, om retteprosessen og eventuelle hensyn du har tatt for å få informasjon om hver enkelt elevs kompetanse når de leverer et felles produkt. Jeg tar lydopptak og notater fra intervjuet. Du kan bli kontaktet med forespørsel om et kort oppfølgingsintervju i etterkant av intervjuet, dersom jeg opplever at det er behov.

Det er frivillig å delta

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

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

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

- Det er kun jeg, Jenny Engh, og min veileder, Alexander J. V. Selling som vil ha tilgang til dine opplysninger.
- Alle data pseudoanonymiseres, det vil si at jeg erstatter navnet ditt med en kode som lagres på en egen navneliste adskilt fra øvrige data, i henhold til UiOs lagringsguide.

- Lyddopptakeren tilhører behandlingsansvarlig institusjon, ILS. Det er kun jeg som skal håndtere lyddopptakeren. Lyddopptakene transkriberes av meg og slettes når prosjektet avsluttes.
- Lyddopptak skal kun behandles på sikker maskinvare tilhørende behandlingsansvarlig institusjon, ILS, i henhold til UiOs lagringsguide.

Ingen personidentifiserende informasjon vil publiseres. Utdrag fra oppgavesettet vil kunne publiseres i oppgaven. Det samme gjelder korte, ikke-personidentifiserende sitater fra intervju.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er i juni 2022. Personopplysninger og opptak slettes ved prosjektslutt.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra ILS ved UiO har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Dine rettigheter

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

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

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

- ILS ved Jenny Engh på epost (jennyen@uv.uio.no) eller telefon: 452 98 724; eller Alexander Jonas Viktor Selling på epost (a.j.v.selling@ils.uio.no).
- Vårt personvernombud: Roger Markgraf-Bye (personvernombud@uio.no, 908 22 826)

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

- NSD – Norsk senter for forskningsdata AS på epost (personvertjenester@nsd.no) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Alexander Jonas Viktor Selling
Veileder

Jenny Engh
Masterstudent

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Vurdering av matematisk kompetanse gjennom parprøver», og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i intervju
- at utdrag fra mitt oppgavesett publiseres i masteroppgaven

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

(Signert av prosjektdeltaker, dato)

Vedlegg 4: Informasjonsskriv og samtykkeskjema, elev

6 Vil du delta i forskningsprosjektet

«Vurdering av matematisk kompetanse gjennom parprøver»?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å komme nærmere en måte å bruke parprøver på, slik at lærer får informasjon om elevenes kompetanse i matematikk. Parprøver er i dette prosjektet definert som vurderingssituasjoner hvor elevene jobber i par og leverer et felles skriftlig produkt. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Opgaven handler om parprøver som vurderingssituasjon. Gjennom skoleåret gjennomføres det små og store vurderingssituasjoner i matematikk. Disse kalles underveisvurdering og målet er blant annet at eleven og læreren skal få forståelse for hva eleven mestrer. Det kan være vanskelig å vurdere hver enkelt elevs kompetanse i matematikk når elevene jobber i par og leverer et felles skriftlig produkt. Gjennom denne oppgaven vil jeg undersøke noen av mulighetene parprøver gir, og hvilke hensyn lærer må ta i oppgavedesign og retting for å få informasjon om hver enkelt elevs kompetanse i matematikk.

Det er viktig at det er samsvar mellom undervisning og vurdering, og også mellom det som foregår på skolen og det elevene møter etter endt skolegang. Med et økende fokus på samarbeid på skolen og i arbeidslivet synes jeg det er interessant og relevant å undersøke hvordan samarbeid kan inkluderes i vurdering i skolen.

Den overordnede problemstillingen prosjektet skal forsøke å besvare er

Hvilke hensyn må lærere ta for å sikre informasjon om den matematiske kompetansen til hver enkelt elev i en prøvesituasjon hvor elevene jobber i par?

Videre ønsker jeg å finne ut:

4. Hvilke kompetanser er det ønsket at parprøvene skal måle og hvilke kompetanser kan parprøvene måle?
5. Hvilke hensyn tar lærer i oppgavedesignet og retteprosess for å sørge for informasjon om elevenes læring gjennom en parprøve?

Problemstillingene skal besvares ved oppgaveanalyse av parprøver, observasjon av elever som jobber med parprøver og intervju av lærere.

Prosjektet er en masteroppgave og er avslutningen på min (Jenny Engh) femårige utdanning ved Lektorprogrammet i realfag – matematikk og kjemi.

Hvem er ansvarlig for forskningsprosjektet?

Institutt for lærerutdanning og skoleforskning (ILS) ved Universitetet i Oslo (UiO) er ansvarlig for prosjektet. Masterstudent er Jenny Engh, og prosjektet er veiledet av stipendiat ved ILS Alexander Jonas Viktor Selling.

Hvorfor får du spørsmål om å delta?

I dette prosjektet er det ønsket at læreren som deltar kjenner elevene som deltar. Læreren din i matematikk har takket ja til å delta i prosjektet, og derfor blir du spurt om å delta. Alle i matematikklassen din får denne forespørselen.

Hva innebærer det for deg å delta?

Hvis du velger å delta i dette prosjektet innebærer det at du skal jobbe med et oppgavesett som læreren din har laget. Du skal samarbeide med en annen elev fra klassen din. Jeg tar lydopptak og notater mens dere jobber. Oppgavesettet og noen av svarene dine vil bli brukt i et intervju med læreren din senere.

Jeg vil også be læreren din om å oppgi hvilket nivå du vanligvis ligger på i matematikk.

Det er frivillig å delta

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

Å delta i prosjektet vil ikke påvirke din karakter i faget eller din relasjon til læreren din.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

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

- Det er kun jeg, Jenny Engh, og min veileder, Alexander J. V. Selling som vil ha tilgang til dine opplysninger.
- Alle data pseudoanonymiseres, det vil si at jeg erstatter navnet ditt med en kode som lagres på en egen navneliste adskilt fra øvrige data, i henhold til UiOs lagringsguide.
- Lydopptakeren tilhører behandlingsansvarlig institusjon, ILS. Det er kun jeg som skal håndtere lydopptakeren. Lydopptakene transkriberes av meg og slettes når prosjektet avsluttes.
- Lydopptak skal kun behandles på sikker maskinvare tilhørende behandlingsansvarlig institusjon, ILS, i henhold til UiOs lagringsguide.

Ingen personidentifiserende informasjon vil publiseres. Transkriberte utdrag fra din besvarelse av oppgavesettet og eventuelle kladdemark vil kunne publiseres i oppgaven. Det samme gjelder sitater fra samtalen mellom deg og den du løser prøven med.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er i juni 2022. Personopplysninger og opptak slettes ved prosjektslutt.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Dine rettigheter

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

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

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

- ILS ved Jenny Engh på epost (jennyen@uv.uio.no); eller Alexander Jonas Viktor Selling på epost (a.j.v.selling@ils.uio.no).
- Vårt personvernombud: Roger Markgraf-Bye (personvernombud@uio.no, 908 22 826)

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

- NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Alexander Jonas Viktor Selling
Veileder

Jenny Engh
Masterstudent

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Vurdering av matematisk kompetanse gjennom parprøver», og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i observasjon hvor det tas lydopptak
- at lærer kan gi opplysninger om meg til prosjektet
- at transkripsjon av utdrag fra min besvarelse og mitt kladdemark publiseres i masteroppgaven

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

(Signert av prosjektdeltaker, dato)

Vedlegg 5: Intervjuguide

Intervjuguide

Intervju 7. februar 2022

Formål:

Komme nærmere prosedyrer for bruk av parprøver til å måle elevers matematiske kompetanse.

Intro

Kan du fortelle om din bakgrunn og tidligere erfaring med parprøver?

Er det kultur for bruk av parprøver på denne skolen?

Hva synes du om bruk av parprøver?

Muligheter

Hvilke muligheter gir parprøver?

Når egner det seg med parprøve?

Hvilke kompetanser egner den seg til å måle?

Hvilke temaer egner den seg for?

Hvilke muligheter ligger i parprøven sammenliknet med andre vurderingsformer?

Prøvedesign

Hva er viktig for å lage en god parprøve?

Hva slags spørsmål fungerer/fungerer ikke

Tenker du annerledes når du lager en parprøve sammenliknet med en individuell prøve?

Oppgave 1c / oppgave 3 / oppgave 4

Hva er hensikten med denne oppgaven?

Hva skal elevene vise?

Hva tror du skjer hos elevene når de får denne oppgaven?

Hva i oppgaven skal sørge for at elevene viser det du ønsker?

Er oppgaven god – er du fornøyd med oppgaven?

Er det noen utfordringer med denne oppgaven fra et parprøveperspektiv?

Er det tatt hensyn til at elevene jobber i par?

Begrensinger

Har parprøven noen begrensninger?

Finnes det ulemper ved at elevene jobber i par?

Etterarbeid

Hvordan gikk denne prøven synes du?

Hva er forskjellen på å rette en parprøve sammenliknet med en individuell prøve?

Tar du hensyn til at de har jobbet i par når du retter?

Hva ser du etter?

Får du tilgang på elevenes prosesser i besvarelsene? Hvordan/hvorfor ikke?

Får du informasjon om hver enkelt elevs kompetanse?

Eventuelt hvordan? Er det viktig?

Hva er utfordrende i retteprosessen?

Går info om elevenes kompetanse tapt? I så fall, gjør det noe?

Hvordan kan parprøven brukes til å hjelpe eleven videre?

Avslutningsvis

Har du lært noe gjennom din parprøvepraksis som du ønsker å dele?