

# Elevenes fremvisning av naturvitenskaplige praksiser og tenkemåter i møtet med PISA- oppgaver

Alexander Michelsen

Veileder: Professor Stein Dankert Kolstø



Masteroppgave i fysikkdidaktikk

Institutt for fysikk og teknologi

UNIVERSITETET I BERGEN

3. juni 2019

© Alexander Michelsen

2019

Elevers fremvisning av naturvitenskaplige praksiser og tenkemåter i møtet med PISA-  
oppgaver

Alexander Michelsen

<https://bora.uib.no/>

## Abstrakt

Hensikten med denne masteroppgaven er å skaffe innsikt i hvordan elever viser frem kunnskap om naturvitenskaplige praksiser og tenkemåter, når de jobber med to frigitte oppgavesett i PISA-undersøkelsen fra 2015, og hvorvidt disse oppgavene klarer å måle elevers fremvisning av kunnskap innenfor naturvitenskaplige praksiser og tenkemåter. Oppgaven forsøker å formulere et svar på dette gjennom å observere praksisen til 13 elever i 15-års alderen, som var fordelt på fire grupper, når de jobbet med de frigitte oppgavesettene «Løping i varmt vær» og «Undersøkelse av dalsider». Til dette ble det sett på de underliggende kunnskapsdimensjonene for vitenskapelig resonnering, som Kind (2013) og Osborne (2013) beskriver, samt en modell om vitenskapelig resonnering Osborne (2013) og tidligere beskrivelser av elevers praksiser (Schauble et al., 1991) og tenkemåter (Kind, 2013; Larkin, 1981) Ut ifra analyse av lydopptak, ble det etablert en modell, som beskriver tre hovedstrategier elevene brukte når de jobber med oppgaver, som går ut på simulering av datasett. I tillegg ble det identifisert åtte tenkemåter, som beskriver elevenes epistemiske tenkemåter i møtet med oppgaver, som testet elevenes evner til å tolke data i lys av teori, og deres epistemiske forståelse rundt målinger av én variabel. Resultatet av analysen tyder på at de to frigitte oppgavene er i stand til å måle resultatet av elevers praksiser og tenkemåter, men ikke helt klarer å måle nyansene i disse.



## **Abstract**

The aim of this master thesis is to provide an insight into how students show knowledge of science practices and way of thinking when working with two published exercises of the PISA-study from 2015 and whether these exercises measure the students' knowledge in these areas. This thesis is answering these questions by observing 13 students in the age of 15, who were divided into groups of four, while they worked with the published exercises "Running in Hot Weather" and "Slope-Face Investigation". For this reason, I looked at the underlying knowledge dimensions for scientific reasoning, as Kind (2013) and Osborne (2013) describe. In addition, previous descriptions of student's practices (Schauble et al., 1991) and ways of thinking (Kind, 2013; Larkin, 1981) were examined. Based on analysis of sound recordings, a model that describes three main strategies students use when working with the exercises was established. This involved simulation of data sets. Furthermore, eight ways of thinking were identified, which describe students' epistemological ways of thinking when confronted with the exercises. These epistemological ways of thinking test students' abilities to interpret data in light of theory and their epistemological knowledge about measurements of one variable. The result of this analysis suggests that while the two published exercises are able to measure the results of students' practices and ways of thinking, they are not able to measure nuances in these results.



## Sammendrag

Temaet for denne masteroppgaven er elevers fremvisning av naturvitenskaplig praksis og tenkemåter, og hvordan frigitte oppgaver fra PISA-studien i 2015 måler dette. Oppgavens fokus ligger på de kunnskapsdimensjoner som inngår i naturvitenskaplig praksis og tenkemåte, nemlig innholds- prosedyre- og epistemisk kunnskap, som også betegnes som Nature of Science (NOS) (Lederman, 2013; Osborne, 2013). Deretter skal det se på rammeverk som beskriver praksisene og tenkemåtene til forskere innen naturvitenskap, og hva som kjennetegner elevenes praksiser og tenkemåter i møtet med praktiske naturvitenskaplige oppgaver.

Oppgaven undersøker elevenes fremvisning av naturvitenskaplige praksis og tenkemåter gjennom samtaler, når de jobber med to frigitte PISA-oppgaver fra undersøkelsen i 2015. Dette er en kvalitativ intervju studie, der totalt 14 elever, som var 15 år gamle, ble fordelt på 4 grupper, som jobbet med oppgavene. Datamaterialet ble innsamlet ved hjelp av lydopptak.

Analyse av datamaterialet ga opphav til tre løsningsstrategier, når elevene jobbet med oppgaver, som krevet simulering og tolkning av data. Disse ble satt i en modell, som beskriver elevenes løsningsprosesser i møtet med PISA-oppgavene. I tillegg blir det presentert åtte kategorier, som beskriver elevenes tenkemåter innenfor NOS.

Disse løsningsstrategiene ble sammenliknet med modellene om naturvitenskaplig praksis og tenkemåter, og PISAs eksempler på godkjente svar til å avgjøre om hva PISA måler av elevsvarene. Her viste det seg at PISA måler elevenes epistemisk kunnskap, når det gjelder vurdering av hypoteser med hensyn på datamaterialet, men dataene peker på at oppgavene ikke fanger inn den epistemiske siden ved planleggingen av kvantitativt datainnsamling. I tillegg tyder kategoriseringen av elevenes tenkemåter på at PISAs oppgaver om NOS plaget av de samme problemene, som nesten hver implisitt testing av NOS har som begrensning: Den viser at elevene anvender epistemisk kunnskap, men har sine svakheter når det kommer til måling av de epistemiske oppfatningene som ligger bak denne anvendelsen (Abd-El-Khalick, 2012).





## **Takk**

Jeg vil først og fremst takke min veileder, Professor Stein Dankert Kolstø for konstruktive tilbakemeldinger, verdifulle tips, og rike samtaler gjennom hele oppgaven. Dine tilbakemeldinger og anbefalinger har gitt meg svært stor innsikt, som jeg kommer til å huske lenge når jeg skal begynne som lærer.

Lektorutdanningen hadde ikke den samme uten mine medstudenter. Jeg følger meg så lykkelig og heldig til å ha blitt kjent med alle dere. Et særskilt takk til alle mine medstudenter på rom 367. Det var alltid en meget hyggelig stemning gjennom hele oppgaveskrivingen, takket være dere. Jeg takker også for dannete og kritiske innspill underveis.

Jeg takker min familie for støtten gjennom hele utdanningen min. Særlig vil jeg også takke samboeren min, Sasi. Din omtanke og støtte har alltid vært en stor motivasjon for meg, særlig gjennom det siste semesteret.

Sist, men ikke minst, vil jeg takke alle elever som valgte til å delta, og lærere, som var en stor hjelp til å organisere intervjuene. Dette prosjektet hadde ikke vært mulig uten dere.



# Innholdsfortegnelse

Abstrakt.....	II
Sammendrag.....	VI
Takk .....	VIII
Figurer .....	XIV
1 Innledning.....	1
1.1 Motivasjon .....	1
1.2 Utforming av forskningsspørsmål.....	2
1.3 Oppgavens bidrag .....	3
2 Teoretisk bakgrunn.....	4
2.1 PISA-undersøkelsen.....	4
2.1.1 PISA-undersøkelsens definisjon av naturfag.....	5
2.2 Naturvitenskaplige praksiser og tenkemåter .....	7
2.2.1 Kunnskap som ligger til grunn.....	7
2.2.2 Hvordan tenker og arbeider man i naturvitenskap? .....	9
2.2.3 Testing innenfor NPT.....	12
2.3 Elevers praksiser og oppfatninger innenfor naturvitenskap.....	18
2.4 Et rammeverk for å beskrive elevenes tenke og arbeidsmåter .....	20
3 Metode.....	23
3.1 Kvalitativ Metode.....	23
3.2 Validitet og Reliabilitet .....	24
3.3 Utvalg av PISA-oppgaver.....	26
3.3.1 Kort beskrivelse av PISA-oppgavene.....	26
3.4 Utvalg av deltakere.....	28
3.5 Fokusgruppe .....	29
3.6 Kontekstbeskrivelse.....	31
3.7 Transkribering og koding.....	32
3.8 Dataanalyse .....	33

3.9	Etiske betraktninger .....	35
3.10	Begrensninger .....	36
4	Funn .....	38
4.1	Kategorier innenfor elevers epistemiske løsningsstrategi .....	38
4.1.1	Simulerer prøvende .....	38
4.1.2	Simulerer målrettet .....	40
4.1.3	Tolkning av datasett .....	43
4.1.4	Tester egen tolkning .....	45
4.1.5	Konsistensvurdering .....	45
4.1.6	Modell av elevenes epistemiske løsningsstrategi .....	47
4.2	Underliggende sammenhenger .....	50
4.2.1	Elever bruker teoretisk kunnskap .....	50
4.2.2	Korrelasjon med tidslinje som årsak .....	52
4.2.3	Årsaksforklaring krever teoretisk innsikt .....	53
4.3	Epistemiske oppfatninger .....	54
4.3.1	Måleinstrumenter viser enten rett eller feil .....	54
4.3.2	Usikkerhet som sikkerhet innen et intervall .....	56
4.3.3	Usikkerhet som variasjon i observert fenomen .....	58
4.3.4	Flere uavhengige målinger minker usikkerhet .....	60
4.3.5	De største numeriske forskjellene gir en årsak .....	61
5	Diskusjon .....	63
5.1	Elevenes løsningsstrategier .....	63
5.1.1	Løsningsvei L1 – Praktisk tilnærming .....	63
5.1.2	Løsningsvei L2 – Praktisk-strategisk tilnærming .....	64
5.1.3	Løsningsvei L3 – Vitenskapelig tilnærming .....	65
5.1.4	Oppsummering og sammenlikning med eksempelsvar til PISA .....	66
5.2	Elevenes tenkemåter .....	68

5.2.1	Måling av én variabel og inngående usikkerhet.....	68
5.2.2	Teoriens rolle i en forklaring.....	71
5.2.3	Oppsummering og sammenlikning med eksempelsvar til PISA.....	73
5.3	Forslag til videre forskning.....	77
Bibliografi.....		79
Vedlegg.....		83



## Figurer

Figur 1: Oversikt over kompetansene innenfor "Scientific literacy", hentet fra Kjærnsli og Jensen (2016b) .....	6
Figur 2: Gieres (1991) modell om vitenskapelig modellering. ....	10
Figur 3: Figur 2, utvidet med prosesser fra psykologiske modeller, hentet fra Osborne (2013) .....	11
Figur 4: Utdrag av spørsmål i VNOS-testen, hentet fra Abd-El-Khalick et al. (2001). ....	17
Figur 5: Rammeverk for å måle elevenes vitenskapelige resonnering, hentet fra Kind (2013). ....	21
Figur 6: Illustrering av elevenes kunnskapsnivå ('knowledge' på figuren) og kognitive ferdigheter ('cognitive behaviour' på figuren), hentet fra Kind (2013). ....	21
Figur 7: Eksempel for en oppgave fra settet "Løping i varmt vær", hentet fra Kjærnsli og Jensen (2016c).....	27
Figur 8: Oppgave 2 fra settet "Undersøkelse av dalsider", hentet fra Kjærnsli og Jensen (2016c). ....	28
Figur 9: Oversikt over gruppenes løsningsstrategi for oppgavesettet "Løping i varmt vær". .....	48





# 1 Innledning

En grunnleggende måte for et land til å evaluere sitt utdanningssystem er sammenlikningen med andre land sine utdanningssystemer. I nyere tid er derfor har OECDs PISA-undersøkelsen som tester elevenes kompetanser innenfor egendefinerte kompetansemål av lesing, matematikk og naturfag, blitt ett av de mest verdifulle verktøyene til å sammenlikne forskjellige lands utdanningssystemer (Grek, 2009). I tillegg har det blitt en trend blant OECD-landene, også Norge, til å legge større vekt på naturvitenskaplige praksiser og tenkemåter (Osborne, 2013, 2014). Denne oppgaven vil fokusere på norske elevers fremvisning av naturvitenskaplige praksiser og tenkemåter, når de jobber med oppgaver i PISA-undersøkelsen fra 2015, som tester innenfor dette.

## 1.1 Motivasjon

En av de større reformene i norsk grunnskoleutdanning i nyere tid, Kunnskapsløftet, ble preget av «PISA-sjokket», som har hatt stor innflytelse på utformingen av debatten om norsk skole (Sjøberg, 2018). Man har innført nasjonale prøver og deltar i flere internasjonale prøver, det faglige har kommet mer i fokus, og alle fag har fått krav om lærere med utdanning innenfor egne undervisningsfag (Sjøberg, 2014). Liknende har skjedd i Sverige, og Serder (2015) har, i hennes doktorgradsavhandling, stilt seg spørsmålet om hvordan man skal forstå svenske elevers resultater innenfor naturfag i PISA-undersøkelsen.

Med Kunnskapsløftet kom Forskerspiren som eget hovedområde i naturfag, som tydeliggjorde krav om undervisning innenfor naturvitenskapelig praksis og tenkemåter (Knain & Kolstø, 2011). Osborne (2013) nevner utviklingen av en modell om, og måling av vitenskapelig resonnering som en av de største utfordringene innenfor opplæring i naturfag på 2000-tallet. Elevene skal ikke lenger kun repetere praksiser som forskere bruker, men burde istedenfor få en opplæring av naturvitenskapens egenart. Hverdagen i naturfagundervisning er derimot preget av at man kommer, i lukkede forsøk, frem til allerede etablert kunnskap, og heller etterlikner forskningsprosessen enn å eksplisitt snakke om den (Osborne, 2014). Dette gjelder også for norsk skole (Ødegaard & Arnesen, 2010). Utkast til den nye læreplanen i Fagfornyelsen (Utdanningsforbundet, 2018) tar også hensyn til dette, i kjerneelementet *naturvitenskaplige praksiser og tenkemåter (NPT)*. Etter å ha fullført obligatorisk skolegang skal elevene blant annet kunne kritisk vurdere

andres forskning, og selv kunne lage og teste hypotese. Dette krever kunnskap om naturvitenskapens innhold, prosedyre og epistemologi (Allchin, 2011; Kind, 2013; Osborne, 2013). PISA-undersøkelsen fra 2015 har etablert, ifølge Osborne (2014), et lovende rammeverk for beskrivelsen av praksiser innfor NPT, som tar hensyn til vesentlige praksiser i naturvitenskapen.

Særlig innenfor epistemisk kunnskap har det foregått en debatt om dette bør testes eksplisitt, for å avgjøre om elevene har reflekterte oppfatninger, eller implisitt, for å avgjøre om elevene kan drøfte andre forskningsdesign kritisk og utvikle egne prosedyrer (Abd-El-Khalick, 2012). Innenfor den implisitte tilnærmingen har det etablert seg en retning, som mener at epistemisk kunnskap ikke kan testes isolert fra vitenskapelig resonnering generelt (Abd-El-Khalick, 2012; Allchin, 2011; Osborne, 2013). PISA-undersøkelsen fra 2015 kan også tolkes som at den er tilhenger av en implisitt testemåte i at den nevner at det ikke er nok til å bare gjengi kunnskap, men man også kunne anvende den for å løse oppgavene. I tillegg er en av de frigitte oppgavene om epistemisk kunnskap sterk tilknyttet en bestemt kontekst, uten noen henvisning til at elevene må eksplisitt nevne epistemisk kunnskap (Kjærnsli & Jensen, 2016c).

Fokuset på NPT, og epistemisk kunnskap innenfor dette, og diskusjonen om hvordan disse bør måles generelt, og hvordan vurderingen av elevers prestasjoner i PISA-undersøkelsen fungerer, peker på et behov for å vite mer om hvordan elever tenker i møte med oppgaver som skal teste NPT. Som et innspill til denne debatten vil denne oppgaven undersøke elevers praksis og tenkemåter i møte med to frigitte oppgavesett i PISA 2015.

## **1.2 Utforming av forskningsspørsmål**

Oppgavens formål er å få innsyn i hvordan elever, i 15-års alderen, forholder seg til PISA-oppgaver om NPT. Etter å ha lest deler av avhandlingen til Serder (2015), der hun beskriver svenske elevers møte med PISA-undersøkelsen fra 2006, ble det formulert følgende overordnede spørsmål for dette masterprosjektet: *Hvordan viser elever frem sin kunnskap om naturvitenskapelige praksiser og tenkemåter, når de jobber med PISA-oppgaver om naturvitenskapelige praksiser og tenkemåter?*

Ut ifra det overordnede spørsmålet ble det formulert to mer spesifikke spørsmål, som vil ligge til grunn for konkrete analyser og diskusjoner.

- 1) Hvilke praksiser og tenkemåter viser elevene når de jobber med frigitte PISA-oppgaver om naturvitenskapelige praksiser og tenkemåter?
- 2) I hvilken grad måler de frigitte oppgavene i PISA-undersøkelsen fra 2015 naturvitenskapelige praksiser og tenkemåter hos elevene?

Det første spørsmålet vil bli besvart gjennom en empirisk undersøkelse. Denne vil først greie ut hvilke løsningsstrategier elevene bruker, når PISA-oppgavene tester planlegging og gjennomføring av forskningsdesign. Deretter skal den forsøke å kartlegge tenkemåter som ligger til grunn for strategiene, og tenkemåtene som elevene viser når de jobber med epistemiske spørsmål. Det andre spørsmålet vil bli besvart gjennom en diskusjon av de empiriske resultatene opp mot eksempelsvar i PISA-undersøkelsen (Kjærnsli & Jensen, 2016c).

### **1.3 Oppgavens bidrag**

Den empiriske undersøkelsen i denne oppgaven resulterte i en modell som beskriver og visualiserer elevers løsningsstrategi i møtet med oppgaver som krever vurdering av data innsamlet vha. datasimulering. Denne modellen kan brukes som utgangspunkt til å vurdere utforming av digital testing av NPT, som byr på mange muligheter (Osborne, 2013). Et eksempel på dette kan være å utforme oppgaver, som krever at elevene tolker datamateriale gjennom at oppgavene lages slik at de ikke legger til rette for at elever kun vurderer konsistensen til svarforslagene i lys av data. I tillegg kan de ulike kategoriene i denne modellen støtte kategoriseringen av elevenes naturvitenskapelige praksiser og tenkemåter i forbindelse med muntlige prøver.

Oppgavens spørsmål bidrar også til feltet om hvordan man skal forstå elevenes prestasjoner innenfor PISA-undersøkelsen, særlig når det kommer til fremvisning av epistemologisk kunnskap. Den kan bidra til avklaring om hvordan PISA-undersøkelsen tester den epistemiske forståelsen til elevene.

## 2 Teoretisk bakgrunn

### 2.1 PISA-undersøkelsen

For å få innsikt hva dette prosjektet handler om, vil jeg gi en beskrivelse av hva PISA-undersøkelsen er, og hvordan naturfag er definert i PISA. Til dette tar jeg utgangspunkt i den norske hovedrapporten av Kjærnsli og Jensen (2016a). Grunnen til at det finnes en rapport for hvert deltakerland er at deltakerlandene finansierer OECD, organisasjonen bak PISA-undersøkelsen, og derfor forventer en egen rapport (Meyer & Benavot, 2013). Videre blir det kort sett på kritikk til PISA-undersøkelsen.

Kjærnsli og Jensen (2016a) bruker følgende beskrivelse av PISA-undersøkelsen:

«Programme for International Student Assessment (PISA) er en internasjonal undersøkelse av 15-åringers kunnskaper og kompetanser i noen sentrale fagområder – lesing, matematikk og naturfag. For å løse oppgavene i PISA-prøven er det ikke nok å gjengi fakta, elevene må også bruke sin kunnskap og kompetanse i en kontekst. Fagområdene er valgt ut, fordi de er sentrale i skolen i alle land, og fordi de lar seg måle på tvers av land. [...] Den overordnede ideen bak PISA er å vurdere hvor godt skolesystemene i ulike land forbereder elevene til videre skolegang, studier, arbeidsliv og aktiv og reflektert deltakelse i samfunnet.» (s. 11)

Deltakere er undersøkelsen 15 år gamle elever. Elever i denne aldersgruppen går mot slutten av den obligatoriske skoleutdanningen i de fleste deltakerland. Videre er det formulert et generelt rammeverk for testingen istedenfor at man tar utgangspunkt i læreplanen til hver deltaker, slik som TIMSS gjør. Dette kan gi innsikt i at over potensielle endringer i læreplaner, og trend i utdanningssektoren til disse landene (Sjøberg, 2018).

Ved å teste elevene i samme alder på et internasjonalt nivå gjennom de samme oppgavene, er PISA er verdifullt verktøy for den komplekse sammenlikningen av deltakerlandenes utdanningssystemer. Å kunne evaluere sitt nasjonale utdanningssystem krever sammenlikning med andre utdanningssystemer. Ved å sammenlikne utdanningssystemene uavhengig fra landets kulturelle kontekst, kan man gi et bedre bilde av den faglige kvaliteten til et utdanningssystem, som igjen kan gi et solid grunnlag til å forbedre utdanningen til hvert land. Dette har positive konsekvenser på deltakerlandenes økonomi, siden et godt utdanningssystem anses som en fordel i møtet med utfordringer som en stadig mer globalisert verden byr på seg (Grek, 2009).

Resultatene blir presentert og sammenliknet med andre land (Kjærnsli & Jensen, 2016a), gjerne grafisk og i en rangliste (Sjøberg, 2018), der det også blir sett på trender, sammenliknet med PISA-undersøkelser fra tidligere år. Dette er ett av de største kritikkpunktene mot PISA-undersøkelsen. Når presentasjonene til de forskjellige landene settes opp på en så forenklet måte, kan dette betraktes som en del av en trend, der økonomiske konsepter er på vei inn i utdanningssektoren. En skarp formulert kritikk av denne trenden, som PISA-undersøkelsen muligens også bedriver, er at denne blir beskrevet som en global utdanningskonkurrans (eng. «global education race»). Tegn av denne konkurransen kan observeres gjennom PISA-sjokkene land som f.eks. de skandinaviske landene og Tyskland opplevde. Som reaksjon ble læreplanen bearbeidet til å gjøre det bedre de neste gangene. Dermed blir skoler, som skal også reflektere kulturen i til ethvert land, utsatt for denne konkurransen (Sjøberg, 2018).

### **2.1.1 PISA-undersøkelsens definisjon av naturfag**

PISA tar utgangspunkt i begrepet «scientific literacy» for å beskrive elevenes kompetanser istedenfor naturfag. Likevel brukes ordet «naturfag» som et synonym for «scientific literacy» (Kjærnsli & Jensen, 2016b). Det understrekes at elevene må ha kunnskap om naturvitenskap som prosess og produkt. For å definere begrepet «scientific literacy», blir dette delt i fire aspekter: *Kontekster, kunnskaper, kompetanser og holdninger*.

Kontekster angir sammenhengen til oppgavene, og er skal være relevante for 15-år gamle elever. Derfor er de ikke kun begrenset til kontekster fra skolefaget naturfaget. Holdninger beskriver elevens interesse og vilje til å engasjere seg i naturvitenskaplige og teknologiske tema. Dette innebærer generell interesse, som også kan være kritisk ovenfor bestemte områder i naturvitenskapen (Kjærnsli & Jensen, 2016b). Selv om kontekst- og holdningsaspektet har innflytelse på fremvisning av kompetanse og kunnskap, betrakter jeg dette ikke som så relevant som kunnskaps- og kompetanseaspektet. Grunnen til dette er at disse er underliggende for tenkemåtene, men beskriver disse ikke.

Kunnskapsaspektet har gått gjennom en del forandring gjennom årene i PISA-undersøkelsen. I undersøkelsen fra 2015 er denne delt i tre former for naturvitenskapelig kunnskap: innhold, metode og epistemologi. Innholdskunnskap dreier seg om fakta, begreper og teorier innen naturvitenskapen. Metodekunnskap handler om hvilke naturvitenskapelige metoder som forskere benytter til å genere kunnskap.

Epistemologisk kunnskap handler om hvorfor de vitenskapelige metoder generer kunnskap, og hvordan modeller har innflytelse på naturvitenskapelige metoder (Kjærnsli & Jensen, 2016b). Disse tydeligere beskrevet gjennom kunnskap eleven forventes å ha for å gjelde kompetent innenfor hver kunnskap.

Kompetanseaspektet er selve definisjonen på «scientific literacy», ifølge Kjærnsli og Jensen (2016b), som alle andre aspekter dreier seg om. Denne er delt i tre kompetanser, som elevene skal kunne beherske: (1) Forklare fenomener på en naturvitenskapelig måte, (2) vurdere og planlegge naturvitenskapelige undersøkelser, og (3) tolke data og evidens på en naturvitenskapelig måte. Den mer detaljert beskrivelse er gitt i Figur 1. I tillegg til denne beskrivelsen, er det også en del praksiser som inngår i hver kompetanse. Siden en beskrivelse av disse praksisene ville sprengte rammen her, holder jeg meg til denne generelle beskrivelsen.

«Scientific literacy» er evnen og viljen til å engasjere seg i og delta i diskusjoner om naturfagsrelaterte temaer. Dette innebærer følgende tre kompetanser:

**1. Forklare fenomener på en naturvitenskapelig måte**

- Gjenkjenne, gjengi og vurdere forklaringer på en rekke naturlige og teknologiske fenomener.

**2. Vurdere og planlegge naturvitenskapelige undersøkelser**

- Beskrive og vurdere naturvitenskapelige undersøkelser og foreslå hvordan spørsmål kan besvares på en naturvitenskapelig måte.

**3. Tolke data og evidens på en naturvitenskapelig måte**

- Analysere og vurdere data, påstander og argumenter i en rekke ulike framstillinger og trekke riktige naturvitenskapelige konklusjoner.

*Figur 1: Oversikt over kompetansene innenfor "Scientific literacy", hentet fra Kjærnsli og Jensen (2016b)*

Å si at aspektene kontekster, kunnskaper og holdninger dreier seg om kompetanseaspektet betyr at for å fremvise kompetanse må elevene kunne noe om konteksten til oppgaven, og hvordan elevene benytter kompetansene avhenger av kunnskap og holdninger (Kjærnsli & Jensen, 2016b).

## 2.2 Naturvitenskaplige praksiser og tenkemåter

### 2.2.1 Kunnskap som ligger til grunn

På lik linje med PISA-undersøkelsen, beskriver litteraturen om vitenskapelig resonnering, og testingen av denne, tre kunnskaper som er grunnlaget for vitenskapelig resonnering (Kind, 2013; Osborne, 2013):

- Innholdskunnskap
- Prosedyrekunnskap
- Epistemisk kunnskap

Jeg kommer til å bruke begrepet *naturvitenskaplige praksiser og tenkemåter (NPT)* synonymt med vitenskapelig resonnering. Innholdskunnskap, eller konseptuell kunnskap, omfatter resultatet av naturvitenskapelige prosesser, teoriene, og er preget av tentative formuleringer som et resultat av logiske begrensninger (Allchin, 2011). Disse logiske begrensningene går ut på at forskere ser etter gjengående mønstre i enkelte fenomener og formulerer en generalisering ut ifra disse. Men det finnes ingen basis for å anta at fremtiden, som man forsøke å forutsi gjennom generaliseringen, vil oppføre seg som fortiden, altså mønstrene som man har generalisert. Denne begrensningen er kjent som det klassiske induksjonsproblemet (Godfrey-Smith, 2003, s. 39-40). For å takle dette problemet, bruker naturvitenskapen gjerne den hypotetisk-deduktive metoden, som har mål om å falsifisere generaliseringen, eller teorien, enn å verifisere den. Man bruker teorien til å lage testbare forutsigelser, hypoteser. Dersom hypotesene bekreftes av testingen, støttes teorien. Dersom hypotesene avkreftes, blir også teorien avkreftet, altså falsifisert (Godfrey-Smith, 2003, s. 69-70). Derfor inkluderer innholdskunnskap også kunnskap om hvordan disse teoriene bygges ut ifra observasjoner fra den reelle verden, og testingen av disse (Giere, 1991; Kind, 2013).

Den andre kunnskapsdimensjonen, prosedyrekunnskap, handler om hvordan disse teoriene testes. Dette omhandler forskjellige eksperimentdesign, og siden den handler om teori- og hypotesetesting, er tett tilknyttet kunnskapsutvikling (Kind, 2013). Ut ifra eksperimentene får man datamateriale, som krever argumentering for deres relevans og validitet. Dermed er begrepet usikkerhet til datamaterialet og hvordan man forholder seg til dette også en del av metodekunnskap, og blir derfor også kalt «konsepter om bevismateriale» (Osborne, 2013). Med dette i bakhodet, kan denne kunnskapen deles inn i *kunnskap om analyse og interpretasjon av data, og kunnskap om evaluering av*

*forskningsdesign*. Osborne (2013) utdyper dette ved å nevne følgende kunnskap som inngår i hver av disse inndelingene:

<b>Prosedyre kunnskap</b>	
<b><i>Kunnskap om interpretasjon og analyse av data</i></b>	<b><i>Kunnskap om evaluering av forskningsdesign</i></b>
Konsepter om målinger, f.eks. kvalitative, kvantitative, skalaer, variabler	Variabelkonseptet og skillet mellom avhengige og uavhengige variabler
Teste og minimere usikkerhet	Strategier for kontroll av variabler og dets rolle i ett eksperiment
Mekanismer som muliggjør reproduserbarhet av data	Egenskapene til et egnet design for å svare på vitenskapelige spørsmål
Abstrahere og representere data, f.eks. tabeller eller grafer	

Tabell 1: Kunnskaper som inngår i metodisk kunnskap (Osborne, 2013).

Det kan legges til at disse kunnskapene går igjen, til og med ordrett, i PISA-undersøkelsens utforming av metodekunnskap (Kjærnsli & Jensen, 2016b). Også kompetansene 2. og 3. i Figur 1 besitter en sterk likhet med de to inndelingene av metodekunnskap, gitt i Figur 1. Årsaken til dette kan være at Jonathan Osborne, som jeg referer til i utformingen av kunnskapsformene i NPT, var hovedansvarlig for å definere «scientific literacy» i PISA-undersøkelsen fra 2015 (Sjøberg, 2018). Selv referer Osborne (2013) til OECD, organisasjonen bak PISA-undersøkelsen, når han lister opp kunnskapene innenfor metode- og epistemisk kunnskap. Likevel bruker han kunnskapene som en oppsummering av andre artikler innenfor dette feltet. Derfor anså jeg artikkelen likevel som en god oppsummering av kunnskapene som inngår i NPT.

Mens metodekunnskap svarer på hvordan man gjennomfører eksperimenter og undersøkelser, må man fortsatt spør seg hvordan og hvorfor disse eksperimentene generer ny kunnskap. For å svare på dette, kreves det epistemisk kunnskap (Kind, 2013). Den tredje, og siste kunnskap inngående i NPT, er denne «hvorfor»-dimensjonen til naturvitenskapen. Her inngår det også iboende verdier til naturvitenskapen og hvordan de utvikles, og NOS brukes ofte som et synonymt begrep (Lederman, 2013). Kunnskapen innebærer å kunne vurdere vitenskapelige påstander og krever å kunne skille mellom epistemiske konstrukt i vitenskapelige teorier og hvordan disse testes, f.eks. konstruktene «hypotese», «beskrivelse av observasjon», «fakta», osv. (Kind, 2013; Osborne, 2013) For å kunne gjennomføre en slik vurdering, har Osborne (2013) utpekt følgende typer kunnskap som inngår i epistemisk kunnskap:

- Egenskaper til vitenskapelige observasjoner, fakta, hypoteser, modeller og teorier

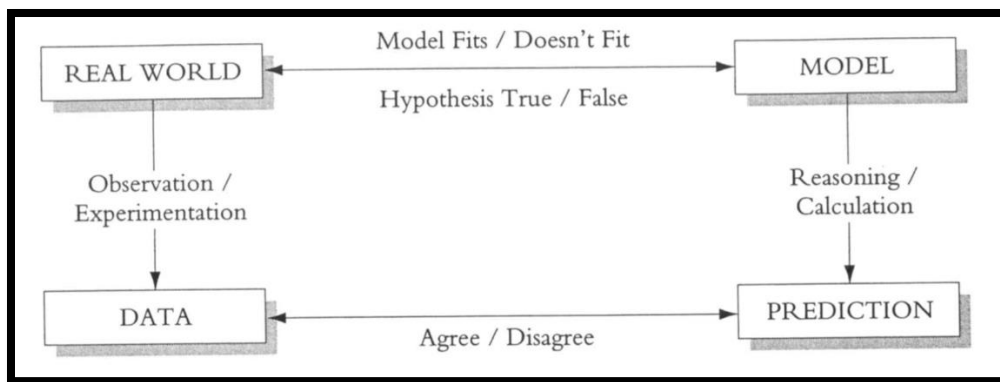


- Hvordan vitenskapelige påstander er underbygget av data og vitenskapelig resonnering
- Funksjonen til hypoteser i å muliggjør testbare forutsigelser
- Hvordan målefeil påvirker sikkerheten til vitenskapelig kunnskap
- Bruk og betydning av fysiske, systematiske og abstrakte modeller, samt deres begrensinger
- Betydning av samarbeid, kritikk og fagfellevurdering i å etablere troverdighet i vitenskapelige påstander

Igjen, denne konkretiseringen av kunnskap går også igjen i PISA-undersøkelsens definisjon av epistemisk kunnskap (Kjærnsli & Jensen, 2016b), men inkluderes, fordi forfatteren beskriver dette som en oppsummering av litteratur.

### **2.2.2 Hvordan tenker og arbeider man i naturvitenskap?**

Etter å sett på de grunnleggende kunnskapstypene som inngår i naturvitenskapelig resonnering, kreves det en utvidelse av rammeverket. Dette begrunnes i at man oftest anvender en kombinasjon av disse kunnskapsdimensjonene. Samtidig er inndelingen også begrenset i at denne ikke henviser til den vitenskapelig praksis som eleven skal benytte seg av. For å gjøre kategoriseringen av oppgaver oversiktligere, og belyse hvilke aspekter fra NPT som inngår i oppgavene, vil jeg derfor tilskrive dem den vitenskapelige praksisen de operer i. På lik linje med Kind (2013) vil jeg utgangspunkt i selve praksisen naturvitere benytter til å komme frem til nye modeller som beskriver fenomener i den fysiske verden, og hvordan disse undersøkes. Utgangspunktet for kategoriseringen av praksisen er modellen til Giere (1991) om hvordan naturvitere utvikler modeller fra den fysiske verden og hvordan disse modellene undersøkes. Denne tar utgangspunkt i fire hjørnesteiner og forholdene mellom dem. Modellen er visualisert i Figur 2.

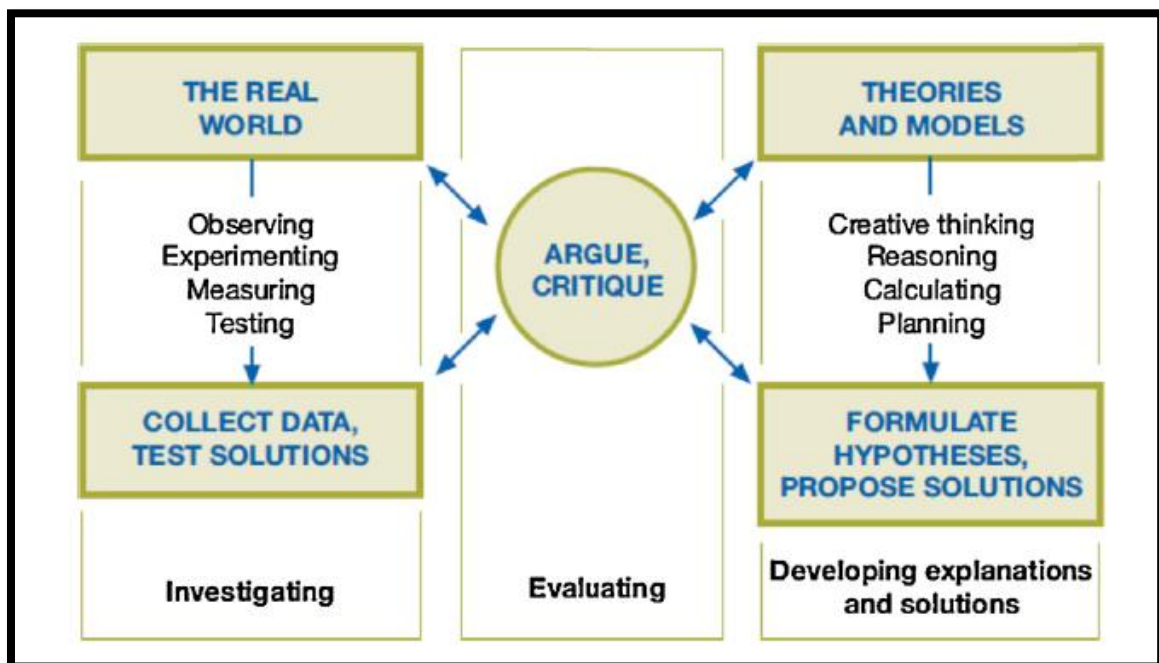


Figur 2: Gieres (1991) modell om vitenskapelig modellering.

Øverst har man den fysiske verden og modellen. Modellen utvikles gjennom kreative gjetninger av den fysiske verden, men det eksakte forholdet mellom dem kan ikke bestemmes direkte. Jeg bruker ordet *gjetning* for å beskrive prosessen for å komme frem til en modell fra den virkelige verden. Det betyr ikke at forskere fullstendig gjetter ut av det blå når de ser på naturen, men ser etter gjengående mønstre. Siden også forskerens kreativitet er sentralt i denne prosessen, skal dette også uttrykkes, og derfor bruker jeg ordet *gjetning* til å beskrive dette. Derfor må man resonnerer seg frem til forutsigelser fra modellen (høyre side), og observere/utføre eksperimenter fra den fysiske verden (venstre side), som kan sammenliknes (nede). Disse praksisene tilsvarer praksisene som brukes i vitenskapelig resonnering, der modell- og hypotesebygging oftest fattes sammen (Kind, 2013; Nowak et al., 2013).

Begrensningen med denne filosofiske tilnærmingen er at den viser et normativt bilde av hvordan naturvitenskapen fungerer uten å henvise til en faktisk beskrivelse av arbeids- og tenkemåten til forskere. Likevel er det paralleller mellom psykologiske, aktivitetsorienterte modeller og filosofiske, normative modeller. Osborne (2013) henviser til en utvidelse av denne modellen, der det ble tilført prosesser, som henviser til fasene i utforskningen, i tillegg til de filosofiske stadiene, og konkluderte med at det finnes tre sentrale praksiser, som er vist i Figur 3. Fasen i forskningen, der man med utgangspunkt i den fysiske verden, bygger modeller derifra og formulerer hypoteser vha. abduktiv og deduktiv resonnering, er på høyre side, og betegnes som «Utvikle forklaringer og løsninger». For å kunne utlede gode modeller og hypoteser, kreves det god innholdskunnskap. Når disse ideene skal testes, må man komme opp med et forskningsdesign. Dette designet gjennomføres og gir data, som skal analyseres. Osborne (2013) setter denne prosessen som venstre side til modellen, «Utforskning», og krever

god innsikt i metodekunnskap, som igjen skal gi svar på det epistemologiske spørsmål om hvordan man kan vite/være sikker på kunnskap. Den tredje prosessen i midten, «Evaluering», har sitt opphav i at alle observasjoner og data er teoriladde, dvs. at data ikke kan betraktes uten et teoretisk rammeverk. Dermed består det et krav for å argumentere for validiteten til teorier, hypoteser og forskningsdesign ut ifra bevismateriale, og står sentralt i hvordan naturvitenskapelig resonnering er bygget opp. Man må f.eks. argumentere for at hypoteser gir en god representasjon av virkeligheten og data som støtter denne modellen, som hypotesen er dedusert fra. Det bemerkes at alle prosessene som er beskrevet ovenfor er generaliseringer for et mangfold av praksiser som naturvitenskapen benytter seg av i deres resonnering.



Figur 3: Figur 2, utvidet med prosesser fra psykologiske modeller, hentet fra Osborne (2013)

Som tidligere nevnt, finnes det et mangfold av praksiser, som inngår i naturvitenskapelige praksiser og tenkemåter. Disse varierer mellom de vitenskapelige disipliner (Dagher & Erduran, 2016; Nowak et al., 2013). Likevel finner man en liste hos Osborne (2014), der det oppføres åtte praksiser som går igjen i all vitenskapelig forskning. En oversikt over praksisene er gitt i Tabell 2. Rammeverket til PISA-undersøkelsen 2015 bruker selv begreper som går igjen i denne listen (Kjærnsli & Jensen, 2016b; Osborne, 2014) til å definere hva som inngår i naturfaglig kompetanse. Likevel består det et krav om å tilpasse praksisene til de forskjellige naturfaglige disiplinene, siden det finnes flere bekymringer ved å fremstille et syn om enighet om metoder og deklarativer påstander (Allchin, 2011;

Dagher & Erduran, 2016). Denne tilpasningen skal også gi et mer konkret innblikk i metodene og hvilke metoder mulige oppgaver bruke, og derfor kan «Family Resemblance Approach» (FRA), beskrevet hos Dagher og Erduran (2016). Idéen i FRA-rammeverket er at man har et generelt sett med egenskaper/praksiser som tilskrives forskjellige disipliner, fremfor å etablere et generelt sett med egenskaper. Denne tar hensyn til bekymringene beskrevet overfor og gir et brukbart utgangspunkt til å kategorisere de fremstilte praksisene i oppgavene nøyaktigere fremfor å bruke mer generelle formuleringer.

<b>Praksis</b>	<b>Beskrivelse</b>
<b>Spør vitenskapelige spørsmål</b>	Dette er den drivende kraften bak all vitenskap, siden de setter grunnlaget for forklaringer og gjør innholdet mer begripelig
<b>Utvikle og bruke modeller</b>	Modeller brukes til å beskrive ting som er for komplekse til å kun forestille seg. Kunnskap om dette medfører også epistemisk kunnskap.
<b>Konstruere forklaringer</b>	Sterkt tilknyttet til modeller, siden man bruker disse til å forklare modeller.
<b>Argumentere fra bevismateriale</b>	Argumentere for validiteten til hypoteser ut ifra data. Dette gir også innsikt i epistemiske entiteter.
<b>Planlegge og gjennomføre utforskning</b>	Evaluere forskningsdesign og -praksis. Krever en blanding av prosedyre- og epistemisk kunnskap
<b>Analysere og interpretere data</b>	Data har ikke mening i seg selv, men krever tolkning. Dette innebærer også mulige feil ved innsamlingen av disse.
<b>Bruke matematisk og beregnende tenkning</b>	Å kunne tallfeste vitenskapelige ideer for å muliggjøre representasjonen gjennom variabler.
<b>Innsamling, evaluering og kommunisere informasjon</b>	Det er fundamentalt å kunne skrive og lese vitenskap, som er kjerneaktiviteter for å gjøre vitenskap.

Tabell 2: : Generelle vitenskapelige praksiser, hentet fra Osborne (2014).

For å oppsummere: Naturvitere tar utgangspunkt i den fysiske verden og utvikler teorier fra denne. Disse teoriene må testes, og derfor formuleres det testbare hypoteser fra disse. Deretter vender man tilbake til den fysiske verden og samler inn data, som hypotesene forutsier. Data vil derfor være preget av teori, altså være teoriladde. Konkretere praksiser innenfor naturvitenskapelig praksis, er gitt i Tabell 2. Sentralt her er argumentering, siden man må f.eks. overbevise andre over dataenes validitet, at en modell er en god representasjon av virkeligheten, eller at datamateriale støtter en modell.

### 2.2.3 Testing innenfor NPT

Testing innenfor de første to kunnskapsdimensjonene alene er forholdsvis simpel. Når man skal teste innholdskunnskap, krever man av elevene at de viser forståelse av vitenskapelige teorier og kunne anvende disse på forskjellige områder. Det er denne anvendelsen man er interessert i, når det kommer til testing av vitenskapelig resonnering

(Kind, 2013). Dette er en veldig forenklet beskrivelse av testing innholdskunnskap, og jeg kommer ikke til å utdype denne, siden oppgaven handler om prosessene som leder til denne teoretiske kunnskapen fremfor teoretisk kunnskap selv.

Metodekunnskap handler om testing av kunnskap om vitenskapelige data og praksiser. Men denne kunnskapen handler også om kunnskapsbygg, og har dermed en epistemisk side ved seg. Den som skal testes må vite hvordan et forslag til forskningsdesign er vel- eller dårlig egnet til generering av kunnskap. Kravet for å teste denne kunnskapen utspringer fra debatten om vitenskapelig resonnering skulle testes med utgangspunkt i hva man vet, eller hvordan man resonnerer. Siden begge standpunktene er sentrale deler i forskerens arbeid, ble inndelingen i innholds- og prosedyrekunnskap ansett som mer fruktbart enn å fortsette debatten rundt ett sentralt element i testingen av vitenskapelig resonnering (Kind, 2013).

Testing av NOS er et mer omstridt felt. Både i undervisningen og i testingen finnes det to dominerende retninger: en implisitt og en eksplisitt tilnærming. Den implisitte tilnærmingen tar utgangspunkt i selve deltakelsen i den vitenskapelige prosessen. Epistemisk grunnlag vil her utvikle seg implisitt underveis, og derfor kalles dette synet også ofte for «the living perspective». En eksplisitt tilnærming krever refleksjon rundt metodene og deres epistemologiske egenskaper. Denne tilnærmingen er kjent som «the reflective perspective» (Abd-El-Khalick, 2012). Jeg kommer til å bruke begrepene implisitt og eksplisitt tilnærming i denne oppgaven. Begge tilnærmingene har sine fordeler og begrensningene når det kommer til testing av elevers kunnskap om NOS. For å kunne belyse egenskapene bedre vil jeg ta utgangspunkt i Blooms taksonomi (Lasley, 2010), siden både representanter av den implisitte (Allchin, 2011) og eksplisitte tilnærmingen (Schwartz et al., 2012) benytter seg av denne i sin diskurs.

Fordelen med en eksplisitt tilnærming fremfor en implisitt ligger i undervisning av NOS. Ifølge denne tilnærmingen er den epistemologiske dimensjonen av den vitenskapelige praksisen er noe som skal som man skal kunne forklare språklig, og derfor reflekteres over. Dette er skillet mellom dette og det implisitte perspektivet, der man går ut ifra at elevene tilegner seg epistemologisk kunnskap idet de driver med vitenskapelig praksis. McDonald & McRobbie (2012) henviser til studier som tilsier at det kan gi en økt forståelse av NOS, i forhold til ingen undervisning i dette emnet dersom de engasjeres i argumentering innenfor vitenskapelige kontekster, selv uten eksplisitt opplæring i NOS.

Likevel nevnes at en eksplisitt undervisning i NOS og engasjering i argumentering vil oftere lede til en økt forståelse av NOS. I tillegg til at implisitt opplæring ikke nødvendigvis øker forståelsen av NOS, er en begrensning til denne tilnærmingen usikkerheten angående elevenes faktiske epistemiske kunnskap. Eleven vet at en slik metode gir rike resultat, men vil ikke være i stand til å forklare epistemologien bak metoden (Abd-El-Khalick, 2012). Så, selv om begge tilnærminger gir en økt forståelse av vitenskapelig praksis, skjer økningen i forståelse av eksplisitt opplæring oftere og mer entydig enn forhold til implisitt undervisning (Abd-El-Khalick, 2012; McDonald & McRobbie, 2012).

Når det gjelder testing av elevers kunnskap om NOS, vil mange hevde at det skjer en omvending av hvilken tilnærming egner seg bedre (Osborne, 2013). Idet man eksplisitt spør om f.eks. hva et eksperiment er, får man i beste fall innsikt i hva eleven vet, men dette tilsvarer kun første nivå av Blooms taksonomi. Med andre ord vet man ikke om eleven er i stand til å anvende kunnskapen eller evaluere vha. denne (Allchin, 2011; Lasley, 2010). Forkjempere av den eksplisitte tilnærmingen argumenterer at kritikerne unnlater å vise frem spørsmål som tilsvarer høyere kognitive nivå (Schwartz et al., 2012). Eksemplene som brukes til å underbygge deres påstand består av at testing innenfor den eksplisitte tilnærmingen bruker oppfølgingsspørsmål som f.eks.: «Krever utviklingen av vitenskapelig kunnskap eksperimenter? Argumenter for svaret ditt.» Dette oppfølgingsspørsmålet Schwartz et al. (2012) bruker beveger seg ikke på det laveste nivået i Blooms taksonomi, men vil ikke kunne nå de to høyeste heller. Her etterspørres det forholdet mellom vitenskapelig kunnskap og eksperiment. Det kreves av eleven å kunne sammenlikne epistemologiske elementer i eksperimentet med kjennetegn til vitenskapelig kunnskap. Dermed beveges det seg bort fra de laveste nivåene i Blooms taksonomi og mot nivå fire, analyse-nivået (Lasley, 2010). Dersom man brukte andre taksonomier som f.eks. SOLO-taksonomien, kunne man henvise til at svarene vil kreve en bedre forståelse, men siden SOLO-taksonomien har grunnlaget i elevenes svar enn oppgavens oppbygging og planlegging, tviles det på relevansen når det gjelder drøftingen av oppgaver (Biggs & Collis, 1982). Flere eksempler på eksplisitt testing, som er på høyere nivå innenfor Blooms taksonomi, er oppført i Figur 4.

En mer radikal implisitt tilnærming innenfor implisitt testing av NOS er «Whole Science» rammeverket til Allchin (2011) (Abd-El-Khalick, 2012). Denne tar ikke kun utgangspunkt i å teste NOS-kunnskapen til elevene, men betrakter også elevenes innholds- og

prosesskunnskap. I tillegg er et annet fundamentalt skille at den betrakter NOS mer som et analytisk verktøy enn et sett med kunnskap, slik forkjempere av eksplisitte tilnærmingen betrakter NOS (Schwartz et al., 2012). I eksempeloppgavene til Allchin (2011) blir dette klart i at elevene settes i, ifølge Allchin, autentiske situasjoner, der elevene skal komme frem med sine egne konklusjoner rundt et tema, f.eks. vaksiner eller global oppvarming. Til dette vedlegges noen artikler som eleven skal ta utgangspunkt i. En fordel med denne tilnærmingen er at oppgavene er designet til at elevene evaluerer mulige svar til åpne spørsmål. Oppgavene skal kreve at elevene har klare forestillinger om bevisets og påstandenes rolle og kunne bruke disse til å be- eller avkrefte påstander. Dermed beveger oppgavene seg på evalueringsnivået til Blooms taksonomi (Lasley, 2010). I tillegg scorer den også høyt på rammeverk som er tilpasset naturvitenskapelig resonnering (Kind, 2013). PISA-oppgavene fra 2015 operer innen dette rammeverket, siden studien leter etter om og hvordan elever bruker sin kunnskap fremfor gjengivelse av etablert kunnskap (Kjærnsli & Jensen, 2016b).

Kritikk av den radikale tilnærmingen til Allchin (2011) kan fattes sammen til følgende sentrale punkt (Schwartz et al., 2012):

- NOS er ikke et sett med analytiske ferdigheter, men kunnskap.
- Å vite er mer relevant enn å kunne anvende
- «Whole Science» er en omformulering av begrepet «scientific literacy»
- Det settes for høye kognitive krav dersom man tester alle aspekter av naturvitenskap i én oppgave
- Eksplisitte testemåter tilsvarer også høyere nivå på Blooms taksonomi

Jeg kommer til å ikke diskutere det første punktet, siden dette omfatter definisjonen til NOS fremfor å teste denne direkte. Det samme kan også sies om det tredje punktet, siden dette omhandler hva som skiller begrepene «Whole Science» og «Scientific Literacy» fra hverandre.

«Å vite er mer relevant enn å kunne anvende» innebærer, ifølge Schwartz et al. (2012), at det er mer ønskelig at elever besitter en større innsikt om vitenskapelig praksis enn konkrete ferdigheter innenfor dette. NOS er ikke en analytisk ferdighet, men et sett med kunnskap. Problemet med denne kritikken er at Allchin (2011) aldri hevdet at NOS i seg selv er en analytisk ferdighet, men heller at forståelse innenfor NOS bidrar til en bedre

kritisk analyse av forskningsdesign. Likevel har Schwartz et al. (2012) et poeng her, når det gjelder testing av NOS som analytisk redskap, i at man nesten ikke kan si noe om elevenes kunnskap og holdninger innenfor holdninger, dersom man tester dette som analytisk redskap, eller implisitt i oppgaver som omhandler alle aspekter av naturvitenskapen (Abd-El-Khalick, 2012).

«Det settes for høye kognitive krav dersom man tester alle aspekter av naturvitenskap i én oppgave» går ut ifra at man ikke kan forvente fra elever at disse kan sette seg inn i alle naturvitenskapelige aspekter når de jobber med en oppgave. Det blir gitt en sammenlikning, der man bruker cellebegrepet fra biologien. Man forventer ikke at elevene skal beherske alt angående celler. Allchin (2011) forventer heller ikke at elevene skal beherske absolutt alle dimensjoner av naturvitenskap, men han kritiserer at oppgaver om NOS kun fokuserer på faste kriterier innenfor epistemologi, istedenfor å se på NOS som et sett med dimensjoner om hvordan reliabilitet kan oppnås. Derfor må disse settes i en kontekst, der elevene kan vurdere forskjellige forskningsdesign. Dette krever mer enn å bare spør eksplisitt om forestillinger, men oppgaver kan likevel velges, slik at de samsvarer læreplanens krav for måloppnåelse.

I siste kritikkpunkt har Schwartz et al. (2012) rett i at Allchin (2011) utelatelse av oppfølgingsspørsmål sørger for at rammeverker for eksplisitt testing av NOS fremstilles som kun tilsvarende de laveste nivåene på Blooms taksonomi. Dette er allerede beskrevet tidligere, der det ble snakket om eksplisitt og implisitt testing (s. 9-10). Nedenfor er det vist en figur av oppgaver, hentet fra VNOS-testen (Abd-El-Khalick et al., 2001), som også Allchin (2011) brukte i sin artikkel. Jeg kommer til å belyse kort på hvilket nivå de befinner seg på Blooms taksonomi.



6. Science textbooks often represent the atom as a central nucleus composed of protons (positively charged particles) and neutrons (neutral particles) with electrons (negatively charged particles) orbiting that nucleus. How certain are scientists about the structure of the atom? What specific evidence **do you think** scientists used to determine what an atom looks like?
7. Science textbooks often define a species as a group of organisms that share similar characteristics and can interbreed with one another to produce fertile offspring. How certain are scientists about their characterization of what a species is? What specific evidence **do you think** scientists used to determine what a species is?
8. It is believed that about 65 million years ago the dinosaurs became extinct. Of the hypotheses formulated by scientists to explain the extinction, two enjoy wide support. The first, formulated by one group of scientists, suggests that a huge meteorite hit the earth 65 million years ago and led to a series of events that caused the extinction. The second hypothesis, formulated by another group of scientists, suggests that massive and violent volcanic eruptions were responsible for the extinction. How are these **different conclusions** possible if scientists in both groups have access to and use the **same set of data** to derive their conclusions?

(figure continues)

Figur 4: Utdrag av spørsmål i VNOS-testen, hentet fra Abd-El-Khalick et al. (2001).

Oppgave 6 er delt i to oppgaver. Den første oppgaven handler om hvor sikker forskere er om strukturen til et atom, og den andre om hvilket bevismateriale forskere anvender til å bestemme strukturen til et atom. Disse oppgavene skal måle elevenes kunnskap om rollen slutninger til de beste forklaringer og kreativitet innen naturvitenskap (Abd-El-Khalick et al., 2001). Denne oppgaven er holdt til en konkret, fra undervisningen kjent kontekst, og ber elevene om å anvende abstrakt kunnskap om hvordan vitenskapelige teorier bygges innen denne konteksten. Derfor vil jeg påstå at denne oppgaven beveger seg på forståelses- eller anvendelsesnivået av Blooms taksonomi (Lasley, 2010). Forståelsesnivå innebærer at man kan organisere informasjon til at den gir mening for elevene selv. På anvendelsesnivået kan elevene betrakte abstrakte prinsipp i konkrete situasjoner.

Det samme gjelder for oppgave 7. Oppgaven spør elevene om hvor sikre forskere er på karakterisering av biologiske arter, og hvilket bevismateriale de bruker. Igjen kreves kunnskap om hvordan kreativitet og slutninger til den beste forklaringen inngår i bygging av teorier, som man må bruke innen denne konteksten.

I oppgave 8 burde elevene vite om teoriladde data for å kunne forklare hvorfor to forskjellige konklusjoner kan bli dratt ut fra det samme datasett. Dette krever også en viss forståelse av et abstrakt konstrukt. Konstruktet består av at data er preget av teori. Man

trenger ikke å undersøke visse elementer og forholdet mellom dem. Ut ifra Blooms taksonomi befinner denne oppgaven seg da på mer anvendelsesnivået (Lasley, 2010). Eksemplene antyder at noen oppgaver beveger seg på de høyere nivåene i Blooms taksonomi, men de klarer aldri å nå evalueringsnivået, slik rammeverket til Allchin (2011) gjør.

Grunnen til at jeg nevner denne debatten er at det frigitte oppgavesettet fra PISA 2015, som belyser epistemiske spørsmål (Kjærnsli & Jensen, 2016c), kan interpreteres som den beveger seg innenfor rammeverket til Allchin (2011), selv om dette ikke blir nevnt eksplisitt. Dette oppgavesettet er «Undersøkelse av dalsider», og er delt i to underoppgaver. I den første skal man begrunne hvorfor man velger flere instrumenter til å måle én variabel, f.eks. to vannkolber til å måle nedbør. Dette legger vekt på det kunnskapsgererende, det epistemiske, aspektet ved et bestemt forskningsdesign. Den andre underoppgaven spør om forståelse rundt hvordan usikkerhet påvirker nøyaktigheten til data, og elevene skal etablere en årsakssammenheng ut ifra dette. Her kombineres igjen prosedyre- med epistemisk kunnskap, i at elever skal vurdere kvaliteten på data, som påvirker sikkerheten til vitenskapelig kunnskap. Begge er innenfor lokale kontekster, som kan ha en betydning for eleven, som Allchin (2011) legger stor vekt på i utforming av oppgaver. Men siden rammeverket til Allchin ble omtalt som kontroversielt blant tilhengere av den eksplisitte testingen (Abd-El-Khalick, 2012), anså jeg som nødvendig til å fremvise dette gjennom å belyse kritikken av dette rammeverket.

### **2.3 Elevers praksiser og oppfatninger innenfor naturvitenskap**

Når elever skal drive med eksperimentering, er det mulig å skille mellom to forskjellige tilnærminger som preger elevenes måte å jobbe og resonnerer på. En av disse tilnærmingene blir betegnet som *den praktiske tilnærmingen*, mens den andre kalles for *den vitenskapelige tilnærmingen*. Det nevnes at dette er modeller som gjelder elever og ikke beskriver tenkemåten til forskere eller ingeniører (Schauble et al., 1991).

Elever som bruker en praktisk tilnærming, eller ingeniør-tilnærming som Schauble et al. (1991) kaller dette (eng. «engineering model»), er mest interessert i å optimalisere et ønsket resultat. Dette gjøres ved å sammenlikne kontrastrike observasjoner, som gjøres ved å lete etter variabler som viser fenomenet tydeligst, til man finner variabelen som synes til å være årsaken til fenomenet. Slutninger ved bruk av denne tilnærmingen kjennetegnes av at elevene er interessert i å inkludere variabler i kausalpåstander.

Dersom det ønskete resultatet blir oppnådd, er målet oppnådd og eksperimenteringen stopper (Schauble et al., 1991).

Den andre tilnærmingen, som betegnes som vitenskapelig tilnærming (eng. «science model»). Hovedmålet til denne er å oppnå forståelse om relasjoner mellom årsaker og deres virkninger fremfor å produsere det mest ønskelige resultatet. Til dette brukes alle mulige variabler i alle mulige kombinasjoner av variabler, og man er interessert i å bestemme hvordan disse innspringer i et fenomen. Slutningene kjennetegnes av at de forsøker å utvikle påstander om kausalitet og er heller villige til å anerkjenne uvitenhet om en variabel sin rolle fremfor å ekskludere denne. Elever med en vitenskapelig tilnærming anser forskningen som avsluttet når hvert utfall av hver manipulerbar variabel er testet (Schauble et al., 1991).

Schauble et al. (1991) tilføyer at disse to ikke er de eneste strategiene elever bruker for å løse praktiske naturfaglige oppgaver. Hvilken av de to tilnærmingene ovenfor elever velger å benytte, avhenger også av hvilke oppgaver elevene jobber med, og dermed hvorvidt oppgaven er i stand til å fortelle elevene at de må gå slipp av sine hverdagsforestillinger i løsningsstrategien deres.

Det finnes samsvar i litteraturen for elevenes oppfatninger når de er konfrontert med problemer om epistemologi. Disse skiller mellom lavere, naivere nivåer, og høyere, reflekterte nivåer (Kind, 2013; Larkin, 1981). De naive nivåene er kjennetegnet av at individet tenker naivt positivistisk, og relaterer konkrete eksempler til en vitenskapelig idé uten form for abstraksjon. De reflekterte nivåene kjennetegnes av at individer kjenner til og bruker generelle vitenskapelige modeller, løsrevet fra en hverdagslig kontekst, når de konfronteres med problemer (Kind, 2013; Larkin, 1981). Forskjellen ligger i hvordan disse nivåene blir oppdelt i. Kind (2013) foreslår en tredelt skala, der det laveste og det høyeste nivået samsvarer med det som er beskrevet ovenfor, men at det også finnes et mellomstadium. I mellomstadiet vet individet om at modeller eksisterer, men feiler ved at de ikke ser at noen modeller er bedre egnet for noen problemer enn andre. Larkin (1981) derimot skiller mellom to representasjonsnivåer innenfor fysikalske problemer. Disse representasjonene vil jeg overføre til å dekke naturvitenskapelige idéer, istedenfor fysikalske prinsipper.

<b>Naive representasjon</b>	<b>Vitenskapelig representasjon</b>
Hverdagslige entiteter/objekter	Vitenskapelige entiteter/objekter
Egne idéer utspringer fra tidligere, egne idéer	Egne idéer utspringer fra entiteter
Ett grunnlag for meningsgjøring	Flere grunnlag for meningsgjøring

Tabell 3: Representasjoner av epistemisk-vitenskapelige entiteter, med utgangspunkt i representasjoner, beskrevet hos Larkin (1981)

Jeg kommer først til å snakke om de naive representasjonene. Denne bygger på kjente, ofte visuelle, entiteter, som man selv har erfart tidligere. Disse er kjennetegnet av tidsavhengige formuleringer, f.eks. man setter opp to vannkolber, fordi det kan hende at noen dyr drikker fra disse. Dette er tett tilknyttet til konkrete situasjoner, som man er kjent med. I tillegg følger denne påstanden også en konkret tidslinje, der man først setter opp måleinstrumentet, og så blir den påvirket av ytre, kjente, omstendigheter. I tillegg er det kun én begrunnelse på hvorfor man setter opp flere måleinstrumenter, nemlig minimering av feilkilder. Man tar f.eks. ikke utgangspunkt i at målinger av ett fenomen innebærer alltid en usikkerhet, som er en vitenskapelig forestilling. I stedet tar man utgangspunkt i egne forestillinger, at f.eks. en bestemt kontekst ikke tillater helt nøyaktige målinger. Det brukes heller ytre påvirkninger i naive representasjoner.

En vitenskapelig representasjon går ut over egne erfaringer og tar utgangspunkt i vitenskapelige entiteter. Dette kan f.eks. utspringe fra idéen om at målinger alltid innebærer en usikkerhet, som man ønsker å minimere, eller at man tar utgangspunkt i minimering av feilkilder. I tillegg til at man har flere grunnlag til å forklare et slikt forskningsdesign, er disse også tidsuavhengig. Man tar utgangspunkt i en tidsuavhengig idé og resonnerer ut ifra denne, fremfor å etablere kjente kontekster, som følger en tidslinje.

## **2.4 Et rammeverk for å beskrive elevenes tenke og arbeidsmåter**

Jeg kommer til å ta utgangspunkt i et rammeverk, som Kind (2013) presenterer til å diskutere hvilke kunnskapstyper som inngår i elevenes praksis når de jobber med PISA-oppgavene som står i fokus i denne studien. Rammeverket beskriver kognitive prosesser elever kan utføre, nemlig hypotesebygging, eksperimentering, og evaluering av bevismateriale. I den sistnevnte inngår det også hvordan elever evaluerer forskningsdesignet til å komme frem til bevismaterialet. Fordelen til dette er at den viser hvordan de tre kunnskapene, som ble diskutert i 2.2.1, inngår i praksisene fremfor å

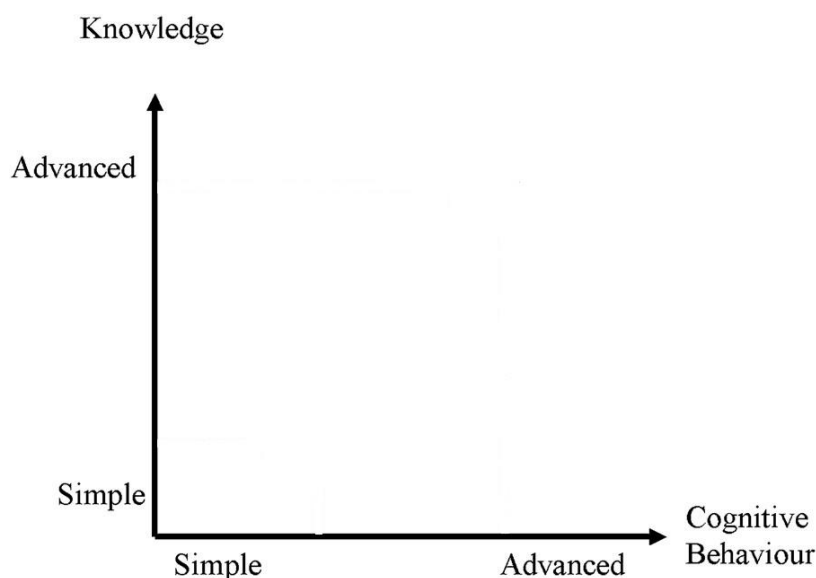
assosiere én praksis med én kunnskapstype. En visualisering av dette er oppgitt i tabellen nedenfor.

Kunnskapstype	Praksis		
	<i>Hypotesebygging</i>	<i>Eksperimentering</i>	<i>Evaluering av bevismateriale</i>
Innhold			
Prosedyre			
Epistemisk			

Figur 5: Rammeverk for å måle elevenes vitenskapelige resonnering, hentet fra Kind (2013).

Her går det også an å beskrive hvor mye en av de tre kunnskapstypene inngår i hver av de praksisene. Kind (2013) bruker f.eks. forskjellige markeringer til å merke hvilken av kunnskapstyper som inngår mest i praksisene. F.eks. dersom man viser mye innholdskunnskap under hypotesebyggingen, kan man bruke et stort «X», mens man bruke et små «x» hos de andre kunnskapene.

Begrensningen med dette er testingen elever som ikke viser eksplisitt kunnskap, men likevel bedriver med praksisene på en fornuftig måte. Dermed vil jeg tilføye en dimensjon som handler om kognitive ferdigheter, foreslått av Kind (2013). Han argumenterte for denne utvidelsen gjennom at den viser hvor vidt elevene anvendte kunnskapene i forskjellige oppgaver. Disse kognitive ferdighetene er bestemt ut ifra elevenes evne til å utføre kognitive prosesser, som Krathwohl (2002) nevner. Disse ble delt inn i 6 hovedkategorier: Huske, forstå, anvende, analysere, evaluere og skape.



Figur 6: Illustrering av elevenes kunnskapsnivå ('knowledge' på figuren) og kognitive ferdigheter ('cognitive behaviour' på figuren), hentet fra Kind (2013).

Figur 6 viser hvordan kognitive ferdigheter og kunnskap har innflytelse på elevenes praksiser. Dersom elevene har god kunnskap innenfor innhold, prosedyrer eller epistemologi, vil det vises lengre opp langs «Knowledge»-aksen. Denne avhenger av hvor god kunnskap elevene besitter innenfor de forskjellige kunnskapene, beskrevet i kapittel 2.2.1. Dersom praksisen deres er kjennetegnet av kognitive prosesser, vil det vises lengre til høyre langs «Cognitive Behaviour»-aksen. Tanken bak denne inndelingen er at selv utføring av enkle kognitive ferdigheter forutsetter, og innebærer bruk av, relevant kunnskap. Denne kunnskapen kan være bevist og kan deklarerer eksplisitt, eller mer ubevisst og implisitt. Så, selv om elevene ikke formulerer kunnskapen eksplisitt, vil høyere grads kognitive prosesser likevel peke på implisitt kunnskap hos elevene. Graden til kognitive prosessene her avhenger av kognitive praksiser er underliggende i vitenskapelig resonnering. Eksempler på dette er problemløsningsferdigheter, reflektert drøfting av utsagn på deres sannhetsverdi, induktive og deduktive slutninger, kausal tenkning, eller bruk av analogier (Dunbar & Fugelsang, 2005).

Når det gjelder diskusjon av elevenes oppfatninger og tenkemåter, vil jeg også ta utgangspunkt i tabell 3, siden denne anses som et velegnet verktøy til å knytte elevenes utsagn til epistemisk kunnskap, når de evaluerer data eller forskningsdesign.

## 3 Metode

### 3.1 Kvalitativ Metode

Innenfor sosiale vitenskap skilles det mellom to typer forskningsmetoder, kvantitativ og kvalitativ. Et kvantitativt forskningsdesign kjennetegnes av hypotesetesting, variabelkontroll, randomisert utvalg og en viss systematikk som opprettholdes under hele forskningsprosessen. Man er interessert i å fremstille empiri matematisk, som muliggjør en god fremvisning av korrelasjoner (Hoy, 2010, s. 2). Mens kvantitativ forskning er drevet av hypoteser til modellutvikling, ser kvalitativ forskning etter å gi forklaringer på hvorfor mennesker oppfører seg som de gjør. Dette bygger på den ontologiske antakelsen om at hvert menneske oppfatter virkeligheten på hver sin måte. Hvert menneske konstruerer en egen kompleks og stadig forandrende oppfatning av virkeligheten. Det er denne oppfatningen den kvalitative forskeren vil ha tak i til å bygge kunnskap. Kunnskapen blir da bygget i samspill mellom forskeren og forskningsdeltakeren (Nilssen, 2012, s. 25). Mange vil hevde at disse forskningsmetodene ikke står i motsetning til hverandre, men heller utfyller hverandre (Hoy, 2010, s. 2). Siden dette prosjektet omhandler elevers responser på PISA-oppgaver om NPT og er ute etter hvordan elever tenker når de jobber med disse oppgavene, vil et kvalitativt forskningsdesign være egnet. Dette begrunnes i at prosjektet ønsker å få innsikt i elevenes fremvisning av tenkemåter i lyset av oppgavene, og dermed elevenes oppfatning. Til dette anses kvalitativ forskning til å være bedre egnet enn kvantitativ forskning.

Dette valget krevet refleksivitet og bevissthet rundt forskerens rolle i kvalitativ forskning. Nilssen (2012, s. 31-33) skriver at forskerens tilstedeværelse kan påvirke deltakernes oppførsel, der hun bruker et eksempel hvor deltakere kan være mer engasjert i en aktivitet enn de vanligvis er. I tillegg til deltakere, er det også data som blir påvirket av forskerens forforståelse. Dette innebærer at forskerens kunnskap, erfaringer, og verdier påvirker hvordan et forskningsprosjekt designes, hva som skal observeres, hvordan data tolkes og diskuteres, og hvilke data som forskeren ønsker å presentere (Johannessen et al., 2010, s. 38-39; Nilssen, 2012, s. 68). Denne forforståelsen må forskeren være seg bevisst på under hele forskningsprosjektet.

Min forforståelse ble først og fremst preget av innsamlet teori. Der fikk jeg innsikt i en del modeller og teori som behandlet naturvitenskapelig tenkemåte og tilnærminger til hvordan elever tenker i naturvitenskapelige kontekster. Dette har hatt innflytelse på

hvilke tenkemåter jeg ønsket å observere, som ytret gjennom spørsmålene jeg valgte å stille.

### **3.2 Validitet og Reliabilitet**

Validiteten til et kvalitativt prosjekt dreier seg om hvorvidt den brukte metoden egner seg til å få frem data som har en god sammenheng med fenomenene som skal undersøkes (Johannessen et al., 2010, s. 230; Kvale & Brinkmann, 2009, s. 250). Dette krever at presenterte teori og innsamlet data, gir fornuftige og overbevisende slutninger (Kvale & Brinkmann, 2009, s. 250). Johannessen et al. (2010, s. 230) betegner dette kravet som intern validitet, og Kvale og Brinkmann (2009, s. 254) argumenterer for validitetens kontrollerende rolle i at forskeren skal alltid være seg bevisst på mulige feilkilder og kontrollere funnenes pålitelighet, sannsynlighet og troverdighet.

For å øke dette prosjektets indre validitet var valget av fokusgrupper som innsamling av data. Fokusgrupper er gruppeintervjuer, der et sentralt tema står i fokus som driver samtalen, og samtale mellom deltakere er ønsket (King & Horrocks, 2010, s. 65-66). Å ha et konkret tema i fokus under intervjuet oppleves ofte som mer komfortable for deltakerne å snakke om et konkret tema enn å bli konfrontert med forskjellige spørsmål. Gruppemedlemmene kan også snakke med hverandre, som igjen øker validiteten til dataene, siden deltakere sjekker og balanserer hverandre sine utsagn (Patton, 2002, s. 386). I tillegg ble alle fokusgruppene etter hver oppgave spurt om å repetere deres fremgangsmåte for å løse denne oppgaven. Ved å gjenta løsningsprosessen minimeres tolkningsrommet til forskeren, siden dette kan gi avklaring om eventuelle feiltolkninger fra forskerens side. Intervjueren kunne ha oppsummert sin hele tolkning av løsningsprosessen og kunne deltakerne la be- eller avkrefte denne, men dette ville ha blitt forholdsvis lange spørsmål. Derimot kunne deltakerne korrigert hverandre, dersom oppsummeringen til den spurte deltakeren ikke eller kun delvis reflekterer tankegangen til de andre deltakerne. Det blir også stilt noen ledende spørsmål om elevenes påstander som f.eks. «Altså det du sier, er at ...» for å overprøve om forskeren fortolkning overensstemte med elevenes utsagn. Dette gir deltakerne muligheten til å motsi eller verifisere intervjuerens tolkninger, som kan være positivt for validiteten (Kvale & Brinkmann, 2009, s. 182-183) Når forskningsdesignet utdypes, blir det sett på kilder og valg som muligens kunne ha innskrenket resultatenes validitet, som er uttrykk for validitetsspørsmålets kontrollerende funksjon.



En annen form for validitet er det Johannessen et al. (2010, s. 230-232) betegner som ekstern validitet, som går ut på dataene overførbarhet til andre situasjoner. Ekstern validitet finnes også i kvantitative studier som statistisk generalisering, der man slutter fra et utvalg på en populasjon. Kvalitative studier har også et ønske om å generalisere, men bruker uttrykket «overføring» fremfor «generalisering», siden dette gir henviser heller til statistisk generalisering av kvantitative studier.

I dette prosjektet ble det satt vekt på å holde kodene og kategoriene så generelle som mulig for å gjøre overførbarheten til andre situasjoner, der elevene driver med oppgaver som tester innenfor naturvitenskapelig praksiser og tenkemåter. Disse kategoriene er utarbeidet ved å intervjuer grupper, som hadde forskjellig faglige nivåer. Det faglige nivået i selve gruppene var homogent. Dette valget skal styrke prosjektets overførbarhet i at kategoriene kan anvendes på tvers av elevenes faglige nivå. Et annet grep som ble gjort er å gi en tjukk beskrivelse (eng. 'thick description') av konteksten. Dette gjør det mulig for fremtidige lesere å avgjøre om oppgavens funn kan overføres til flere kontekster, altså at man legger grunnlaget for en naturalistisk generalisering (Firestone, 1993). Videre nevner Firestone (1993) at en økning av diversitet blant deltakere er en annen metode for å øke overførbarheten til funnene. For å etterkomme dette kravet, ble lagt vekt på geografisk diversitet gjennom å ha med forskningsdeltakere fra Bergensområdet og Vest-Agder, og diversitet i elevenes kunnskap, gjennom å be lærere til å også velge elever som de anså til å middels til lav måloppnåelse, sammenliknet med læreplanmålene.

Kvale og Brinkmann (2009, s. 250) bruker ordet reliabilitet når det er snakk om studiens konsistens og troverdighet, og behandler hvorvidt data kan reproduseres. Dette er sterkt tilknyttet til valget av data og deres nøyaktighet, innsamlings-, og bearbeidingsmetode (Johannessen et al., 2010, s. 40). Likevel har kvalitative studier, i motsetning til kvantitative studier, vansker til å oppfylle kravet om reproduserbarhet i at observasjonene er sterkt kontekstavhengig og forskjellige forskere har forskjellig erfaringsbakgrunn. En metode å styrke reliabiliteten til en studie er å gi en grundig kontekstbeskrivelse (Johannessen et al., 2010, s. 229-230). Dette styrker reliabiliteten til et prosjekt, fordi dette gir leseren muligheten til å selv vurdere om funnene gir mening i lys av innsamlet materiale (King & Horrocks, 2010) For å etterkomme dette kravet om reliabilitet, vil dette prosjektet gi en detaljert beskrivelse av hele forskningsprosessen.

### 3.3 Utvalg av PISA-oppgaver

Kjærnsli og Jensen (2016c) har vedlagt oppgaver fra PISA 2015, der de kategoriserte oppgavene i en tabell etter oppgavetype, kompetanse, kunnskap – system, kontekst, og vanskegrad. Kategoriene er hentet fra deres definisjon på naturfag gitt i tidligere kapittel (Kjærnsli & Jensen, 2016b) og følger etablert teori om naturvitenskapelig resonnering til en viss grad (Osborne, 2014). For eksempel er oppgavetype i oppgaven «Løping i varmt vær: Oppgave 1» «kompleks flervalg» (se Tabell 4). Dermed er mitt eneste kriterium for utvalg av oppgaver følgende: Kompetanse- og kunnskapskategorien skal ikke være henholdsvis «Forklare fenomener på en naturfaglig måte» og «Innhold» samtidig. Selv om kunnskapene og kompetansene innenfor naturfaglig resonnering påvirker hverandre til en viss grad interagerer med hverandre (Kind, 2013), er dette prosjektet ikke interessert i hvor godt elevene kan forklare selve innholdet, men heller hvordan dette kan påvirke deres kompetanse innenfor NPT. Dermed falt utvalget på to oppgaver: «Løping i varmt vær» og «Undersøkelse av dalsider», som behandlet stort sett metodekunnskap og epistemologisk kunnskap (Kjærnsli & Jensen, 2016c). Et unntak her er oppgave 3B fra «Løping i varmt vær», som ikke oppfyller kriteriet nevnt overfor. Likevel er denne oppgaven inkludert, siden den er del av en annen oppgave, som tar for seg metode, og kan gi innsikt i hvordan elevene ser på spillet mellom påstand og bevis.

<b>Oppgavetype</b>	Kompleks flervalg
<b>Kompetanse</b>	Tolke data og evidens på en naturvitenskapelig måte
<b>Kunnskap – system</b>	Metode – Levende
<b>Kontekst</b>	Personlig – Helse og sykdom
<b>Vanskegrad</b>	495 – Nivå 3

Tabell 4: Kategorisering av en oppgave fra PISA 2015, hentet fra Kjærnsli and Jensen (2016c)

#### 3.3.1 Kort beskrivelse av PISA-oppgavene

Opgavesettet «Løping i varmt vær» gikk ut på at elevene brukte en simulasjon av biologiske prosesser hos en løper, og består av fem deloppgaver. Der måtte man stille inn på variablene «Lufttemperatur», «Luftfuktighet» og om løperen drikker vann. Når alt var stilt inn, kjøres simulasjonen og man får ut datasett, som inkluderer de tre variablene nevnt ovenfor og verdier om løperens kroppstemperatur, væsketap og svettmengde. Disse datasettene skulle brukes til å svare på spørsmål angående verdier som ikke gikk an å simulere, eller å etablere forholdet mellom avhengige og uavhengige variabler.

PISA 2015

**Løping i varmt vær**  
Spørsmål 3 / 6

► **Hvordan kjøre simuleringen**

Bruk informasjonen nedenfor, og kjør simuleringen for å få fram data. Svar på spørsmålet ved å klikke på et av alternativene. Velg deretter data i tabellen, og skriv en begrunnelse for svaret ditt.

Hvordan påvirkes svettemengden etter en times løping når lufttemperaturen øker og luftfuktigheten er 60 %?

Svettemengden øker  
 Svettemengden minker

★ Velg to rader med data i tabellen for å støtte svaret ditt.

Hva er den biologiske årsaken til dette?

Svette-  
mengde (liter)

Væske-  
tap (%)

Kropp-  
temperatur (°C)

Lufttemperatur (°C)  20 25 30 35 40

Luftfuktighet (%)  20 40 60

Driker vann  Ja  Nei

**Kjør**

Lufttemperatur (°C)	Luftfuktighet (%)	Driker vann	Svettemengde (liter)	Væsketap (%)	Kroppstemperatur (°C)

Figur 7: Eksempel for en oppgave fra settet "Løping i varmt vær", hentet fra Kjærnsli og Jensen (2016c).

Det andre oppgavesettet het «Undersøkelse av dalsider», og består av to deloppgaver. I den første oppgaven må elevene gi en forklaring på hvorfor man velger å sette opp to av hvert måleinstrument til å måle data som gir informasjon om en dalsides vegetasjon. Den andre oppgaven gir elevene data av målingene, inkludert usikkerhet, og spurte om forskjell i solstråling eller nedbør var årsaken til forskjellen i jordfuktighet.<sup>1</sup>

<sup>1</sup> Dette er en link, som henviser til alle eksempeloppgaver, inkludert eksempelsvar <https://www.idunn.no/sto-kurs-pisa-2015/vedlegg-eksempeloppgaver-i-naturfag-fra-pisa-2015>

PISA 2015

**Undersøkelse av dalsider**  
Spørsmål 3 / 3

Les "Dataanalyse" til høyre. Svar på spørsmålet ved å klikke på et av alternativene, og begrunn deretter svaret ditt.

To elever er uenige om hvorfor jordfuktigheten i de to dalsidene er forskjellig.

- Elev 1 mener at forskjellen i jordfuktighet skyldes en forskjell i solstråling mellom de to dalsidene.
- Elev 2 mener at forskjellen i jordfuktighet skyldes en forskjell i nedbør mellom de to dalsidene.

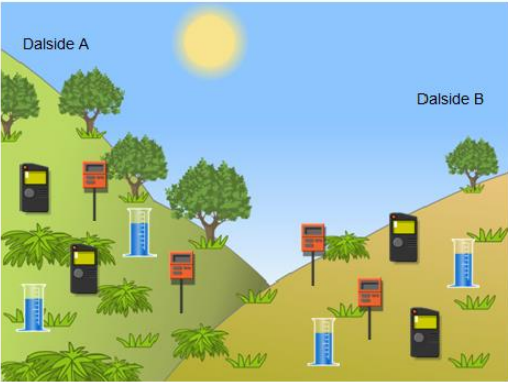
Hvilken elev har rett, ifølge dataene?

Elev 1  
 Elev 2

Begrunn svaret ditt.

**UNDERSØKELSE AV DALSIDER**  
**Dataanalyse**

Elevene tar gjennomsnittet av målingene som hvert instrument har registrert i hver dalside over en gitt tidsperiode, og beregner usikkerheten i disse gjennomsnittene. Resultatene vises i tabellen nedenfor. Usikkerheten er gitt etter tegnet "±".



	Gjennomsnittlig solstråling	Gjennomsnittlig jordfuktighet	Gjennomsnittlig nedbør
Dalside A	3800 ± 300 MJ/m <sup>2</sup>	28 ± 2 %	450 ± 40 mm
Dalside B	7200 ± 400 MJ/m <sup>2</sup>	18 ± 3 %	440 ± 50 mm

Figur 8: Oppgave 2 fra settet "Undersøkelse av dalsider", hentet fra Kjærnsli og Jensen (2016c).

### 3.4 Utvalg av deltakere

PISA-studiens deltakere er elever på 15 år, som nesten alle gikk på 10. trinn. Skolene og elevene ble utvalgt etter kriterier om representativitet og tilfeldighet (Kjærnsli & Jensen, 2016a). Dermed var inkluderingskriteriene for dette prosjektet at deltakerne skulle være 15 år gamle og gikk i 10. klasse. Etter å ha definert utvalgspopulasjonen (eng. 'sampling universe') (Robinson, 2014), ble selve utvalget bestemt etter prinsippet om kriteriebasert utvalg (eng. 'criterion sampling') (Patton, 2002, s. 238), der dette prosjektet hadde tre kriterier. Lærerne fikk informasjon om at de skulle helst velge elever som likte å snakke om hvordan de løser oppgaver, som ikke deltok kun for å slippe undervisning, og deltok frivillig. Grunnen til at det ble overlatt til læreren å velge ut deltakere, var at hen har innsikt i dynamikken mellom elevene og kan derfor velge ut elever, slik at intervjuet oppleves som mest komfortabel for dem, siden dette kan øke kvaliteten på innsamlet data (Patton, 2002, s. 386). Fordelen med disse kriteriene er at det skulle øke sannsynligheten for at intervjuene ga rike data.

Begrensningen til denne måten å velge ut elever på denne måten, er det Robinson (2014) betegner som selvutvalget sitt forutinntatthet (eng. 'self-selection bias'). Dette referer til at frivillige deltakere er forskjellig fra de ikke deltakende pga. grunner som er andre enn inkluderingskriteriene. Den spesifikke utfordringen med selvutvalget sitt forutinntatthet i dette prosjektet likner på beskrivelsen av ideelle intervjupersoner hos Kvale og Brinkmann (2009, s. 175-176). En ideell intervjuperson karakteriseres som konsistent og presis, og liker å gi lange og levende beskrivelser. Denne beskrivelsen passer oftest på en mann fra øvre middelklasse med høyere utdanning. I dette prosjektet ville det være en høyere sannsynlighet for å få deltakere, som gjør det godt i faget, siden deltakelse var frivillig (Robinson, 2014). Pga. dette ble lærerne i etterkant bedt om å gi en kort beskrivelse av deltakernes ferdigheter innenfor naturfag. Dette var mulig, siden elevene har gitt skriftlig samtykke til at læreren kan gi mer utfyllende informasjon om dem. Lærerne karakteriserte deltakerne som faglig sterke eller som del av det øvre middelmål. Derfor ble kriteriene for senere utvalget endret til at elevene skulle delta frivillig, og prestere middelmådig til dårlig i faget. Å like å snakke om oppgaver var fortsatt et ønske, men ikke et krav. Denne vendingen har egenskapene til et opportunistisk utvalg, dvs. forskeren tilpasser seg til uforutsette begivenheter som oppstår under feltarbeidet (Ritchie et al., 2003). Gruppene selv var fortsatt homogene, men med de nye kriteriene finnes det variasjoner mellom gruppene. Dermed ble utvalgsstrategien til slutt stratifisert målrettet utvalg (eng. 'stratified purposive sampling'), som er betegnelsen på denne strategien (Patton, 2002, s. 240). Likevel ble disse elevene også karakterisert til å ha gode leseferdigheter, selv om de hadde middels måloppnåelse med hensyn til læreplanen i naturfag, som igjen kan være en form for forutinntattheten til selvutvalget.

### **3.5 Fokusgruppe**

Utfordringen til prosjektet er å oppnå innsyn i elevenes tanker når de jobber direkte med oppgavene, uten at intervjuet kunne oppfattes for styrt av intervjueren. For å takle denne utfordringen, ble fokusgruppe som intervjumetode ansett som best egnet. To spørsmål som raskt fremstilte seg i forbindelse med utformingene av intervjuene var da: «Hvor mange elever skal delta i studien?», og: «Burde en intervjuer/moderator være tilstede?» Grunnlaget til begge spørsmålene ligger i hvordan deltakerne kommer til å snakke om oppgavene. Både (King & Horrocks, 2010, s. 66-69) og Serder (2015) understreker betydningen av at deltakere må være i et miljø de er komfortable til å snakke i, og at dette kan oppnås til en stor grad gjennom å sette dem i autentiske situasjoner.

Med hensyn til det første spørsmålet ble det valgt grupper på tre elever som skulle jobbe med oppgavene. Dette gir elevene mulighet til å diskutere seg frem til et svar selv i en mer naturlig kontekst enn hos ett 'vanlig' intervju, der de intervjues alene (King & Horrocks, 2010, s. 61-62). Likevel foregår intervjuer, også fokusgrupper, alltid utenfor omgivelsene, der sosiale hendelser vanligvis foregår (Patton, 2002, s. 387-388), og dette må forskeren være seg bevisst på for å oppnå en komfortabel følelse hos deltakerne.

Det andre spørsmålet omfatter en asymmetri om når en intervjuer er tilstede under et intervju med elever. I en vanlig intervjusetting er det deltakerens oppfatninger av virkeligheten forskeren ber om å få innsikt i til å bygge kunnskap (Nilssen, 2012, s. 28-31). Med andre, enklere ord: Forskeren ber forskningsdeltakeren om å få kunnskap, som forskeren ikke har. Når man skal intervjuere elever, ber man dem også om å få innsikt i deres tanker, men i en kontekst, der elevene skal jobbe med oppgaver, må man være seg bevisst på en asymmetri, som også spiller inn. Denne asymmetrien består av at intervjueren, den voksne, er den som sitter med kunnskap, mens den intervjuede, eleven, er den som ikke eier den kunnskapen. Elevene kommer da i en situasjon, der de heller kun vil velge innspill som føles relevante for konteksten de befinner seg i enn å snakke fullstendig fritt (Serder, 2015, s. 80-82). Med andre ord: Elevene vil heller si noe som oppleves sant enn å faktisk formulere sine helt egne tanker, dersom konteksten oppleves som mer offisiell. Serder (2015, s. 80) taklet dette problemet ved å ikke være tilstede under selve oppgavejobbingen, men tok videoopptak for å kunne observere det non-verbale når elevene satt i klasserommet og jobbet med oppgavene, før gruppeintervjuene ble gjennomført. Et slikt design, med videokamera og uten intervjuer, ble også vurdert i dette prosjektet, men dette ble vurdert til å være unødvendig.

Grunnene for dette valget er at dette prosjektet er interessert i om elever med kunnskap om NPT viser frem denne kunnskapen når de løser PISA-oppgaver om temaet. Dette krever eksplisitte formuleringer av konstruktene fra elevene selv. Når det jobbes med oppgavene kan det derfor føles mer autentisk for deltakerne når det snakkes om konstruktet idet man løser oppgavene fremfor å ta en oppsummerende runde etterpå. Dette vil det også gjøre det mulig å plukke opp tankesett til enkeltelever som oppstår under oppgavejobbingen, men som kan bli tapt etter et lengre tidsrom, som f.eks. fra oppgaven er avsluttet til intervjueren spør om en generell løsningsstrategi. Derfor ble det

valgt at en intervjuer/moderator er tilstede under oppgavejobbingen, og kommer med spørsmål etter elevene har avgitt svar til en oppgave, før det går videre til neste oppgave.

Etter å ha gjennomført det første intervjuet, viste det seg i tillegg at dette designet var bedre egnet i denne oppgaven, pga. tekniske utfordringer i deloppgavene om «Løping i varmt vær». Utfordringene bestod av at når trykket på knappen som ledet videre til neste deloppgave, ble datasettene, som man skulle underbygge svaret sitt med, slettet. For å kunne referere direkte til elevenes svar, uten behov for at man skulle sette seg inn i oppgaven igjen, og gjøre oppgavejobbingen så autentisk som mulig, er det også fordelaktig at intervjueren er tilstede fremfor å forandre på presentasjonen av oppgavene og kreve en ny gjennomgang av disse.

Det ble også valgt å bruke lydopptak fremfor videoopptak, siden fordelene til videoopptaket, å få en god observasjon av den non-verbale interaksjonen til deltakerne (King & Horrocks, 2010, s. 46-47), ikke spiller en stor rolle for dette prosjektets formål. Å innsamle non-verbal kommunikasjon kan oftest være verdifullt, men siden dette handler primært om å få tak i elevenes formuleringer, som heller gir uttrykk for deres tenkemåter fremfor deres non-verbale samspill.

### **3.6 Kontekstbeskrivelse**

Prosjektet består av fire fokusgrupper. Alle deltakere var 15 år gamle. Av de fire gruppene var to rene guttegrupper to grupper hadde både jenter og grupper. Tre grupper bestod av tre elever, og én guttegruppe bestod av fire elever, pga. naturfagslæreren og elevenes ønske. Tre grupper er forskjellige skoler fra Bergensområdet, og én gruppe er fra Listerregionen i Vest-Agder. Det nevnes at hverken deltakernes kjønn eller hvor i landet skolene befinner seg anses som relevant i prosjektet, men blir nevnt her for å gi en tydeligere kontekst.

Elevene og intervjueren ble satt i et eget, avskilt rom, enten et klasse- eller et møterom på skolen. Dette rommet ble utvalgt av naturfagslæreren til elevene Jobbing med oppgaven foregikk på en PC medbrakt av intervjueren, som hentet uvalgte PISA-oppgave fra internett. Elevene fikk først informasjon, både muntlig og skriftlig, om hensikten til prosjektet og hva intervjueren ser etter under oppgavejobbingen. Dette ble fulgt av informasjon om intervjuets struktur, dvs. at det jobbes med oppgaver og at det blir laget lydopptak, som blir transkribert og deltakerne anonymisert, samt deltakernes

personvernrettigheter. Deretter ble det avsatt tid til spørsmål om ytterligere informasjon til prosjektet.

Før oppgavejobbingen begynte ble elevene bedt om å si fornavnet sitt høyt, for å enklere kunne skille mellom elevene under transkriberingen. Denne delen av lydopptaket ble slettet med en gang når transkriberingen var avsluttet. Elevene fikk informasjon om dette.

Det ble alltid først jobbet med oppgavene fra oppgavesettet «Løping i varmt vær», og deretter med oppgavene fra settet «Undersøkelse av dalsider», der rekkefølgen til deloppgavene er gitt i oppgaven. Etter hver oppgave var gjennomført ble elevene spurt om detaljer innenfor deres løsningsvei, eller hvordan de forstod bestemte ord. Til slutt ble det spurt om hva som var vanskelig forstå i oppgaven og om én kunne gi en kort oppsummering av gruppens løsningsmetode.

Avslutningen til intervjuene, etter lydopptaket var avsluttet, ble markert med en kort «debriefing» (Kvale & Brinkmann, 2009, s. 142), der elevene ble spurt om deres opplevelse av intervjuet og om de hadde noen flere kommentarer eller spørsmålet til prosjektet. Av og til spurte intervjueren om noen generelle inntrykk han fikk av elevenes løsningsvei og ga elevene mulighet til å be- eller avkrefte denne tolkningen. Dette ble gjort for å oppnå en entydig tolkning av utsagnet til deltakeren.

### **3.7 Transkribering og koding**

Totalt ble omtrent fire timer lydopptak transkribert, som resulterte i 95 sider, som inneholdt 19.239 ord totalt, med datamateriale. Alt ble transkribert, siden jeg anså alle innspill som elevene hadde beveget seg innenfor oppgavekonteksten. Det ble transkribert ordrett i den forstanden at ordstilling blir fullstendig transkribert, men ord, som er særegne for forskjellige dialekter, ble oversatt til vanlig bokmål. Pauser og egne uttrykk ble også transkribert, der det skilles mellom korte pauser, som markeres som to punktum etter et ord, og lange pauser, som ble markert som tre punktum. Dersom noen la stort vokalt trykk på et ord, ble dette understreket.

Jeg valgte å transkribere selv for å oppnå bedre innsikt i materialet, siden dette er en viktig del av analysen. Under transkriberingen dukker det opp nye idéer og tanker til kodingen, som ble skrevet ned. (Nilssen, 2012, s. 47-49). I tillegg kjenner jeg da bedre til konteksten til enkelte sitat, som gjør det enklere å avklare hvilke tenkemåter som lo bak et bestemt utsagn. En tilfeldig utvalgt side av transkripsjonen er vedlagt.



I tillegg til forskningsnotatet ble programvaren NVivo 12 brukt som kodingsverktøy for å kunne bedre holde oversikt over kodene i datamaterialet.

### **3.8 Dataanalyse**

Det ble transkribert omtrent fire timer lydopptak, som resulterte i 95 sider med datamateriale. Datamaterialet ble analysert ned utgangspunkt i konstant komparativ analyse, beskrevet hos Nilssen (2012, s. 79-80). Praksisene til Osborne (2014) (se Tabell 2) ble brukt som forhåndsbestemte kategorier. Dette ble gjort for å minimere datamengden og unngå en stor mengde med ord og koder, som kan forekomme under en åpen kodingsfase (Nilssen, 2012, s. 82). Åpen koding er kjennetegnet gjennom at forskeren gir slipp på teoriene og kun tar utgangspunkt i hva datamaterialet forteller uten forhåndsdefinerte koder og kategorier (Nilssen, 2012, s. 65). Derfor vil jeg ikke betegne dette som åpen koding, siden jeg hadde forhåndsdefinerte kategorier, som pekte på hva som skulle anses som relevant. Likevel hjalp dette grepet med å redusere mengden av datamateriale, og samtidig gjennomføre en begrunnet en grov identifisering av praksiser.

Deretter vendte jeg tilbake til datamaterialet og åpnet kodingen opp gjennom at jeg valgte å innføre underkategorier til enhver av Osbornes (2014) kategorier. Dette ble gjort, fordi de opprinnelige kodene ikke fanget alle nyansene til elevenes utsagn. En åpnere koding ville derfor gi et mer nyansert bilde av praksisene til elevene. Det, som styrte denne kodingsprosessen, var hvordan elevene kjørte simuleringen i oppgavesettet «Løping i varmt vær» og hvordan de tolket datasettene som ble simulert. I oppgavesettet «Undersøkelse av dalsider» så jeg etter hvilke argumenter elevene brukte i sine svar på de epistemiske spørsmålene, og forsøkte å identifisere like tankesett i forskjellige ytringer. Jeg pendlet mellom å lese enkeltutsagn om igjen, sammenlikne kategorier med underkategorier, og se etter liknende tenkemåter i forskjellige samtaler. Forskjellige koder ble prøvd ut og sammenliknet med forskjellige rammeverk innenfor vitenskapelig resonnering og NOS. Her var det viktig å konstant relatere koder til datamaterialet og omvendt for å ikke bevege seg for langt bort fra det som faktisk ble observert.

Først forsøkte jeg å henvende meg til Giere (1991) sitt filosofiske rammeverk om fasene i naturvitenskapelig modellutvikling, siden elevene jobbet med en modell om hvordan kroppen fungerer under løping med forskjellige parametere. I begynnelsen virket dette som en oversiktlig måte til å beskrive datamaterialet, men etter hvert antydte det seg at denne normative oversikten ikke helt fanget nyansene til hva elevene faktisk gjorde og sa.

Likevel inspirerte dette meg til å se på elevenes løsningsmåter i oppgavesettet «Løping i varmt vær» som en prosess i flere faser, der jeg brukte underkategoriene som enkelte steg i prosessen. Under denne utformingen så jeg på hvordan Schauble et al. (1991) beskrev elevers tenkemåter når de jobbet med åpne oppgaver som enten krevet praktiske eller vitenskapelige løsningsmetoder. I dette rammeverket ble flere løsningsveier til elever karakterisert, som inspirerte meg til å beskrive flere løsningsveier ved siden av hverandre uten å karakterisere den ene som mer rett enn den andre. Dette førte til sammenslåing av noen underkategorier, der jeg forsøkte å beskrive nøyaktigere hvordan elever kjørte simuleringen og tolket dataene. Med disse nye kategoriene så jeg på datamaterialet en gang til for å undersøke om dette gir en brukbar beskrivelse av elevenes løsningsprosess. Denne gangen viste det seg at disse tok hensyn til forskjellige nyanser til elevenes tenkemåte, som f.eks. at noen kjørte simuleringen målrettet, mens noen ikke gjorde dette. Denne konstant komparative analysen resulterte i 5 kategorier, som ble satt inn i en modell om elevenes løsningsstrategier i oppgavesettet.

Et unntak i analysen av dette oppgavesettet var en del av oppgave 3, der elevene skulle forklare hvorfor mennesker svetter mer når lufttemperaturen økte. Ut ifra dette, leste jeg elevenes forklaringer på nytt og så etter hvordan de brukte teori, uavhengig om den var rett eller ikke, og hvordan de forstod teoriens rolle i en forklaring. Derfra ble det utviklet tre hovedkategorier, som først ble sammenliknet praksisen «Konstruere forklaringer» til Osborne (2014), siden oppgaven behandlet dette ved å knytte teori til datasettene, og deretter med igjen med elevenes utsagn, for å teste om de fortsatt samsvarte med datamaterialet.

I den videre kodingsprosessen av datamaterialet fra oppgavesettet «Undersøkelse av dalsider» spurte jeg meg hvilken argumentasjon elevene brukte i oppgavene og på mine spørsmål, og hvilke oppfatninger som lo til grunn for dette. Her viste seg mine foreløpige underkategorier til kategoriene av Osborne (2014) som svært nyttige, siden disse allerede ga et forholdsvis beskrivende bilde over dette. Jeg pendlet ofte frem og tilbake mellom de underkategoriene jeg hadde og datamaterialet for å finne likheter i tenkemåtene, som kunne oppsummeres i forskjellige kategorier. Her virket Kind (2013) sine kriterier om målinger av en enkel variabel som et rimelig utgangspunkt, siden den oppgavesettet settet dreiet seg om dette. Dette ble brukt til å lete etter forståelse rundt målinger av enkelte variabler og spørsmålene som ledet her var: «Hvilken forståelse viser elever forståelse

rundt målinger?» Utfordrende i denne prosessen var å beskrive oppfatningene til elevene uten å evaluere dem med hensyn på naive eller vitenskapelige oppfatninger. Denne kodingen resulterte i 5 hovedkategorier, som dreier seg om kjennetegn i elevenes resonnering angående målinger av én variabel.

### **3.9 Ethiske betraktninger**

I et intervju finnes det samspill mellom mennesker, som påvirker intervjupersonene og den produserte kunnskapen påvirker vårt syn på menneskets situasjon. Da blir spørsmålet om intervjuets midler og dets mål sentrales, som også gjør dette til en moralsk undersøkelse (Kvale & Brinkmann, 2009, s. 79-80). Dette krever en vurdering av prosjektets etikk.

Intervjustudier inneholder vanligvis direkte sitater fra mennesker som inngår i offentlige rapporter. Da må forskeren underkaste seg det etiske kravet om å ivareta deltakernes anonymitet på en mulig bekostning av at forskeren kan tolke utsagnene uten å bli motsagt (Kvale & Brinkmann, 2009, s. 90). Den etiske utfordringen i å bruke lydopptak er at deltakerne kan bruke navnet til meddeltakerne når de snakker til dem, og at alle deltakerne kan identifiseres ut ifra stemmene deres. Med hensyn på dette ble alle deltakerne anonymisert i transkriberingen, også når de referer til meddeltakere.

Anonymitetens begrensning ved at deltakerne ikke har mulighet til å motsi en mulig offentliggjøring av en ikke ønskelig tolkning er også en utfordring, som blir kort nevnt ovenfor. Denne begrensningen kan unngås i dette prosjektet ved å repetere noen resonnementer til elevene for å gi dem mulighet å motsi, som nevnt i avsnittet om validitet og reliabilitet (side 24). I tillegg har alle deltakerne mulighet til å få sine utsagn slettet, dersom de ønsker dette. Dette skal gi deltakerne noe innflytelse, dersom de mener at et utsagn kan lett feiltolkes-.

Som kvalitativ forsker vil man få tak i deltakernes oppfatning av virkeligheten (Nilssen, 2012, s. 25). Å tilbakeholde informasjon eller, som et enda mer ekstremt eksempel, å forske på mennesker som ikke er seg bevisst på at de er med i en kvalitativ studie, kan muligens resultere i mer kunnskap (Kvale & Brinkmann, 2009, s. 89-90), men går på sterk bekostning av etikk. Derfor er det et krav om at deltakerne vet hva forskningens formål er, og hvilke rettigheter de besitter, før de innvilger til frivillig deltakelse (King & Horrocks, 2010, s. 110; Nilssen, 2012, s. 145). Dette prinsippet kalles for informert

samtykke. Det kan tilføyes at samtykke er en pågående prosess, som medfører at deltakerne alltid har rettigheten til å trekke seg fra prosjektet, selv etter intervjuene er gjennomført (King & Horrocks, 2010, s. 115).

For å tilfredsstille kravet om informert samtykke, ble det utarbeidet et skriftlig samtykkeskjema ut ifra en mal fra NSD sitt nettsted (se vedlegg). Deltakerne ble også informert muntlig om deres rettigheter og studiens formål. Samtykke er en pågående prosess. Derfor fikk elevene også informasjon om at det er mulig å trekke seg i etterkant av intervjuene. Det ble sagt at dette kunne gjøres ved å si ifra til læreren, eller kontakte meg eller min veileder. I tillegg fikk de informasjon om at de kunne få innsikt om alle lagrete personopplysninger. Siden deltakere var elever under 16 år, ble de også bedt om å ta hjem et samtykkeskjema til sine foresatte. For å unngå mulige misforståelser av prosjektets formål var det også mulig å få innsikt i intervjuguiden og oppgavene på forhånd.

Dette prosjektet er vurdert av Norsk senter for forskningsdata AS (NSD) og behandlingen av personopplysninger står i samsvar med personregelverket.

### **3.10 Begrensninger**

En begrensning til utvalg av deltakere til fokusgruppene var at deltakerne kjente hverandre fra før av, som endrer måten deltakerne påvirker hverandre (Patton, 2002, s. 387). For eksempel kan én elev bli ansett som dyktigere i naturfag og dermed godtar resten av gruppen elevens mening uten diskusjon. Dette har negative konsekvenser for datakvaliteten, siden diskusjon mellom deltakerne gir bedre kvalitet på data (King & Horrocks, 2010; Patton, 2002, s. 386). I tillegg kan deltakere, som anser sin mening til å være i mindretall, velge å ikke dele den med andre pga. frykt av negative konsekvenser (Patton, 2002, s. 387). Dette kan muligvis intensiveres hvis deltakerne kjenner, siden disse mulige konsekvensene kan kjennes over et lengre tidsrom. Derfor ble det forsøkt å gjenta resoneringene til elevene med litt forandret ordbruk, for at elevene skal kunne få muligheten til å motsi intervjueren, som også kan øke datakvaliteten (Kvale & Brinkmann, 2009), og ikke hverandre.

En annen utfordring med oppsettet til fokusgruppene var at oppgavejobbingen tok rundt én times tid, uten introduksjon og debriefing. Å holde på så lenge uten pauser, i en forholdsvis fremmed omgivelse, kan være utmattende for deltakerne, særlig i løpet av de

siste deloppgavene. Dette kan ha hatt innflytelse på hvordan elevene anvendte sin kunnskap i at, selv om elevene hadde forestillinger om et konstrukt, valgte de ikke å anvende dette, siden de krevde mer kognitive ressurser enn det de var villige å bruke på dette tidspunktet.

Som tidligere nevnt, kan forskerens tilstedeværelse ha innflytelse på deltakernes oppførsel, beskrevet av Nilssen (2012, s. 31-33) og Serder (2015, s. 80-82), der sistnevnte beskriver denne påvirkningen i konteksten av oppgavejobbing. I tillegg til forskeren, kunne deltakerne selv påvirke hverandre. Dette innebærer at noen deltakere, som ikke hadde noen konkrete løsninger for eller tanker om oppgaver, kunne ha blitt påvirket av deltakere som gjorde, og kunne overta deres språk fremfor deres egen. På lik måte kunne en omvendt innflytelse kunne ha funnet sted, altså deltakere med tanker om en oppgave kunne ha blitt av deltakere som ikke gjorde det.

## 4 Funn

Ut ifra dette ble det etablert 9 kategorier, der 4 av dem er satt i en modell, som skal beskrive elevenes tenkemåte når de jobbet med oppgaver innenfor oppgavesettet «Løping i varmt vær». De resterende kategoriene dreiet seg om elevenes tenkemåter i forbindelse med hvordan de tolket begrepet «årsak», og når de vurderte data og et forskningsdesign i oppgavesettet «Undersøkelse av dalsider». Kategoriene er oppført nedenfor:

- Elevenes epistemiske løsningsstrategi
  - Simulerer prøvende
  - Simulerer målrettet
  - Tolkning av datasett
  - Tester egen tolkning
  - Konsistensvurdering
- Underliggende sammenhenger
  - Korrelasjon med tidslinje som årsak
  - Årsaksforklaring krever teoretisk innsikt
- Epistemiske oppfatninger
  - Måleinstrument viser enten rett eller feil
  - Usikkerhet som sikkerhet innen et intervall
  - Usikkerhet som variasjon
  - Flere uavhengige målinger minker usikkerhet

Det blir gitt en grundigere forklaring av kategoriene når funnene oppføres.

### 4.1 Kategorier innenfor elevens epistemiske løsningsstrategi

#### 4.1.1 Simulerer prøvende

Noen fokusgrupper valgte å sette inn forskjellige verdier i simuleringen, som ikke var oppgitt i oppgaven og så deretter å kjøre simulasjonen uten en målrettet løsningsstrategi for oppgavene. Simuleringen skal da føre til datasett som sammenliknes med svarforslagene.

Elever prøver å få til et svar som passer med svaralternativene fremfor å finne en strategi som vil gi dem et rett svar.

	PERSON	YTRING
1	B1	Og så prøver vi på..
2	B2	Det er jo kanskje den som er uten [uforståelig]
3	B3	Kan jo stille på 20, så 25 og så videre.
4	B1	Trykket du på kjør?
5	B2	Åja, det blir 30 da.
6	B3	Du kan jo se på 35 også, da.
7		[...]
8	B2	Mh, da må vi ta heteslaget. Så må vi ta den, og det under heteslaget, som er 35.
9	B1	Ja.
10	B2	35 var det høyeste, for når vi kom opp på 40 grader, så fikk personen heteslag?

Samtale 1: Fra fokusgruppe B i oppgave 4, oppgavesett "Løping i varmt vær"

I linje 1 foreslår elev B1 å prøve seg frem til de får et brukbart resultat. Elev B2 viderefører denne tanken i linje 3 ved å foreslå å stille lufttemperaturen inn på det laveste, 20° C, og deretter kjøre simulasjonen med hver mulig innstilling for lufttemperatur. Dette forslag blir gjennomført og man oppdager at ved 40° C lufttemperatur utsettes løperen for heteslag. Ut ifra dette konkluderer elevene at 35° C er den høyeste lufttemperaturen løperen kan ha før den utsettes for heteslag. Når elevene ble spurt om deres fremgangsmåte i denne oppgaven, var dette responsen:

	PERSON	YTRING
1	I	Kunne du ha gitt en kort oppsummering av, eh, det oppgaven vil at dere skal gjøre for å komme frem til svar?
2	B3	Ok. Oppgaven ville at vi skulle.. ville vi skulle finne ut hva det var.. Hva det høyeste lufttemperaturen vi kunne løpe i, uten å få heteslag.
3	I	Mhm.
4	B3	Og så vil den at vi skal forklare, og at, liksom, hva støttedata vi har funnet.
5	I	Mhm. Og hvordan finner dere dataene, som-
6	B3	Vi kjørte simuleringen for å se ... se hvilke resultat vi fikk.

Samtale 2: Fra fokusgruppe B i oppgave 4, oppgavesett "Løping i varmt vær"

I linje 2 og 4 svarer elev B3 på spørsmålet om løsningsmetoden ved å repetere spørsmålet til oppgaven og avbryter en forenkling av spørsmålet, der intervjueren spurte om hvordan de finner støttende data ved å si at de bare kjørte simulasjonen. Dette tyder på mangelen av en konkret løsningsstrategi, men det må bemerkes at spørsmålene var forholdsvis tungvint formulert. Å spørre om hva elevene gjorde for å komme frem til svaret fremfor hva oppgaven ville de skulle, kunne ha vært en bedre fremgangsmåte her.

	PERSON	YTRING
1	A3	[leser] 20, 40 eller 60...
2	A1	[leser] Og lufttemperaturen på 40.
3	A2	Luftfuktigheten er 50 %.
4	A1	Og så må vi bruke 20, 40 og 60 % for å prøve å komme frem til 50 eller om det fortsetter.

Samtale 3: Fra fokusgruppe A i oppgave 5, oppgavesett "Løping i varmt vær"

I linje 1 og 2 repeterer elev A3 og elev A1 repeterer hva oppgaven oppgir som faste innstillinger til simuleringen, og hva som må varieres. Elev A2 sier i linje 3 hva som er målet til oppgaven, nemlig å få informasjon om luftfuktigheten på 50 %. Til slutt foreslår elev A1 i linje 4 å prøve seg frem med alle mulige innstillinger til de får brukbar data til å kunne si noe om hva som skjer ved 50 % luftfuktighet. Denne episoden er også kjennetegnet av at elevene skal prøve å komme frem til femti, fremfor å utvikle en konkret strategi, der de velger data som konkret gir informasjon, som kan tyde på hva som skjer ved 50 % luftfuktighet.

#### 4.1.2 Simulerer målrettet

I noen oppgaver hadde fokusgruppene en konkret strategi for å komme frem til datasett, som de viste kommer til å gi informasjon til å formulere et svar. Denne kategorien ble delt inn i to underkategorier, siden elevene hadde to grunner til å drive med målrettet simulering. Den første grunnen tar utgangspunkt i forestillinger elevene hadde på forhånd om at bestemte innstillinger gir data som vil grunnlag for et svar. Den andre grunnen var at, etter å ha tolket dataene, elevene utviklet en interpretasjon av dataene, som de ønsket å bekrefte.

	PERSON	YTRING
1	A2	Ah, ok. Luftfuktighet på 60..
2	A1	Oooi.
3	A2	Men lufttemperaturen bare økes.. eller?
4	A3	Den kan økes og minkes. Den må tas først på det laveste og så på det høyeste, eller?

Samtale 4: Fra fokusgruppe A i oppgave 3, oppgavesett "Løping i varmt vær"

I linje 1 foreslår elev A2 først å stille luftfuktigheten på den verdien oppgaven oppgir, nemlig 60 %, og sette inn forskjellige lufttemperaturer i linje 3. I linje 4 kommer elev A3 med et mer konkret forslag om å ta den høyeste og den minste verdien til lufttemperaturen for å kunne se den størst mulige forskjellen i variablene, som vil mest



sannsynlig gi et godt synlig resultat. Etter å ha spurt konkret hva elevenes løsningsmetode var for denne oppgaven, svarte de følgende:

	PERSON	YTRING
1	I	Kunne dere gitt en kort oppsummering av hvordan dere gikk frem her?
2	A2	A3, vil du ta den?
3	A3	Nei, det er jo-
4	A1	Vi tok- altså en informasjon vi fikk i oppgaven, altså at lufttemperaturen skulle øke og hva luftfuktigheten var. Og satt det inn og kjørte to forskjellige resultater med mod forskjellig temperatur, og vi fikk.. to.. forskjellige data i tabellen, som viste at svettemengde økte.
5	I	Altså hvorfor akkurat bare de to?
6	A1	Vi tok jo flere, men det var to vi skulle velge. Så derfor tok vi jo flere for å støtte svaret enda mer.

*Samtale 5: Fra fokusgruppe A i oppgave 3, oppgavesett "Løping i varmt vær"*

Elev A1 bekrefter i linje 4 at utgangspunktet var å ta to forskjellige verdier for lufttemperaturen. Dette henviser til en konkret plan for å finne ut om det er en forskjell i svettemengde, når man endrer på lufttemperaturen. I linje 6 sier hen at de simulerte flere enn to datasett, fordi de ønsket deres antakelse bekreftet. Dette kan da tydes som et eksempel for den andre motivasjonen elevene så ut til å ha for å drive med målrettet data, nemlig å bekrefte sin interpretasjon av data.

	PERSON	YTRING
1	D2	Ok. Så da må vi ta forskjellige.
2	D1	Ta bare-
3	D2	-med.. ta forskjellige temperaturer.
4	D4	Ta 20, 30, 40, kanskje.
5	D2	Ja.
6	D1	Kjør på med alle, sikkert. Jeg vet ikke.
7	D2	Vi tar.. hvis vi tar 20, 30, 40, så har vi forskjellig.

*Samtale 6: Fra fokusgruppe D i oppgave 3, oppgavesett "Løping i varmt vær"*

I linje 1 og 3 sier elev D2 om at man må ha med forskjellige temperaturer fremfor å prøve forskjellige temperaturer med samme innstilling. Dette kan tydes i to forskjellige retninger: Enten blir oppgaven interpretert slik at man må prøve seg frem til data som kan interpreteres i lys av en av de to hypotesene, eller eleven så at man måtte ha mer enn to forskjellige temperaturer for å kunne se en forskjell bedre. Følgende replikk kan gi avklaring om elevene tok et bevisst valg fremfor å prøve seg frem på oppgavens instruksjer:

	PERSON	YTRING
1	D3	Da kan vi ta den første og den siste. [20 og 40]
2	I	Hvorfor den første og den siste?
3	D3	Fordi det er mest forskjell mellom dem. Så du ser mer tydelig hvordan det støtter svaret vårt.
4	D1	Ser stor økning i svettemengden.
5	I	Altså det hadde ikke vært lurt å ta to temperaturer som ligger rett ve siden av hverandre, f.eks. 20 og 25?
6	D1	Du ser jo det, men-
7	D2	Du ser forskjellen, men det- du ser jo større økning i jo større lufttemperaturen nå blir. Da øker kroppstemperaturen òg.
8	I	Mhm. Altså dere ville ha sagt at det er bedre hvis vi har stor forskjell i temperaturer?
9	D2	For å få frem forskjellen, så.. så du ser du jo det, at jo større forskjell i temperaturen, jo større forskjell blir det jo på svettemengden.

Samtale 7: Fra fokusgruppe D i oppgave 3, oppgavesett "Løping i varmt vær"

I linje valgte elevene datasettene som viste størst forskjell mellom både lufttemperatur og svettemengde, og begrunnet dette med at dette gjør forskjellen tydeligere og dermed kvaliteten på argumentet bedre, i linje 3, 4, 7 og 9. Dette gjenspeiler seg i elevenes løsningsmåte, i at de aktivt valgte temperaturer som var 10° C fra hverandre, fremfor å ta temperaturer som er 5° C fra hverandre, som simulasjonen også tillater. Dermed kan man konkludere med at gruppen hadde en målrettet plan for å løse denne oppgaven fremfor å prøve seg frem. Den neste dialogen er fra oppgave 4, der alle gruppene drev med prøvende planlegging, som vist i Samtale 1. Slik så da en målrettet strategi ut:

	PERSON	YTRING
1	D1	Skal vi bare-
2	D2	Vi kan jo bare sjekke om det er-
3	D1	Alle sammen?
4	D2	Til du ser-
5	D1	At det går over?

Samtale 8: Fra fokusgruppe D i oppgave 4, oppgavesett "Løping i varmt vær"

Denne dialogen mellom elev D1 og elev D2 står i kontrast til den prøvende planleggingen i Samtale 1. Planleggingen til gruppe B var preget av at de skulle prøve seg frem og mente selv at de ikke hadde en konkret strategi. I denne dialogen derimot formuleres det tentativt av elev D1 og D2 en strategi, der de kjører alle simuleringen til løperen utsettes for en helsefare, som da vil gi elevene et grunnlag for svar.

### 4.1.3 Tolkning av datasett

Det bemerkes her at alle episodene er funnet hos grupper som har kjørt simulasjonen målrettet. Deretter foregikk det en generell tolkning av data uten en konkret henvisning til de foreslåtte hypotesene. Når elevene senere vurderer konsistensen til data med hensyn til en hypotese, krever dette tolkning av data, men disse tolkningene er sterkt tilknyttet til hypotesen fremfor at elevene selv utarbeider en tolkning.

	PERSON	YTRING
1	C2	Skal vi ser forskjellen da... Ja, vi ser jo at, vi ser jo at det er forskjell på væsketap, at det eneste forskjellen med å drikke vann, da. Så når man drikker vann, så får man ikke væsketap, og når man ikke drikker vann, så får man væsketap.
2	C3	Ja.
3	C2	Ja, kroppstemperaturen.. er like.
4	C1	Også drikke forandres.
5	C2	Siden det blir..
6	C3	Ved å drikke så unngår han..
7	C3	At det reduserer risikoen for dehydrering, men ikke for heteslag.

Samtale 9: Fra fokusgruppe C i oppgave 1, oppgavesett "Løping i varmt vær"

I linje 1 og 3 tolker elev C2 data med hensyn på forskjellene som oppstår av å drikke vann, nemlig at væsketapet blir påvirket av dette, mens kroppstemperaturen forblir konstant. Elev C1 tilfører dette i linje 4 med at også inntak av vann forandres, fra «Nei» til «Ja». Fokuset til elevene ligger på selve datasettene og de forsøker å danne en egen tolkning fra disse. En støtte til denne tolkningen er utsagnet til elev C1 i linje 4, siden inntaket av vannet blir allerede forandret i selve simuleringen. Likevel velger denne eleven kun å se på dataene til å etablere en egen interpretasjon. I linje 7 er elev C1 allerede inne på å vurdere konsistensen med de foreslåtte hypotesene i at hen overtar ordene til en hypotese i sitt eget utsagn.

De neste to samtalene omhandler tolkning av data i oppgave 3 fra settet «Løping i varmt vær». Begge gruppene kommer frem til tolkninger som likner på hverandre, men forskjellen ligger i hvordan disse tolkninger er formulert og konsekvensene av denne, nemlig at den ene formulerer en sammenheng som en kausalitet, mens den andre kun ser på dette som en korrelasjon.

	PERSON	YTRING
1	D1 og D4	Svettemengden øker.
2	D3	Ja.
3	D2	Temperaturen øker svettemengden.

Samtale 10: Fra fokusgruppe D i oppgave 3, oppgavesett "Løping i varmt vær"

Her går elevene også bort fra formuleringen til de foreslåtte hypotesene, og sier at en økning i temperaturen medfører en økning i svettemengde. Elevene formulerer dette som en kausalitet, der temperaturen får en aktiv rolle i økningen av svettemengden, fremfor oppgavens formulering, der det sies at: «Svettemengden øker», (Kjærnsli & Jensen, 2016c) når det spørres etter hva som hender når lufttemperaturen øker. En mer passiv formulering av denne sammenhengen kan observeres i gruppe A:

	PERSON	YTRING
1	A1	Ja, per 5 grad.
2	A3	0,3 liter, ja.
3	A1	Tre desiliter ja. Per- hver gang vi øker temperaturen med 5 grader da.
4		[...]
5	A3	Ja, svettemengden-
6	A1	Og kroppstemperaturen òg endres.
7	A2	Ja, kroppstemperaturen øker.
8	A1	Med 0,5 grader.

Samtale 11: Fra fokusgruppe A i oppgave 3, oppgavesett "Løping i varmt vær"

Her foregår det en mattematisering, der elevene forsøker å etablere et konstant forhold mellom lufttemperatur og svettemengde. Blir dette forholdet formulert som en samvariasjon i linje 3, 6 og 7. Igjen, dette er løslatt fra de foreslåtte hypotesene og elevene forsøker å interpretere dataene selv uten en konkret henvisning til hypotesene.

	PERSON	YTRING
1	D1	Da er det 40 %.
2	D2	Nei, 35. 35 grader. Kan han gå uten å få-
3	D1	Ja, jeg tror det.
4	D2	Uten å få heteslag. Hvis han går over på 40, så.. så får han jo heteslag.
5	D1	Hvis vi tar de to, da.
6	D2 og D3	Mhm.
7	D2	Det forklarer jo ganske bra.

Samtale 12: Fra fokusgruppe D i oppgave 4, oppgavesett "Løping i varmt vær"

Dette er fortsettelsen til planleggingen vist i Samtale 8. Episoden viser hvordan tolkningen av data og konsistensvurderingen samspiller, når simuleringen er målrettet. Ut ifra dataene selv etablerer elevene i linje 1-4 at den høyeste ufarlige lufttemperaturen for løperen er 35° C, siden 40° C, fører til heteslag. I linje 5 knytter elev D1 denne tolkningen til hypotesen ved å si at det er disse to datasettene som er mest konsistente med hypotesen om av 35° C er den høyeste lufttemperaturen løperen kan løpe i.

#### 4.1.4 Tester egen tolkning

Etter å ha sett og tolket dataene, har elever en første tolkning over hva dataene viser. Noen elever ville bekrefte denne tolkningen på nytt. Så selv om elevene simulerer med konkret mål, som i forrige kategori, simulerer de for å teste en tolkning. De simulerer ikke, fordi de mener at dataene vil gi et svar.

	PERSON	YTRING
1	B3	Går det an å putte den på 50, mellom 60 og 40?
2	B2	Mh, nei. Men vi kan kjøre han én gang på <u>begge</u> .
3	B1	Der ... Det du ser du får heteslag.
4	B2	Men så-
5	B3	Tar på 20 òg, da. Så ser vi, liksom, om det sånn en konstant økning.

Samtale 13: Fra fokusgruppe B i oppgave 5, oppgavesett "Løping i varmt vær"

I linje 1 og 2 blir det planlagt strategien. Elevene etablerer at det ikke er mulig å innstille simulasjonen på 50 % usikkerhet, og deretter kjører simuleringen på kun 40 % og 60 % luftfuktighet for å finne ut om dette er helsefarlig for løperen. Etter å ha kjørt simulasjonen to ganger, ser elev B1 at dette samsvarer med hypotesen om at løperen utsettes for en helsefare, nemlig heteslag, men elev B3 krever i linje 5 å bekrefte hens antakelse om at kroppstemperaturen økes konstant, eller om dette kun var tilfellet for mellom 40 % og 60 %.

Elev A1 sier i Samtale 5 at hen også simulerte flere datasett enn det man trengte for å bekrefte hens tolkning om at en økning i lufttemperatur førte til en økning i svettemengde. Felles for begge samtalene er at elevene ser etter en bekreftelse for tolkningen sin, fremfor å forholde seg til oppgaven og finne ut at simuleringen gi data som er relevant for en svarformulering.

#### 4.1.5 Konsistensvurdering

Som tidligere nevnt, må det foregå i prinsippet en tolkning av data, hvis man skulle knytte dem til en foreslått hypotese, selv om dette ikke er synlig. Forskjellen mellom kategoriene «Tolkning av data» og denne ligger i egenheten av tolkningen, der konsistensvurderingen er tydeligere knyttet til en av oppgaven foreslått hypotese.

Følgende episode viser en diskurs om hvilke datasett støtter hypotesen om at å løpe under 50 % luftfuktighet er helsefarlig. Elevene har rask etablert at dette er tilfellet og deretter diskuterer en tolkning av data som støtter denne påstanden.

	PERSON	YTRING
1	C2	Ja, det er egentlig et godt spørsmål. Det er litt vanskelig å få, syns jeg, hvilke rader man skal bygge det på.
2	C1	Den tenker jeg vi må ha.
3	C3	Ja, i hvert fall.. I hvert fall den. Der er jo der, der den personen har fått heteslag. Jeg tenker vi burde ta den nederste for å vise at, liksom, personen ikke har heteslag, så at man kan se, liksom.. at personen ikke har heteslag.
4	C1	[utydelig tale] det er jo 50 % de spør etter.
5	C2	Vi må få vist at, em, kroppstemperaturen øker når luftfuktigheten øker. Det må vi få vist, fordi det hadde ikke holdt med én, for da kunne jo man bare tenkt at: «Ja, hvis luftfuktigheten går opp ...» Sant, vi må jo kunne bevise at da går også kroppstemperaturen opp.
6	C3	Ja.
7	C2	Sånn jeg tror at det har ikke så mye å si hvilken av de to vi underbygger det med.
8	C3	Jeg tenker med å ha den nederste, sånn at man kan se at personen har heteslag og at personen ikke har. Men det er bare min mening.

Samtale 14: Fra fokusgruppe C i oppgave 5, oppgavesett "Løping i varmt vær"

Karakteristisk for dette er at alle replikkene er drevet av påstanden om at å løpe under 50 % luftfuktighet er helsefarlig. Elev C2 mener i linje 5 at det er sentralt med å vise en økning og ha med 40 % luftfuktighet, der det er irrelevant hvilke andre datasettene som brukes, ifølge det eleven mener i linje 7. Elev C3 vil, i linje 3 og 8, påvise denne økningen ved å velge 20 % og 40 % luftfuktighet, siden å gå fra ikke helsefarlig til helsefarlig viser en tydelig økning.

Neste samtale er hentet fra gruppe B, som jobbet med oppgave 4 fra datasettet «Løping i varmt vær».

	PERSON	YTRING
1	B1	Det uttrykket, hva skal vi gjøre med det? 35.
2		[velger datasett]
3	B3	Men det var <u>med</u> å drikke.
4	B2	Ja, men heteslaget hadde jo ingenting med-
5	B3	Ok.
6		[...]
7	B2	35 var det høyeste, for når vi kom på 40 grader, så fikk personen heteslag?
8	B1	Ja.

Samtale 15: Fra fokusgruppe B i oppgave 4, oppgavesett "Løping i varmt vær"

Her er det også forholdsvis raskt etablert at 35° C er den høyeste ufarlige lufttemperaturen. Diskusjonen er ledet av dette og elevene forsøker å finne ut hvilke data som underbygger dette best. Elev B3 tviler i linje 3 på gyldigheten til datasettene, siden

det var mer enn én variabel ble endret, men blir forklart at dette er irrelevant for svaret. I linje 7 foreslår elev B2 å ta datasettene med 35° C og 40° C. Dette står i kontrast til Samtale 8 og Samtale 12, der elevene visste hva de skulle lete etter, nemlig når temperaturen gikk over heteslagsgrensen og støttende data er rask etablert.

Felles for begge episodene er at begge utspringer av at planleggingen var prøvende. Dette kommer til uttrykk her at det er tvil om at det foregår en diskusjon hvilke datasett som støtter deres valgte hypotese best. Som kontrast er det en episode fra gruppe D:

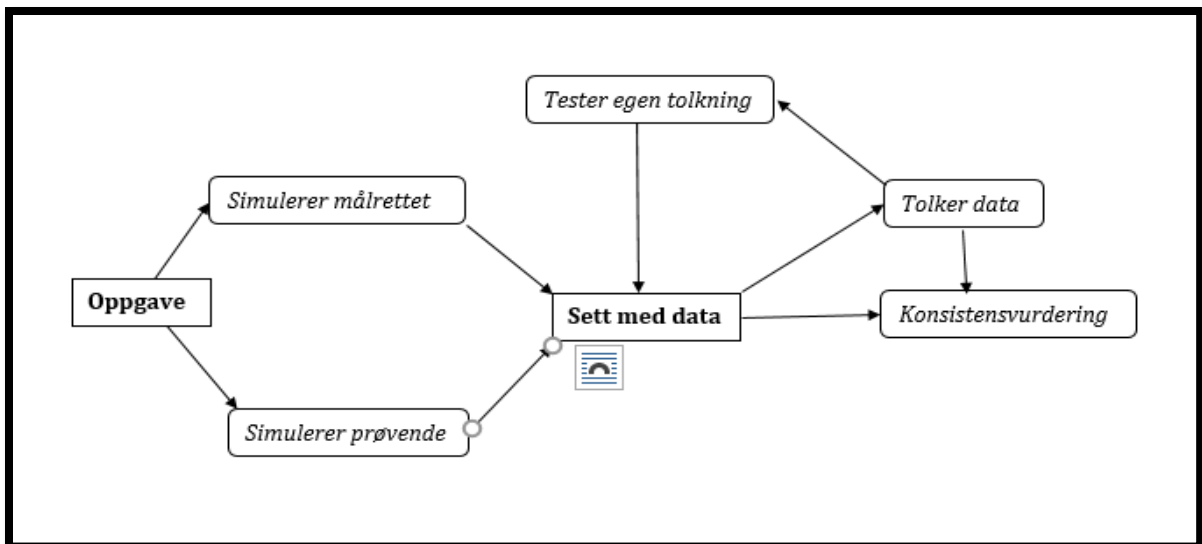
	PERSON	YTRING
1	D2	Da må vi ta.. én av hver av dem [40 % og 50 % luftfuktighet]
2	D1	Den her? [peker på en hypotese]
3	D3	Jo, fordi..
4	D1	Ja ... Ja.
5	D4	[Utydelig tale]
6	D2	Bare skriv, én av dere.
7	D1	Jeg vet ikke hvordan jeg skal skrive det.
8	I	Men kunne du ha forklart det da, hvordan det støtter, mens han skriver.
9	D4	Ja, fordi når du tar en med 40 og en med 60, begge to er over heteslagsgrensen. Og hvis du da.. for du kan ikke velge 50, men vi ser det er imellom, så er 50 fortsatt over heteslagsgrensen.

*Samtale 16: Fra fokusgruppe D i oppgave 5, oppgavesett "Løping i varmt vær"*

Dette var hele løsningsprosessen for oppgave 5 fra settet «Løping i varmt vær». Dette er et tilfelle, der elever simulerer målrettet, men ikke tolker dataene synlig for intervjueren. Elevene velger kun to rader, så spørsmålet over hvilke data som støtter hypotesen stiller seg ikke for elevene, men var klart på forhånd. Så konsistensvurderingen begynner i linje 2 av elev D1, der han spør om dette er hypotesen, som samsvarer med observasjonen de gjorde på forhånd. I linje 10 utdyper elev D4 tydeligere konsistensen mellom hypotesen og de utvalgte datasettene.

#### **4.1.6 Modell av elevenes epistemiske løsningsstrategi**

Følgende illustrasjon beskriver tenkemåtene elevene brukte for å komme frem til ett svar i oppgavesettet «Løping i varmt vær».



Figur 9: Oversikt over gruppens løsningsstrategi for oppgavesettet "Løping i varmt vær".

Bokser i en ramme med faste kanter og med bokstaver i tjukke font illustrerer hva elevene ser på skjermen når de jobber en med oppgave fra settet «Løping i varmt vær». De skal gi referansepunkter, som er utgangspunkt for elevenes aktiviteter.

Hva skjermen viser	Bekrivelse
<b>Oppgave</b>	Elevene viser oppgaven, som elevene leser i begynnelsen.
<b>Sett med data</b>	Etter å ha kjørt simuleringen noen ganger har elevene viser skjermen datasett som elevene skal interpretere.

Tabell 5: Oversikt over det skjermen i oppgavesettet "Løping i varmt vær" viser.

Bokser i rammer med avrundete kanter og med bokstaver i kursiv font er strategier som elever gjennomfører for å nå stadiene. Pilene viser hvilket stadium prosessene leder til. En beskrivelse av disse er gitt i tabellen nedenfor.

Prosess	Beskrivelse
<i>Simulerer målrettet</i>	Elevene har en konkret plan om hvilke simuleringer gir relevante data, som muliggjør egen tolkning.
<i>Simulerer prøvende</i>	Elever setter verdier, gitt fra oppgaveteksten, i simuleringen og varier kun verdien, som oppgaven ikke sa noe om.
<i>Tolker data</i>	Elevene tolker dataene til å komme med egne utsagn om hva de viser.
<i>Tester egen tolkning</i>	Elevene har kommet frem til en tolkning av data og ønsker derfor å teste denne tolkningen ved å kjøre simuleringen på nytt, men med andre innstillinger.
<i>Konsistensvurdering</i>	Elevene vurderer hvordan dataene passer med svarforslagene.

Tabell 6: Oversikt over prosessene, som inngår i elevenes løsningsstrategier når de jobbet med oppgavesettet "Løping i varmt vær".



Grunnen til at «Simulerer målrettet», «Tolker data» og «Tester egen tolkning» ligger i øvre del av Figur 9 er i at elever som tolket dataene til å komme frem til en egen interpretasjon, drev med en målrettet simulering til å komme frem til datasett. Elever som derimot simulerte prøvende, vurderte dataene direkte med hensyn på svaralternativene.

Ut ifra en helhetlig analyse av dataene og kategorier antyder det seg tre løsningsstrategier. Den første løsningsmåten (L1) tar utgangspunkt i at elevene simulerer prøvende. Språket er preget av begrepet «prøver», og ser hva resultatene blir. Når datasettene vises, pendler elevene frem og tilbake mellom svaralternativene og svaralternativene i oppgaven. Interessant for elevene er visuelt oppsiktsvekkende data, altså når løperen utsettes for en helsefare, som blir markert av oppgaven selv i at en rød farge markerer en helsefare. Når et svarforslag er valgt, velges to datasett, som samsvarer med dette svarforslaget. Et eksempel på denne er Samtale 15. Det blir spurt hva som skal man gjøre med svaret og det blir snakket om hvordan datasettene med 35° C og 40° C samsvarer med svarforslaget «35° C».

I den andre løsningsmåten (L2) simulerer elevene målrettet, altså de gir uttrykk for bevissthet om hvilke datasett man er ute etter. Deretter vurderer man konsistensen på et svarforslag med målesettene, og kommer raskt til en konklusjon. Grunnen til dette er at elevene allerede vet hva på forhånd hvilke datasett oppgaven etterspør. Dette vises i Samtale 16, der elevene begynte på simuleringen i linje 1 kom frem til svaret allerede i linje 4, uten lang diskusjon.

Den tredje, og siste, løsningsveien (L3), tar også utgangspunkt i at elevene simulerer målrettet. Her er man ute etter et bestemt forhold mellom to variabler, istedenfor bestemte datasett. Når datasettene kommer opp på skjermen, tolker elevene dataene selv, uten at de konkret henviser til et svarforslag. Dette vises Samtale 4, der elevene velger variabler av samme enhet som er lengst fra hverandre til å se på forholdet mellom lufttemperatur og svettmengde. Når datasettene er tolket, tar elevene et valg. Enten kjøres simuleringen igjen, men med andre innstillinger for å teste om tolkningen er rett, eller man vurderer konsistensen med et svarforslag. Her vurderes konsistensen mellom elevenes tolkning av datasettene og svarforslagene, istedenfor at man pendler mellom data og svarforslag. Å velge to datasett viser seg til å være ganske enkelt, siden alle bekrefter denne tolkningen, men elevene tenderer til å velge datasett som viser

tolkningen best, altså datasett med en tydelig forskjell, som elevene i Samtale 3 og Samtale 7 påpeker.

## 4.2 Underliggende sammenhenger

I oppgave 3 fra settet «Løping i varmt vær» blir elevene spurt om å etablere en biologisk årsakssammenheng mellom lufttemperatur og svettmengde. Dette krever at elevene har teoretisk kunnskap om kroppens indre prosesser, særlig at kroppens avkjølingsmekanisme er svetting. De neste to kategoriene dreier seg om hvordan elevene etablerte denne årsakssammenhengen.

### 4.2.1 Elever bruker teoretisk kunnskap

I oppgave 3 av settet «Løping i varmt vær» ble elevene bedt om å knytte en biologisk årsak til at svettmengden øker når lufttemperaturen øker. Denne og den neste kategorien behandler hvordan elevene svarte på dette spørsmålet.

	PERSON	YTRING
1	A2	Ja. Så viser at han går oppover.. «Biologiske årsaken»..
2	A3	Hmm.
3	A2	Fordi.. nei.
4	A1	Altså menneskets måte å kjøle seg ned på er ikke gjennom å svette da?
5	A3	Er det kanskje. Det er det.
6	A1	Og hvis det blir varmere, så svetter vi sikker mer? [utydelig tale]
7	A2	Jo.
8	A3	Ja, det er sant. Ja, der blir varmere og man svetter mer for man blir varmere. Og det er jo den biologiske årsak. Det er jo at det blir varmere, eller ikke? Er det ikke det?

*Samtale 17: Fra fokusgruppe A i oppgave 3, oppgavesett "Løping i varmt vær"*

I linje 1 gjentar elev A2 frasen «biologiske årsaken» fra oppgaven, og elev A1 foreslår i linje 4 at svetting er menneskets måte å kjøle seg ned på. Til slutt, i linje 8, forsøker elev A3 til å knytte de simulerte datasettene til teorien. Her viser elev A1 teoretisk kunnskap, når hen blir spurt om å gi en årsaksforklaring. Elev A3, som muligens kunne ha overtatt elev A1s teoretiske innslag i sin tenkning, som linje 5 antyder, forsøker å etablere en tidslinje. I denne tidslinjen trigger en høyere lufttemperatur kroppens behov for å kjøle seg ned. Denne slutningen kom også frem fra gruppe 2, der de også forsøkte å etablere en sammenheng mellom teorien og datasettene.

	PERSON	YTRING
1	I	Mhm. Har dere noen andre biologiske årsaker da?
2	D4	Tror ikke det.
3	D3	Jo raskere du løper, jo mer energi du bruker. Så, eh..

4	I	Energi, hvordan henger det sammen med svette?
5	D3	Det.. vet ikke hvordan..
6	D1	Du svette- jo mer energi du bruker jo mer vil du svette, liksom.

Samtale 18: Fra fokusgruppe D i oppgave 3, oppgavesett "Løping i varmt vær"

Denne gruppen kom først ikke på en teoretisk forklaring på den økende svettemengden, når de snakket om en biologisk årsak. Istedenfor gjentok de at lufttemperaturen øker svettemengden som en påstand om kausalitet. Likevel bruker elev D3 i linje 3 og elev D1 i linje 6 en teoretisk idé om energi til å gi en alternativ forklaring. Dette tyder på at elevene vet at en årsaksforklaring krever konseptuell forståelse. Her blir det også etablert en tidslinje, men elevene bruker et annet utgangspunkt, der elev D3 begynner å snakke om jo raskere man løper, og elev D1 fullfører denne tanken i linje 6 med at da bruker kroppen mer energi, som fører til mer svette.

Felles for begge replikkene er at de er sterkt tilknyttet elevenes hverdagslige erfaringer. I Samtale 17 blir det referert til at svetting er kroppens måte å kjøle seg ned på, som kan være hentet fra at man føler seg nedkjølt når man svetter. I Samtale 18 blir en økt svettemengde tilknyttet erfaringen om at når man løper lengre og/eller raskere svetter man mer, siden man bruker mer energi. En samtale, der man går bort fra personlige erfaringer ble funnet i fokusgruppe B.

	PERSON	YTRING
1	B2	[...] Den biologiske årsaken?
2	B1	Ja, som gjør at.. at man svette mer når det blir varmere?
3	B2	Kjøle seg ned.
4	B2	Ja,
5	B2	What you think?
6	B3	Akkurat det samme.
7		[...]
8	B3	Altså, bare.. la oss begynne på nytt da.
9	B2	Ja, men når temperaturen stiger-
10	B1	Så kvitter kroppen seg med væske og sånn-
11	B2	Så svetter man ut gjennom huden for å holde temperaturen mest mulig jevn.

Samtale 19: Fra fokusgruppe B i oppgave 3, oppgavesett "Løping i varmt vær"

I linje 3 foreslår elev B2 at kroppen kjøler seg ned som det teoretiske leddet i tidslinjen, introdusert av elev B1 i linje 2. Tidsforløpet er da følgende: Når det blir varmere vil kroppen kjøle seg ned og produsere svette. Dette viser elevene seg enige i linje 4-6. Dette er igjen sterkt tilknyttet til elevenes erfaringer om at svette medførte en nedkjøling. Senere blir et etablert en ny tidslinje, i linje 8-11, der elev B2 sier igjen at temperaturen stiger, men denne gangen fører dette til mer svette, siden kroppen ønsker å holde

kroppstemperaturen på et jevnt nivå. Denne årsaken er mer løslatt fra erfaringene, i at eleven impliserer at kroppen ønsker en konstant temperatur, fremfor å kun si at den ønsker å kjøles ned.

#### 4.2.2 Korrelasjon med tidslinje som årsak

Fokusgruppe C og D brukte en påstand om kausalitet mellom lufttemperatur og svette mengde når de ble spurt om å gi en biologisk årsak.

	PERSON	YTRING
1	C1	Du svetter jo automatisk mer når det blir varmere.
2	C2	Ja. Der er jo en sammenheng med kroppstemperaturen.. Ja.. Vanskelig å formulere det, kanskje.
3	C1	Ja.
4	I	Altså vanskeligheten er å formulere et svar eller å finne en biologisk årsak til dette?
5	C2	Altså jeg tror at vi tenker at årsaken er at det er varmere ute, så da blir kroppen varmere og da svetter man mer. Ehm. Jeg tror det er enklest å skrive det jeg sa, så.. Eller dere har jo innspill, vet ikke.
6	C1	Nei.
7	C3	Det hørtes rett ut, det du sa.

*Samtale 20: Fra fokusgruppe C i oppgave 3, oppgavesett "Løping i varmt vær"*

I linje 1 innleder elev C1 en påstand om at høyere lufttemperatur er årsaken til økt svette mengde. Elev C2 tar dette videre i linje 2 og 5, der hen forsøke å formulere et svar ut ifra denne påstand. Her blir det ikke trigget frem en konseptuell forståelse, men heller en påstand om kausalitet, som elevene etablerte tidligere. Denne påstanden følger en klar tidslinje, i linje 5, der elev C2 først sier at en økt lufttemperatur fører til en økt kroppstemperatur, som til slutt i en økt svette mengde, men uten et teoretisk element.

Den andre gruppen som oppga en korrelasjon med tidslinje som en årsak, var gruppe D.

	PERSON	YTRING
1	D1	«Hva er den biologiske årsaken til dette?»
2	D3	Jo høyere kroppstemperatur, jo mer må kroppen, eh, støtte ut vann?
3	D1	Ja.

*Samtale 21: Fra fokusgruppe D i oppgave 3, oppgavesett "Løping i varmt vær"*

I linje 1 repeterer elev D1 spørsmålet, og elev D2 kommer i linje med en forklaring, der hen beskriver korrelasjonen mellom lufttemperatur som en årsakssammenheng. Også her er det en klar tidslinje, der det sies at en økt kroppstemperatur fører til økt svette mengde, men uten et teoretisk innslag.

### 4.2.3 Årsaksforklaring krever teoretisk innsikt

Selv om noen grupper ga en årsaksforklaring uten teoretiske element, var de seg bevisst på teoriens rolle i en årsaksforklaring. I Samtale 18 viser gruppe D allerede implisitt om teoriens rolle, og i tillegg er det en mer eksplisitt formulering her:

	PERSON	YTRING
1	I	Mhm. Og.. det med biologisk årsak: Eh, da lurer jeg på hva dere legger i ordet «biologisk årsak», eller uttrykket?
2	D4	I dette tilfellet det som skjer med kroppen, og så det som naturen rundt han gjør da, siden det er så varmt. Så.. blir det jo- så skjer det jo.. så blir han jo varmere og da er det jo- det er noe som gjør at kroppen svetter jo varmere det blir, siden du.. siden han må utstøtte svette.
3		[...]
4	I	Kunne jeg da spurt dere, hvis vi er inne på ordet «årsak»: Kan dere gi meg en forskjell mellom ordet «årsak» og ordet «beskrivele»?
5	D4	«Årsak» er grunnen til noe, mens «beskrivelse», du forteller, liksom, hvordan det skjer.
6	D2	Hva som skjer, liksom.
7	I	OK. Og så, vil dere si da at det dere skrev her: «Det her er en årsak, ikke en beskrivelse»?
8	D1	Ja, det er en årsak, fordi det beskriver hvordan det blir sånn og ikke hva det blir.

Samtale 22: Fra fokusgruppe D i oppgave 3, oppgavesett "Løping i varmt vær"

På spørsmålet hva elevene legger i begrepet «biologisk» årsak, svarer elev D4 i linje 2 om at det må hende noe med kroppen, som sørger for en økt svette mengde. Dette tyder på en bevissthet rundt at det må finnes et teoretisk element i tidslinjen, mellom den økte temperaturen, og den økte svette mengden. I linje 4 og 7 blir gruppen spurt om hva en årsak er og om deres svar karakteriseres som årsak. Det ble valgt et ledende spørsmål i linje 7 for å gi elevene mulighet til å motsi og dermed være sikrere på mulige tolkninger. Svarene til elev D4 i linje 5, og elev D1 i linje 8 tyder på at deres forskjell mellom «årsak» og «beskrivelse» er tilstedeværelsen av en tidslinje, der årsaken følger denne tidslinjen, mens beskrivelsen foregår for en isolert hendelse. Dette blir tydelig i linje 8, siden elev D1 sier hvordan det blir istedenfor hva det blir.

	PERSON	YTRING
1	C1	Altså vi gir ikke så mye informasjon. Vi bare sier, liksom, at vi svetter mer pga. det blir varmere.
2	C2	Vi gjengir egentlig bare resultatene-
3	C1	Ja.
4	C2	-på det vi har kjørt- fått fra simulasjonen, da. Ja, vi har vel ikke så mye kunnskap om prosessene som skjer inn i kroppen når det blir varmere. Så det er kanskje derfor det blir vanskeligere å komme med en god, biologisk årsak på dette. Jeg vet ikke, men.

I denne dialogen sier elev C1, i linje 1, at de ikke gir mye ny informasjon, og elev C2 utdyper dette, i linje 2 og 4, at de trenger mer kunnskap for å gi en tilfredsstillende årsaksforklaring. Igjen, i linje 4 gjengir elev C2 en tidslinje, som begynner med en økt kroppstemperatur og at dette sørger for at en teoretisk effekt finner sted i kroppen.

### 4.3 Epistemiske oppfatninger

Opgavesettet «Undersøkelse av dalsider» gikk ut på å tolke et fiktivt forskningsdesign og deretter tolke data til å drøfte påstander om en årsakssammenheng. Dette gikk ut på å vurdere de epistemiske egenskapene til forskningsdesign og prinsippet om måleusikkerhet. De følgende kategoriene skal beskrive elevenes tenkemåte innenfor denne oppgaven.

#### 4.3.1 Måleinstrumenter viser enten rett eller feil

I den første oppgaven av oppgavesettet «Undersøkelse av dalsider» blir elevene bedt om å forklare hvorfor det blir satt opp to av hvert måleinstrument når man skulle undersøke mengden av sollys, nedbør og jordfuktighet i en dalside. En grunn som noen fokusgrupper nevnte var at noen av måleapparatene kunne gi data, som man ikke kunne stole på, mens noen formulerte dette enda klarere ved å si at de viser rett hvis de ikke er beheftet med feil. Begge formuleringer kan derfor ses på som uttrykk for at måleapparater måler enten rett, eller så fører ytre påvirkninger til en feil måling.

	PERSON	YTRING
1	A2	Hvis den andre hadde velte, sånn som, det ser ut som en kolbe her, ikke sant? Eller?
2	A1	Ja.
3	A2	Ja, så hadde den velte. Ikke sant? Det kan være feil, men..
4	A3	Ja, vi er alle enig her.
5	I	Ja, altså du ville ha sagt.. Ville det ha vært like bra med ett instrument, for eksempel?
6	A3	Det kunne vært litt.. Det kunne ha vært feil, fordi at.. La oss si at det har gått noen på tur der og helt masse vann ned i det
7	A1	Ja, eller bare kjøddet noe til eller gjort noe greier.
8	I	Men hvis det ikke hadde skjedd noe med målingene i det hele tatt, eller med instrumentene: Ville det ha vært like bra med ett instrument, eller?
9	A1	Ja. Målingene hadde vært like bra. Det er så klart hvis du- det er ikke om å måle med flest instrumenter. Hvis du setter det på riktig sted, og ingen kjødder det til, så skal de jo kunne få like målinger.

Denne diskursen handlet også om hvorfor det brukes to måleinstrumenter. Elev A2 sier i linje 1 at kolben kunne ha veltet, og elevene A1 og A3 foreslår at noen kunne sabotert måleren, i linje 6-7. Det blir kun gitt konkrete eksempler uten å knytte disse til hverken det vitenskapelige begrepet om usikkerhet eller feilkilde. Istedenfor beskriver elevene scenarier som er resultatet av ytre påvirkninger. Til slutt, i linje 8 og 9, blir elev A1 spurt om hva skjer hvis det ikke skjer noen uhell, og hen svarer med at dersom det ikke skjer noe med måleinstrumentene, vil de gi lik og dermed rett data. Dette tyder ikke på en oppfatning om at det finnes usikkerhet forstått som variasjon i måleverdier.

Disse forklaringene kan være uttrykk for et tankesett, der målingene anses generelt som sikre. Dersom det hender noen ytre uhell med apparatene, vil dette påvirke målingene og dermed ville måledataene blitt feil, men det blir sett på målinger som at de ikke besitter usikkerhet.

Også elev B3 viser tegn av denne tenkemåten. Hen sier at dersom ett måleinstrument gir en muligens upålitelig verdi, hadde man «back-up» for å se om dette stemmer. Måleinstrumentene besitter her en kontrollerende funksjon i at dersom begge viser samme verdi, måler de rett. Dette peker på samme grunnidé, nemlig at begge måleinstrumenter vil måle likt, dersom det ikke er skjedd en feil med ett instrument.

Funnene kan også tydes som at eleven mener at målinger er enten rett eller feil, og at flere måleinstrumenter kan gi avklaring over hva som kan anses som rett måledata. I motsetning til diskursen fra fokusgruppe A går denne eleven ikke ut ifra at ett godt oppsett kan føre til rette målinger uten krav for overprøving.

Neste replikk er hentet fra fokusgruppe D, der elev D1 snakker om hvorfor usikkerheten kan oppstå.

	<b>PERSON</b>	<b>YTRING</b>
<b>1</b>	D1	[...] Men jeg vet ikke hvordan de der vannmåletingene funker, men det kan jo være også at sola lyser på dem, at det damper opp da. Er det det som skaper usikkerheten da, eller er det, liksom...
<b>2</b>	I	Det kan være en mulig årsak. Har dere noen andre årsaker til at usikkerheten skjer?
<b>3</b>	D1	Nei, vi vet jo aldri hva som kan skje. Det kan være hvis det står åpent, så kan det jo kanskje være at noen dyr drikker fra det vannet.

*Samtale 25: Fra fokusgruppe D i oppgave 2, oppgavesett "Undersøkelse av dalsider"*

I samtalen knyttes ytre påvirkninger til usikkerhetsbegrepet, introdusert i oppgaven. Forskjellen mellom fokusgruppe A, B, og D er at gruppe A sier at disse feilene kan unngås,

og når dette skjer vil måleinstrumentene vise rett. Elev B3 påpekte at dersom alle måler likt, vil de måle rett. Gruppe D mener derimot at ute i naturen vil feil være umulig å unngå. Dette blir tydelig gjenspeilet i innspillet til elev D1 i linje 3 av Samtale 25, der hen sier: «Nei, vi vet jo aldri hva som kan skje.» Måleinstrumentet kan i seg selv måle rett, men blir påvirket av de ytre omstendighetene.

Felles for alle gjengitte samtaler i denne kategorien at det kun brukes konkrete eksempler når det er snakk om hvorfor man bruker flere måleinstrument og hva som er opphavet til usikkerhetsverdiene. Måleinstrumentene kan ha vært ødelagt, eller noe andre ytre påvirkninger som førte til at ett av måleinstrumentene viste feil tall fremfor at målingene selv blir betraktet som usikre. I tillegg beskriver elevene kun ytre påvirkninger som har innflytelse på målingene. Et begrep om usikkerhet er synligere i Samtale 25, siden de formulerer påvirkningene som uunngåelige, i motsetning til Samtale 24, der feil kan unngås. En mulig årsak til dette kan være at begrepet «usikkerhet» ikke ble introdusert før oppgave 2 i settet «Undersøkelse av dalsider».

#### 4.3.2 Usikkerhet som sikkerhet innen et intervall

Denne kategorien omhandler tenkemåter, som beskriver hvordan elevene så på usikkerhet som et intervall rundt en gjennomsnittsverdi. Dette intervallet kaller jeg for usikkerhetsintervall. I denne kategorien finnes det to nyanser: (1) Elever som ser på hele usikkerhetsintervallet som et sikkert intervall, og (2) elever som mener at det finnes én sikker verdi i intervallet. Felles for begge tenkemåter er at det finnes en eller annen form for sikkerhet i usikkerhetsintervallet.

Fokusgruppe B hadde en annen tilnærming når det gjaldt løsning av oppgave 2 i settet «Undersøkelse av dalsider». Fremfor å se på størst forskjell mellom verdiene til datasettene, så gruppen på usikkerheten til å bestemme hva som kunne ha vært årsaken, og kom frem til følgende slutning:

	PERSON	YTRING
1	B2	Ok. Kanskje trenger vi en helt ny tenkemåte. De to blir jo det samme, hvis man plusser de, sant? Fordi denne er 440 med 50 usikkerhet, og den andre er 450 med 40 usikkerhet.
2	B1	Ja.
3	B2	Men hvis vi ser på solstrålingen.. Så blir ikke de det samme, hvis vi plusser de. For det blir firetusen-ett og syvtusen-seks.
4	B1	Ja.



5	B2	Vet ikke. Det kan hende at vi er helt på feil spor. Så er det der greien med prosenten da.
---	----	--

*Samtale 26: Fra fokusgruppe B i oppgave 2, oppgavesett "Undersøkelse av dalsider"*

Elev B2 foreslår i linje 1 at man burde gå bort fra de forrige tenkemåtene. Hen ser at intervallene, som usikkerhetene danner rundt gjennomsnittsverdiene til nedbørmengden, overlapper. I linje 3 bemerker hen at dette ikke er tilfellet for solstrålingen. Likevel tviler hen på denne resonneringen. Denne resonneringen tar utgangspunkt i at usikkerhetsintervallet gir en viss sikkerhet, og når to intervall overlapper, vil måleinstrumentene muligens vise det samme.

Alle fokusgruppene ble spurt om på hvilken dalside det regnet mest på. Det bemerkes at før dette ble det snakket om usikkerhetsbegrepet og –angivelsene i tabellen. Dette kan har hatt innflytelse på hvordan elevene strukturerte svaret sitt. Likevel kan det muligens gi innsyn i hvordan elevene tenkte om usikkerhetsbegrepet. Alle gruppene brukte samme fremgangsmåte til å svare på dette spørsmålet,

	PERSON	YTRING
1	C2	Gjennomsnittet er høyere på A, men..
2	C3	Slingringsområdet er jo høyere-
3	C1	Men, eh-
4	C2	Men altså det høyeste kan være likt.
5	C3	Ja. For ingen av dem går over 490.
6	C2	Ja, det er det høyeste det kan være på begge.. Men den laveste er en forskjell på. Så det- på B, så er det laveste det kan være 390, mens på A er det laveste det kan være 410, så derfor blir gjennomsnittet litt høyere da. Så det er større sannsynlighet for at det kommer på A.

*Samtale 27: Fra fokusgruppe C i oppgave 2, oppgavesett "Undersøkelse av dalsider"*

Elev C2 sier i linje 3 at gjennomsnittet var høyere på A. Dette tyder på at *gjennomsnittsverdien anses som den verdien, som er mest riktig innenfor usikkerhetsintervallet*. Deretter henviser elev C2 og elev C3 i henholdsvis linje 4 og 5, at den øvre grensen for begge usikkerhetsintervallene er 490. Den nedre grensen er derimot, som elev C2 påpeker i linje 6, høyere på dalside A enn på dalside B. Derfor er det høyere sannsynlighet at det regnet mer på dalside B.

Fokusgruppene A, C og D antydte at gjennomsnittsverdien var den verdien som var mest rett. Én elev i hver av gruppene C og D brukte begrepene «sannsynlighet» og «sjanse» når de argumenterte for at det var mer nedbør på dalside A. Denne sannsynligheten oppstod gjennom at de lave tallene i usikkerhetsintervallet til nedbøren på dalside B ikke overlappet med det tilsvarende usikkerhetsintervallet på dalside A. Resten av elevene,

med to unntak som jeg kommer til å snakke om senere, formulerte dette uten bruk av ordet «sannsynlighet» eller «sjanse». Altså, siden den nedre grensen til intervallet på dalside A var mindre enn intervallet på dalside B, regnet det mest på dalside A<sup>2</sup>.

Alle disse tenkemåtene kan betraktes som former av nyanse (1) i denne kategorien. Gjennom at usikkerhetsintervallet anses som noen lunde sikkert, vil kun en fullstendig overlapp mellom to intervaller kunne anses som ett og det samme resultat. Siden intervallene ikke overlapper fullstendig, vil det alltid være rom for at verdiene henviser til forskjellige fenomener.

Den andre nyansen, (2), i denne tenkemåten vil jeg vise med følgende utsagn fra elev A1:

	PERSON	YTRING
1	A1	Ja, men du kan på en måte ikke bare plusse de, for det kan være -40 på den ene og +50 på den andre, så er du plutselig 80 ifra hverandre, ikke sant?
2	A3	Tror det regnet mest på A.
3	A1	Men det skal jo ha regnet mest på mest på A, sånn som oppgaven sier. Men samtidig, så kunne det vært. Det kunne gått begge veier, på en måte.

Samtale 28: Fra fokusgruppe A i oppgave 2, oppgavesett "Undersøkelse av dalsider"

Elev A1 nevner i linje 1 at det finnes en mulighet for at det kunne ha vært mye mer på den ene dalsiden og mye mindre på den andre, siden begge verdier ligger i de tilsvarende usikkerhetsintervallene. Denne tenkemåten kan oppsummeres gjennom når hen sier at: «Det kunne ha gått begge veier, på en måte.» I linje 2 sier hen at oppgaven sier at det regnet mest på A, som antyder at hen også primært anser gjennomsnittet som den muligens rette verdien. En slik tankegang ble også funnet i en replikk til elev D2.

Denne tenkemåten preget av at det *finnes én rett verdi innenfor usikkerhetsintervallet, og at denne er mest sannsynlig gjennomsnittsverdien*. Likevel finnes det en mulighet for at denne rette verdien befinner seg i usikkerhetsintervallet fremfor at det er gjennomsnittsverdien, som linje 1 i Samtale 28 antyder.

#### 4.3.3 Usikkerhet som variasjon i observert fenomen

Denne kategorien beskriver tenkningen til elevene om hvorfor usikkerhet oppstår. Noen elever peker på variasjoner i omgivelsene til måleinstrumentene, som f.eks. allerede antydnet i Samtale 25 til gruppe D. Kategorien står ikke i kontrast til at måleinstrumenter

<sup>2</sup> Hvem? Unntaket?

kan enten vise rett eller feil, men kan eksistere ved siden av denne. Jeg påstår dette, fordi at Gruppe A, i Samtale 24, og gruppe B antyder at måleinstrumentet enten viser rett eller feil, uten å tilføye at ytre variasjoner er uunngåelige og må tas hensyn til i målingene.

Elev B1 henviser til variasjoner i et observert fenomen uten å knytte dette til måleinstrumenter som viser enten feil eller rett. Der skiller hen seg fra elev D1, som også snakker om variasjoner i målinger, men mener at dette fører til at måleinstrumentet viser feil.

**TALENDE YTRING**

<b>1</b>	B1	Jeg må bare se. På bildet ser det som ut som de har plassert noen høyere og noen litt lavere. Kanskje for å se.. forskjeller.
<b>2</b>	I	Mhm. For å se forskjeller hvor? Hva mener du?
<b>3</b>	B1	Hvor mengde nedbør og sol som treffer på de forskjellige høydene.

*Samtale 29: Fra fokusgruppe B i oppgave 1, oppgavesett "Undersøkelse av dalsider"*

Elev B1 snakker om forskjeller, som må tas hensyn til i linje 1. Forskjellene er beskrevet i linje 3, der hen nevner mengde sollys og nedbør som treffer på de forskjellige høydene. Dette tyder på at eleven har forståelse for at målinger et fenomen, som foregår på et stort område, vil variere. Variasjonene oppstår, fordi et stort område vil ha forskjellige egenskaper i seg, som f.eks. høydeforskjeller.

Elev D2 og elev D2 viste den samme type forståelser for at målinger vil variere og at man må ta hensyn til dette.

**PERSON YTRING**

<b>1</b>	D4	For da ser du det hvordan det er på forskjellige plasser.
<b>2</b>	D2	Det er mer nøyaktig måling i, rundt det området den er satt ned da. Siden du skal ha et stort område, siden da må du ha flere måleinstrumenter for å kunne, ehm, få mest mulig nøyaktig måling over dette området.
<b>3</b>	D4	Også ser du, f. eks. med solmåleren hvor sola treffer hen. Da kan han treffe én, men ikke den andre da.

*Samtale 30: Fra fokusgruppe D i oppgave 1, oppgavesett "Undersøkelse av dalsider"*

Elevene anerkjenner også at det vil finnes forskjeller på forskjellige steder når man gjennomfører målinger. Elev D4 bruker et mer konkret språk i linje 1 og 3, der det er snakk om forskjellige plasser og hvordan dette kan f.eks. påvirke måling av sollys. Elev D2 snakker mer generelt i linje 2 ved å si at man må ha flere måleinstrumenter, for å oppnå et mest mulig nøyaktig bilde over området.

Elev D2 viser en enda tydeligere forståelse for at variasjoner oppstår i fenomener enn elev D4 og B1. Mens de to sistnevnte snakker om at man kan måle på forskjellige steder for å gjøre rede for variasjoner, snakker elev D2 om at man må gjøre dette for å oppnå en nøyaktig måling over området. Grunnen til dette ligger i at ett måleinstrument kan gi sikrere målinger over en del av det store området, men ikke kan ta hensyn til variasjonene over hele området. Tenkningen kan sies til å være preget av en oppfatning, der det man må ta hensyn til variasjoner i et fenomen, fremfor at man kan gjøre dette.

Det pekes på at alle replikker handler om usikkerhet som variasjon i et stort område, ute i naturen. Alle utsagn er preget av denne konteksten, uten at noen har en oppfatning om at usikkerhet er et kjennetegn ved all absolutt, som må tas hensyn til. Istedenfor ser man på variasjoner som automatisk oppstående ytre påvirkninger.

#### 4.3.4 Flere uavhengige målinger minker usikkerhet

Kategorien beskriver utsagn, som tar hensyn til at flere uavhengige målinger minker variasjon og usikkerhet til data. Her kravet for videre utforskning er sterkest, siden den ble kun funnet hos ett utsagn.

	PERSON	YTRING
1	A1	Da får du mye- det er som å si hvis du skal ta sånn, finne ut sjansen, eller sannsynligheten for noe. Hvis sjansen for sånn.. for å få Yatzy, så er det sånn én av 1000, så kan jeg ikke trille ganger og regne resultatet jeg ut av det, sant? Du må jo ta flere målinger for å komme nærmest mulig det riktige.. prosent, på en måte. Andelen, eller jeg vet ikke. Det spørres hva du måler, ikke sant? Måle flere og flere ganger for å komme nærmere det riktige svaret.

*Samtale 31: Fra fokusgruppe A i oppgave 1, oppgavesett "Undersøkelse av dalsider"*

Eleven sammenlikner valget for å ha med flere måleinstrumenter med å spille Yatzy, der målingene er uavhengige fra hverandre. Flere målinger fører til mindre usikkerhet og dermed en større sannsynlighet for å komme nærmere det rette svaret. Ved å bruke flere, uavhengige målinger økes nøyaktigheten, og dermed finnes det et smalere intervall for å finne det rette svaret. I kontekst av Samtale 24, der elev A1 også støtter oppfatningen om at måleinstrumenter kan enten vise feil eller rett, kan dette også oppfattes som at ytre påvirkninger, som forfalsker målingene, er uunngåelige. Likevel kan man nå en rett måling, dersom man måler ofte nok, så det kan argumenteres for at disse oppfatningene kan eksistere ved siden av hverandre, men at det mangles én konkret forestilling om konstruert usikkerhet i målinger.

### 4.3.5 De største numeriske forskjellene gir en årsak

Denne kategorien beskriver tenkemåten til gruppe A, C, og D når det de jobbet med oppgave 2 fra settet «Undersøkelse av dalsider». Elevene skulle bestemme over hva som var årsaken til at dalsidene hadde forskjellig jordfuktighet, og hadde valget mellom mengde solstråling og mengde nedbør. Følgende samtale viser tenkemåten til de tre gruppene nevnt ovenfor.

	PERSON	YTRING
1	C2	Ja, altså. ... Jeg mener at elev én har rett, fordi at det ikke er så veldig stor forskjell mellom nedbør; der er det nesten ingenting, og, eh, den jordfuktighet, det kan også ha litt å si, men det er størst utslag på solstråling, syns jeg da.
2	C1 og C3	Ja.
3	C2	Det er jo.. det er jo på en måte så mye, så.. ja. Også, for å ta med bildet, så ser det ut som.. nei, jeg vet ikke, nei. Men ut ifra tallene er det størst forskjell på solstrålingene. Så at.. ja. Planter der har lettere for å bli mer vissen, eller sånne ting da.
4		[...]
5	C2	Oi, nei vent! Altså jeg ser- ehm. Ja, siden.. Ja, ok, jeg tror jeg leste litt fort her. Siden.. Ja, jeg tenkte at de mente forskjellen i sånn frodighet eller sånn, på en måte; at det var det de skulle frem til, men- altså det er jo jordfuktigheten de skal fram til.
6	C3	Ja.
7	C2	Og jeg lest litt kjapt, så jeg tenkte at elev to [svaralternativ fra oppgaven] mente at det var forskjell på nedbør og jordfuktigheten som var årsaken, men.. ja.. jeg gjør ikke så stort utsalg da fortsatt- altså at det er relativt likt, det er sånn- det er bare 10 millimeter forskjell.

Samtale 32: Fra fokusgruppe C i oppgave 2, oppgavesett "Undersøkelse av dalsider"

I linje 1 og 2 sier elev C2 at det ikke stor forskjell i nedbørmengden på de to dalsidene, og størst forskjell i solstråling. Det nevnes også at det er stor forskjell i jordfuktigheten, som kan også være en mulig årsak, men hen påpeker i linje 5 og 7 at dette utsagnet oppstod på grunn av en lesefeil. Deretter korrigerer hen begrunnelsen i linje 7, der mengde solstråling er årsaken til forskjellen mellom jordfuktigheten i dalsidene, fordi det er større forskjell der enn hos mengde nedbør.

Fokusgruppe A og D brukte det samme argumentet, at forskjellen i mengde solstråling er større enn forskjellen i mengde nedbør. Gruppe A brukte også en matematisk metode ved å beregne forskjellen i både mengde nedbør og solstråling i prosent. Gruppe D så på usikkerhetsbegrepet, men kjernen i argumentet ble ikke påvirket av dette.

	PERSON	YTRING
1	I	Du løftet litt på skuldrene når jeg spurte om du- når jeg spurte om noen mente at elev to var rett.
2	D1	Men nei, det var fordi jeg ikke helt forstod det tegnet der, men så leste jeg på det og så skjønte jeg det.
3	I	Kan du gi en kort forklaring hva du skjønte av det?
4	D1	At det er, liksom, det er jeg i hvert fall tror at det tallet er som står etterpå det er det som er, liksom, usikkerhet. Det kan på en måte være så mye, eller at det er så mye mer eller så mye mindre, eller.. ja.
5	I	Altså kan du da gi et konkret eksempel med det, når du snakker om usikkerhet?
6	D1	Hæ?
7	I	Altså du snakket at det var så mye mer eller så mye mindre. Hva mente du med det?
8	D1	Nei, at jeg tror at det er, liksom, det er usikkerheten på hvor mye mer eller mye mindre det kan være. Så på den 450 kan det være alt mellom 410 til 450, eller noe sånt.
9	D2	Men, eh, du ser fortsatt at det er en mye mindre forskjell i nedbør på de forskjellige at det er mye mindre forskjell i nedbør på de forskjellige sidene, uansett med usikkerheten, enn det du ser på solstrålinga.
10	D1	Det var det jeg sa først.
11	D2	Men uansett. Mye, mye større forskjell mellom de her solstrålinga på de forskjellige områdene.

*Samtale 33: Fra fokusgruppe D i oppgave 2, oppgavesett "Undersøkelse av dalsider"*

Her ser elev D1 i linje 4 og 8 på usikkerhet som variasjonen til en gjennomsnittsverdi i at hen sier at den kan være alle mulige verdier i intervallet fra 410 til 450. Et slikt tankesett kan også antas for elev D2 i at hen ikke korrigerer elev D1 i hens tolkning av usikkerhet, men sier i linje 9 og 11 at dette ikke er relevant for svaret, siden forskjellen i solstrålingen er større uansett om man tar hensyn til usikkerhetene eller ikke.

Alle gjengitte samtaler antyder at elevene anså verdiene med størst differanse som årsak til andre verdier som oppstår. Verdier som er nærme hverandre, f.eks. 450 og 440 i oppgaven, blir likestilt gjennom en forestilling om at verdiene er uttrykk for at verdiene viser det samme. Det betyr ikke at elevene mener at verdiene er identiske. Det var kun elev D2 som mente at nedbøren er lik. Gruppe A og C mente at det var forskjell, men det hadde ikke noe å si for jordfuktigheten. Dette tyder på dette tankesettet om at stor forskjell er forbundet med en årsakssammenheng.

## 5 Diskusjon

Oppgavenes formål er å karakterisere elevenes praksis og tenkemåter når de jobber med PISA oppgaver. Denne karakteriseringen betraktes i lys av teori om NPT for å finne ut om elevene er i stand til å svare på frigitte oppgaver fra PISA-undersøkelsen. Innenfor NPT ble det identifisert tre kunnskapstyper, som resonneringen bygger på: innholds-, metode- og epistemisk kunnskap (Kind, 2013; Osborne, 2013). To modeller om vitenskapelig resonnering ble også belyst (Giere, 1991; Osborne, 2013). I tillegg ble det sett på elevers praksiser og tenkemåter i møte med praktiske naturfaglige oppgaver. Praksisene ble delt inn i en vitenskapelig, som hadde som mål å oppnå forståelse, og en praktisk, som hadde som mål å identifisere variabelen som ga ønsket resultat. Oppfatninger ble skilt mellom en naiv og vitenskapelig representasjon av kunnskapene. Naive representasjoner er kjennetegnet gjennom at de er tett tilknyttet og utspringer fra elevenes hverdagslige forestillinger, og følger en konkret tidslinje. Vitenskapelige representasjoner kommer derimot hovedsakelig fra konkrete, vitenskapelige begrep, som ikke følger en tidslinje. Man kan bruke flere vitenskapelige begrep for å argumentere for en tolkning (Larkin, 1981).

For å oppnå innsyn om elevenes praksis og tenkemåter i møte med PISA-oppgaver, ble det studert lydopptak av elever som jobbet med oppgavene «Løping i varmt vær» og «Undersøkelse av dalsider». Under oppgavejobbingen ble det stilt spørsmål for å klargjøre elevenes løsningsstrategier. Derifra ble det laget en modell, som beskriver elevenes løsningsstrategier når de jobbet med settet «Løping i varmt vær» og koder som karakteriserer elevenes epistemiske oppfatninger. Denne karakteriseringen av elevenes praksis og tenkemåter skulle gi innsyn til å svare på forskningsspørsmålene om hvilke tenke- og arbeidsmåter elevene viste når de jobbet med PISA-oppgaver om NPT, og hvordan PISA måler dette.

### 5.1 Elevenes løsningsstrategier

#### 5.1.1 Løsningsvei L1 – Praktisk tilnærming

Funnene som ble oppsummert i kapittel 4.1.6 ga opphav til tre hovedstrategier elevene brukte til å generere kunnskap gjennom å vurdere hypoteser når de jobbet med oppgavesettet «Løping i varmt vær». Den første (L1) var preget av at elevene *simulerte prøvende* for å få datasett med alle mulige innstillinger, som oppgaveteksten tillot, og

sammenliknet datasettene med svaralternativene til å finne ut hvilket svarforslag som var mest konsistent med datasettene.

Når elevene brukte denne løsningsmetoden, ble det ikke vist noe konkret prosedyrekunnskap. Grunnen til dette er at elevene fulgte det oppgaven oppgav som en kokebokoppskrift, altså de simulerte datasettene uten å ha en tanke om hvorfor dette kommer til å generere et svar, eller kunnskap, på oppgavespørsmålet. Dette epistemiske aspektet med prosedyren ble først etablert etter konsistensvurderingen mellom data og svarforslagene, når elevene skulle svare på hvorfor datasettene støttet svaret deres. Elevene vurderer datasettene med hensyn på hypotesene, fremfor å fremvise prosedyrekunnskap, som begrunnes med epistemisk kunnskap. Argumenteringen var preget av at man argumenterte for at to bestemte datasett viste hypotesen best, siden oppgaven spurte om to datasett til å underbygge svaret. Dermed vil jeg kategorisere strategien L1 som en praktisk tilnærming, dersom man skal sette det i lys av strategiene som Schauble et al. (1991) beskriver, beskrevet i kapittel 2.3. Målet til elevene var å oppnå et ønsket utfall av data, som kommer til å hende dersom man kjører simuleringen med alle mulige innstillinger som deduseres ut ifra oppgavestillingen. Deretter leter man etter svarforslaget som samsvarer best med dataene. Elever testet alle mulige kombinasjoner, som samsvarer med en vitenskapelig tilnærming, men ikke for å teste en tolkning, men for å raskt kunne avgjøre om svaret er rett.

### **5.1.2 Løsningsvei L2 – Praktisk-strategisk tilnærming**

Ved løsningsstrategi L2 simulerte elevene målrettet. Hensikten bak simuleringen var preget av elevene hadde en antakelse om at datasettene vil føre dem til svaret. De hadde ikke en egen fase, der de tolket data på skjermen, men elevene vurderte oppgitte svarforslag i lys av data som simuleringen resulterte i. Dette skiller seg fra konsistensvurderingen i L1 gjennom at disse vurderte konsistensen mellom svarforslagene og alle datasett.

Gjennom hele L2 viser elevene prosedyrekunnskap, som har en epistemisk begrunnelse. Elevene hadde en forestilling om at deres simuleringer vil gi ønsket utfall, fordi de vet hva de leter etter. Derfor blir variablene manipulert med utgangspunkt i elevenes strategi, og ikke fordi dette er den eneste variabelen oppgaven ikke sier noe om. Siden den første forskjellen mellom L1 og denne strategien er at elevene har en bestemt strategi, karakteriserer jeg denne løsningsmetoden som praktisk-strategisk. Elevene er ute etter å



produsere et bestemt resultat som hjelper dem til å gi dem et bestemt svar, og når dette resultatet er oppnådd, avsluttes simuleringsfasen. Deretter leter man etter svaralternativet, som overensstemmer med datasettene. Forskjellen mellom L1 og denne strategien er at elevene er seg allerede bevisst på hvilke epistemiske egenskaper dataene har, dvs. de vet på forhånd hvordan dataene kommer til å støtte et svaralternativ. Derfor foregår letingen etter dette temmelig raskere.

### **5.1.3 Løsningsvei L3 – Vitenskapelig tilnærming**

Når man brukte den tredje strategien (L3), simulerte man først målrettet med intensjonen om at datasettene vil muliggjør tolkning. Deretter tolket elevene dataene med hensyn på spørsmålet i oppgaven og sammenliknet denne tolkningen med svaralternativene. Dersom elevene ønsket bekreftelse for tolkningen sin, kjørte de simuleringen en gang til, men med andre innstillinger og vurderte om disse fulgte forventet tolkning.

På lik linje med L2, viser elevene frem prosedyrekunnskap gjennom at de planlegger en simuleringsstrategi, som gir dem utgangspunkt for et svar, og vet at disse dataene kommer til å gi elevene nok kunnskap til å formulere et svar. I tillegg til dette foregår det en evaluering av datamaterialet som forekommer. Dette kontrasterer strategi L2, der dette var allerede innbakt i selve utviklingen av simuleringsstrategien. Ut ifra denne evalueringen formuleres det en tolkning av data, der elevene vet hvordan dataene kommer til å støtte mulige tolkninger – noe som er en dimensjon av epistemisk kunnskap (Osborne, 2013). Deretter hadde elevene valget mellom to muligheter: enten å finne svaralternativet som samsvarte med elevenes tolkning, eller kjøre simuleringen en gang til for å teste tolkningen. Det her skillet går mellom den praktiske og vitenskapelige tilnærmingen (Schauble et al., 1991). Dersom elevene sammenlikner tolkningen sin direkte med svaralternativene, tydes dette på en praktisk strategi. Elevene anser simuleringen som avsluttet når en tolkning er etablert og handler videre derifra. Elever med strategien L3 derimot, som valgte å teste sin tolkning av data, kan heller karakteriseres som utøvere av den vitenskapelige tilnærmingen. Man forsøker å forstå sammenhengen mellom variablene, fremfor å komme så raskt som mulig frem til et svar. Dette tyder også på en dypere kunnskap innenfor prosedyre og epistemologi, siden man først forsøker å etablere en tolkning og deretter gjennomfører simuleringer, som støtter denne tolkningen. Når denne testingen er avsluttet, altså når tolkningen er bekreftet, går elevene videre til å sammenlikne denne med svar.

### 5.1.4 Oppsummering og sammenlikning med eksempelsvar til PISA

Nedenfor er det oppgitt en tabell, der elevenes løsningsstrategier er oppført. Det ble beskrevet hvorvidt prosedyre- og epistemisk kunnskap kjennetegnet de enkelte løsningsstrategiene.

<i>Fokusgruppe</i>	<i>Gruppe A</i>	<i>Gruppe B</i>	<i>Gruppe C</i>	<i>Gruppe D</i>
<i>Oppgave 1</i>	X	X	X	X
<i>Oppgave 2</i>	L3	L2	L1	L2
<i>Oppgave 3A</i>	L3T	L1	L1	L3T
<i>Oppgave 4</i>	L1	L1	L1	L3
<i>Oppgave 5</i>	<b>L1</b>	L3T	<b>L1</b>	L2

*Tabell 7: Oversikt over løsningsstrategier til fokusgruppene i oppgavesettet "Løping i varmt vær". Skrift i tykk font betyr at man ikke går ut ifra at elevenes svarforslag ville ha gitt score, når de ble sammenliknet med eksempelsvarene hos Kjærnsli og Jensen (2016c).*

Grunne til at til det står «X» på alle grupper ved oppgave 1, er at denne ikke krevde en egen tolkning av flere datasett for å kunne bli besvart. Man skulle se hvilken måling tilhørte hvilken enhet, og hvilken enhet var ansvarlig for hvilken helsefare. Derfor vil jeg ikke si at denne oppgaven testet prosedyre- eller epistemologiske kunnskaper.

Bortsett fra gruppe A og C på oppgave 5 har alle gruppene svart rett slik jeg vurderer det gjennom en sammenlikning med eksempelsvarene i PISA (Kjærnsli & Jensen, 2016c). En mulig forklaring på dette er at det var et gruppearbeid. Dermed kan elever som hadde dypere innsikt i NPT ha påvirket svaret til dem som ikke hadde det. Denne forklaringen underbygges av at elevene i gruppene kjenner hverandre, slik at elevene har en viss forstilling om kompetansene til hverandre. Før det går dypere inn, nevnes det at å skille mellom praktiske og vitenskapelige tilnærminger betyr ikke at en tilnærming anses som bedre, eller mer rett, enn den andre. Oppgavene i dette settet bærer trekk av lukkede oppgaver og derfor er det ikke overraskende at elever prøver å simulere til å oppnå et ønsket resultat så raskt som mulig fremfor å bedrive vitenskapelig uttesting og tolkning til å svare.

Elevene viste lite til ingen prosedyrekunnskap, dersom de brukte L1. Tegn på epistemisk kunnskap skinte gjennom når de skulle vurdere svarforslagenes konsistens med alle datasettene. Da er det positivt å se at gruppene fikk full score, når deres besvarelser sammenliknes med eksempelsvarene på Oppgave 5. Dette tyder på at denne oppgaven tester hvor dyp kunnskap elevene har innenfor prosedyrekunnskap og epistemisk kunnskap.

Et annet interessant resultat er at oppgave 2, 3A og 4 fra settet ble besvart rett også av elever som brukte L1. En mulig årsak av dette kan ligge i utformingen av oppgave 2. Dersom elevene svarer rett på denne oppgaven, kommer de til å vite at inntak av vann kun påvirker væsketapet, men hverken kroppstemperatur eller svette mengde. De resterende oppgaver, bortsett fra oppgave 5 spurte eksplisitt om disse to størrelsene. Dermed hadde elevene en mindre variabel å ta hensyn til, når det gjaldt simuleringen. Dette kan ha gjort det enklere å finne et rett svar også med L1. I tillegg til dette, gjorde oppgavene tydelig oppmerksom på når løperen ble utsatt for en helsefare gjennom å la en rød pil lyse opp på skjermen. Dette gjorde det enklere for elevene å se etter oppsiktsvekkende resultat, som oppgave 1, 2 og 4, krevet, enten for å løse, eller innstillingene i oppgaveteksten førte til et slikt resultat.

I alle andre oppgaver, bortsett fra oppgave 5, var strategiene blandet for å komme frem til det rette svaret. Oppgave 3A krevet av elevene å finne ut forholdet mellom den uavhengige variabelen «Lufttemperatur» og en av dens avhengige variabel, «Svette mengde». Dersom elevene ikke brukte strategi L1, ble L3 brukt, der elevene også testet tolkningen sin. Dette tyder på at man ikke nødvendigvis trenger å ha eksplisitt prosedyrekunnskap for å kunne løse denne oppgaven, men at oppgaven likevel kan trigge elevene til å oppnå en forståelse av hvordan forholdet mellom variablene i denne oppgaven er. Det samme gjelder for de andre oppgavene, bortsett fra oppgave 5.

Hvilke epistemiske tenkemåter viser elevene frem da, når de jobbet med oppgavene? Jeg vil først se på oppgave 3A. Oppgaven gikk ut på å etablere at en økt lufttemperatur ville medføre i en økt svette mengde. Der var det to slutninger gruppene som tydeliggjorde seg. En slutning gikk ut på å formulere en kausalsammenheng mellom kroppstemperatur. Dette ble tydelig i formuleringer som «Lufttemperaturen øker svette mengden», og tyder på mer overfladisk epistemisk kunnskap. Man kommer direkte til konklusjoner om at lufttemperaturen er årsaken til den økte svette mengden. Her gir man mer kunnskap enn det man kan, ut ifra datasettene, gi. Interessant her er at gruppene, som etablerte en kausalsammenheng, ikke fikk full score på oppfølgings spørsmålet, sammenliknet med eksempelsvarene (Kjærnsli & Jensen, 2016c). Den andre slutningen som ble synlig i datamaterialet var å si at svette mengden øker hver gang man øker lufttemperaturen. Her er man nærmere empirien av de simulerte datasettene, siden man kun gjengir forholdet mellom variablene «lufttemperatur» og «svette mengde», uten å gi en påstand om

kausaltet. Dette tyder på en dypere epistemisk forståelse, siden denne formuleringen tillater årsaksforklaringer i fremtiden, fremfor at man begrenser seg til at en økt lufttemperatur forårsaker økt svette mengde. Også en interessant påpekning her er at gruppene, som ikke brukte en kausalformulering i denne oppgaven, fikk full score på oppfølgingsspørsmålet ved sammenlikning av eksempelsvarene. Oppgaven kunne ikke måle forskjellene mellom disse tenkemåtene, siden den var en flervalgsoppgave med to alternativer, som begge er formulert som en korrelasjon fremfor en årsakssammenheng (Kjærnsli & Jensen, 2016c).

Dette kan man hevde for alle oppgaver innen dette settet, som direkte hadde å gjøre med simulering. Elevene var stilt ovenfor faste svaralternativer i oppgavene, som innskrenket muligheten til å teste elevenes epistemiske kunnskap som inngikk direkte i prosesskunnskapen. Oppgavene inkluderte et spørsmål, der elevene skulle argumenterte hvorfor to datasett, som elevene velger ut selv, støtter svaret de valgte. Dette gir innblikk i elevenes vurdering av data med hensyn på hypoteser, som er innenfor epistemisk kunnskap (Osborne, 2013). Likevel ble det ikke funnet noe i datamaterialet, som entydig pekte på at oppgavene testet epistemisk kunnskap, som inngår i planleggingen av undersøkelser, selv om dette kan tydes til å ha blitt vist av noen elever, som brukte L2- og L3-strategien.

En oppsummering av diskusjonen hittil nå er at resultatene antyder at noen elever ikke trengte å målet å finne ut om variablenes samspill, for å kunne oppnå full score i sammenlikning med eksempelsvarene (Kjærnsli & Jensen, 2016c). Dette var ikke tilfellet med oppgave 5. I tillegg peker resultatene også på at, dersom noen elever hadde som mål å finne ut om variablenes samspill, ville de kunne få full score på oppgavene. Ut ifra dette kan man si at det består en viss tvil om at oppgavesettet «Løping i varmt vær» måler prosedyrekunnskapens epistemologiske dimensjon, bortsett fra oppgave 5.

## **5.2 Elevenes tenkemåter**

### **5.2.1 Måling av én variabel og inngående usikkerhet**

Dersom man bruker karakteristikkene fra Tabell 3 til å beskrive tenkemåtene til fokusgruppene i oppgave 1 fra settet «Undersøkelse av dalsider», oppstår det et interessant bilde. Alle elever brukte forholdte seg til konkrete eksempler, som var tilknyttet til oppgavesettet, selv om spørsmålene som jeg stilte dem ble formulert mer generelle, som f.eks. «Hvorfor gir flere måleinstrumenter en mer nøyaktig måling?». Her

kan det nevnes at elevene tolket spørsmålet slik at den kun var ment for denne konteksten. Likevel var det ingen av elevene som eksplisitt kom innom idéen om at alle målinger i seg selv er preget av usikkerhet og baserte sine resonneringer ut ifra dette. Noen av elevene kan til og med karakteriseres til å ha naive forestillinger om data, nemlig de som tok som utgangspunkt at dersom man fjerner alle feilkildene, vil måleinstrumentene vise rett. Disse inngår i grunntanken om at man vil ha så nøyaktige målinger som mulig, men har ingen tydelig forestilling om at det består usikkerhet i målingene selv. Dette kan tydes til at elevene har implisitte forestillinger om NOS i at elevene mente at flere måleinstrumenter bidro til mer nøyaktige målinger, men mislykket i å peke direkte på dette. Derimot ble de gitt konkrete eksempler, f.eks. at dyr drikker fra vann måleren eller sabotasje av måleinstrumentene, på hva som kan skje i naturen og at dette blir vanskelig å unngå. Disse følger en tidslinje i at man først setter opp måleinstrumentet og så skjer det noe uventet med dette, uten konkret henvisning til et vitenskapelig konstrukt. Derfor vil jeg karakterisere dette som naive oppfatninger innenfor målinger av én variabel.

Det var to elever som derimot pekte på geografiske variasjoner, som er uunngåelige når man skal gjennomføre målinger for å beskrive et stort område. Derfor trengte man flere måleinstrumenter. Elev D2 formulerte dette til og med som noe man må gjøre for å oppnå større nøyaktighet i måling av et stort område, fordi flere uavhengige målinger kan ta hensyn til de geografiske variasjonene til området. Et eksempel på dette kan være at det regnet mer på ett sted. Dette tyder på en oppfatning, som fortsatt er tilknyttet til dalsidekonteksten, men man begynner å se visse generaliseringer angående målinger av én variabel. Nemlig at i store områder må man regne med mer usikkerhet i enkelte målinger. Slike oppfatninger gir er også mer uavhengig av tid, Man sier at usikkerhet i målinger er noe man må ta hensyn til uansett, fremfor at man sier at det kunne regne mer på ett sted enn på et annet og derfor skal man sette opp måleinstrumentene på forskjellige steder. Man befinner seg fortsatt innenfor en naiv representasjon, siden utsagnene fortsatt er tilknyttet geografiske variasjoner, og det mangles en eksplisitt formulering av den generelle idéen, som står bak denne. Derfor vil jeg karakterisere denne tenkemåten som implisitt, siden eleven har visse epistemiske forestillinger, selv disse fortsatt besitter en tilknytting til elevenes individuelle erfaringer.

En bemerkning her at elever, som anså målinger til å enten være feil eller rett, kom også til slutningen at feilkilder var umulige til å unngå, som f.eks. vist i Samtale 25. Likevel kan spørsmålet, som førte til utsagnet om at man aldri vet om hva som kommer til å skje, oppfattes som ledende gjennom at årsaker til usikkerheten skjer kun peker på ytre omstendigheter. Jeg vil argumentere for at eleven kunne hatt avvist denne letingen etter bestemte årsaker og kunne ha pekt på usikkerhet i målinger generelt, dersom en slik idé hadde eksistert. Derfor kan utsagnet heller tolkes som at antall mulige feilkilder er uendelige fremfor at det finnes en generell usikkerhet i alle målinger. For å synliggjøre dette tydeligere, kan man sammenlikne denne tenkemåten med tenkemåten som pekte på uunngåelige variasjoner. Gjennom å påpeke på uunngåelige variasjoner i et stort område, utelukker man fullstendig at ett måleinstrument kan måle helt nøyaktig innenfor et stort område. Dersom man tenker på uendelige muligheter på feilkilder, kan måleinstrumentene måle korrekt innenfor området, dersom man fjerner feilkildene, selv om dette kan anses som en fullstendig utopisk tanke innenfor dette tankesettet.

Én elev brukte en analogi, der hen sammenliknet målinger av én variabel med å spille Yatzy. Det ble ikke nevnt en generell strategi, løsrevet fra erfaringer, for å øke nøyaktigheten til måldata. Denne tenkemåten er likevel et uttrykk for en idé at flere målinger øker nøyaktigheten til data, og derfor vil jeg karakterisere denne tenkemåten som at den har nyanserte implisitte epistemiske oppfatninger. Pga. den sterke tilknyttingen til noe observerbart, og at den ikke utspringer fra en eksplisitt formulert vitenskapelig idé, er denne tenkemåten fortsatt preget av naive representasjoner.

I oppgave 2 av settet «Undersøkelse av dalsider» blir usikkerheten tallfestet gjennom at den blir oppgitt som usikkerhet et intervall rundt en gjennomsnittsverdi. Kun én fokusgruppe tok hensyn til dette under jobbingen med selve oppgaven, mens resten brukte kun gjennomsnittsverdien som utgangspunkt. Pga. dette ble alle grupper spurt på hvilken dalside det regnet mest, for å kunne gi avklaring om elevene hadde en oppfatning av hva usikkerheten har å si for nøyaktigheten til målingene. Det tilføyes at ingen elev visste fra før av, hva skrivemåten, altså gjennomsnittsverdi  $\pm$  usikkerhetsverdi, betydde. Elevenes tenkemåter var kjennetegnet av at de så på hele usikkerhetsintervall som sikker eller en verdi innenfor intervallet er rett, mest sannsynlig gjennomsnittsverdien. Dette viste seg gjennom at alle fokusgrupper kom til slutningen at det regnet mer på én dalside pga. en manglende overlapp av usikkerhetsintervall hos de lave verdiene i intervallene.

Elevenes evaluering av datamaterialet i oppgaven bygger i noen grad på en epistemisk oppfatning om at usikkerhet påvirker nøyaktigheten til måleverdier. Likevel viste elevene ikke noen tegn av at dette kan ha innflytelse på kunnskapsbyggingen, som kreves for å svare på spørsmålet om hvilken dalside som hadde mest nedbør. Dersom man hadde slik kunnskap kunne man kommet frem til at, hvis usikkerhetsintervallene til to målinger har for mye overlapp, vil man ikke kunne si om dette skyldes til to forskjellige fenomener. Elevene valgte derimot å si at det regnet mer på en av dalsidene, siden intervallene ikke overlapper. Derfor er noen form for epistemiske idéer tilstede her, men disse er temmelig implisitte.

Et spørsmål som da merker seg: Hva kjennetegnet tenkemåten til de tre resterende fokusgruppene når de faktisk jobbet med oppgave 2 fra «Undersøkelse av dalsider»? I kapittel 4.3.5 blir det beskrevet at disse gruppene kun brukte gjennomsnittsverdiene, og ignorerte usikkerhetstallene. Gjennomsnittsverdiene som hadde størst differanse mellom dem i forhold til deres numeriske verdi, måtte være årsaken til forskjellig jordfuktighet på dalsidene. Denne tenkemåten kan ses som å være preget av elevenes konseptuelle kunnskap fremfor dataene epistemiske egenskaper. Elevene vet at 10 mm nedbør ikke gjør en stor forskjell over et større tidsrom. Problemet med denne evalueringen av data, når man bruker mer konseptuell kunnskap, er at dersom usikkerhetene til alle målinger hadde vært så store at man ikke kunne bestemme en årsak entydig, ville elevene ikke kunne sett dette. Selvfølgelig kan man argumentere for at såpass store usikkerhetsmarginer kanskje ville kunnet virke oppsiktsvekkende for elevene, og føre til inkludering av usikkerheten i resonneringen deres.

### **5.2.2 Teoriens rolle i en forklaring**

Før jeg kommer til å diskutere elevenes svar, vil jeg oppgi eksempelsvarene til PISA (Kjærnsli & Jensen, 2016c) til å påpeke de teoretiske idéene som en inngår i disse.

«For å få et godkjent svar på denne oppgaven må elevene forklare at svette bidrar til å kjøle ned eller regulere kroppstemperaturen, for eksempel:

- Svette fordamper og kjøler ned kroppen ved høye temperaturer
- En økning i svettemengde ved høye temperaturer forhindrer at kroppen blir for varm
- Svette bidrar til å holde kroppstemperaturen på et trygt nivå

Oppgave 3B krever at elevene bruker kunnskap innenfor biologi (innholdskunnskap) til å forklare at svette bidrar til å kjøle ned kroppen ved høye temperaturer.» (s. 214)

Tenkemåten, som vises i selve eksempelsvarene er en vitenskapelig representasjon i at den tar utgangspunkt i kunnskap, som går ut over erfaringer og formulerer dette generelt, uten henvisning til en prosess, som følger en tidslinje. Dette inkluderer både termiske prosesser i første kulepunkt, og kroppens behov for å holde kroppstemperaturen på et konstant nivå. Likevel tolket jeg utsagnet om at elevene må forklare at svette bidrar til å kjøle ned eller regulere kroppstemperaturen som at også svar, som innehar en mer naiv representasjon, kan få full score.

Alle fokusgrupper viste tegn av epistemisk kunnskap om når det skulle svare på hvorfor en økt lufttemperatur førte til en økt svettemengde i oppgave 3. Den mest implisitte formen av denne kunnskapen ble observert når elevene forklarte en påstand vha. teori. Alle elever, som besvarte oppgaven på denne måten, viste en implisitt forståelse for at teori kreves i en årsaksforklaring. Elevene som ikke inkluderte teori i sin besvarelse, så ut til å vite at de trengte mer kunnskap om kroppens funksjoner for å kunne gi en tilfredsstillende besvarelse. Dette vil jeg også karakterisere som implisitte forestillinger, siden ingen av disse elevene ga uttrykk for at teori alltid kreves i en forklaring, selv om dette er idéen bak elevenes tankegang. Å si at man ikke har nok kunnskap om kroppen til å gi en årsak er tett tilknyttet den epistemiske idéen om at teori kreves generelt for å gi en årsaksforklaring, selv om elevene kun snakker innenfor en konkret kontekst. Derfor vil jeg karakterisere denne representasjonen som mer vitenskapelig enn naiv. Denne slutningen kan fortsatt diskuteres, siden elevene ikke eksplisitt nevner den rollen teori har i en årsaksforklaring.

Jeg vil nå se litt nærmere på elevene som ga en forklaring, som inneholdt teoretiske elementer i oppgave 3 fra «løping i varmt vær». Forklaringen er at kroppen svetter mer for å kjøle seg ned. Tankegangen her følger en konkret tidslinje, gjennom at lufttemperaturen øker, blir kroppen varmere og derfor svetter mer. I tillegg kan utsagnet om at man svetter for å kjøle seg ned være tilknyttet til elevenes egne erfaringer. Så selv om dette er fremvisning av innholdskunnskap og man bruker data til å underbygge denne kunnskapen, karakteriseres dette, ifølge kriteriene i Tabell 3, som en naiv tilnærming.



En elev som oppga en teoretisk forklaring, som ikke gikk på at kroppen kjøler seg ned, er elev B2 i Samtale 19. Hen sier først at man svetter for å kjøle seg ned, men deretter belyser kroppens krav for å holde temperaturen mest mulig jevn. Dette kan tydes som at eleven har en konseptuell forståelse, som er preget av tidsuavhengighet, og anvender denne forståelsen i en kontekst fra erfaringen. Altså at man svetter for å kjøle seg ned, er et resultat av kunnskapen om at kroppen har et behov for å holde kroppstemperaturen på et konstant nivå. Denne eksplisitte fremvisningen av kunnskap kan derfor tydes som en vitenskapelig representasjon, siden slutningen om at man svetter når man blir varmere er sterkt tilknyttet generell kunnskap. I motsetning står forklaringen, som er diskutert ovenfor, der man kun sier at kroppen svetter for å oppnå nedkjøling, som ikke har en henvisning til en tidsuavhengig formulering.

Det var også en elev, som ga en teoretisk forklaring, der hen henviste til energi som forklaring. Hen mente at dersom kroppen bruker mer energi, så vil den svette mer. Her blir igjen konseptuell forståelse fremvist, idet at hen vet at mer løping fører til et større forbruk av energi. Likevel er denne igjen tidsavhengig. Man løper mer, bruker energi og som konsekvens svetter man mer. Derfor er denne representasjonen også naiv.

### 5.2.3 Oppsummering og sammenlikning med eksempelsvar til PISA

For å gjøre tabellene oversiktligere, tilskrives i denne oppgaven hver oppfatning med en tallkode, som tilsvarende kapittelnavnet i funn, som beskriver oppfatningen. En oversikt over tallkodene er gitt nedenfor.

Tallkode	Epistemisk oppfatning
2.1	Elever bruker teoretisk kunnskap
2.2	Korrelasjon med tidslinje som årsak
2.3	Årsaksforklaring krever teoretisk innsikt
3.1	Måleinstrumenter viser enten rett eller feil
3.2	Usikkerhet som sikkerhet innen et intervall
3.3	Usikkerhet som variasjon i observert fenomen
3.4	Flere uavhengige variabler øker sikkerhet
3.5	De største numeriske forskjellene gir en årsak

Tabell 8: Oversikt over tallkodene, som brukes i Tabell 9.

Neste tabell viser hvilke oppfatninger elevene i gruppene viste. I tillegg viser den om formuleringen, som utspringer denne oppfatningen, ville ha gitt et rett svar når sammenliknes med eksempelsvarene fra PISA (Kjærnsli & Jensen, 2016c). I tillegg viser tabellen oppfatningen til elevene, når de ble spurt om på hvilken dalside det regnet mest, ut ifra dataene fra oppgave 2 i settet «Løping i varmt vær». Dette er vist som «Spørsmål»

tabellen. Der ble det valgt å se bort fra rett eller feil svar, siden dette spørsmålet gi dypere innsikt i elevenes tenkemåter.

	<i>Gruppe A</i>	<i>Gruppe B</i>	<i>Gruppe C</i>	<i>Gruppe D</i>
<i>Oppgave 3B</i>	2.1	2.1	<b>2.2</b> <b>2.3</b>	<b>2.1</b> <b>2.2</b> <b>2.3</b>
<i>Oppgave 1</i>	3.1 3.5	3.3	3.1	3.1 3.3
<i>Oppgave 2</i>	3.5	3.2	3.5	3.5
<i>Spørsmål</i>	3.2	3.2	3.2	3.2

*Tabell 9: Elevers oppfatninger, når de jobbet med oppgave 3B fra settet "Løping i varmt vær", oppgavesettet "Undersøkelse av Dalsider", og spørsmålet om på hvilken dalside det regnet mest. Svar, som ikke fikk full score er markert gjennom tykk font.*

Jeg vil først se på elevenes svar på oppgave 3B fra settet «Løping i varmt vært», der man skulle forklare årsaken til økt svette mengde ved høyere temperaturer. Der var to grupper, som brukte ulik teoretisk kunnskap om at svette bidrar til å holde kroppstemperaturen på et trygt nivå, i sin forklaring. Gruppe A visste at kroppen svetter, når den blir varmere, så de besitter viss konseptuell kunnskap. Denne følger en tidslinje, tett knyttet til deres hverdagslige forestillinger. Derfor var det utfordrende å bestemme om denne besvarelsen ville ha fått full score. Eksempelsvaret som kommer nærmest dette er «Svette fordampner og kjøler ned kroppen ved høye temperaturer» (Kjærnsli & Jensen, 2016c), men den går bort fra erfaringer og er innom de termiske prosessene som inngår ved fordampning. Likevel nevner PISA at dersom man forklarer at svette bidrar til å senke kroppstemperaturen, vil man oppnå full score (Kjærnsli & Jensen, 2016c). Derfor ble dette svaret markert som et scorende svar.

En løsning fra gruppe D, som mente at energi var en mulig årsak, ble markert i tabellen som et ikke scorende svar. Igjen, dette følger den tidslinje og bærer preg av elevens erfaringer, men viser ikke at svette bidrar til regulering av kroppstemperaturen. Likevel består det en viss usikkerhet om at elever, som hadde den konseptuelle kunnskapen, som krevdes til å få full score på oppgaven, oppnådde denne scoren, dersom denne er preget av naive representasjoner. Også elever, som ikke brukte konseptuell kunnskap, brukte en forklaring, som fulgte en tidslinje. Lufttemperaturen øker, noe som medfører en økning i kroppstemperatur og dermed en økning i svette mengde.

Hvorfor brukte så mange elever en årsaksforklaring, som fulgte ei tidslinje? En mulig forklaring på dette kan være at elevene besitter naive representasjoner av denne

konseptuelle kunnskapen (Larkin, 1981). En annen mulig forklaring kan ligge i formuleringen av oppgaven. Formuleringen er «Hva er den biologiske årsaken til dette?» (Kjærnsli & Jensen, 2016c). Ordet «dette» brukes, siden elevene måtte tidligere svare om svettemengden øker eller minker, dersom lufttemperaturen øker. Jeg velger å se på uttrykket «årsak». Det kan hevdes at bruken av ordet «årsak» er egnet til å skille ut elever med naive representasjoner, men kan også lede til misforståelser. Gjennom å bruke ordet «årsak», kan en del elever tenkes å oppfatte at å løse denne oppgaven krever en forklaring, som følger en lineær tidslinje (Grotzer & Basca, 2003), selv om de har kunnskap om f.eks. kroppens krav for å holde temperaturen konstant.

Noen elever valgte å knytte datasettene til deres årsaksforklaring, når de argumenterte for deres teoretiske tolkning, f.eks. at kroppen i praksis ikke kjøles ned, noe datasettene viste, men at svette kunne ha bremsset oppvarmingen. Dette viser både noen av at elevene har epistemisk kunnskap om at teori må støttes av data, og at både tolkningen av data og teorien må argumenteres for (Osborne, 2013, 2014). I tillegg viser dette hvordan besvaring av oppgaven innebærer samspill mellom alle tre kunnskapsdimensjonene, gjennom at elevene evaluerer dataene i lys av teori og bruker disse til å argumentere for denne teorien. Dette samspillet mellom kunnskapene er sentralt for NPT (Kind, 2013). Dette kom ikke til uttrykk i de eksempelsvarene, gjengitt i 5.2.2, som viste konseptuell kunnskap, som var jo isolert fra denne konteksten. I lys av PISA-undersøkelsens formulering, om at det ikke er nok å bare gjengi kunnskap, men man må også kunne anvende denne (Kjærnsli & Jensen, 2016a, 2016c), virker dette uheldig. Oppgave 3B fra «Undersøkelse av dalsider» kan derfor synes til å ikke få frem alle elevenes evne til å argumentere i naturvitenskapelige kontekster, selv om dette står sentralt i vitenskapelig praksis (Osborne, 2013).

Alle elever, som viste naive epistemiske representasjoner ved oppgave 1 og tilnærmet ingen epistemisk kunnskap ved oppgave 2, fikk full score, ved sammenlikning med eksempelsvarene. Dette kan ses som et resultat av oppgavens enkle og implisitte testemåte, som Abd-El-Khalick (2012) også påpekte når han skrev om den implisitte tilnærmingen av NOS. Elevene synes til å ha kompetanser som samsvarer med dyp epistemisk kunnskap, men man vet ikke noe om oppfatningene som ligger bak elevenes svar og løsningsmetoder. I oppgave 1 viste nesten alle elever tegn på en naive representasjon gjennom tenkemåten at måleinstrumenter enten kan vise rett eller feil.

Dette førte til utsagn som «I tilfellet ett av instrumentene blir ødelagt» eller «Man får et mer nøyaktig svar». Disse svarene er også representert i eksempelsvarene til PISA (Kjærnsli & Jensen, 2016c) (se vedlegg). Dette kan støtte påstanden til Allchin (2011) om at dersom man tester NOS implisitt, må man teste innenfor alle kunnskapsdimensjoner til vitenskapelig resonnering for å oppnå et helhetlig bilde av både elevenes kunnskap og hvordan elevene anvender kunnskap. Oppgave 1 i «Undersøkelse av dalsider» tegner kun et bilde av hvordan elevene bruker sin kunnskap, uavhengig om den er naiv eller vitenskapelig.

Det samme kan påstås om oppgave 2 av settet «Undersøkelse av Dalsider». Kun én gruppe brukte usikkerhetsverdiene i tabellen til å si noe om dataenes nøyaktighet ut ifra spørsmålet oppgaven stilte. Resten valgte å ignorere dette og lette etter største differansen i hvert sett av gjennomsnittsverdier. Man kan hevde at dette er en form for kunnskap om egenskapene til vitenskapelige observasjoner når man skal bygge kunnskap om vegetasjon. Ut ifra dette kan man hevde at 10 mm forskjell i nedbør ikke har en stor effekt over et lengre tidsrom. Likevel tar en tenkemåte, som utelukkende ser på gjennomsnittsverdiene, ikke hensyn til spredning i dataene når de besitter en høy usikkerhet. Hvis alle datasettene i denne PISA-oppgaven ville hatt en så stor usikkerhet, ville man ikke kunne hevde noe om årsaken til forskjellen i vegetasjon mellom begge dalsidene. Derfor vil jeg hevde at å kun se på gjennomsnittet tyder på lite anvendelse av epistemisk kunnskap om usikkerhet. I tillegg er det preget at oppfatninger tett tilknyttet egne erfaringer, som karakteriseres som naive. Oppgaven har, i likhet med oppgave 1 fra det samme settet, vansker med å avsløre kunnskapsnivået, eller oppfatningen, som ligger bak argumenteringen til elevene.

Denne slutningen støttes gjennom tenkemåten elevene fremviste når de ble spurt på hvilken dalside det regnet mest. Der viste elevene noen epistemiske oppfatninger rundt oppgitte usikkerhetstallene gjennom at de ikke kunne entydig si på hvilken dalside det regnet mest, og at hvordan usikkerhetsintervall påvirket nøyaktigheten til data. Likevel var utsagnene preget av en oppfatning om at det måtte ha vært mer nedbør på én dalside., så elevenes epistemiske kunnskap var preget av noen mangler. Her bemerkes det at elevene ble spurt om hva de visste om usikkerhet og hva de mente tallene bak plussminus symbolet var. Dermed var elevene allerede i en kontekst, der man snakket om usikkerhet, noe som kan ha påvirket deres resonnering.

Oppsummert kan man si at når elevene viste en vitenskapelig representasjon av konseptuell kunnskap i oppgave 3B fra settet «Løping i varmt vær», var oppgaven i stand til å måle dette. Likevel kan formuleringer i oppgaven ha medført for at elevene valgte å bruke en mer naiv representasjon, selv om de egentlig var i en besittelse av en mer vitenskapelig representasjon. Oppgavene fra settet «Undersøkelse av dalsider» er derimot preget av et generelt problem til implisitt tilnærming ved testing av NOS, der elever, som viste en naiv representasjon av epistemisk kunnskap, kunne få full score på oppgavene, sammenliknet med eksempelsvarene til PISA.

### **5.3 Forslag til videre forskning**

Denne oppgaven identifiserte tre hovedstrategier innenfor elevers løsningsmetoder av oppgaver innenfor prosedyrekunnskap. Disse kan utdypes videre ved å studere elevers møte med utforskende oppgaver utenfor PISA-undersøkelsen. I tillegg vil videre kvalitativ forskning kunne identifisere flere løsningsstrategier, på lik linje med Schauble et al. (1991) som pekte på at det stadig dukket opp flere strategier enn praktisk og vitenskapelig i møtet med åpne oppgaver.

Elevene jobbet i grupper når de i denne masteroppgaven jobbet med oppgavene fra PISA-undersøkelsen. Dette kan ha hatt innflytelse på elevenes svar ved at elever med høyere kompetanse innenfor NPT ble påvirket av elevene med lavere kompetanser, og omvendt. Dette understreker den indikerende naturen til slutningene i denne oppgaven, om hva PISA-oppgavene måler. Derfor kan det være aktuelt å gjennomføre intervjuer med enkeltelever, når de jobber med oppgavene, for å be- eller avkrefte funnet om at elever som besitter naive representasjoner av konsepter innen NPT, i noen tilfeller kan få full score på PISA-oppgavene.



## Bibliografi

- Abd-El-Khalick, F. (2012). Nature of Science in Science Education: Toward a Coherent Framework for Synergistic Research and Development. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 1041-1060). Dordrecht: Springer Netherlands.
- Abd-El-Khalick, F., Lederman, N. G., Bell, R. L., & Schwartz, R. S. (2001). Views of Nature of Science Questionnaire (VNOS): Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science.
- Allchin, D. (2011). Evaluating Knowledge of the Nature of (Whole) Science. *Science Education*, 95(3), 518-542. doi:10.1002/sce.20432
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning : the SOLO taxonomy (structure of the observed learning outcome)*. New York: Academic Press.
- Dagher, Z. R., & Erduran, S. (2016). Reconceptualizing the Nature of Science for Science Education: Why Does it Matter? *Science & Education*, 25(1), 147-164. doi:10.1007/s11191-015-9800-8
- Dunbar, K., & Fugelsang, J. (2005). Scientific thinking and reasoning. *The Cambridge handbook of thinking and reasoning*, 705-725.
- Firestone, W. A. (1993). Alternative arguments for generalizing from data as applied to qualitative research. *Educational researcher*, 22(4), 16-23.
- Giere, R. N. (1991). *Understanding scientific reasoning* (3rd ed.). Fort Worth: Holt, Rinehart, and Winston.
- Godfrey-Smith, P. (2003). *Theory and reality : an introduction to the philosophy of science*.
- Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy*, 24(1), 23-37. Retrieved from <https://doi.org/10.1080/02680930802412669>. doi:10.1080/02680930802412669
- Grotzer, T. A., & Basca, B. B. (2003). How does grasping the underlying causal structures of ecosystems impact students' understanding? *Journal of Biological Education*, 38(1), 16-29.
- Hoy, W. K. (2010). *Quantitative Research in Education: A Primer*. In Thousand Oaks, California: SAGE Publications, Inc.

- Johannessen, A., Christoffersen, L., & Tufte, P. A. (2010). *Introduksjon til samfunnsvitenskapelig metode* (4. utg. ed.). Oslo: Abstrakt.
- Kind, P. M. (2013). Establishing Assessment Scales Using a Novel Disciplinary Rationale for Scientific Reasoning. *Journal of research in science teaching*, 50(5), 530-560. doi:10.1002/tea.21086
- King, N., & Horrocks, C. (2010). *Interviews in Qualitative Research*. London: SAGE Publications Ltd.
- Kjærnsli, M., & Jensen, F. (2016a). 1 PISA 2015 – gjennomføring og noen sentrale resultater. In *Stø kurs* (pp. 11-31).
- Kjærnsli, M., & Jensen, F. (2016b). 2 Naturfag i PISA: definisjon og oppgaver. In *Stø kurs* (pp. 32-48).
- Kjærnsli, M., & Jensen, F. (2016c). Vedlegg: Eksempeloppgaver i naturfag fra PISA 2015. In *Stø kurs* (pp. 209-226).
- Knain, E., & Kolstø, S. D. (2011). Utforskende arbeidsmåter—en oversikt. *Elever som forskere i naturfag*, 127-163.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212-218.
- Kvale, S., & Brinkmann, S. (2009). *Det kvalitative forskningsintervju* (2. utg. ed.). Oslo: Gyldendal akademisk.
- Larkin, J. H. (1981). *The Role of Problem Representation in Physics*. Paper presented at the Mental Models Conference, University of California, San Diego.
- Lasley, T. J. (2010). Bloom's Taxonomy. In J. C. C. Thomas C. Hunt, Thomas J. Lasley II & C. Daniel Raisch (Ed.), *Encyclopedia of Educational Reform and Dissent* (pp. 106-109). Thousand Oaks: SAGE Publications, Inc.
- Lederman, N. G. (2013). Nature of science: Past, present, and future. In *Handbook of research on science education* (pp. 845-894): Routledge.
- McDonald, C. V., & McRobbie, C. J. (2012). Utilising Argumentation to Teach Nature of Science. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 969-986). Dordrecht: Springer Netherlands.
- Meyer, H.-D., & Benavot, A. (2013). *PISA, power, and policy: The emergence of global educational governance*.
- Nilssen, V. L. (2012). *Analyse i kvalitative studier : den skrivende forskeren*. Oslo: Universitetsforl.



- Nowak, K. H., Nehring, A., Tiemann, R., & Upmeyer zu Belzen, A. (2013). Assessing students' abilities in processes of scientific inquiry in biology using a paper-and-pencil test. *Journal of Biological Education*, 47(3), 182-188. Retrieved from <https://doi.org/10.1080/00219266.2013.822747>.  
doi:10.1080/00219266.2013.822747
- Osborne, J. (2013). The 21st century challenge for science education: Assessing scientific reasoning. *Thinking Skills and Creativity*, 10, 265-279. Retrieved from <Go to ISI>://WOS:000326906800026. doi:10.1016/j.tsc.2013.07.006
- Osborne, J. (2014). Scientific practices and inquiry in the science classroom. In *Handbook of Research on Science Education, Volume II* (pp. 593-613): Routledge.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed. ed.). Thousand Oaks, Calif: Sage Publications.
- Ritchie, J., Lewis, J., & Elam, G. (2003). Designing and selecting samples. In J. Ritchie & J. Lewis (Eds.), *Qualitative Research Practice: A Guide for Social Science Students* (pp. 77-108). London: Sage.
- Robinson, O. C. (2014). Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative research in psychology*, 11(1), 25-41.
- Schauble, L., Klopfer, L. E., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of research in science teaching*, 28(9), 859-882.
- Schwartz, R. S., Lederman, N. G., & Abd-El-Khalick, F. (2012). A series of misrepresentations: A response to Allchin's whole approach to assessing nature of science understandings. *Science Education*, 96(4), 685-692.  
doi:10.1002/sce.21013
- Serder, M. (2015). *Möten med PISA : kunskapsmätning som samspel mellan elever och provuppgifter i och om naturvetenskap*. (Doctor), Malmö högskola,
- Sjøberg, S. (2014). PISA-syndromet. Hvordan norsk skolepolitikk blir styrt av OECD. *Nytt norsk tidsskrift*, 31(1), 30-43.
- Sjøberg, S. (2018). The power and paradoxes of PISA: Should Inquiry-Based Science Education be sacrificed to climb on the rankings? *Nordic Studies in Science Education*, 14(2), 186-202.
- Utdanningsforbundet. (2018). Læreplanen i naturfag. Retrieved from <https://hoering.udir.no/Hoering/v2/277?notatId=531>

Ødegaard, M., & Arnesen, N. (2010). Hva skjer i naturfagklasserommet?–resultater fra en videobasert klasseromsstudie; PISA+. *Nordic Studies in Science Education*, 6(1), 16-32.

## Vedlegg

### Samtykkeskjema

#### **Vil du delta i forskningsprosjektet**

#### ***”Elevs møte med PISA-oppgaver om naturvitenskapelig tenke- og arbeidsmåter”?***

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å få innsyn hva PISA-spørsmål tester om, og hvordan elever tolker disse oppgavene. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

#### **Formål**

PISA er en internasjonal studie som har blant annet målet å kartlegge elevs prestasjon i blant annet naturfag. For å teste dette stiller PISA oppgaver om også blant annet naturvitenskapelig tenke- og arbeidsmåte. Denne oppgaven skal dreie seg om disse oppgavene. Det skal drøftes hvilke områder av dette temaet blir dekket av oppgavene, og hvordan du som elev fortolker oppgavene. Informasjonen som samles inn her skal brukes i en masteroppgave, og kan gi en bedre innsikt om hvordan elever tolker slike oppgaver.

#### **Hvem er ansvarlig for forskningsprosjektet?**

Universitetet i Bergen er ansvarlig for prosjektet og gjennomføres på Institutt for Fysikk og Teknologi.

#### **Hvorfor får du spørsmål om å delta?**

Du ble utvalgt til å delta i dette prosjektet, fordi du er i aldersgruppen som PISA tester. Veilederen min (Professor Stein Dankert Kolstø) har, i samarbeid med læreren din, bestemt at elever fra din klasse kan bli spurt om å delta.

#### **Hva innebærer det for deg å delta?**

- Hvis du velger å delta, innebærer det at du skal jobbe i gruppe med noen medelever, der dere jobber med 1-2 oppgaver, i forbindelse med et gruppeintervju.
- Oppgavene er hentet fra tidligere PISA-spørreundersøkelser og handler om naturvitenskapelig tenke- og arbeidsmåte. Disse oppgavene er kun ment til å sette i gang samtalen under intervjuet. Oppgavene besvares skriftlig og svarene dine vil være fullstendig anonyme.
- Under gruppeintervjuet vil du bli spurt om hvordan du tenker om oppgavene, og om du har hatt erfaringer som likner på situasjonene, som er beskrevet i oppgavene, i naturfagsundervisningen din. Jeg kommer til å ta lydopptak og notater under intervjuet.
- Hvis ønskelig, kan dine foreldre/foresatte se på oppgavene/intervjuguide på forhånd.

**Det er frivillig å delta** Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykke tilbake uten å oppgi noen grunn. Alle opplysninger om deg vil da bli anonymisert. 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 dine opplysninger til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Det er kun veilederen min og meg som kommer til å ha tilgang til lydopptakene og informasjon om hvem som deltok i intervjuet. Alle dine opplysninger blir erstattet med en kode, som lagres på en egen navneliste adskilt fra øvrige data. All innsamlet informasjon blir lagret på en passordbeskyttet dataservert knyttet til Universitetet i Bergen. I selve masteroppgaven vil all informasjon om deg bli anonymisert.

### **Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Prosjektet skal etter planen avsluttes 30. juni 2019. Alle personopplysninger vil bli slettet og lydopptakene vil bli anonymisert innen denne dato.

#### **Dine rettigheter**

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

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

### **Hva gir oss rett til å behandle personopplysninger om deg?**

Vi behandler opplysninger om deg basert på ditt samtykke.

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

#### **Hvor kan jeg finne ut mer?**

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

- Alexander Michelsen, masterstudent ved Institutt for Fysikk og Teknologi, Universitetet i Bergen, epost: [fjernet], tlf: [fjernet]
- Professor Stein Dankert Kolstø ved Institutt for Fysikk og Teknologi, Universitetet i Bergen, epost: [fjernet], telefon: [fjernet]
- Vårt personvernombud: Janecke Helene Veim, epost: personvernombud@uib.no
- NSD – Norsk senter for forskningsdata AS, på epost (personverntjenester@nsd.no) eller telefon: 55 58 21 17.

Med vennlig hilsen

Prosjektansvarlig Student

Professor Stein Dankert Kolstø Alexander Michelsen

-----  
---

## **Samtykkeerklæring**

Jeg har mottatt og forstått informasjon om prosjektet *Elevens møte med PISA-oppgaver om naturvitenskapelig tenke- og arbeidsmåte*, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i gruppeintervju med PISA-oppgaver
- at lærer kan gi opplysninger om mitt navn, klasse, og skole til prosjektet (i tilfellet læreren samler inn samtykkeerklæringen)

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. 30. juni 2019

-----  
(Signert av prosjektdeltaker, dato)

## Tilfeldig side fra transkript

I: Mhm. Og noen valgte faktisk å ha 20 og 40 istedenfor 40 og 60. Ville dere ha sagt at disse resultatene som de fikk, eller den begrunnelsen som de gjorde, var bedre enn deres? Eller var deres bevis bedre enn de gjorde?

3: Hvis de tar 20 og 40, så kommer de aldri opp til 50% i luftfuktighet.

2: Du kommer jo uansett frem til at det er farlig, så lenge det er en fare med luftfuktighet på 40%. Så er jo 50% høyere enn det og det kommer fortsatt til å være over heteslagsgrensen. Men du- når du har både 40 og 60 da, så- så får du jo sett litt mer- kan- eh, ja, litt mer nøyaktig tenke deg fram til hva han er da på 50%. I hvert fall imellom 41,2 og 40,7.

I: Mhm. Ja, altså da er vi vel inne igjen på den diskusjonen: «Burde vi ha én istedenfor to», hvis vi vet at 40 er allerede farlig.

1: Man trenger jo egentlig bare én.

I: Er du enig i det? At man bare trenger én i dette tilfellet, eller?

3: Mhm.

I: Og hvorfor tror dere at oppgaven ville da ha to?

2: For å finne mer ut nøyaktig hvordan kroppstemperaturen blir på 50% luftfuktighet?

3: Jo mer svar du har jo mer nøyere blir det.

I: Hva mener du med her, hvis vi anvender på den oppgaven?

3: Hvis du har én så ser du jo: «Okay, det er svaret, det går greit», men hvis du har to, så ser du- så kan du få, eh, et bedre svar, for å si det sånn. Fordi du har mer informasjon på det.

I: Og ville du da ha sagt at det gjør ingen forskjell på beviset vårt, om vi har én eller datasett? Eller gjør det en forskjell på kvaliteten på dette beviset?

3: Det gjør ikke en forskjell på svaret egentlig, men.. det er jo bedre kvalitet jo mer svar du har, men ett svar er greit nok her.

4: Det blir jo bedre med to, fordi du kan jo se selv at 40 er- nei, hva var det? Ja, at ført-luftfuktigheten på 40 er fortsatt høyere enn heteslag, og at det er und- nei, hvordan skal jeg forklare det? Jeg vet ikke hvordan jeg skal forklare det.

I: Altså du sa nå at det var bedre med faktisk to, å ha med to.

4: Ja, på den syns jeg det, hvert fall. For du kan jo se at den ligger imellom de to.

2: Jeg ville ha sagt at det gjør ikke noe, men du får bare mer nøyaktig svar da. Så hvis.. Det er jo ikke noe skade i å.. å ha med to rader. Det gjør det jo ikke verre i det hele tatt.