

Nivådifferensierte oppgaver og mestringsforventning i matematikkfaget

*En studie av elever på 9. trinn i møte med
nivåmarkerte oppgaver*

Maria Klaussen Herset



Masteroppgave i matematikdidaktikk
Institutt for lærerutdanning og skoleforskning
Utdanningsvitenskapelig fakultet

UNIVERSITETET I OSLO

Våren 2014

Nivådifferentierte oppgaver og mestringsforventning i matematikkfaget

En studie av elever på 9. trinn i møte med nivåmarkerte oppgaver

© Maria Klaussen Herset

Våren 2014

Nivådifferensierte oppgaver og mestringsforventning i matematikkfaget

Maria Klaussen Herset

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

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

«Tilpasset opplæring» er ett av satsningsområdene i Kunnskapsløftet 2006. Et tiltak som ofte blir benyttet i matematikkundervisningen er å differensiere oppgavene etter elevenes faglige nivå, såkalt «nivådifferensiering». Hensikten er at elevene skal oppnå optimal læringseffekt og økt mestringsopplevelse. Valg av tema for masteroppgaven begrunnes med at nivådifferensiering og mestringsforventning står sentralt i et didaktisk perspektiv. Formålet med oppgaven er å undersøke om matematikkoppgavens nivåmarkering både har en effekt på elevenes valg av oppgave og om markeringen påvirker elevenes resultater. Oppgavens problemstilling er: *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes valg av oppgave og prestasjon på oppgaven?*

For å måle effekten av nivåmarkeringen gjennomføres en empirisk test. Utvalget til testen består av 289 elever fra 9. trinn ved praksisskoler i Oslo, som bruker nivådifferensierte lærebøker. Testen inneholder blant annet seks matematikkoppgaver med tilnærmet lik vanskelighetsgrad. For å måle effekten av nivåmarkeringen blir elevene inndelt i en eksperimentgruppe og en kontrollgruppe. I motsetning til kontrollgruppen får eksperimentgruppen oppgavene påført med ulike nivåmarkeringer, selv om oppgavene har tilnærmet *samme* vanskelighetsgrad. Nivåmarkeringen er tilfeldig fordelt og varierer innad i gruppen. Analysemetoden er «binomisk logistisk regresjon», der resultatene fremstilles grafisk.

Funn fra analysen viser at nivåmarkeringen har en prediktiv effekt på elevenes valg av oppgave. Selv om oppgavene har forholdsvis lik vanskelighetsgrad, velger elevene med lave terminkarakterer hovedsakelig oppgaven som er markert «lett», mens elevene med høye terminkarakterer velger oppgaven markert som «vanskelig». I tillegg antyder resultatene fra analysen at selv om elevene har like terminkarakterer, karakterene 1, 2, 3 eller 4, er det større sannsynlighet å få riktig svar på oppgaven når den er markert «lett», sammenlignet med elevene som får den *samme* oppgaven markert som «vanskelig». Resultatene viser en *motsatt* effekt for elevene med karakterene 5 eller 6. Selv om de har lik karakter i faget er det altså en mindre sannsynlighet for å få riktig resultat på oppgaven for elevene som får den markert «lett», sammenlignet med de som får den samme oppgaven markert som «vanskelig».

Forord

Høsten 2009 startet jeg på Lektorprogrammet i realfag ved Universitetet i Oslo. Det har både vært lærerikt og krevende å være student på Blindern. Jeg vil takke til foreleserne som har fanget min interesse for matematikdidaktikk og som har inspirert meg til å fullføre studieløpet. Studenttilværelsen blir uforglemmelig, og jeg vil anbefale alle å studere på Blindern.

De første tankene rundt masteroppgaven startet under et praksisopphold på en ungdomsskole våren 2012. Jeg retter en stor takk til praksisskolen og medstudenter som har hatt tro på mine idéer rundt oppgaven. Uten deres støtte og interesse hadde trolig ikke masteroppgaven blitt realisert. En spesiell takk går til praksisveilederne som var behjelpelig både under og etter praksisperioden.

Det har vært en lang og vanskelig prosess å få tak i informanter til undersøkelsene. Tusen takk til alle lærerne som var positive til å la meg utføre undersøkelsen, samt elever som deltok på testen. Besvarelsene deres har gitt viktig og interessant informasjon.

En spesiell takk går til min hovedveileder Arne Hole og biveileder Christian Brandmo. Dere har utfyllt hverandre optimalt og bidratt med gode tilbakemeldinger. Samlet har dere vist en særdeles god fagforståelse innenfor matematikdidaktikk, pedagogikk og metode.

Jeg vil også takke familier og venner for at dere har støttet og motivert meg til å gjennomføre masteroppgaven. Jeg setter stor pris på all den omsorg, oppmuntring og hjelp dere har gitt underveis. Ironisk nok har jeg slitt med egen mestringsforventning under masterprosessen, men takket være dere har jeg virkelig fått troen på meg selv.

Til slutt vil jeg takke Birger Rückstein, Kjetil Albertsen og Ingvild Spilling som har lest korrektur og kommet med gode innspill underveis. I tillegg vil jeg takke mamma og pappa – Dere er fantastiske på alle måter!

Innholdsfortegnelse

1	Innledning.....	1
1.1	Bakgrunn for valg av tema	2
1.2	Formål og problemstilling	3
1.3	Teoretisk og metodisk avgrensning	4
2	Teorier og perspektiver	5
2.1	Tilpasset opplæring.....	5
2.1.1	Differensiering	5
2.1.2	Nivådifferensiering.....	6
2.1.3	Elevmedvirkning i opplæringen	9
2.2	Mestringsforventning.....	9
2.2.1	Definisjon av mestringsforventning	10
2.2.2	Kilder til mestringsforventning	10
2.2.3	Viktigheten av mestringsforventning	12
2.3	Resultater fra tidligere empiriske studier.....	15
2.4	Hypoteser.....	16
3	Data og metode.....	19
3.1	Troverdighet	19
3.2	Forskningsdesign	19
3.3	Utvalg	21
3.4	Oppgavemateriale og prosedyrer	21
3.4.1	Datainnsamling.....	22
3.4.2	Utforming av oppgavematerialet.....	24
3.4.3	Forskningsetikk	26
3.4.4	Pilotundersøkelse	27
3.4.5	Variabler og målenivå	27
3.4.6	Begrepsvaliditet.....	29
3.5	Metode for analyse	30
3.5.1	Binomisk logistisk regresjonsanalyse og statistiske begreper.....	30
3.5.2	Forutsetninger for logistisk regresjonsanalyse	31
3.5.3	Tolkning av den logistiske regresjonslikningen.....	32
4	Analyse og resultater	35

4.1	Analyse 1 – Elevenes valg av oppgave	35
4.1.1	Valg av oppgave A	36
4.1.2	Valg av oppgave B	39
4.1.3	Valg av oppgave C	43
4.1.4	Oppsummering av hovedfunn fra analyse 1	46
4.2	Analyse 2 – Elevenes prestasjon på oppgavene	47
4.2.1	Prestasjon på oppgave D	48
4.2.2	Prestasjon på oppgave E.....	51
4.2.3	Prestasjon på oppgave F	55
4.2.4	Oppsummering av hovedfunn fra analyse 2.....	58
5	Diskusjon.....	61
5.1	Oppsummering	61
5.2	Hovedfunn	62
5.3	Konklusjon.....	63
5.4	Kritisk refleksjon rundt nivåddifferensierte oppgaver.....	64
5.5	Ytre validitet	66
5.6	Forslag til videre forskning.....	67
	Litteraturliste	69
	Vedlegg	73

1 Innledning

På 1960-tallet ble grunnskolen i Norge utvidet fra syv til ni år, noe som resulterte i et større behov for *differensiering* (Imsen, 1997, s. 274). Differensiering er en «bevisst forskjellsbehandling av elevene til beste for hver enkelt» (Solvang, 1992, s. 181). Hensikten med differensiering er å tilpasse undervisningen slik at elevenes forutsetninger og behov imøtekommes (Solvang, 1992, s. 181). Et tiltak for å differensiere undervisningen ble derfor å innføre kursplanssystemet (Kunnskapsdepartementet, 2011). I matematikkfaget ble elevene delt inn i tre grupper basert på ferdighetsnivåene lav, middels og høy. Elevens ulike ferdigheter definerte hvilket nivå undervisningen skulle ligge på. Hvilken gruppe de tilhørte hadde betydning for deres videre utdanningsmuligheter. Kursplansystemet ble avskaffet, og det ble konkludert med at det ikke var skolens oppgave å forutse elevenes fremtid i arbeidshierarkiet (Ibid.). Erfaringene fra kursplansystemet har bidratt til at flertallet ønsker en skole som bygger på enhetsskoleprinsippet. Det betyr at elevene skal få et likeverdig undervisningstilbud, der tanken er å gi alle en lik mulighet til å delta og utvikle samfunnet (Solvang, 1992, s. 7).

I dag har vi en 10-årig obligatorisk grunnskole. Skolen består av gjennomsnittsklasser, hvor elevene inndeles etter samme årskull. Behovet for differensiering har derfor aldri vært så sentralt som nå, og anses som et viktig tiltak. Dette som en følge av at det er obligatorisk for alle elevene å lære matematikk på grunnskolen (Solvang, 1992, s. 7). Matematikkfaget er oppbygd hierarkisk, noe som betyr at faget må læres trinn for trinn. Elevene må for eksempel lære addisjon før multiplikasjon, da addisjon er en forutsetning for å forstå hvordan gjentatt addisjon beregnes. Det vil derfor være viktig å tilrettelegge undervisningen etter elevens nivå slik at de mestrer oppgavene og oppnår god læringseffekt. En god læringseffekt oppstår når oppgavens utfordring er tilpasset elevenes ferdigheter, slik at elevene arbeider med oppgaver som er litt krevende (Csikszentmihalyi, 2005). Hvis oppgavene blir for lette eller for vanskelige kan elevene oppleve lav mestring i faget. En lav mestring kan bidra til at elevene kjeder seg, blir frustrerte og/eller opplever angst (Ibid.).

1.1 Bakgrunn for valg av tema

«Tilpasset opplæring» er en viktig del av Kunnskapsløftet 2006, K06 (Utdanningsdirektoratet, 2012, s. 4). Et tiltak som ofte blir benyttet for å imøtekomme kravene fra K06 er å differensiere oppgavene etter elevenes faglige nivå, såkalt nivådifferensiering (Dale & Wærness, 2003, s. 90). I følge Mathiassen (2009, s. 129) er det vanlig å bruke nivådifferensierte lærebøker i matematikkfaget, der oppgavene markeres etter vanskelighetsnivå. For eksempel er lett -, middels - og vanskelig oppgavenivå markert henholdsvis blå, rød og grønn. Tanken bak fargeinndeling er at elevene selv skal kunne velge oppgaver som bidrar til optimal læringseffekt og økt mestringsopplevelse (ibid.).

Observasjoner rundt bruk av nivådifferensierte oppgaver fra eget praksisopphold ga noen interessante refleksjoner. Skolen opererte med lærebøker og matematikkprøver som var nivådifferensierte. Prøvene var utformet slik at elevene selv fikk velge oppgaver. Eksempelvis kunne det velges mellom oppgave 1a som ga 1 poeng, eller oppgave 1b som ga 3 poeng. Det var flere elever som fikk *alt* riktig på prøven, men ettersom de kun besvarte de letteste oppgavene, fikk de lave karakterer. Observasjonen var interessant og førte til følgende eksperimentelle spørsmål:

1. Anta at oppgave 1a og 1b byttet nivåmarkering, slik at den letteste oppgaven ble markert vanskelig og omvendt. Kan det tenkes at de faglig svake elevene hadde valgt oppgaven som *nå* var markert «lett», til tross for at den i realiteten var den vanskeligste oppgaven?
2. Økes sannsynligheten for å få riktig svar på oppgaven hvis den vanskeligste oppgaven ble markert «lett», sammenlignet med den opprinnelige markeringen (vanskelig)?

I løpet av praksisperioden foretok jeg en enkel matematikktest av elevene på skolen. Testen besto av å gi den *samme* matematikkoppgaven til alle elevene. Øverst på oppgaven var det påført en tilfeldig nivåmarkering. Det betyr at det var noen elever som fikk oppgaven markert «lett», mens andre fikk den samme oppgaven markert «middels» eller «vanskelig».

Resultatene fra undersøkelsen viste at det var størst sjanse å få riktig svar på oppgaven hvis den var markert «lett», sammenlignet med når oppgaven var markert «vanskelig».

Resultatene fra testen var både interessante og overraskende, og er bakgrunn for valg av tema i masteroppgaven.

1.2 Formål og problemstilling

Denne masteroppgaven tar for seg nivådifferensiering og *mestringsforventning* i matematikkfaget. Mestringsforventning handler om hvilken oppfatning en har til å mestre en spesifikk oppgave (Bandura, 1997, s. 3). Som nevnt ble temaet valgt etter interessante observasjoner fra praksis. Temaet ble også valgt fordi det anses som viktig og relevant for undervisning i matematikkfaget, ettersom det blir brukt som tiltak for å imøtekomme kravene fra Kunnskapsløftet 2006 (Utdanningsdirektoratet, u.d.a). Nivådifferensiering står derfor sentralt i et didaktisk perspektiv, spesielt med tanke på at det brukes nivådifferensierte lærebøker i grunnskolen.

På bakgrunn av dette stilles følgende spørsmål: Kan det tenkes at nivådifferensierte oppgaver bidrar til at det er *elevne* som tilpasser seg oppgavenivået og ikke at *oppgavenivået* tilpasses eleven? På bakgrunn av dette ble følgende problemstilling valgt: *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes valg av oppgave og prestasjon på oppgaven?*

Problemstillingen gir kun svar på *om* oppgavens markering har en effekt, og forklarer ikke *hvorfor* den har en eventuell effekt. Problemstillingen er todelt og behandles separat i analysedelen. Oppgavens to deler er som følger:

1. I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes *valg av oppgave*?
2. I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes *prestasjon på oppgaven*?

Undersøkelsen kan forhåpentligvis bidra til ny teori innenfor feltet, fordi forskning på bruk av nivådifferensierte oppgaver i matematikkundervisningen er beskjedent. Ettersom det ikke er selvsagt at denne form for differensiering har en positiv effekt, er det mulig man kan kritisere den nåværende bruken av nivådifferensiering. I tillegg er det behov for forskning innenfor mestringsforventning og matematikk (Hannula, 2006, s. 211). Besvarelsen av problemstillingen kan gi viktig informasjon til forskningsfeltet, og videre skape grunnlag for forskning som kan styrke kvaliteten i skolesystemet.

1.3 Teoretisk og metodisk avgrensning

For å besvare problemstillingen foreligger det en teoretisk tilnærming som tar for seg begrepene differensiering og mestringsforventning i matematikkfaget. Differensiering avgrenses til nivådifferensiering, der hensikten er å oppnå mestring for økt læringseffekt. Behovet for nivådifferensierte oppgaver forklares ved bruk av Csikszentmihalyis (1975) «flytsonemodell» (Mathiassen, 2009, s. 129). Modellen illustrerer nødvendigheten av å tilpasse oppgavens vanskelighetsgrad etter elevenes ferdighet. Teori om mestringsforventning presenteres med utgangspunkt i Bandura (1997) sosial-kognitive teori om læring. Banduras læringsteori ble valgt, da han er anerkjent innenfor forventningstradisjonen (Imsen, 2008, s. 380). Mestringsforventning kan inndeles i generelle og spesifikke mestringsforventninger (Bandura, 1997). En elev kan for eksempel ha generelle mestringsforventninger om å prestere i matematikkfaget og spesifikke mestringsforventning som er knyttet til matematikkoppgaver i faget. I denne oppgaven vektlegges elevenes spesifikke mestringsforventninger. En begrunnelse er at det i større grad gir informasjon om hvordan forventningen påvirker personens atferd under bestemte oppgaver. Dette støttes av Pokay (1996) som mener at det er de spesifikke mestringsforventningene som må studeres i matematikkfaget, ettersom faget inneholder forskjellige matematiske temaer (referert i Throndsen, 2005, s.78). Det blir redegjort for fire kilder som, ifølge Bandura, konstruerer mestringsforventningen, nemlig egenopplevde mestrings erfaringer, informasjon via andres mestrings erfaringer, verbal overtalelse og den fysiologiske og affektive tilstanden (1997, s. 79). Deretter beskrives mestringsforventnings betydning for de kognitive -, affektive -, selektive - og motivasjonelle prosessene (Bandura, 1997, s. 116).

Forskningsdesignet er eksperimentelt. Det er gjennomført en tverrsnittstudie, som gir et øyeblikksbilde av situasjonen. Utvalget består av 289 elever fra 9. trinn ved praksisskoler i Oslo, som bruker nivådifferensierte lærebøker. Datainnsamlingsmetoden er kvantitativ og foregår ved bruk av en empirisk test. Datamaterialet analyseres ved bruk av «binomisk logistisk regresjon». Resultatene presenteres ved hjelp av tabeller og grafer. Problemstillingen besvares etter en tolkning av datamaterialet sammen med det teoretiske rammeverket.

Avslutningsvis gis en oppsummering av analysens hovedfunn, etterfulgt av kritisk refleksjon rundt bruk av nivådifferensierte oppgaver. Til slutt presenteres forslag til videre forskning på området.

2 Teorier og perspektiver

Kapitelet inneholder teori som benyttes til å belyse oppgavens problemstilling. I teoridelen behandles temaene *differensiering* og *mestringsforventning* i matematikkfaget separat. I tillegg presenteres det empiri fra tidligere studier. Avslutningsvis blir de teoretiske perspektivene samt resultatene fra tidligere forskning anvendt til å danne hypoteser. Hypotesene antyder egne forventninger til den empiriske analysen.

2.1 Tilpasset opplæring

I læreplanverket for Kunnskapsløftet 2006, LK06, er det fastsatt krav om tilpasset opplæring: «alle elever skal i arbeid med fagene få møte utfordringer de kan strekke seg mot, og som de kan mestre på egen hånd eller sammen med andre» (Utdanningsdirektoratet, 2012, s. 4).

Hensikten med tilpasset opplæring er at elevene inkluderes i et felles læringsmiljø, opplever mestring for en optimal læringseffekt, samt medvirkning av opplæringen (Utdanningsdirektoratet, u.d.a, s. 22).

I denne oppgaven avgrenses tilpasset opplæring til mestring for økt læringseffekt, da dette er intensjonen ved bruk av nivådifferensierte oppgaver. Elevenes medvirkning av opplæringen blir likevel nevnt, fordi bruk av nivådifferensierte lærebøker tilrettelegger for at elevene selv kan velge matematikkoppgavene.

2.1.1 Differensiering

Et tiltak for å tilpasse opplæringen er å differensiere undervisningen. Ongstad (1979, s. 110) definerer differensiering som: «en bevisst forskjellsbehandling til beste for hver enkelt, og som ikke fører til at en opphever det positive som binder mennesker sammen, eller at en skjuler eller beskytter det negative som setter skiller mellom dem» (referert i, Mathiassen, 2009, s. 125). I læreplanen presiseres det at hensikten med differensiering er å bidra til at elevene får muligheten til å oppnå optimal læring. Dette forutsetter at opplæringen er tilpasset mestringsnivået til eleven (Utdanningsdirektoratet, u.d.a, s. 23).

Utdanningsdirektoratet uttaler at differensiering oppnås gjennom variert undervisning: «Tilpasset opplæring for den enkelte elev kjennetegnes ved variasjon i bruk av lærestoff,

arbeidsmåter, læremidler, samt variasjon i organisering av og intensitet i opplæringen» (2012, s. 5). Det skiller mellom pedagogisk og organisatorisk differensiering (Solvang, 1992, s. 182). Organisatorisk differensiering foregår ved å dele elevene i grupper etter interesse, nivå eller evner. Et eksempel, som ble nevnt i innledningen, er kursplansystemet der elevene ble inndelt i grupper etter ferdighetsnivå (Kunnskapsdepartementet, 2011). Organisatorisk differensiering er den eldste formen for differensiering, men den brukes ikke like mye i dag (Mathiassen, 2009, s. 128).

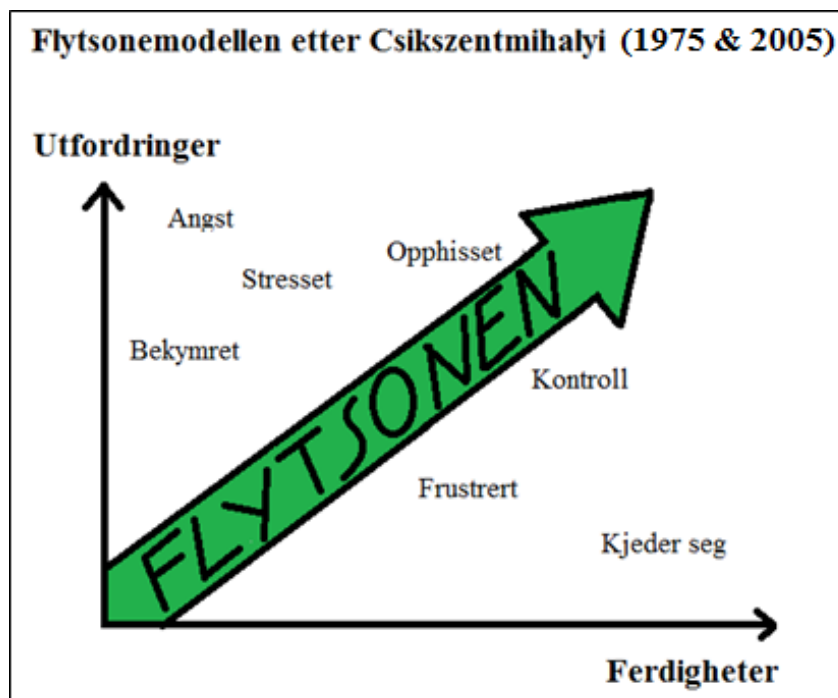
I dag er det et økt fokus på pedagogisk differensiering. Pedagogisk differensiering ble innført som obligatorisk i Mønsterplan 1974, og iverksettes ved å differensiere undervisningen i en samlet klasse (Imsen, 1997, s. 275-276). Denne formen for differensiering ble ikke vektlagt før i Reform 1997, noe som antageligvis skyldes mangel på konkrete differensieringstiltak. Årsaken til et større fokus på pedagogisk differensiering ligger i et ønske om en felles skole, samt et endret læringssyn. Dagens læringssyn går i større grad ut på at læring oppstår i fellesskap (Mathiassen, 2009, s. 127-128).

Pedagogisk differensiering skiller mellom faglig og metodisk differensiering. Metodisk differensiering kjennetegnes ved å presentere faget ved bruk av forskjellige metoder. For eksempel kan læreren variere undervisningen ved å bruke praktiske øvelser eller teoretiske tilnærminger. Faglig differensiering iverksettes ved å variere fagstoffet, enten etter nivå eller tempo (Solvang, 1992). Imsen (1997) presenterer tempodifferensiering som foregår ved å tilpasse undervisningen slik at elevene bruker forskjellig tid til å nå det samme målet. Hensikten med en slik form for differensiering er at elevene får den treningsmengden som trengs for å oppnå mestring. Læreren kan også tilrettelegge undervisningen ved å variere vanskelighetsgraden, kalt nivådifferensiering (Imsen, 1997, s. 277-278). For eksempel kan læreren variere det faglige nivået under muntlig kommunikasjon, eller ved å tilrettelegge undervisningen slik at elevene arbeider med oppgaver som er tilpasset eget ferdighetsnivå (Mathiassen, 2009).

2.1.2 Nivådifferensiering

«Flytsonemodellen», lansert av Csikszentmihalyi (1975), kan brukes til å illustrere behovet for nivådifferensierte matematikkoppgaver (Mathiassen, 2009, s. 129). Modellen viser i hvilken sone eleven bør befinne seg i for å oppnå «flyt», se figur 1. Det fins flere ulike

versjoner av «flytsonemodellen». I denne oppgaven ble det valgt å ta utgangspunktet i den originale modellen, for deretter å påskrive konsekvenser Csikszentmihalyi (2005) mener eleven kan få ved å havne utenfor sin egen flytzone (se figur 1).



Figur 1 Flytsonemodellen etter Csikszentmihalyi (1975, referert i Mathiassen, 2009) og Csikszentmihalyi (2005)

Csikszentmihalyi (2005, s. 46) forklarer «flyt», oversatt fra «flow», som en *optimal opplevelse* som oppstår i en positiv situasjon. Elever vil da føle glede og engasjement, glemme følelsen av tid og rom, og oppmerksomheten rettes mot den bestemte aktiviteten. «Flytsonemodellen» tar utgangspunkt i relasjonen mellom oppgavens utfordring og elevens ferdigheter (Ibid.) Mathiassen (2009) forklarer at elevene befinner seg innenfor «flytsonen» når matematikkoppgavens utfordring er tilpasset elevens ferdigheter, noe som bidrar til optimal læring og økt mestringsopplevelse (s. 129-130). En sterkere grad av «flyt» erverves dets lengre oppover i flytsonen elevene befinner seg. De største mestringsopplevelsene oppstår når elevene arbeider med oppgaver som er så utfordrende at de selv tror det ikke er mulig å prestere. Det er slike oppgaver som bidrar til et økt læringsutbytte (Botten, Daland, & Dalvang, 2008).

Elever med lavt ferdighetsnivå og som får tilpasset oppgavenivået etter flytsonemodellen «befinner» seg i den nederste delen av flytsonen. Csikszentmihalyi (2005, s. 43-44) skriver at

elevene ikke nødvendigvis opplever en *optimal* følelse av «flyt», selv om oppgavenivået er tilpasset eleven. Dersom egne ferdigheter og oppgavens vanskelighetsgrad oppfattes som lave, kan en forholde seg mer likegyldig til oppgavene, dette til tross for at de befinner seg innenfor sin egen flytsone (Csikszentmihalyi, 2005, s. 43). Det vil derfor være viktig å hjelpe elevene oppover i sonen, slik at en positiv opplevelse av arbeidsprosessen kan forekomme. Dette oppnås ved at elevene tilegner seg høyere ferdigheter, samtidig som oppgavene blir mer utfordrende.

Mathiassen (2009, s. 130) bemerker at elever som arbeider med lite krevende aktiviteter ender under sin egen flytsone. Resultatet blir at elevene får et lavt læringsutbytte. I tillegg blir elevene frustrerte og kan oppfatte matematikkoppgavene som kjedelige (Mathiassen, 2009, s. 130). Elever med høy kontroll over aktiviteten oppnår ikke «flyt» om oppgavens vanskelighetsgrad er for lett. Kontroll gir en tilfreds følelse og oppfattes som en positiv tilstand. Hvis elevene ikke arbeider med mer utfordrende oppgaver kan det føre til at elevene mister konsentrasjonen og energinivået (Csikszentmihalyi, 2005, s. 46).

Videre påpeker Mathiassen (2009) at arbeid med matematikkoppgaver som er for krevende for elevene vil resultere i at de «havner» ovenfor sin egen flytsone. Det blir derfor vanskelig å oppnå læring, fordi elevene ikke har de tilstrekkelige ferdighetene som skal til. Svake elever vil derfor føle seg bekymret, stresset og kan oppleve angst (Mathiassen, 2009, s. 130). Csikszentmihalyi (2005) poengterer at hvis oppgavenivået blir for høyt må det reduseres, da slike følelser kan resultere i at elevene blir motløse og derfor gir opp situasjonen. Elever med et over gjennomsnittlig ferdighetsnivå og som arbeider med oppgaver som er litt for utfordrende, vil ifølge modellen oppnå en «opphisset tilstand». Det betyr at de er aktive, fokusert, og involvert i aktiviteten. For at elevene skal «bevege» seg inn i «flytsonen», uten at oppgavenivået senkes, må de øke innsatsen og/eller tilegne seg mer kunnskap (Csikszentmihalyi, 2005, s. 45-46).

Differensieringstiltak

I Norge inndeles elevene etter trinnets gjennomsnitt (Imsen, 1997, s. 275). Av den grunn vil det være problematisk å velge like matematikkoppgaver til klassen, ettersom elevene befinner seg på forskjellige nivåer. Valg av oppgaver er kun formålstjenlig når oppgavene er tilpasset elevenes evner og forutsetninger (Utdanningsdirektoratet, u.d.a, s. 23).

En matematikkbok bestående av oppgaver med lik vanskelighetsgrad vil dermed være lite hensiktsmessig, for bruk av pedagogisk differensiering. Et differensieringstiltak som læreboken *Grunntall* benytter, er å markere matematikkoppgavene med ulike vanskelighetsnivåer. Fargene blå, rød og grønn, står for henholdsvis lett -, middels -og vanskelige oppgavenivå. Lærebokforfatterne begrunner bruk av markeringen med at «elevene våre sier at de vil se hvilke oppgaver de må få til for å komme opp på et høyere nivå» (Grunntall, u.d., s. 1). Boken har i tillegg laget forslag til nivådifferensierte matematikkprøver, der elevene selv kan velge oppgavenivå (Grunntall, u.d., s. 1). Det er flere som mener at elevenes matematikkprøver bør inneholde oppgaver med forskjellige vanskelighetsnivå, slik at elevene selv kan velge oppgaver for å vise hva de kan (Botten, Daland, & Dalvang, 2008).

2.1.3 Elevmedvirkning i opplæringen

I læreplanen står det at tilpasset opplæring også skal bidra til at elevene får muligheten til å ta del i planlegging, gjennomføring og vurdering av opplæringen. På denne måten vil de tilegne seg kunnskap om egne evner og forutsetninger knyttet til læring av faget. Dette hjelper elevene til å bli i stand til å ta egne valg (Utdanningsdirektoratet, u.d.a, s. 22-23). I tillegg presiseres det, i den generelle delen av læreplanen, at en «god undervisning skal gi elevane erfaring for å lukkast i sitt arbeid, gi tru på egne evner og utvikle ansvar for eiga læring og eige liv» (Utdanningsdirektoratet, u.d.b).

For å oppfylle kravene om at elevene skal ha medvirkning i opplæringen kan et differensieringstiltak være å gi arbeidsplaner til elevene. På denne måten må eleven planlegge og velge oppgaver selv. I læreplanen poengteres det at bruk av arbeidsplaner kun er formålstjenlig for de elevene som er i stand til å organisere læringsarbeidet. Dette blir begrunnet med tidligere forskning som påpeker at det er elever som får *for* mye ansvar for egen læring (Utdanningsdirektoratet, u.d.a, s. 24).

2.2 Mestringsforventning

Som nevnt i innledningen legges det vekt på mestringsforventninger knyttet til bestemte matematikkoppgaver. Dette ble begrunnet med at det ofte er de spesifikke mestringsforventningene som er situasjonsbestemte, og som derfor har innvirkning på

elevenes atferd (Teigen, 1997, s. 444). Selvvurdering, som handler om elevens egen vurdering av generelle egenskaper, vil derfor ikke være i fokus, etter som det har mindre betydning for atferden (Pajares & Kranzler, 1995).

2.2.1 Definisjon av mestringsforventning

Innenfor forventningstradisjonen er Banduras sosial-kognitive læringsteori om «self-efficacy» anerkjent (Evans, 1997, s. 327-328). I denne oppgaven oversettes «self-efficacy» til mestringsforventning. Bandura definerer mestringsforventning som «beliefs in one's capabilities to organize and execute the courses of action required to produce a given attainments» (1997, s. 3). Mestringsforventning kan derfor forstås som ens oppfatning om å mestre en spesifikk oppgave eller aktivitet. Videre påpeker Bandura viktigheten av å ha positiv tro på egen mestring: «beliefs of personal efficacy constitute the key factor of human agency» (Bandura, 1997, s. 3). Det vil altså være avgjørende at eleven har den oppfatningen om det er mulig å påvirke læringssituasjonen ved å utføre bestemte handlinger.

I følge Bandura kan mestringsforventningene variere langs flere forskjellige dimensjoner (1997, s. 42). Som nevnt i innledningen kan forventningen variere i generalitet, som beskriver hvor generelle eller spesifikke mestringsforventninger en har. Mestringsforventningen kan også variere i styrke, som beskriver mestringsforventningen på en skala fra lav til høy. I tillegg presenterer Bandura en tredje dimensjon som han kaller nivå (1997, s. 42-43). Det betyr at oppgavens vanskelighetsgrad kan påvirke grad av mestringsforventningen i forhold til oppgavenivået.

2.2.2 Kilder til mestringsforventning

I følge Bandura er det fire kilder som gir informasjon til de kognitive prosessene som bygger og strukturerer mestringsforventningen. Kildene er: (1) egenopplevde mestringserfaringer, (2) informasjon via andres mestringserfaringer, (3) verbal overtalelse og (4) den fysiologiske og affektive tilstanden (1997, s. 79). Personer som møter en bestemt oppgave eller situasjon vil vurdere informasjon fra kildene, for deretter å bedømme om de er i stand til utføre handlingen eller oppgaven.

Bandura (1997, s. 80) påpeker at *egenopplevde mestringserfaringer* er den mest

betydningsfulle kilden som bygger mestringsforventningen, da disse erfaringene ofte oppleves som ekte for personen. I en skolesammenheng kan elever som arbeider med matematikkoppgaver konstruere eller rekonstruere egne mestringsforventninger i forhold til egne prestasjoner på oppgaver. Elever som over en lengre periode erfarer at de ikke mestrer oppgavene vil oppleve fremtidige oppgaver som en trussel. Dette vil resultere i at elever unngår krevende oppgaver. Motsatt kan elever som alltid oppnår en rask mestring under oppgaveløsning forvente at resultatet på neste oppgave også skal komme like lett. En konsekvens av dette vil være at når elever opplever å «stå fast» ved en lignende oppgave, vil de føle seg motløse. En stabilisering av egenopplevde mestringserfaringer skjer ved at eleven, over tid, mestrer utfordrende oppgaver som samsvarer med egne ferdigheter (Bandura, 1997, s. 80).

Prestasjon alene er ikke nødvendigvis tilstrekkelig for å bygge egneopplevde mestringserfaringer. Bandura (1997, s. 81-82) lister opp flere faktorer som påvirker kilden, hvilket er personens forestilling om egne evner, oppgavens oppfattet vanskelighetsgrad, innsats, grad av selvstendighet, arbeidsforhold, tidligere prestasjonsmønster, samt hvordan de kognitive erfaringene organiseres og restruktureres i hukommelsen. Elever som arbeider med matematikkoppgaver der vanskelighetsgraden oppleves som lett, vil nødvendigvis ikke ha behov for å revurdere sine mestringssevner. Mestringsforventningen konstrueres først i møte med krevende oppgaver, ettersom et økt behov for informasjon bidrar til en høyere forventning (Bandura, 1997, s. 81-82).

Bandura (1997, s. 86) hevder at det ofte kan være vanskelig å bedømme ferdighetsnivået på matematikkoppgavene en mestrer, og det vil derfor være nødvendig å tilegne seg informasjon gjennom *andres mestringserfaringer*. Med andre ord oppstår det ofte en sosial sammenligning innad i grupper. Gruppens nivå og karakteristika får derav betydning for om de kan indentifisere seg med hverandre (Bandura & Jourden, 1991; Wood, 1989, ref. i Bandura, 1997, s. 87). Det betyr at eleven må gjenkjenne seg med andre personer for å få en optimal økning av mestringsforventningen (Bandura, 1997, s. 87). For eksempel vil observasjon av en klassekamerat som presterer på en matematikkoppgave, gi økt mestringsforventning til de elevene som er på samme faglige nivå, og/eller for de elevene som opplever at de har lik personlighetsbeskrivelse.

I tillegg kan mestringsforventning konstrueres gjennom *verbal overtalelse* i realistiske situasjoner. Men dersom eleven allerede har gitt opp troen på mestring, vil ikke positive kommentarer *alene* være tilstrekkelig for å øke mestringsforventningen (Bandura, 1997, s. 101). *Hvem* som oppmuntrer eleven vil være en avgjørende faktor for å tilegne seg en større mestringsforventning. Dersom læreren gjentatte ganger gir eleven urealistiske forventninger om å mestre, vil ikke en slik oppmuntring bidra til en økt mestringsforventning. Resultatet blir at overtalelsen undervurderes, fordi lærerens troverdighet har blitt redusert (Bandura, 1997, s. 101).

Fysiologiske og affektive tilstander er den fjerde kilden som Bandura (1997, s. 106) hevder påvirker konstrueringen av mestringsforventningen. Spesielt ved stressede og krevende situasjoner vil bruk av informasjon om egen fysiologiske tilstand vurderes. Elever som kontrollerer stressreaksjonene som oppstår vil ikke oppleve frykt som en faktor som er med på å redusere mestringsforventningen. Elever som derimot ikke klarer å kontrollere dette, vil for eksempel under eksamen, oppleve situasjoner som truende. Den affektive tilstanden vil også påvirke mestringsforventningen. Et eksempel vil være elever som føler seg i godt humør eller opplever en god dagsform. Den økte mestringsforventningen skyldes at oppmerksomheten er forbedret, noe som igjen påvirker de kognitive prosessene (Bandura, 1997, s. 106).

2.2.3 Viktigheten av mestringsforventning

I dette underkapitlet vil mestringsforventningens betydning være i fokus. Pajares og Miller (1995) hevder at det er viktig med en positiv mestringsforventning i møte med oppgaver som er vanskelige. Dette blir begrunnet med at forventningen påvirker utholdenhet, innsats og valg av oppgave. Zimmerman og Martines-Pons (1990) fastslår også viktigheten av høy mestringsforventning. De viser til en undersøkelse der mestringsforventningene hadde en signifikant sammenheng med elevprestasjoner i matematikk. I følge Bandura & Locke vil en positiv mestringsforventning være vesentlig:

They affect whether individuals think in self-enhancing or self-debilitating ways, how well they motivate themselves and persevere in the face of difficulties, the quality of their emotional well-being and their vulnerability to stress and depression, and the choices they make at important decisional points (2003, s. 87).

Dette innebærer at mestringsforventningen har en viktig betydning for de affektive -, kognitive -, motivasjonelle - og selektive prosessene (Bandura, 1997, s. 116). Med andre ord har mestringsforventningen betydning for elevens følelser, tanker, motivasjon, samt de valgene som tas.

De affektive prosessene påvirkes av mestringsforventningen. For eksempel kan forventningen om å mestre en matematikkoppgave påvirke om eleven får angst og/eller blir stresset under oppgaveløsning. Elever som ikke har tro på at de mestrer, og derfor ser på situasjonen som truende, kan oppleve angst og stressreaksjoner, som igjen påvirker prestasjonene (Bandura, 1997, s. 137).

Gjennom de kognitive prosessene påvirker mestringsforventningen optimistisk tankegang. Bandura hevder at personer med høy mestringsforventning har en tendens til å strukturere livet med positive mål (1997, s. 116). Graden av mestringsforventning har derfor betydning for hvor høye mål elever setter seg, samt hvor forpliktet de forholder seg til målet. Elever med høy mestringsforventning vil derfor ha større ambisjoner i faget. I tillegg vil de føle seg forpliktet til å nå målet (Bandura & Wood, 1989; Locke & Latham, 1990, referert i Bandura, 1997, s. 117).

Elevens mestringsforventning har også betydning for prestasjonene, ettersom de påvirkes kognitivt gjennom motivasjonelle prosesser (Bandura, 1997, s. 123). Innenfor psykologien kan motivasjon forstås som drivkraften til å utføre handlinger (Solvang, 1992, s. 213). Kognitiv motivasjon inndeles ofte i tre ulike teorier: forventningsverditeori, målteori og attribusjonsteori (Bandura, 1997, s. 122). I forventningsverditeorien vil «outcome expectation», forventningsutfallet, være en viktig motivasjonsfaktor. Høy forventning om at handlingen som utførers bidrar til et ønsket utfall, gir økt motivasjon til å gjennomføre handlingen (Bandura, 1997, s. 125). Bandura (1997) mener at motivasjonen påvirkes både av mestringsforventningen og forventningsutfallet. Atkinson (1958, referert i Teigen, 1997, s. 207) forklarer teori om «prestasjons-motivasjon» med at atferd påvirkes av *forventing* og *verdi*, og at det har innvirkning på hvilke oppgaver som velges. Det betyr at handlingen påvirkes ut fra egen vurdering av sannsynligheten for å lykkes, samt hvilken verdi utfallet får for eleven. Under valg av oppgave vil oppgavens verdi påvirkes av sjansen for å prestere. Dette fører til at oppgaven kan få større verdi for eleven hvis sjansen for å få riktig svar er

mindre, sammenlignet med å velge en oppgave der det er stor sannsynlighet for å prestere (Teigen, 1997, s. 207).

«Cognized goals», som inngår i målteori, handler om elevenes målsetting. Bandura påpeker at utfordrende mål kan gi økt motivasjon (1997, s. 128). Personer med høy mestringsforventning vil, i møte med oppgaver som er utfordrende, se på situasjonen som en mulighet for å oppnå resultater. Motsatt vil personer med lav mestringsforventning vurdere situasjonen som en fare for nederlag (Krueger & Dickson, 1994, referert i Bandura, 1997, s. 117).

Attribusjonsteori handler om en persons forklaring på hvorfor en bestemt prestasjonen førte til et spesifikk utfall (Skaalvik & Skaalvik, 2005, s. 107). Årsakssammenhengene kan inndeles i indre- og ytre mekanismer. De indre mekanismene knyttes til personen selv, og er innsats, strategi og evner. Flaks, vanskelighetsgraden på oppgaven, undervisningens kvalitet og dagsform er eksempler på ytre mekanismer. I tillegg inndeles forklaringsmekanismene i kontrollerbare -og ukontrollerbare mekanismer. Ofte vil innsats, strategi og oppgavens vanskelighetsgrad vurderes som kontrollerbare, i motsetning til evner, flaks, undervisningens kvalitet og dagsform, som vurderes som ukontrollerbare (Skaalvik & Skaalvik, 2005, s. 107). Disse forklaringsmekanismene påvirkes av elevers mestringsforventninger og forståelse av hvilke mekanismer som er kontrollerbare.

Jenter har en tendens til å refererer til manglende evner ved nederlag, og gutter på den andre siden forklarer årsaken med en lav innsats (Imsen, 2008, s. 458). Dette kan sees i sammenheng med Pajares og Miller (1994) som hevder at gutter har høyere mestringsforventning enn jenter i matematikkfaget til tross om de er på samme ferdighetsnivå. En forklaringen kan være at jenter ofte undervurderer sin matematikkunnskap. En lavere selvtillit kan være grunnen til at de velger oppgaver som er mindre krevende enn gutter, til tross for at de er på samme faglige nivå. Denne kjønnsforskjellen viser seg å øke med elevenes alder (Pajares & Miller, 1994, s. 196).

Oppsummert påvirker de affektive -, motivasjonelle -og kognitive prosessene elevenes mål, utholdenhet og innsats. Dette påvirker den selektive prosessen som får betydning for valg av oppgave (Bandura, 1997, s. 161). Elever med høye mestringsforventnien vil, ifølge Meyer (1987) være elever som velger oppgaver med et høyt ferdighetsnivå. Bandura legger til at

disse elevene vil være mer utholdende under oppgaveløsning (1997, s. 160). Til sammen vil disse prosessene ha innflytelse på elevenes sjanse til å oppnå ønsket resultat (Bandura, 1997). Disse prosessene har relevans for mitt studie, ettersom *mestringsforventningen* påvirker den *reelle mestringen*. Resultatene på matematikkoppgavene vil altså være avhengig av om eleven blir stresset, får angst, grad av utholdenhet og innsats (Skaalvik & Skaalvik, 2005, s. 93). Den *reelle mestringen* påvirker elevens *opplevde mestring*. Opplevd mestring konstrueres gjennom elevens attribusjon, kriterier for mestring og oppgavevalget. Opplevelsen av mestring blir for eksempel styrket hvis eleven i ettertid fikk høre at oppgaven var vanskelig (Ibid.). Av den grunn kan det imidlertid tenkes at oppgavens nivåmarkering har hatt betydning for elevens opplevde mestring. Elevenes tidligere mestringserfaringer kan, som nevnt, konstruere fremtidige mestringsforventninger, og som senere kan bli benyttet ved liknende anledninger (Bandura, 1997). Denne sirkulære prosessen har betydning for elevenes oppgavevalg samt prestasjon, og det argumenteres derfor for at oppgavene må tilpasses elevenes forutsetninger (Skaalvik & Skaalvik, 2005, s. 94).

2.3 Resultater fra tidligere empiriske studier

I Norge er det utført relativt lite forskning på bruk av nivådifferensierte matematikkoppgaver i lærebøker. Dette er overraskende med tanke på at det er mange lærebøker, i den norske grunnskolen, som opererer med å markere matematikkoppgavens vanskelighetsgrad med ulike fargekoder.

Som nevnt er det flere studier som viser sterk sammenheng mellom mestringsforventning og prestasjoner i matematikk. Et annet eksempel er et studie som ble gjennomført av Collins (1982, referert i Bandura, 1993, s. 119). Collins målte elevenes kunnskap etter en vurdering av noen matematikkoppgaver som elevene hadde fått beskjed om å løse. Etterpå ble elevene inndelt i tre ulike grupper basert på deres kunnskapsnivå. Videre ble hver gruppe inndelt i to undergrupper; lav -og høy mestringsforventning. Elevene fikk deretter beskjed om å løse noen matematikkoppgaver. En sammenligning av de to undergruppene viste at elevene som hadde høyest mestringsforventning klarte flest oppgaver (Bandura, 1993). Med andre ord tydet resultatene på at elever som var på samme faglig nivå i matematikk, hadde større sjanse for å oppnå gode resultater i faget hvis de hadde en høy mestringsforventning. Undersøkelsen antydte at årsaken til at det var elever som fikk dårlige resultater på oppgavene var lave

ferdigheter i matematikk og/eller lav mestringsforventning (Bandura, 1993, s. 119).

Forskjellen mellom masteroppgavens undersøkelse og Collins studie, er at denne undersøkelsen antyder at *nivåmarkeringen* har en effekt på resultatene, gitt at elevene har samme terminkarakter. Collins resultater viser derimot at *mestringsforventningen* har betydning for resultatene, til tross for at elevene er på samme faglige nivå. En kan derfor undres om det er en sammenheng mellom oppgavens nivåmarkeringen og elevens mestringsforventning.

2.4 Hypoteser

Avsnittet vil anvende de teoretiske perspektivene samt resultatene fra tidligere forskning for å utforme hypoteser som testet senere i den empiriske analysen. Hypotesene vil også baseres på egne forventninger og erfaringer som matematikklærer i skolen.

Oppsummert er målet med studiet å undersøke om oppgavens nivåmarkering har en effekt på elevenes oppgavevalg, samt om nivåmarkeringen påvirker elevenes oppgaveløsning. Det forekommer ikke en måling av elevenes mestringsforventning, men på bakgrunn av overnevnt teori tas det for gitt av mestringsforventning har en betydning for denne undersøkelsen. En antakelse er at oppgavemarkeringen aktiverer kunnskap basert på tidligere erfaringer og/eller elevens mestringsforventninger, og at dette påvirker oppgavevalget samt prestasjonene. I tillegg antas det at elevens tidligere terminkarakterer i faget påvirker mestringsforventningen, som således har en påvirkning på valg og prestasjon.

Som nevnt påpeker Bandura (1997) at elevens mestringsforventning påvirker de selektive prosessene. Det betyr at elever med lav forventning om å mestre ofte unngår krevende oppgaver, og motsatt. På bakgrunn av dette er det trolig å anta at *hvis* elevene faktisk fikk oppgaver bestående av *forskjellige* vanskelighetsnivåer, ville elevene ha valgt ut fra motivasjonelle prosesser.

I denne undersøkelsen er matematikkoppgavenes vanskelighetsgrad tilnærmet like, selv om matematikkoppgavene er påført en tilfeldig vanskelighetsnivåer. Det er usikkert hvordan elevene velger når oppgavene har nokså lik vanskelighetsgrad, ettersom det ikke fins konkret

forskning på området. Erfaringer fra praksisoppholdet tydet på at elevene selv ikke vurderte oppgavens vanskelighetsgrad, og at de derfor vurderte oppgavenivået ut fra informasjon om nivåmarkeringen. Hvis elevene som deltar på undersøkelsen ikke vurderer oppgavens faktiske vanskelighetsgrad, vil trolig den selektive prosessen påvirkes av om de *tror* de vil klare oppgaven, basert på informasjon om nivåmarkeringen, samt elevens vurdering av oppgavens verdi. Disse antakelsene støttes av teori om *prestasjons-motivasjon*, som sier at valg av oppgave påvirkes av forventning og verdi (Teigen, 1997, s. 207). På bakgrunn av dette antas det at elevene med høy mestringsforventning velger oppgaven med høy nivåmarkering, og motsatt. Forskning viser at mestringsforventningen varierer med elevenes karakterer, og at elevene med høyest matematikkarakterer oftest har større forventning om å mestre (Zimmerman & Martinez-Pons, 1990). Av den grunn er en videre antakelse at effekten av nivåmarkeringen vil variere med forskjellige terminkarakterer. På bakgrunn av dette er følgende hypotese formulert:

Hypotese 1: Matematikkoppgavens nivåmarkering og elevens terminkarakter har en effekt på oppgavevalget

Neste hypotese begrunnes ut fra tidligere forskning, utført av Collins (1982), som viser at selv om elevene er på samme faglig nivå, vil prestasjonen variere i trå med elevenes mestringsforventninger (referert i, Bandura, 1993). Det antas derfor at to elever med lik terminkarakter vil prestere ulikt, da forskjellig nivåmarkering kan påvirke forventningene om å klare oppgaven. I tillegg antydte matematikktesten fra praksisoppholdet at nivåmarkeringen påvirket elevenes resultater. Trolig vil resultatene, også her, variere med elevens terminkarakter. Dette blir begrunnet med at det er større sannsynlighet for at eleven presterer hvis det faglige nivået til eleven er høyt. På bakgrunn av dette er følgende hypotese formulert:

Hypotese 2: Matematikkoppgavens nivåmarkering og elevens terminkarakter har en effekt på oppgaveprestasjonen

3 Data og metode

Dette kapitlet beskriver valg og vurderinger av forskningsmetoden, samt en redegjørelse av datainnsamlingen. I tillegg blir analysemetoden present, der også statistiske begreper defineres.

3.1 Troverdighet

For å vurdere troverdigheten av forskningsrapporten blir validiteten og reliabiliteten av datamaterialet drøftet. Reliabilitet angår dataens pålitelighet og nøyaktighet, og gir en vurdering av hvorvidt tilfeldige målingsfeil har påvirket datamaterialet (Kleven, 2011c, s. 89). Validitet vedrører datamaterialets relevans og gyldighet, og gir en samlet vurdering av datakvaliteten (Kleven, 2011a, s. 23).

Validiteten vurderes i tre deler, «begrepsvaliditet», «indre validitet» og «ytre validitet». Begrepsvaliditet defineres som «grad av samsvar mellom begrepet slik det er definert teoretisk, og begrepet slik vi lykkes med å operasjonalisere det» (Kleven, 2011c, s. 86). En vurdering av begrepsvaliditeten foretas etter en redegjørelsen av oppgavematerialet og prosedyren for undersøkelsen. Indre validitet handler om hvorvidt datamaterialet kan forklare årsaksammenhenger (Skog, 2007, s. 107). En forklaring av sammenhengen mellom årsak og virkning blir gjort ved å vurdere forskningsdesignet. Ytre validitet beskriver vurderingen av dataresultatene gyldighetsområde, både *for* personer og *i* situasjoner (Kleven, 2011e, s. 123-134). Den ytre validiteten blir drøftet i oppgavens avslutning.

3.2 Forskningsdesign

I pedagogisk forskningssammenheng skilles det ofte mellom kvantitativ og kvalitativ metode. I følge Kleven (2011a, s. 19-20) kjennetegnes kvantitativ datainnsamlingsmetode først og fremst med at utvalget er stort, og at distansen mellom forskeren og informanten er større enn ved en kvalitativ metode. Et annet kjennetegn er at kvantitativ tilnærming ofte fokuserer på kausale årsakssammenhenger, i motsetning til en kvalitativ metode, som ønsker en helhetsvurdering av enkeltsammenhenger. Begge tilnærmingene har sterke og svake sider, og det er derfor blitt mer vanlig å bruke begge tilnærmingene, såkalt «mixed methods» (Kleven, 2011a, s. 19-20). På denne måten kan forskeren få et rikere datamateriell som kan analyseres

og tolkes ved bruk av flere analysemetoder.

Som nevnt tidligere er formålet med oppgaven å måle effekten av nivåmarkerte oppgaver i matematikkfaget, både når det gjelder prestasjon og valg av oppgave. På bakgrunn av dette ble følgende problemstilling formulert: *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes valg av oppgave og prestasjon på oppgaven?*

Etter en vurdering av studiets formål, problemstilling og tidsramme, ble det vurdert at en kvantitativ tilnærming var tilfredsstillende. Valget begrunnes med at metoden gir slutninger om mange elever og eventuelle slutninger av generell karakter, og bidrar derfor til at problemstillingen besvares (Cohen, Manion, & Morrison, 2011, s. 256-257). For å svare på problemstillingen ble det valgt et eksperimentelt forskningsdesign. Det betyr at det ble gjennomført en empirisk test bestående av en eksperimentgruppe og en kontrollgruppe. Studien er begrenset til en tverrsnittstudie, og variablene som ble målt gir derfor *kun* et øyeblikksbilde av utvalget (Cohen, Manion, & Morrison, 2011, s. 256-257).

Forskningsdesignet ble valgt da det bidrar til å styrke den indre validitet. En vurdering av validiteten foretas fordi det er ønskelig å gi en intensjonal forklaring av resultatene. «En intensjonal forklaring av en handling Y sier at handlingen ble foretatt fordi den var tilpasset handlingens oppfatning og ønsker» (Føllesdal, Walløe & Elster 1986, s. 156, referert i Kleven, 2011d, s. 106-107). En intensjonal forklaring gir altså svar på intensjonen bak handlingen. Fordelen med et eksperimentelt forskningsdesign er at eksperimentgruppen kan sammenlignes med kontrollgruppen (2011d, s. 115-116). En sammenligning mellom gruppene er fordelaktig siden kontrollgruppen kontrollerer for andre variabler som påvirker effektvariabelen. Dette er med på å styrke den indre validiteten. Den indre validiteten styrkes også ettersom elevene hadde like stor sjanse til å delta i kontrollgruppen som eksperimentellgruppen (Skog, 2007, s. 27).

«Binomisk logistisk regresjon» ble valgt som analysemetode. Analysemetoden styrker den indre validiteten, siden den kan regulere for andre variabler som kan påvirke effektvariabelen, samt teste om det forekommer mellomliggende variabler. Når det er sagt, er det viktig å nevne at analysemetoden kun kontrollerer for variablene som er tatt med i analysen, og ikke eventuelle andre bakenforliggende variabler (Skog, 2007, s. 107-108).

3.3 Utvalg

I denne oppgaven utgjør populasjonen ungdomsskoleelever fra 9. trinn ved praksisskoler i Oslo, som bruker nivå-differensierte lærebøker. Trinnet ble valgt siden elevene ofte har mer erfaring fra nivå-differensierte lærebøker, sammenlignet med elever fra 8. trinn. I tillegg er 9. trinn mindre travle enn 10. trinn, da de ikke har avsluttende eksamener.

Utvalgsprosedyren ble gjennomført ved å sende en forespørsel på e-mail til alle praksisskolene i Oslo. Av de 17 ungdomsskolene var det to skoler som responderte positivt til å delta på undersøkelsen. I utgangspunktet var det ønskelig med et tilfeldig utvalg, ettersom et tilfeldig utvalg kunne styrke datamaterialets gyldighetsområde med hensyn til generalisering (Kleven, 2011e). På bakgrunn av en lav respons fra skolene ble det istedenfor foretatt det som karakteriserer et bekvemmelighetsutvalg. Det betyr at skolene som først meldte seg ble valgt. Skolene som deltok på undersøkelsen var de to skolene som svarte ja på e-posten, samt en skole som ble kontaktet via bekjentskap. Det er derfor ikke et sannsynlighetsutvalg (Kleven, 2011e, s. 130).

Undersøkelsen ble gjennomført av tolv 9. klasser med til sammen 289 elever. Den første ungdomsskolen bidro med seks 9. klasser, totalt 147 elever. På den andre skolen besvarte fire klasser skjemaet, med et samlet antall på 95 elever. Ved den tredje skolen deltok to klasser med til sammen 47 elever. Den første skolen brukte læreboken *Grunntall*, mens de andre ungdomsskolene brukte boken *Tetra* som læremiddel. De 289 elevene ble tilfeldig delt inn i to grupper, en eksperimentgruppe og en kontrollgruppe. Inndelingen var uavhengig av hvilken klasse elevene tilhørte. Forskjellen mellom gruppene og hvordan denne inndelingen foregikk, blir presentert i neste avsnitt.

3.4 Oppgavemateriale og prosedyrer

Datainnsamlingen foregikk ved at elevene fikk tildelt et skjema, der de skulle løse noen matematikkoppgaver, samt svare på noen spørsmål. Materialet ble utarbeidet ved hjelp av to forundersøkelser. Den første undersøkelsen var en kalibrering av matematikkoppgavens vanskelighetsgrad. Dette ble gjort ettersom det var ønskelig at materialet skulle bestå av matematikkoppgaver hvor vanskelighetsnivået på oppgavene var tilnærmet like. Den andre undersøkelsen ble utført for å teste hvordan materialet fungerte i praksis, en såkalt

pilotundersøkelse.

Tidsperioden for utførelsen av undersøkelsen ble lagt til månedsskiftet januar/februar 2014. Dette fordi det var ønskelig å få kjennskap til elevens terminkarakter i matematikkfaget. I tillegg ble innsamlingen begrenset til en kort periode, siden en lengre periode kunne ha påvirket resultatene og bidratt til feilkilder.

3.4.1 Datainnsamling

Innsamlingen av data foregikk først ved å gi elevene informasjon om undersøkelsen. Elevene ble fortalt at undersøkelsen skulle besvares anonymt, individuelt og at ingen hjelpemidler var tillatt – kun penn og papir. Forskerrollen under datainnsamlingen var å være til stede for å svare på eventuelle praktiske spørsmål. Spørsmål relatert til matematikkoppgavene ble ikke besvart. Faglærerne hadde ingen betydelig rolle i undersøkelsen.

Materialet som ble brukt i undersøkelsen bestod av et skjema som var tredelt. Del 1 og 2 besto av matematikkoppgaver og den siste delen inneholdt 10 spørsmål. Til sammen ble det utarbeidet ni forskjellige varianter av samme skjema, selv om elevene fikk *nøyaktig* de samme matematikkoppgavene og spørsmålene. Forskjellen mellom oppgaveskjemaene var at rekkefølgen og nivåmarkeringen varierte. For at undersøkelsen skulle bli så tilfeldig som mulig, ble de ulike versjonene av skjemaene blandet sammen i en bunke, før de ble delt ut i klassene. På denne måten var det tilfeldig hvilke elever som ble tilhørende kontrollgruppen og eksperimentgruppen.

Del 1

I den første delen av skjemaet fikk elevene beskjed om å løse tre matematikkoppgaver, oppgave D – F. Hensikten var å måle om elevene fikk riktig eller galt svar på oppgavene. Eksperimentgruppens tre oppgaver var påført forskjellige nivåmarkeringer, selv om oppgavene hadde tilnærmet samme vanskelighetsgrad. Markeringen var tilfeldig fordelt mellom oppgavene, og varierte innad i gruppen. En tilfeldig nivåmarkeringen betyr at det var noen elever som fikk oppgave D markert «lett», mens det var andre elever som fikk den samme oppgaven markert «middels» eller «vanskelig». I tillegg varierte rekkefølgen på oppgavene, noe som resulterte i at det til sammen var seks forskjellige varianter av skjemaet. Tabell 1 illustrerer hvordan nivåmarkeringen og rekkefølgen på oppgavene varierte. For

eksempel inneholdt skjema 4 (se tabell 2) oppgave D, E og F markert henholdsvis «vanskelig», «lett» og «middels». Videre gir tabellen opplysning om at skjemaet hadde oppgave D plassert sist (nr. 3), oppgave E var plassert først (nr. 1), og oppgave F var plassert i midten (nr. 2).

Tabell 1 Oversikt over de ni ulike versjonene av skjemaene

Skjema:	1	2	3	4	5	6	7	8	9
Oppgave D	Lett nr. 1	Lett nr. 3	Vanskelig nr. 1	Vanskelig nr. 3	Middels nr. 2	Middels nr. 2	nr. 1	nr. 3	nr. 2
Oppgave E	Middels nr. 2	Middels nr. 2	Lett nr. 3	Lett nr. 1	Vanskelig nr. 1	Vanskelig nr. 3	nr. 2	nr. 1	nr. 3
Oppgave F	Vanskelig nr. 3	Vanskelig nr. 1	Middels nr. 2	Middels nr. 2	Lett nr. 3	Lett nr. 1	nr. 3	nr. 2	nr. 1

Selv om det i realiteten kan utarbeides 36 forskjellige varianter av oppgaveskjemaet, valgte en kun å brukes de ni utgavene som vises i tabellen. Dette blir begrunnet med at det ville vært vanskeligere å analysere datamaterialet med så mange versjoner, samt at utvalgsstørrelsen var forholdsvis liten. Et utvalg på 300 elever, med 36 forskjellige oppgaveskjemaer, hadde resultert i åtte elevbesvarelser per skjema.

Ved å la den midterste oppgaven alltid være markert «middels», og samtidig sørge for at alle oppgavene vekselvis var markert som «lett», «middels» og «vanskelig», ble antall nivåmarkerte oppgaveskjemaer begrenset til seks. For å kontrollere effekten av nivåmarkeringen ble det i tillegg utarbeidet tre varianter av skjemaet som inneholdt oppgaver *uten nivåmarkering*, der kun oppgavens rekkefølge varierte. Derav ni skjema. Elevene som tilfeldigvis fikk utdelt skjema 7, 8 eller 9 (se tabell 2) ble altså tilhørende kontrollgruppen.

Del 2

I den andre delen av skjemaet ble eksperimentgruppen bedt om å velge én av tre tilsynelatende nivådifferensierte matematikkoppgaver, oppgave A – C. Kontrollgruppen fikk lik beskjed, hvor oppgavene ikke hadde påført forskjellige vanskelighetsnivåer. Formålet var

å finne ut hvilken oppgave elevene valgte. For å undersøke om nivåmarkeringen hadde en effekt på oppgavevalget ble de forskjellige gruppene sammenlignet. På samme måte som i del 1 varierte nivåmarkeringen og rekkefølgen på oppgavene (se tabell 2).

Del 3

I den siste delen av skjemaet ble elevene bedt om å svare på ti spørsmål ved å krysse av på ett svaralternativ. Det var bare to av spørsmålene som ble brukt som variabler i analysen, spørsmål 1 og spørsmål 10:

1. Hvilket kjønn er du?

Jente

Gutt

10. Hvilken matematikkarakter fikk du på terminkortet? (Sett ett kryss)

Karakter 1

Karakter 2

Karakter 3

Karakter 4

Karakter 5

Karakter 6

Vet ikke

Hensikten med del 3 var å få kjennskap til elevenes kjønn og terminkarakterer. De resterende spørsmålene som blant annet handlet om elevenes tidligere erfaringer ved bruk av nivådifferensierte oppgaver, ble ikke benyttet i den endelige oppgaven. En oversikt over alle spørsmålene fins i vedlegg 2.

3.4.2 Utforming av oppgavematerialet

Ved utformingen av oppgavematerialet ble det foretatt en rekke vurderinger, deriblant valg av spørsmål, rekkefølge på oppgaver og spørsmål, layout og tidsperiode for undersøkelsen. Det ble valgt kvantifiserte spørsmål fordi det var ønskelig å anvende statistikkprogrammet SPSS (Statistical Package for Social Sciences) som analyseverktøy.

Materialets del 1 og del 2

Under utvelgelsen av matematikkoppgavene ble det valgt å bruke tekstoppgaver ettersom oppgavens vanskelighetsgrad ikke skulle gjennomskues ved første øyekast. Det var derfor ikke ønskelig å bruke ferdig oppstilte matematikkstykker. Siden elevenes leseferdighet ikke skulle bli en betydelig feilkilde, var det et poeng at lengden på tekstoppgavene skulle være korte. I tillegg ble det valgt oppgaver med *lik* tekstlengde. Dette var viktig at skjemaene skulle ha samme layout for å unngå at elevene oppdaget at medelevene fikk forskjellige oppgaveskjemaer. Oppgavene fra del 1 og del 2 skulle være uavhengig av hverandre, slik at oppgaveprestasjonen fra del 1 ikke påvirket oppgavevalget i del 2. Det ble derfor valgt matematikkoppgaver fra to ulike temaer, algebra i del 1 og sannsynlighet i del 2. For å få informasjon om elevenes kjennskap til temaene ble faglærerne kontaktet på forhånd. Det ble opplyst om at elevene ikke hadde fått undervisning i sannsynlighetsregning på ungdomsskolen. Av den grunn ble det skrevet en kort introduksjonsdel på skjemaet som elevene fikk beskjed om å lese før de valgte oppgave.

Kalibrering av matematikkoppgavene

For å velge matematikkoppgaver til skjemaet ble det foretatt en kalibrering av matematikkoppgaves vanskelighetsgrad. Hensikten med forundersøkelsen var å velge oppgaver med tilnærmet lik vanskelighetsgrad, og for å unngå en «tak-» og/eller «bunneffekt» på oppgavene. I følge Everitt (2002, s. 62) oppstår det en «takeffekt» hvis det velges for lette oppgaver, noe som kan forårsake reduserte målerverdier. En «takeffekt» på variabelen «prestasjon» ville resultert i høye skårer, ettersom nesten alle elevene ville ha fått riktig svar på oppgaven. Motsatt oppstår en «bunneffekt» ved å velge for vanskelige oppgaver. En kontroll av disse effektene kan gi en større variasjon i gruppen (Everitt, 2002, s. 62).

Det var ca. 100 elever fra 9. trinn, bosatt i Hedmark og Akershus fylke, som kalibrerte matematikkoppgavens vanskelighetsgrad. Skolene som deltok ble kontaktet via bekjentskaper. Kalibreringen foregikk ved å be elevene regne 14 matematikkoppgaver uten nivåmarkering (se vedlegg 1). Deretter ble de seks oppgavene som ca. 60 % av elevene fikk riktig resultat på, valgt til å være med i den opprinnelige undersøkelsen. Tre av oppgavene var sannsynlighetsoppgaver (oppgave A – C), og tre oppgaver var algebraoppgaver (oppgave D – F), se tabell 2.

Tabell 2 Kalibrering av matematikkoppgavene

Oppgaver som ble plukket ut til oppgaveskjemaet. Oppgave A – C inngår i del 2, valg av oppgave, og oppgave D – F inngår i del 1, prestasjon.	Riktig svar (%)	Antall (N)
<u>Oppgave A:</u> I klasse 9 G er det til sammen 27 elever, 11 jenter og 16 gutter. En av elevene er ordenselev. Hvor stor er sannsynligheten for at Ida er ordenselev?	68 %	N = 104
<u>Oppgave B:</u> Per er 23 år og bor i Oslo. Hvor stor er sannsynligheten for at bursdagen til Per er på en søndag i år?	57 %	N = 104
<u>Oppgave C:</u> I en skål ligger det fem kuler, 3 blå og 2 røde. Hanna har bind foran øynene, og trekker en kule opp fra skålen. Hva er sannsynligheten for at Hanna trekker en blå kule?	66 %	N = 53
<u>Oppgave D:</u> Anna fra Bergen er tre år eldre enn søsteren sin. De er 29 år til sammen. Hvor gammel er søsteren til Anna?	51 %	N = 53
<u>Oppgave E:</u> Julie og Torstein er to søsken som bor i Tromsø. Torstein er eldst, og er fire år eldre enn Julie. Til sammen er de 28 år. Hvor gammel er Julie?	65 %	N = 101
<u>Oppgave F:</u> Lise, Lene og Lotte har vært på bærtur. Lise plukket 2 kg mer enn Lene. Lotte plukket dobbelt så mange kg bær som Lise. Lene plukket 3 kg blåbær og 2 kg tyttebær. Hvor mange kg bær plukket Lise?	66 %	N = 101

Materialets del 3

Kleven (2011b) hevder at det er viktig at spørsmålene i en undersøkelse er entydige, har et enkelt språk, ikke er ledende og tar kort tid. På bakgrunn av dette ble materialets lengde forholdsvis kort. Et for langt skjema kunne resultert i at motivasjonen til å fullføre undersøkelsen hadde blitt redusert. I tillegg ble hvert spørsmål i del 3 stilt så enkelt som mulig, og ved påse at det bare ble spurt om en ting av gangen. Videre ble det gitt svaralternativer for å lette forskningsarbeidet, ettersom de kvantifiserte dataene skulle bearbeides i SPSS. En videre begrunnelse er at det var tidsbesparende for eleven, samt at svarene avgis på samme måte (Kleven, 2011b, s. 36-37).

3.4.3 Forskningsetikk

Informasjon fra norsk samfunnsvitenskapelige tjeneste (NSD) tilsier at de forskningsetiske kravene i denne undersøkelsen er tilfredsstillt, ettersom all informasjon om elevene og skolene er behandlet konfidensielt. Elevene som deltok på undersøkelsen fikk utdelt et følgebrev der de fikk en presentasjon av forskeren, undersøkelsens hensikt, samt informasjon om egen anonymitet. Følgebrevet var skrevet forholdsvis kort med tanke på at elevene i størst mulig

grad skulle evne å lese igjennom det. I tillegg var språket forenklet med hensyn til elevenes leseferdigheter. Følgerevet er vedlegg nr. 3.

3.4.4 Pilotundersøkelse

I uke 2 ble det foretatt en pilotundersøkelse ved bruk av oppgaveskjemaet, i fire 9. klasser ved en skole i Nordland. Hensikten med pilotundersøkelsen var å teste skjemaet før det ble brukt i den opprinnelige undersøkelsen, slik at gyldigheten av datamaterialet kunne styrkes. Totalt var det 85 elever som deltok. Ungdomsskolen opererte med læreboken *Nye Mega* som bygger på samme nivåmarkering som *Tetra* og *Grunntall*. Informasjon fra piloteringen tilsa at skjemaet fungerte godt, og at undersøkelsen tok ca. 20 minutter. I tillegg ble det registrert at de samme praktiske spørsmålene ble stilt i alle klassene. Dette var opplysninger som ble tatt med under gjennomføringen av undersøkelsen i Oslo. På denne måten ble undersøkelsen presentert likt i hver klasse, der all praktisk informasjon ble gitt på forhånd. Pilotundersøkelsen førte til noen få korrigeringer før skjemaet ble ferdig.

3.4.5 Variabler og målenivå

Nedenfor forklares variablene og dens målenivå som senere blir brukt i regresjonsanalysen. Variabler som ikke brukes i analysen blir ikke presentert.

Avhengige variabler

Prestasjon: I del 1 ble elevens prestasjon målt ved å bedømme om matematikkoppgavene var riktig eller feil besvart. Variabelen er dikotom fordi den bare har to kategorier, 0 = *feil* og 1 = *riktig* (Eikemo & Clausen, 2012, s. 55). I analysen brukes $Y = 0$ som referanseutfall. Blank besvarelse registreres som feil svar.

Valg av oppgave: I del 2 ble elevenes valg av oppgave målt, og elevenes resultat var derfor ikke av interesse. Den avhengige variabelen er dikotom, og kodes 0 = *ikke valgt* og 1 = *valgt*. I analysen brukes $Y = 0$ som referanseutfall. Blank besvarelse og mer enn to oppgavevalg kodes som ikke besvart.

Uavhengig variabel

Markering: Oppgavens nivåmarkering måles på nominalnivå. Dette forklares med at variabelen består av mer enn to kategorier som ikke kan rangeres, og der avstand ikke er lik mellom kategoriene. Variabelen kodes 1 = *lett markering*, 2 = *middels markering* og 3 = *vanskelig markering*. Siden nivåmarkeringen ikke stemmer med oppgavens faktiske vanskelighetsgrad kan kategoriene ikke rangeres i rekkefølge. Av den grunn er variabelen ikke på ordinalnivå (Eikemo & Clausen, 2012, s. 55), men på nominalnivå. I analysen brukes X = 3 som referansekategori.

Karakter: Elevenes terminkarakter er en variabel som måles på intervallnivå. Dette begrunnes med at variabelen er kontinuerlig og har flere enn to kategorier. I tillegg er det en lik¹ avstanden mellom kategoriene og variabelen har ingen nullpunkt (Eikemo & Clausen, 2012, s. 55). Karakterskalaen strekker seg fra 1 til 6, hvor 1 er laveste og 6 høyeste grad av fagkompetanse. Variabelen kodes 1 = *karakter 1*, 2 = *karakter 2*, ..., 6 = *karakter 6*.

Kontrollvariabler

Som nevnt i teoridelen viser forskning at gutter ofte har en høyere mestringsforventning enn jenter (Pajares & Miller, 1994). I tillegg mener Bandura (1997) at tidligere mestrings erfaringer påvirker forventningen om å mestre. På bakgrunn av dette ble oppgavens rekkefølge og elevens kjønn brukt som kontrollvariabler i analysen, siden variablene kan påvirke resultatene.

Kjønn: Elevenes kjønn kodes 1 = *jente* og 0 = *gutt*. Variabelen er dikotom. I analysen brukes X = 0 som referansekategori.

Rekkefølge: Oppgavens rekkefølge kodes 1 = *først*, 2 = *midterst* og 3 = *sist*. Variabelen er på ordinalnivå, da den kan rangeres. Den midterste rekkefølgen tas ikke med som kontrollvariabel, ettersom den vil ha identisk svarmønster med «middels» markering. I analysen brukes X = 3 som referansekategori.

¹ Karakterskalaen har *egentlig* ikke lik avstand mellom kategoriene. Denne forutsetningen tas, da det gjør analysearbeidet lettere.

Samspillseffekter

En samspillvariabel er en uavhengig variabel som påvirker effekten en annen uavhengig variabel har på den avhengige variabelen (Eikemo & Clausen, 2012, s. 100). Variablene «karakter» og «markering» utgjør en samspillvariabel hvis effekten av årsaksvariabelen «markering» og den avhengige variabelen «valg av oppgave» eller «prestasjon» er avhengig av karakteren (Skog, 2007, s. 52). Samspillvariablene som testes i analysen er «Karakter*Lett», «Karakter*Middels» og «Karakter*Vanskelig», der den siste variabelen brukes som referansekategori.

3.4.6 Begrepsvaliditet

Tidligere i kapitlet ble utformingen av oppgavematerialet presentert. Valgene som ble tatt, for eksempel layout, tekstlengde på oppgavene, matematikktema og tidsperiode, ble gjort for å styrke begrepsvaliditeten. I tillegg ble det utført en kalibreringen av matematikkoppgavene for at måleverdiene skulle variere innad i gruppen.

Tilfeldig -og systematisk målingsfeil kan påvirke begrepsvaliditeten, noe som henholdsvis skyldes relabilitetssvikt og validitetssvikt. I denne undersøkelsen kan elevens dagsform eller flaks/uflaks være eksempler på tilfeldige målingsfeil som reduserer begrepsvaliditeten. Størrelsen på utvalget vurderes som forholdsvis stort, noe som trolig har bidratt til at de tilfeldige målingsfeilene stabiliseres. For eksempel vil elevenes dagsform utjevnes med et utvalg som er stort nok, noe som videre resulterer i at stabilitetsaspektet ved reliabilitet ikke påvirker begrepsvaliditeten betydelig (Kleven, 2011c, s. 88-96). Skjemaet er konstruert slik at besvarelsene kodes som riktig eller feil, valgt eller ikke valgt, samt koding av svaralternativer. Vurdererrelabiliteten vil derfor ikke ha betydning for operasjonaliseringen av begrepene, og bidrar til at de tilfeldige målingsfeilene reduseres (Ibid.).

Som nevnt ble det kun brukt to av spørsmålene fra del 3 som variabler i analysen, «kjønn» og «terminkarakter». Det kan være problematisk å hevde at elevene har like terminkarakterer, ettersom vurderingsgrunnlagene kan variere fra de ulike skolene og innad i trinnet. Det vil av den grunn ikke være riktig å påstå at elevene er på samme faglig nivå i matematikk. I tillegg er det trolig å anta at elevene ikke har lært de samme kompetansemålene, siden læreplanen viser hvilke kompetansemål elevene skal lære *etter* 10. årsteg, og ikke fra hvert av trinnene (Utdanningsdirektoratet, u.d.c).

Systematiske målingsfeil, som for eksempel angst, holdning og uærlighet, kan også ha påvirket begrepsvaliditeten (Kleven, 2011c, s. 96-99). Denne målingsfeilen kan ha en viss betydning i og med at noen av elevene kan ha blitt påvirket av at de var i en prøvesituasjon. Opplysninger om egen anonymitet har muligens begrenset målefeilen, siden besvarelsene av skjemaene ikke fikk konsekvenser for elevene. I tillegg styrkes begrepsvaliditeten ettersom valg av innsamlingsmetode kan ha bidratt til en reduksjon av observatøreffekten, sammenlignet med for eksempel bruk av intervju som metode. Når det er sagt kan effekten ikke utelukkes siden forskerens tilstedeværelse kan ha resultert i at elevene endret atferd (Ibid).

3.5 Metode for analyse

I dette avsnittet presenterer og begrunnes valg av analysemetode. I tillegg forklares statistiske begreper som brukes i analysen.

3.5.1 Binomisk logistisk regresjonsanalyse og statistiske begreper

I følge Eikemo & Claussen (2012, s. 115) brukes binomisk logistisk regresjonsanalyse ved dikotome avhengige variabler. Metoden beregner regresjonskoeffisientene som gir sannsynligheten for at den uavhengige og den avhengige variabelen henger sammen. Binomisk logistisk regresjonsanalyse er egnet til å analysere datamaterialet, for deretter å svare på oppgavens problemstilling. Dette forklares med at metoden tillater en dikotom avhengig variabel, i motsetning til en lineær regresjon som blant annet krever linearitet. Logistisk regresjon løser linearitetsproblemet med *maximum likelihood* metoden, hvilket transformerer den avhengige variabelen fra oddsverdier til naturlige logaritmer (Eikemo & Clausen, 2012, s. 115). Oddsverdien beskriver forholdstallet mellom egenskapen som måles (\bar{Y}) og beregnes ved formelen: $\text{Odds} = \bar{Y}/(1-\bar{Y})$. Ved å transformere oddsverdiene til logaritmeverdier unngås S-formede regresjonskurver, fordi logit-skalaen har et intervall på $[-\infty, +\infty]$ (Skog, 2007, s. 355).

Variablenes sammenheng uttrykkes lineært, $Y = b_0 + b_1 \cdot X$. Y er den avhengige variabelen transformert til «logits», b_0 er verdien når den uavhengige variabelene er null, og b_1 er stigningstallet (Eikemo & Clausen, 2012, s. 115). For å kontrollere om det forekommer

bakenforliggende faktorer som også er med på å forklare sammenhengen, foretas multippel logistisk regresjon. Multippel logistisk regresjon undersøker altså om det er flere uavhengige variabler som har en effekt på den avhengige variabelen (Skog, 2007, s. 405).

3.5.2 Forutsetninger for logistisk regresjonsanalyse

Selv om skjemaet var tilpasset slik at datamaterialet skulle analyseres ved logistisk regresjon, var det flere forutsetninger som måtte testes før analysemetoden kunne tas i bruk. Det ble testet for *multikollinearitet*, *diskrimineringsproblematikk*, *ikke-linearitet* i parameterne og for *innflytelsesrike enheter* (Eikemo & Clausen, 2012, s. 145). Det er viktig å påpeke at det ikke er mulig å teste alle forutsetningene for analysemetoden, men at testene som foretas i SPSS gir en god indikator for å kontrollere om datamaterialet tilfredsstillende en logistisk regresjonsanalyse (Ibid.). For eksempel er det ikke mulig å foreta en test som kontrollerer for at alle de uavhengige variablene er med i modellen.

Multikollinearitet

Eikemo & Clausen (2012) påpeker at en viktig forutsetning for å kunne bruke logistisk regresjonsanalyse er lav korrelasjon mellom x-variablene, kalt fravær av multikollinearitet. Multikollinearitet kan oppstå om de uavhengige variablene som er høyt korrelerte inngår i samme regresjonslikning, og kan føre til skjeve/feil estimater og signifikansverdier (2012, s. 157). Ved binomisk logistisk regresjon, i SPSS, foretas en toleransetest kalt «log likelihood ratio»-test (LR). Testen gir ikke-signifikante verdier ved multikollinearitet dersom det er variabler som er høyt interkorrelerte (Skog, 2007, s. 410).

Signifikanstesten oppgir log likelihood verdien (-2LL) hvilket beskriver hvor godt regresjonsmodellen passer til datamaterialet sammenlignet med andre modeller (Skog, 2007, s. 375). Testen beregner forbedret -2LL verdi ved bruk av formelen $LR = (-2LL_0) - (-2LL_A)$. $-2LL_A$ er den estimerte modellen, og har flere variabler enn $-2LL_0$, som er den foregående modellen. En signifikant differanse tolkes som at modell A er bedre til å forklare variasjonen i den avhengige variabelen (Ibid.).

Diskrimineringsproblem

Diskrimineringsproblem oppstår for kategoriske variabler ved ugunstige X- og Y-verdier som får konsekvenser for hvordan oddsen beregnes og hva som er referansekategori (Eikemo &

Clausen, 2012, s. 160-162). For å unngå diskrimineringsproblemet ble det laget krysstabeller for å få en oversikt over om det var noen verdier som lå nær null.

Ikke-linearitet i parameterne

Ikke-linearitet i parameterne ved bruk av logistisk regresjon kan medføre at estimatene blir skjeve. Dette signifikantestes i SPSS med testen kalt *Hosmer and Lemeshow test*, som er en kji-kvadratfordeling som viser sammenhengen mellom predikerte og observerte frekvenser (Eikemo & Clausen, 2012, s. 162-165). Om signifikansnivået er større enn en sannsynlighet på 5 % tilsier testen at nullhypotesen ikke er signifikant. Det betyr at modellen *ikke* kan avvises, siden avviket mellom predikert og observert frekvens er tilnærmet like (Christophersen, 2012, s. 137). Med andre ord gir en god modell lave Kji-kvadrat-verdier, og en sannsynlighetsverdi (p) som ikke er signifikant på 0,05 nivå (Kinnear & Gray, 2004, s. 395).

Innflytelsesrike enheter

Et datamateriell med innflytelsesrike enheter er problematisk, siden det kan påvirke regresjonsresultatene så mye at hvis de utelates vil resultatene endres betydelig (Eikemo & Clausen, 2012, s. 165-166) . For å kontrollere for fravær av innflytelsesrike enheter ble det foretatt tre ulike målinger i SPSS; Leverage, DfBetas og Cook's D. De måler henholdsvis uvante kombinasjoner av verdier på forskjellige variabler, innflytelse på hver enkelenhet i hver variabel, og enkelenhetens innflytelse på modellen samlet (Ibid.).

3.5.3 Tolkning av den logistiske regresjonslikningen

Logistisk regresjon i SPSS gir tabellen «*variables in the equation*». For å presentere variablene i regresjonslikningen på en enkel og oversiktlig måte, vil kun de viktigste verdiene fra tabellen gjengis i analysedelen; *logit* (B), *signifikantverdi* (Sig.) og *oddsratio* (Exp(B)). Fortegnet til logitverdien indikerer om effekten er positiv eller negativ. Positive verdier tilsier at andelen av den bestemte egenskapen er over 50 %, og negative verdier indikerer at andelen er under 50 %. Oddsratio brukes til å vurdere stryken på relasjonen mellom to variabler (Christophersen, 2013, s. 131). Signifikansverdien angir sannsynligheten for at en sann nullhypotese er forkastet (Eikemo & Clausen, 2012, s. 122-123). Nullhypotesen tilsier at det *ikke* er en sammenheng mellom den avhengige og uavhengige variabelen i populasjonen (Eikemo & Clausen, 2012, s. 88). Testen som utføres kalles Wald-test og er basert på en

tilnærmet kji-kvadratfordeling. Vurderingen av signifikantverdien må derfor tolkes sammen med utvalgets størrelse, ettersom et lite utvalg kan avvike fra fordelingen (Skog, 2007, s. 375). Små utvalgsstørrelser kan derfor bidra til at nullhypotesen beholdes selv om den ikke er riktig, kalt type II-feil (Skog, 2007, s. 103).

For å presentere analysearbeidet kan det foretas sannsynlighetsberegninger og utregninger av den prosentvise endringen i odds. Disse beregningene utføres ved bruk av følgende formler (Eikemo & Clausen, 2012, s. 124 & 126):

(1) Positiv endring i odds: $100 * (\text{Exp}(B) - 1)$

(2) Sannsynlighet: $P = 1 / (1 + e^{-L})$, der $L = \text{logit}$

Resultatene i denne analysen presenteres i et betinget effektplott hvor logit-lingningen transformeres til en sannsynlighetsfunksjon. Mitt valg begrunnes med at den predikerte sannsynligheten er mest informativ ved logistisk regresjon (Eikemo & Clausen, 2012, s. 128). Plottet er betinget da de resterende variablene viser utvalgets gjennomsnittsverdi, og er derfor konstanter i regresjonslikningen.

4 Analyse og resultater

I dette kapitlet presenteres analyse og resultater til følgende problemstilling: *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes valg av oppgave og prestasjon på oppgaven?* Som nevnt tidligere har problemstillingen to delproblemstillinger som behandles separat i kapitlet. Begge analysene består av tre regresjonsanalyser som signifikanstester hypotese 1 og hypotese 2. «Analyse 1» tester den første hypotesen som er følgende.

Hypotese 1: matematikkoppgavens nivåmarkering og elevens terminkarakter har en effekt på oppgavevalget.

«Analyse 2» tester så den andre hypotesen.

Hypotese 2: matematikkoppgavens nivåmarkering og elevens terminkarakter har en effekt på oppgaveprestasjonen.

Regresjonsanalysene er oppsummert i egne tabeller, hvor signifikansverdiene er markert med stjerner. En, to og tre stjerner betyr henholdsvis at det var 5 %, 1 % og 0,1 % sannsynlighet for at en sann nullhypotese ble forkastet. Regresjonsanalysen drøfter ikke konstantleddet fordi karakter 0 og markering 0 ikke gir mening i denne sammenhengen. Begge analysedelene avsluttes med en samlet oppsummering av hovedfunnene. En tolkning og drøfting av funnene foretas i kapittel 5.

4.1 Analyse 1 – Elevenes valg av oppgave

Hensikten med analysen var å undersøke følgende delproblemstilling: *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes valg av oppgave?*

Tabell 3 Oversikt over datamaterialet som ble brukt i analyse 1

	Lett	Middels	Vanskelig	Kontrollgruppe	Totalt
Oppgave A	N = 59	N = 62	N = 60	N = 65	N = 246
Oppgave B	N = 60	N = 59	N = 62	N = 65	N = 246
Oppgave C	N = 62	N = 60	N = 59	N = 65	N = 246

Tabell 3 gir en oversikt over datamaterialet som blir brukt i analysen. Undersøkelsen ble besvart av to grupper: en kontrollgruppe på 65 elever og en eksperimentgruppe bestående av 181 elever. Som beskrevet i metodekapitlet fikk alle elevene de samme valgmulighetene. De fikk beskjed om velge én oppgave av tre mulige: oppgave A, B eller C.

Samspilleffekten er, sammen med kontrollgruppen, fremstilt grafisk i et betinget effektplott. Effektplottet er betinget fordi de resterende variablene viste utvalgets gjennomsnittsverdi, og var derfor konstanter i regresjonslikningen. En sammenligning av resultatene fra «analyse 1» (figur 2-4) antyder like funn. For at de tre regresjonsanalysene (valg av oppgave A-C) ikke skal bli for repeterende blir den første regresjonsanalysen prestert grundigst.

4.1.1 Valg av oppgave A

Hypotese 1 ble testet ved bruk av binomisk logistisk regresjon. Nedenfor er «Likelihood ratio»-testen og Wald-testen analysert for å skape et grunnlag for vurdering av nullhypotesen. I tillegg ble Hosmer-Lemeshow test brukt til å vurdere modellens datatilpasning. En oppsummering av regresjonsanalysen er vist i tabell 4.

Tabell 4 Binomisk logistisk regresjonsanalyse utført ved hjelp av SPSS

N = 181	Modell 1		Modell 2	
Variabler	Logit	Oddsratio	Logit	Oddsratio
Konstant	-0.496	0.609	-3.769*	0.023*
Kontrollvariabler:				
Kjønn:				
- Jente	0.329	1.390	0.257	1.292
Rekkefølge:				
- Først	0.147	1.159	0.079	1.082
Uavhengig variabler:				
Markering:				
- Lett	-1.830***	0.160***	4.803*	121.904*
- Middels	-1.155*	0.315*	2.690	14.735
Karakter	0.087	1.091	0.909**	2.483**
Samspillvariabel				
Karakter*Markering:				
- Karakter*Lett			-1.614**	0.199**
- Karakter*Middels			-0.960*	0.393*
-2LL	200.788***		188.100**	
Relativ endring i -2LL	0.103***		0.160***	
Hosmer-Lemeshow(p)	16.670 (0.034)		4.195 (0.839)	

-2LL = -2*LogLikelihood, * = p < 0.05, ** = p < 0.01, *** = p < 0.001

Modell 1

Modell 1 inneholdt de uavhengige variablene «Kjønn», «Rekkefølge», «Karakter» og «Markering». -2LL-verdien viste at modell 1 var signifikant bedre enn modell 0 til å predikere valg av oppgave A. Hosmer-Lemeshow signifikanstest indikerte at modellen ikke var godt tilpasset data ($p < 0.05$). En videre analyse av modellen ble derfor ikke foretatt.

Modell 2

Modell 2 testet samspillet mellom elevens karakter og oppgavens nivåmarkering. I tillegg ble det kontrollert for kjønn og oppgavens rekkefølge. Etter en tolkning av -2LL-verdiene (se tabell 4) ble modell 2 anslått som signifikant bedre enn modell 0 og modell 1 på henholdsvis 0,01 -og 0,001 nivå. Dette betyr at modell 2 var bedre enn de foregående modellene til å predikere valg av oppgave A. Med andre ord kan ikke samspillvariabelen «Karakter*Markering» utelukkes, siden effekten av forskjellige nivåmarkeringer trolig hadde en sammenheng med høye og lave terminkarakterer under valg av oppgaven.

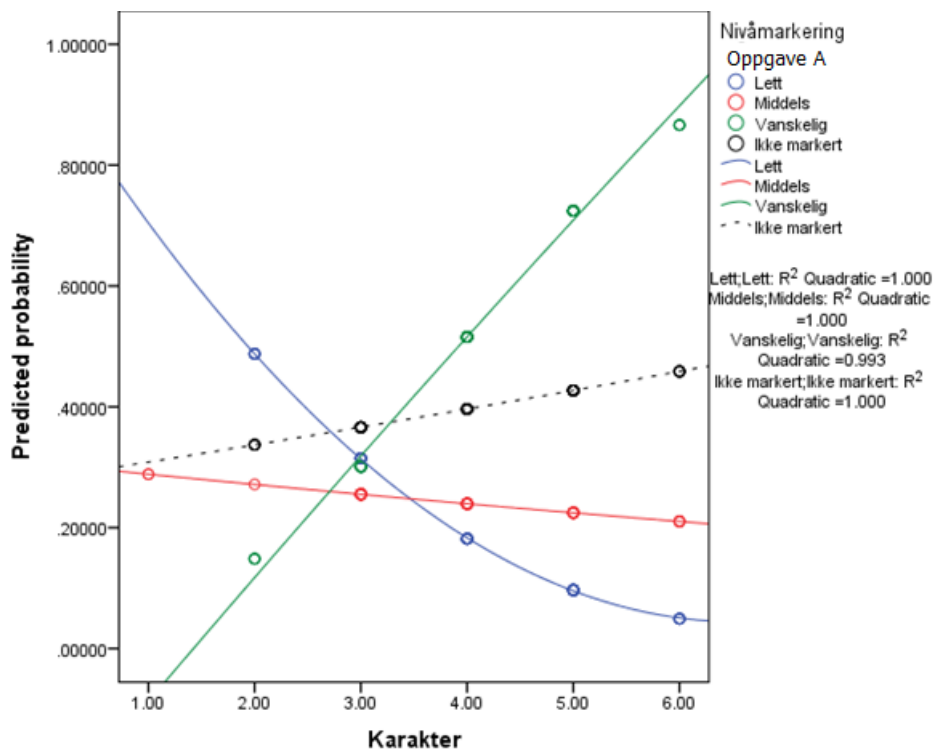
Hosmer-Lemeshow signifikanttest viste en reduksjon av Kji-kvadrat-verdien, og ga en signifikansverdi (p) på 83,9 %. Det betyr at avviket mellom predikerte og observerte frekvenser var tilnærmet like, da $p > 0,05$. Modell 2 vurderes derfor som en god tilpasning til data.

Wald-signifikanttest antydte at kontrollvariablene «Kjønn» og «Rekkefølge» ikke hadde en signifikansverdi på under 0,05 nivå. En tolkning er at elevens kjønn og oppgavens rekkefølge *ikke* hadde en effekt på oppgavevalget. Som nevnt er denne signifikanttesten følsom for små utvalgsstørrelser og testen må derfor tolket med forsiktighet. Det betyr at elevens kjønn og oppgavens rekkefølge ikke kan utelukkes med sikkerhet, ettersom et lite utvalg *kan* ha bidratt til ikke-signifikante verdier, type 2-feil. Signifikanttestene viste også at samspillvariablene var statistisk signifikante på henholdsvis 0,01 -og 0,05 nivå. Det betyr at nullhypotesen forkastes, og det gis grunnlag til å anta hypotese 1. Logit-verdiene til samspillvariablene «Karakter*Lett» og «Karakter*Middels» var negative, da variabelen «Karakter*Vanskelig» bruktes som referansekategori. Det betyr at det var større sannsynlighet for at elever med lave terminkarakterer valgte oppgaven markert «lett» eller «middels», sammenlignet med når oppgaven var markert som «vanskelig».

Oppsummert ga begge signifikanstestene styrke til å anta hypotese 1. Det tilsier at nivåmarkeringen til oppgave A og elevens terminkarakter *har* en effekt på sannsynligheten for å velge oppgaven. En grafisk tolkning av samspilleffekten gis nedenfor.

Betinget effektplott av samspillvariabelen

Figur 2 illustrerer samspilleffekten mellom de uavhengige variablene «Karakter» og «Nivåmarkering». Verdiene på y-aksen er den avhengige variabelen som viser den predikerte² sannsynligheten for at elevene valgte oppgave A. X-aksen viser elevens terminkarakter i faget. Hver av grafene er markert med forskjellige farger for å illustrere de ulike nivåmarkeringene.



Figur 2 Betinget effektplott som illustrerer samspilleffekten til valg av oppgave A

En grafisk tolkning av figuren er presentert ved å først analysere stigningstallet til grafene. Dette ble gjort for å undersøke hvordan sannsynligheten for å velge oppgaven, gitt lik markering, varierte med elevenes terminkarakter. Deretter sammenlignes sannsynligheten for å velge oppgave A, gitt lik terminkarakter, for de ulike nivåmarkeringene.

² Predikert sannsynlighet er sannsynligheten for at eleven valgte oppgaven

Den svarte grafen viser sannsynligheten for at kontrollgruppen valgte oppgaven, gitt ulike terminkarakterer. Som nevnt besto kontrollgruppen av elever som fikk oppgavene uten markert vanskelighetsgrad. En grafisk tolkning av stigningstallet tyder på at sannsynligheten for å velge oppgave A var forholdsvis lik for elever med forskjellige karakterer. «Middels» nivåmarkering (rød graf) har også et lavt stigningstall sammenlignet med blå og grønn graf. Det tolkes, på lik linje med kontrollgruppen, som at sannsynligheten for å velge oppgaven var nokså lik for elever med forskjellige terminkarakterer. En videre analysering indikerte at «lett» nivåmarkering (blå graf), har et høyt negativt stigningstall. Det negative stigningstallet tyder på at var større sannsynlighet for at elever med lave terminkarakterer valgte oppgaven, sammenlignet med de med høye terminkarakterer. Når oppgaven var markert «vanskelig» (grønn graf) viser grafen et høyt positivt stigningstall. Det positive stigningstallet indikerer at sannsynligheten for at oppgaven ble valgt økes med høye terminkarakterer. En sammenligning av størrelsen på grafenes stigningstall tyder på at oppgavevalget hadde størst sammenheng med terminkarakteren når den var markert enten «lett» eller «vanskelig».

En sammenligning av de ulike nivåmarkeringene, gitt elever med like terminkarakterer, viste at sannsynligheten for å velge oppgave A varierte for hver markering. Dette er forklart nærmere i følgende eksempel, der terminkarakter 2 holdes konstant.

Når oppgaven markertes «lett» (blå graf) viste resultatene at det var 50 % sannsynlighet for at elevene med karakter 2 valgte oppgaven. Videre var sannsynligheten litt over 30 % for at oppgaven ble valgt når den ikke hadde markering, og i underkant av 30 % sannsynlighet når oppgaven var påført «middels» nivåmarkering. Dersom oppgaven var markert «vanskelig» (grønn graf), var det 15 % sannsynlighet for at elevene med karakter 2 valgte oppgaven.

4.1.2 Valg av oppgave B

Hypotese 1 ble testet ved bruk av binomisk logistisk regresjon. Nedenfor er signifikanstestene analysert for å skape et grunnlag for vurdering av nullhypotesen. I tillegg ble Hosmer-Lemeshow test brukt til å vurdere modellens datatilpasning. Tabell 5 viser en oppsummering av regresjonsanalysen.

Tabell 5 Binomisk logistisk regresjonsanalyse utført ved hjelp av SPSS

N = 181 Variabler	Modell 1		Modell 2	
	Logit	Oddsratio	Logit	Oddsratio
Konstant	-1.833	0.160	-5.707**	0.003**
Kontrollvariabler:				
Kjønn:				
- Jente	0.148	1.159	0.404	1.498
Rekkefølge:				
- Først	0.136	1.145	0.148	1.160
Uavhengig variabler:				
Markering:				
- Lett	-2.277***	0.103***	9.435**	12520.595**
- Middels	-2.534***	0.079***	4.206	67.099
Karakter	0.368	1.445	1.317**	3.730**
Samspillvariabel				
Karakter*Markering:				
- Karakter*Lett			-3.127***	0.044***
- Karakter*Middels			-1.628**	0.196**
-2LL	145.528***		122.053***	
Relativ endring i -2LL	0.194***		0.324***	
Hosmer-Lemeshow(p)	23.076 (0.002)		6.193 (0.628)	

-2LL = -2*LogLikelihood, * = p < 0.05, ** = p < 0.01, *** = p < 0.001

Modell 1

Modell 1 var signifikant bedre enn modell 0 til å predikere oppgavevalget. Modellen inneholdt de uavhengige variablene «Kjønn», «Rekkefølge», «Karakter» og «Markering». Hosmer-Lemeshows signifikantest tilsa at modellen *ikke* var godt tilpasset data (p < 0.05). På lik linje med modell 1 i forrige analyse ble det ikke foretatt en videre analysering av modellen.

Modell 2

Modell 2 testet samspillet mellom elevenes karakter og oppgavens nivåmarkering. I tillegg ble det kontrollert for kjønn og oppgavens rekkefølge. Regresjonsanalysen anslo modellen som signifikant bedre enn modell 0 og modell 1 på 0,001 nivå. Det betyr at modell 2 var mer egnet enn de foregående modellene til å predikere valg av oppgaven. I likhet med forrige regresjonsanalyse kan ikke samspillvariabelen, «Karakter*Markering», utelukkes, siden effekten av forskjellige nivåmarkeringen trolig hadde en sammenheng med høye og lave terminkarakterer under valg av oppgave B.

Hosmer-Lemeshow signifikanttest viste også her en reduksjon av Kji-kvadrat-verdien, og oppga en signifikansverdi på 62,8 %. Det tolkes som at modell 2 er en god tilpasning til data.

Wald-signifikanstest indikerte at kontrollvariablene, «Kjønn» og «Rekkefølge», ikke hadde en signifikansverdi på under 0,05 nivå. Det kan tolkes som om at elevenes kjønn og oppgavens rekkefølge *ikke* hadde en effekt på oppgavevalget. En videre tolkning av signifikanstesten viste at samspillvariablene var statistisk signifikante på henholdsvis 0,001 - og 0,01 nivå. Det betyr at nullhypotesen forkastes, noe som gir grunnlag til å vurdere hypotese 1. Logit-verdiene til samspillvariablene, «Karakter*Lett» og «Karakter*Middels», var negative når variabelen «Karakter*Vanskelig» ble brukt som referansekategori. På tilsvarende måte med forrige analyse var det altså størst sannsynlighet for at elever med lave terminkarakterer valgte oppgaven, når den var markert «lett» eller «middels», sammenlignet med når oppgaven var markert «vanskelig».

Oppsummert ga begge signifikanstestene støtte til hypotese 1, som tilsier at nivåmarkeringen til oppgave B og elevens terminkarakter har en effekt på sannsynligheten for å velge oppgaven.

Betinget effektplott av samspillvariabelen

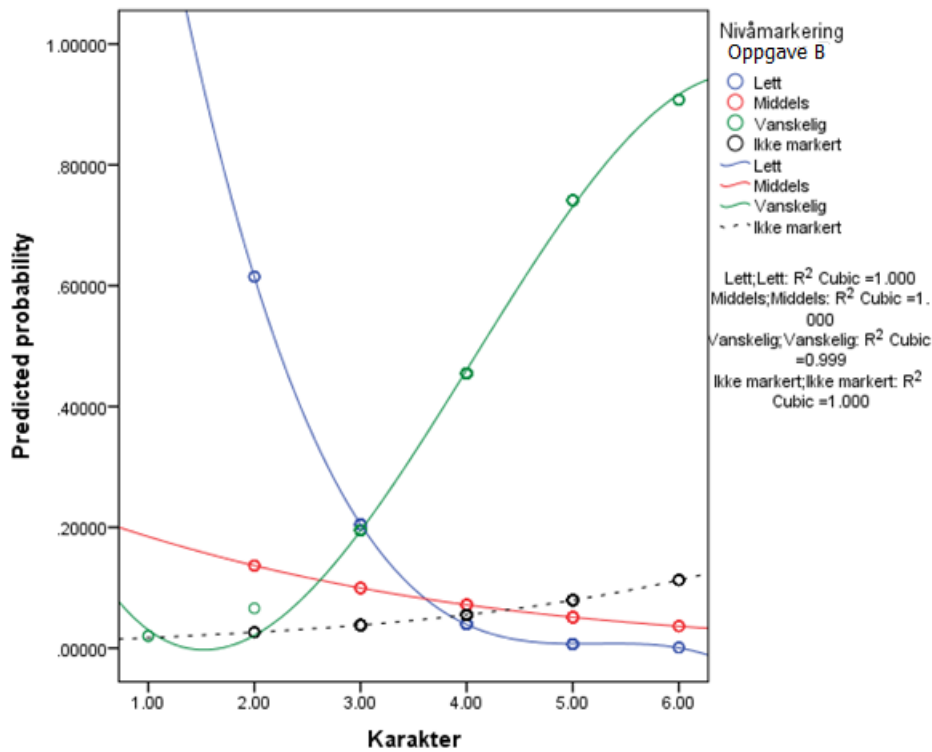
Figur 3 illustrerer samspilleffekten mellom de uavhengige variablene «Karakter» og «Nivåmarkering». Verdiene på y-aksen er den avhengige variabelen som viser predikert sannsynlighet for at elevene valgte oppgave B. X-aksen viser elevens terminkarakter i faget. En grafisk tolkning ble gjort på tilsvarende måte som i forrige analyse.

Den svarte grafen viser sannsynligheten for at kontrollgruppen valgte oppgaven.

Stigningstallet tyder på at sannsynligheten for å velge oppgave B var forholdsvis lik for elever med forskjellige karakterer.

«Middels» nivåmarkering (rød graf) har også et lavt stigningstall sammenlignet med blå og grønn graf. Ut i fra dette kan man si at sannsynligheten for å velge oppgaven var nokså lik for elever med forskjellige terminkarakterer. En videre analyse indikerte at blå graf, «lett» nivåmarkering, hadde et høyt negativt stigningstall. Et negativt stigningstall tyder på at det var større sannsynlighet for at elevene med lav terminkarakterer valgte oppgaven sammenlignet med elevene med høye karakterer. Som den grønne grafen viser har den et høyt positivt stigningstall når oppgaven ble markert «vanskelig». Det positive fortegnet tolkes som at sannsynligheten for å velge oppgaven økte med høye terminkarakterer. En sammenligning av

grafenes størrelse på stigningstallet tyder på at sannsynligheten for å velge oppgaven hadde størst sammenheng med terminkarakteren for «lett» -og «vanskelig» nivåmarkering.



Figur 3 Betinget effektplott som illustrerer samspilleffekten til valg av oppgave B

En sammenligning av de ulike nivåmarkeringene, gitt elever med like terminkarakterer, tyder på at det var en sammenheng mellom oppgavevalget og nivåmarkeringen. Dette kan forklares nærmere i følgende eksempel, der terminkarakter 2 holdes konstant.

Når oppgaven var markert «lett» (blå graf) viser resultatene at det er ca. 65 % sannsynlighet for at elevene med karakter 2 valgte oppgaven. Videre var sannsynligheten litt over 5 % for at oppgaven ble valgt når den ikke hadde markering, og i underkant av 15 % når oppgaven var påført «middels» nivåmarkering. Var oppgaven markert «vanskelig» (grønn graf) indikerer resultatene at det kun var 10 % av elevene som valgte oppgaven.

4.1.3 Valg av oppgave C

Binomisk logistisk regresjon ble på tilsvarende måte som i de foregående regresjonsanalysene testet ved bruk av binomisk logistisk regresjon. Regresjonsanalysen er oppsummert i tabell 6.

Tabell 6 Binomisk logistisk regresjonsanalyse utført ved hjelp av SPSS

N = 181 Variabler	Modell 1		Modell 2	
	Logit	Oddsratio	Logit	Oddsratio
Konstant	2.785**	16.198**	-1.425	0.241
Kontrollvariabler:				
Kjønn:				
- Jente	-0.359	0.699	-0.473	0.623
Rekkefølge:				
- Først	0.181	1.198	0.288	1.334
Uavhengig variabler:				
Markering:				
- Lett	-2.234***	0.197***	6.051**	424.547**
- Middels	-1.758***	0.172***	2.843	17.164
Karakter	-0.317	0.728	0.667*	1.948*
Samspillvariabel				
Karakter*Markering:				
- Karakter*Lett			-2.102***	0.122***
- Karakter*Middels			-1.064*	0.345*
-2LL	216.343***		195.204***	
Relativ endring i -2LL	0.142***		0.226***	
Hosmer-Lemeshow(p)	23.701 (0.003)		1.675 (0.989)	

-2LL = -2*LogLikelihood, * = p < 0.05, ** = p < 0.01, *** = p < 0.001

Modell 1

Modell 1 inneholdt også de uavhengige variablene «Kjønn», «Rekkefølge», «Karakter» og «Markering». Resultatene tyder på at modell 1 var signifikant bedre enn modell 0 til å predikere valg av oppgaven. Det ble ikke foretatt en videre analysing av modellen, etter som Hosmer-Lemeshow signifikanstest indikerte at modellen *ikke* var godt tilpasset data (p < 0.05).

Modell 2

Modell 2 testet samspillet mellom elevens karakter og oppgavens nivåmarkering. Regresjonsanalysen kontrollerte også for variablene «kjønn» og «rekkefølge». Verdiene fra tabellen indikerer at modell 2 var signifikant bedre enn modell 0 og modell 1 på 0,001 nivå, til å predikere valg av oppgave C. Det betyr at samspillvariabelen «Karakter*Markering» ikke kan utelukkes, fordi effekten av forskjellige nivåmarkeringen antakelig har en sammenheng

med høye og lave terminkarakterer under valg av oppgaven.

Hosmer-Lemeshow signifikanttest viste en reduksjon av Kji-kvadrat verdien, og oppga en signifikansverdi på 98,9 %. Dette tolkes som at modell 2 er en god tilpasning til data.

Wald-signifikanttest indikerte at kontrollvariablene ikke hadde en signifikansverdi på under 0,05 nivå. Det tydet på at elevenes kjønn og oppgavens rekkefølge trolig *ikke* hadde en effekt på oppgavevalget. Annen tolkning av signifikanttesten viste at samspillvariablene var statistisk signifikante på henholdsvis 0,01 -og 0,05 nivå. Det betyr at nullhypotesen forkastes, og danner et grunnlag til å vurdere hypotese 1. På lik linje som de forrige analysene var fortegnet til samspillvariablene «Karakter*Lett» og «Karakter*Middels» negative, og det var derfor større sannsynlighet for at elever med lave terminkarakterer valgte oppgaven, når den markeres «lett» eller «middels», sammenlignet med når oppgaven var markert «vanskelig».

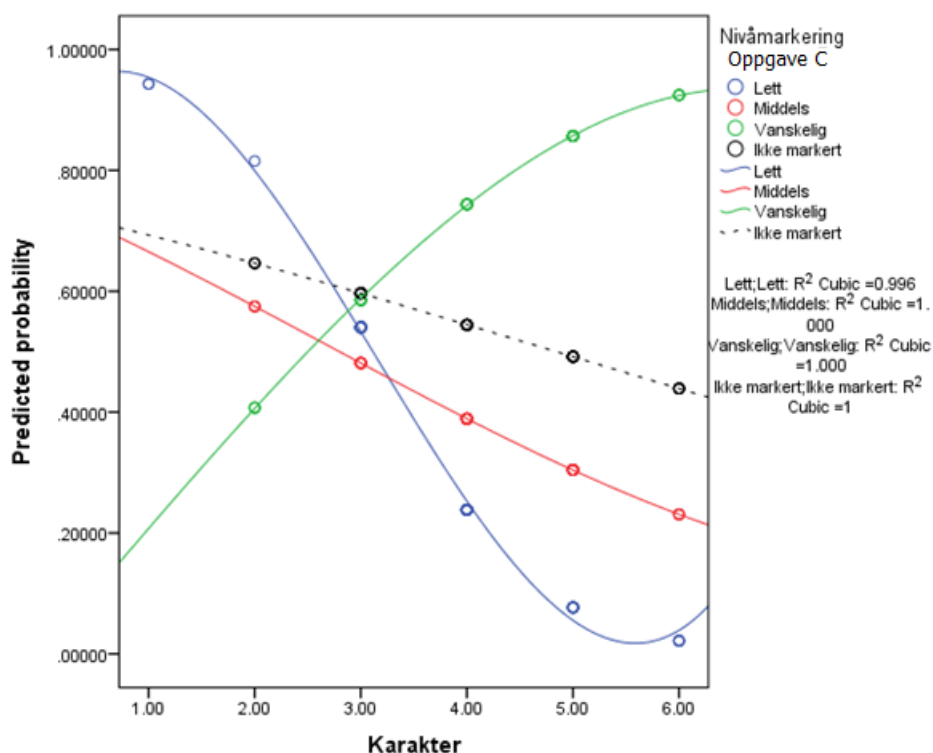
Oppsummert ga begge signifikanttestene grunnlag til å anta hypotese 1, som tilsier at nivåmarkeringen til oppgaven og elevens terminkarakter *har* en effekt på sannsynligheten for å velge oppgaven. Nedenfor gis en grafisk tolkning av samspilleffekten.

Betinget effektplott av samspillvariabelen

Figur 4 illustrerer samspilleffekten mellom de uavhengige variablene «Karakter» og «Nivåmarkering». Effektplottet ble analysert på samme måte som i de to foregående analysene.

Den svarte grafen viser sannsynligheten for at kontrollgruppen valgte oppgave C, gitt ulike terminkarakterer. En grafisk tolkning av stigningstallet tyder på at sannsynligheten for å velge oppgave C var mest lik for elever med forskjellige karakterer, sammenlignet med de andre grafene.

«Middels» nivåmarkering (rød graf) har også et lavt stigningstall sammenlignet med de andre grafene, og var derfor mest lik kontrollgruppen. Både blå og rød graf, henholdsvis «lett» -og «middels» nivåmarkering har et negativt stigningstall. Det kan tolkes som at det er større sannsynlighet for at elever med lave terminkarakterer valgte oppgaven sammenlignet med de med høye terminkarakterer.



Figur 4 Betinget effektplott som illustrerer samspilleffekten til valg av oppgave C

Den grønne grafen viser sammenhengen mellom valg av oppgave, markert «vanskelig», og elevenes terminkarakter. Grafen har et positivt stigningstall. Det positive stigningstallet vurderes til at sannsynligheten for at oppgaven ble valgt økes med høye terminkarakter. En sammenligning av størrelsene på grafenes stigningstall indikerer at sannsynligheten for å velge oppgaven hadde størst sammenheng med terminkarakteren for «lett» -og «vanskelig» nivåmarkering.

En sammenligning av de ulike nivåmarkeringene, gitt elever med lik terminkarakter, viser at sannsynligheten for å velge oppgave C varierte. Dette forklares nærmere i følgende eksempel hvor terminkarakter 2 holdes konstant.

Når oppgaven var markert «lett» (blå graf) viser resultatene at det var ca. 80 % sannsynlighet for at elevene med karakter 2 valgte oppgaven. Videre var sannsynligheten litt over 65 % for at oppgaven ble valgt når den ikke hadde markering, og i underkant av 55 % når oppgaven var

påført «middels» nivåmarkering. Var oppgaven markert «vanskelig» tyder resultatene på at det kun var 40 % av elevene som valgte oppgaven.

4.1.4 Oppsummering av hovedfunn fra analyse 1

For å teste hypotese 1 ble det foretatt tre binomiske logistiske regresjonsanalyser. Hosmer-Lemeshows test indikerte at modellene som ble brukt ga en god beskrivelse av datamaterialet. En vurdering av signifikanstestene, «Likelihood-ratio»-testen og Wald-testen, ga grunnlag til å beholde hypotese 1. For å tolke samspillet mellom karakter og nivåmarkering ble resultatene fremstilt i et betinget effektplott. En sammenligning av alle plottene antydte at *like* nivåmarkeringer hadde tilsvarende effekt på oppgavevalget. I tillegg viste en sammenligning at sannsynligheten for å velge oppgave A, B eller C, gitt lik markering, tenderer til å variere på samme måte for terminkarakterer. Av den grunn gis det en samlet oppsummering nedenfor. Sammenlignbare plott og statistisk signifikante tester styrker vurderingsgrunnlag av følgende delproblemstilling; *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes valg av oppgave?*

Analysens hovedfunn var at matematikkoppgavens markerte nivå *har* en effekt på elevenes valg av oppgave, der effekten i størst grad varierer for høye og lave terminkarakterer, samt ved lave og høye nivåmarkeringer. Resultatene viste at oppgavevalget, for elever med ulike terminkarakterer, hadde størst sammenheng med «lett» og «vanskelig» nivåmarkering. Grafene viste at det var størst sannsynlighet for at elevene med lave terminkarakterer valgte oppgaven når den var markert «lett», og minst sannsynlighet for at den samme oppgaven ble valgt når den var markert «vanskelig». Resultatene viste også at det motsatte forekom når oppgaven var påført «vanskelig» nivåmarkering. Et annet funn fra analysen var at oppgavevalget ikke hadde like stor sammenheng med terminkarakteren for «middels» nivåmarkering sammenlignet med «lett» og «vanskelig» markering. Det samme synes å gjelde for kontrollgruppen, og ble tolket som at elevenes karakterer ikke hadde like stor betydning for om oppgaven ble valgt.

Andre funn fra analysen var at sannsynligheten for å velge oppgaven, gitt elever med *like* terminkarakterer, påvirkes av nivåmarkeringen. Resultatene fra «lett» og «vanskelig» nivåmarkering skilte seg mest fra kontrollgruppen og ble derfor vurdert til å ha størst effekt på oppgavevalget. Sannsynligheten for at oppgaven ble valgt ved «middels» nivåmarkering

avvikte minst fra kontrollgruppen og ble derfor tolket som at markeringen hadde minst effekt på valg av oppgaven.

Etter en sammenligning mellom eksperimentgruppen og kontrollgruppen viste resultatene at effekten ved å tilføre oppgaver med en tilfeldig nivåmarkering påvirket oppgavevalget, og at effektstyrken varierte med elevenes terminkarakterer. Med andre ord kan det se ut som om elevene lot markeringen være bestemmende for oppgavevalget, og at oppgavens faktiske vanskelighetsgrad ikke ble vurdert. Hvordan denne effekten kan ha betydning for matematikkundervisning vil bli drøftet i oppgavens diskusjonsdel.

4.2 Analyse 2 – Elevenes prestasjon på oppgavene

Hensikten med analysen i dette underkapitlet var å undersøke følgende delproblemstilling; *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes prestasjon på oppgaven?*

Tabell 7 Oversikt over datamaterialet som brukes i analyse 2

Oppgaver:	Lett	Middels	Vanskelig	Kontrollgruppe	Totalt
Oppgave D	N = 64	N = 67	N = 64	N = 73	N = 268
Oppgave E	N = 67	N = 64	N = 64	N = 73	N = 268
Oppgave F	N = 64	N = 64	N = 67	N = 73	N = 268

Datamaterialet ble også her besvart av to grupper, en kontrollgruppe på 73 elever og en eksperimentgruppe bestående av 195 elever, se tabell 7. Oppsummert fikk elevene de samme matematikkoppgavene der alle oppgavene skulle besvares.

Samspilleffekten fremstilles også her grafisk, sammen med kontrollgruppen, i et betinget effektplott. Etter en sammenligning av plottene fra analyse 2 (figur 4–7) ble det antydnet at «vanskelig» og «lett» nivåmarkering hadde tilsvarende effekter på oppgaveprestasjonene. En grafisk tolkning av hvert plott presenteres derfor ved å sammenligne sannsynligheten for å riktig svar på oppgaven, gitt «lett» og «vanskelig» markering. For at regresjonsanalysene ikke skal bli for repeterende blir de to siste analysene prestert mindre omfattende.

4.2.1 Prestasjon på oppgave D

Binomisk logistisk regresjons ble brukt for å teste hypotese 2. «Likelihood ratio»-testen og Wald-testen er analysert nedenfor for å skape et grunnlag for vurdering av nullhypotesen. Hosmer-Lemeshow test ble brukt til å vurdere modellens datatilpasning. Regresjonsanalysen oppsummeres i tabell 8.

Tabell 8 Binomisk logistisk regresjonsanalyse utført ved hjelp av SPSS

N = 195 Variabler	Modell 1		Modell 2	
	Logit	Oddsratio	Logit	Oddsratio
Konstant	-1.986	0.014	-3.168*	0.042*
Kontrollvariabler:				
Kjønn:				
- Jente	0.108	1.114	0.129	1.138
Rekkefølge:				
- Først	-0.381	0.683	-0.373	0.689
Uavhengig variabler:				
Markering:				
- Lett	0.282	1.326	1.650	5.208
- Middels	0.039	1.040	2.049	7.763
Karakter	0.727***	2.068***	1.020**	2.773**
Samspillvariabel				
Karakter*Markering:				
- Karakter*Lett			-0.349	0.705
- Karakter*Middels			-0.516	0.597
-2LL	213.794**		212.156	
Relativ endring i -2LL	0.093**		0.099**	
Hosmer-Lemeshow(p)	7.132 (0.522)		6.684 (0.462)	

-2LL = -2*LogLikelihood, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

Modell 1

Modell 1 inneholdt de uavhengige variablene «Kjønn», «Rekkefølge», «Karakter» og «Markering». -2LL-verdien viste at modell 1 var signifikant bedre enn modell 0 til å predikere valg av oppgave D. Hosmer-Lemeshow signifikanstest indikerte at modellen var godt tilpasset data ($p > 0.05$). Wald-testen antydte at elevens terminkarakter var statistisk signifikant, på 0,001 nivå. Fortegnet på logit-verdien var positiv. Det kan dermed anta at elever med høyere karakter hadde større sannsynlighet til å få riktig svar på oppgaven, sammenlignet med elever som hadde lavere karakter.

Modell 2

Modell 2 testet samspillet mellom elevens karakter og oppgavens nivåmarkering. I tillegg ble det kontrollert for kjønn og oppgavens rekkefølge. En tolkning av -2LL-verdien tydet på at modell 2 hadde lavest varians da -2LL verdien var minst sammenlignet med modell 1 og 0. Videre viste signifikanstesten at modell 2 ikke var signifikant bedre enn modell 1 til å forklare riktig eller feil svar av oppgave D. Det betyr at det er usikkert om den predikerte verdien *forbedres* når samspillvariabelen ble tatt med. En sammenligning mellom modell 2 og modell 0 tilsa at den relative endringen i -2LL verdi var statistisk signifikant på 0,01 nivå til å forklare den avhengige variabelen.

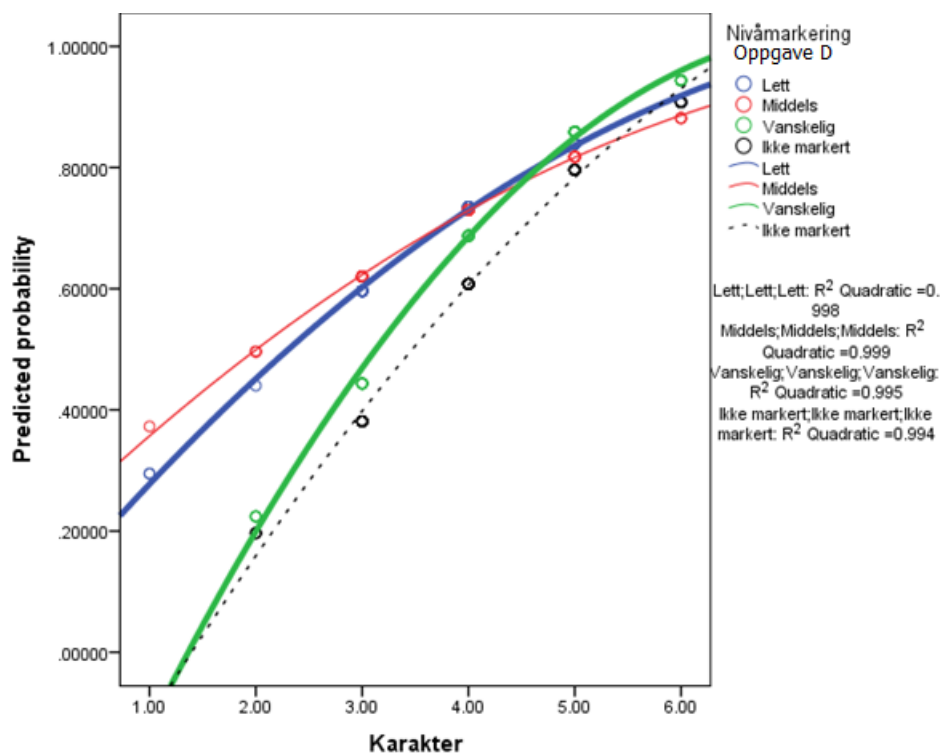
Hosmer-Lemeshow signifikanttest ga en reduksjon av Kji-kvadrat verdien, og viste en signifikansverdi på 46,2 %. Det betyr at avviket mellom predikerte og observerte frekvenser var tilnærmet like, siden $p > 0,05$. Modell 2 vurderes derfor som en god tilpasning til data.

Wald-testen var den svakeste signifikanstesten, og tolkes derfor med forsiktighet, fordi den er følsom for små utvalgsstørrelser. Det påpekes at selv om utvalget hadde en størrelse på 289 elever, var det kun ca. 65 elever som svarte på hver oppgavemarkering. I tillegg ble de 65 elevene gruppert inn i seks undergrupper basert på deres terminkarakter i faget. Denne inndelingen bidrar til en forholdsvis lav utvalgsstørrelse, og kan derfor ha medført at Wald-testen ikke gir signifikante verdier. Utvalgsstørrelsen *kan* derfor ha bidratt til ikke-signifikante verdier selv om nullhypotesen ikke er riktig, type 2-feil. Wald-testen indikerte at elevens terminkarakter var signifikant på 0,01 nivå. Fortegnet på logit-verdien var positiv, og kan tolkes som at elever med høye karakterer hadde større sannsynlighet til å få riktig svar på oppgaven sammenlignet med elever med lavere karakter. En tolkning av de negative fortegnene til samspillvariablene antydte at det var større sannsynlighet til å få riktig svar på oppgaven når den var markert «lett» eller «middels», sammenlignet med når oppgaven var markert «vanskelig», gitt elever med lave terminkarakterer.

Oppsummert viste den første signifikanstesten at modell 2 var signifikant bedre på 0,01 nivå, til å forklare den avhengige variabelen, sammenlignet med en modell uten uavhengige variabler. Wald-testen er svak for små utvalgsstørrelser og det foretas derfor en grafisk tolkning av samspilleeffekten for å unngå en eventuell type 2-feil.

Betinget effektplott av samspillvariabelen

Figur 5 illustrerer samspilleffekten mellom de uavhengige variablene «Karakter» og «Markering». Verdiene på y-aksen er den avhengige variabelen som viser den predikerte sannsynligheten for at elevene fikk riktig svar på oppgave D. X-aksen viser elevens terminkarakter i faget. Hver av grafene er markert med forskjellige farger for å illustrere de ulike nivåmarkeringene.



Figur 5 Betinget effektplott som illustrerer samspilleffekten til prestasjon på oppgave D

Grafene fra figur 5 viser et karakteristisk bilde av sammenhengen mellom elevenes resultater og oppgavens nivåmarkering. En sammenligning mellom blå og grønn graf viser at for elevene med terminkarakter 1–4, var det var størst sannsynlighet for å få riktig svar på matematikkoppgaven når den var markert «lett», sammenlignet med når den *samme* oppgaven var markert «vanskelig». For eksempel viser grafene at for de elevene med karakter 2 i faget, var sannsynligheten ca. 45 % for at elevene fikk riktig resultat på oppgaven når den var markertes «lett». Det var videre 20 % sannsynlig for at eleven presterte når oppgaven var markert «vanskelig».

Skjæringspunktet til blå -og grønn graf viser for hvilken terminkarakter det er like stor sannsynlighet for at elevene fikk riktig resultat på oppgaven. At grafene krysser hverandre i punktet (4.5, 0.75) betyr at det var 75 % sjanse for at elevene med 4,5 i terminkarakter fikk riktig resultat på oppgaven gitt «lett» eller «vanskelig» nivåmarkering. Karakter 4,5 var ikke oppgitt som et svaralternativ på skjemaet, men er likevel med i analysen ettersom terminkarakteren betraktes på intervallnivå. Karakter 4,5 er gjennomsnittsverdien, bestående av like mange elevene med karakterene 4 og 5 i faget.

En videre analysering av grafene indikerer at det er større sannsynlighet for at elevene med terminkarakter 5 eller 6 fikk riktig resultat på oppgave D, for de elevene som fikk oppgaven markert «vanskelig», sammenlignet med de elevene som fikk oppgaven markert «lett».

Signifikanstesten som ble utført i regresjonsanalysen viste at resultatene hadde en sammenheng med terminkarakteren. Effekten vurderes i figuren ved å betrakte grafenes stigningstall. At grafene har positive stigningstall viser at sannsynligheten for at elevene fikk riktig svar på oppgaven økte med høye terminkarakterer.

Som nevnt var ikke oppgavens markering signifikant på 0,05 nivå, gitt lik terminkarakter. En grafisk tolkning av resultatene kan tyde på at oppgavens nivåmarkering, for elever med lik terminkarakter, hadde en effekt på om oppgaven ble besvart riktig. Etter en grafisk vurdering kan det tenkes at ikke-signifikante verdier skyldes en lav utvalgsstørrelse. Når det er sagt kan resultatene skyldes tilfeldigheter, og en kan derfor ikke med sikkerhet fastslå at liknende resultater vil forekomme i et annet utvalg.

4.2.2 Prestasjon på oppgave E

Hypotese 2 ble testet ved bruk av binomisk logistisk regresjon. Nedenfor er signifikanstestene analysert for å skape et grunnlag for vurdering av nullhypotesen. Hosmer-Lemeshow test ble i tillegg brukt til å vurdere modellens datatilpasning. En oppsummering av regresjonsanalysen er vist i tabell 9.

Tabell 9 Binomisk logistisk regresjonsanalyse utført ved hjelp av SPSS

N = 195 Variabler	Modell 1		Modell 2	
	Logit	Oddsratio	Logit	Oddsratio
Konstant	-2.685**	0.068**	-7,053**	0.001**
Kontrollvariabler:				
Kjønn:				
- Jente	0.202	1.223	0.293	1.340
Rekkefølge:				
- Først	0.277	1.319	0.401	1.493
Uavhengig variabler:				
Markering:				
- Lett	0.485	1.623	4.270 ^a	71.512 ^a
- Middels	0.096	1.100	6.603**	737.124**
Karakter	0.969***	2.635***	2.223**	9.230**
Samspillvariabel				
Karakter*Markering:				
- Karakter*Lett			-1.114	0.328
- Karakter*Middels			-7.053**	0.163**
-2LL	168.295***		158.786**	
Relativ endring i -2LL	0.150***		0.198**	
Hosmer-Lemeshow(p)	21.412 (0.006)		9.251 (0.322)	

-2LL = -2*LogLikelihood, * = p < 0.05, ** = p < 0.01, *** = p < 0.001, a = Sig. = 0.100

Modell 1

Modell 1 inneholdt de uavhengige variablene «Kjønn», «Rekkefølge», «Karakter» og «Markering». -2LL-verdien viste at modell 1 var signifikant bedre enn modell 0 til å predikere valg av oppgave E. Hosmer-Lemeshow signifikanstest indikerte, i motsetning til forrige analyse, at modellen ikke var godt tilpasset data (p < 0.05). En videre analyse av modellen ble derfor ikke foretatt.

Modell 2

Modell 2 testet samspillet mellom elevens karakter og oppgavens nivåmarkering. I tillegg ble det også her kontrollert for kjønn og oppgavens rekkefølge. Etter en tolkning av -2LL-verdiene (se tabell 9) anslås modell 2 som signifikant bedre enn modell 0 og modell 1 på 0,01 nivå. Det betyr at modell 2 var mer egnet enn de foregående modellene til å predikere prestasjon på oppgave E. Med andre ord kan ikke samspillvariabelen «Karakter*Markering» utelukkes, da effekten av nivåmarkeringen trolig hadde en sammenheng med høye og lave terminkarakterer ved riktig eller galt svar på oppgaven.

Hosmer-Lemeshows signifikanttest ga en reduksjon av Kji-kvadrat verdien, og anslo en signifikansverdi på 32,2 %. Det betyr at avviket mellom predikerte og observerte frekvenser

var tilnærmet like, da $p > 0,05$, og at modell 2 er en god tilpasning til data (Skog, 2007, s. 384).

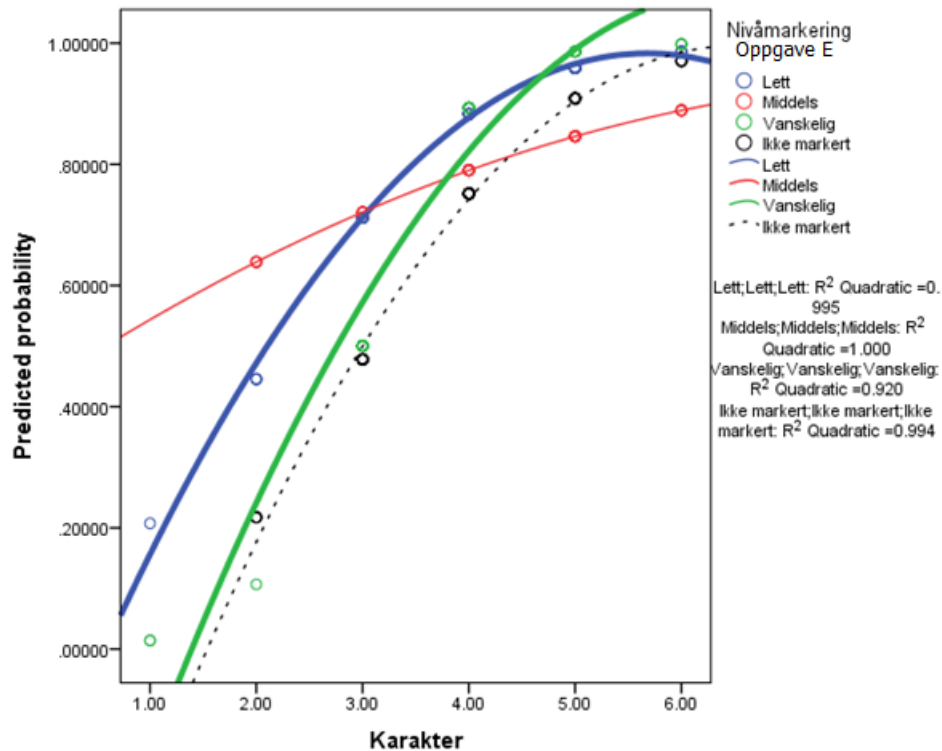
Wald-signifikanstest viste at kontrollvariablene «Kjønn» og «Rekkefølge» ikke hadde en signifikansverdi på under 0,05 nivå. Det betyr at elevens kjønn og oppgavens rekkefølge trolig ikke hadde en effekt på prestasjonen. Signifikanstesten er svak for en lav utvalgsstørrelse og gjør det derfor vanskelig å få signifikante verdier på 0,05 nivå. Testen må derfor tolket med forsiktighet. Videre viste testen at de *negative* samspillvariablene «Karakter*Lett» og «Karakter*Middels» hadde en effekt på prestasjonen. Variabelen «Karakter*Middels» var signifikans på 0,01 nivå. En tolkning av fortegnene indikerte at det var større sannsynlighet for å få riktig svar på oppgaven når den markertes «lett» eller «middels», sammenlignet med når oppgaven var markert «vanskelig», gitt elever med lave terminkarakter.

Oppsummert ga begge signifikanstestene støtte til hypotese 2, som tilsa at nivåmarkeringen til oppgave E og elevens terminkarakter *har* en effekt på sannsynligheten for å få riktig svar på oppgaven. En grafisk tolkning av samspilleffekten gis nedenfor.

Betinget effektplott av samspillvariabelen

Figur 6 illustrerer samspilleffekten mellom de uavhengige variablene «Karakter» og «Markering». Effektplottet blir analysert på tilsvarende måte som forrige analyse.

En tolkning av grafene (se figur 6) tilsier at når oppgaven markertes «lett», blå graf, var det større sannsynlighet for at elevene fikk riktig svar på oppgaven, sammenlignet med når oppgaven markertes «vanskelig», grønn graf. Dette gjaldt for elever med terminkarakter 1-4. For eksempel viser grafene at for elever med karakter 2 i faget var sannsynligheten ca. 45 % for at elevene fikk riktig resultat når oppgaven markertes «lett», og 10 % sannsynlig for at eleven presterer når oppgaven var markert «vanskelig».



Figur 6 Betinget effektplott som illustrerer samspilleffekten til prestasjon på oppgave E

Skjæringspunktet til blå - og grønn graf viser at det er like stor sannsynlighet for at elevene fikk riktig resultat på oppgaven. At grafene krysser hverandre i punktet (4.8, 0.90), betyr at det var 90 % sjans for at elever med litt under 5 i terminkarakter fikk riktig resultat på oppgaven, gitt «lett» eller «vanskelig» nivåmarkering.

En videre analyse antydte at det var større sannsynlighet for at elevene med terminkarakter 5 eller 6 fikk riktig resultat på oppgave E, da den markertes «vanskelig» sammenlignet med oppgaven var markert «lett».

Signifikanstesten som ble utført i regresjonsanalysen viste at resultatene hadde en sammenheng med terminkarakteren. Effekten vurderes i figuren ved å betrakte grafenes stigningstall. At grafen stiger med høye terminkarakterer, tilsier at det var større sannsynlighet for at elevene fikk riktig svar på oppgaven hvis eleven hadde en høy karakter i faget.

Som nevnt var bare den ene samspillvariabelen signifikant på 0,05 nivå, og de resterende samspillvariablene kan derfor skyldes tilfeldigheter. En grafisk tolkning av resultatene tydet imidlertid på at oppgavens nivåmarkeringen, for elever med lik terminkarakter, hadde en

effekt på om oppgaven ble besvares riktig. Etter en grafisk vurdering kan det tenkes at ikke-signifikant verdi *kan* skyldes en lav utvalgsstørrelse.

4.2.3 Prestasjon på oppgave F

Hypotese 2 ble på tilsvarende måte som i de foregående analysene testet ved bruk av binomisk logistisk regresjon. Tabell 10 viser en oppsummering av regresjonsanalysen.

Tabell 10 Binomisk logistisk regresjonsanalyse utført ved hjelp av SPSS

N = 195 Variabler	Modell 1		Modell 2	
	Logit	Oddsratio	Logit	Oddsratio
Konstant	-1.526	0.217	-2.322	0.098
Kontrollvariabler:				
Kjønn:				
- Jente	-0.877*	0.416*	-0.900*	0.407*
Rekkefølge:				
- Først	0.689	1.991	0.719	2.054
Uavhengig variabler:				
Markering:				
- Lett	-0.250	0.779	1.773	5.887
- Middels	0.653	1.922	0.694	2.002
Karakter	0.766***	2.150***	0.984**	2.675**
Samspillvariabel				
Karakter*Markering:				
- Karakter*Lett			-0.531	0.588
- Karakter*Middels			0.009	1.009
Endring i -2LL	178.322***		176.381	
Relativ endring i -2LL	0.144***		0.153***	
Hosmer-Lemeshow(p)	13.183 (0.068)		12.619 (0.082)	

-2LL = -2*LogLikelihood, * = p < 0.05, ** = p < 0.01, *** = p < 0.001

Modell 1

Modell 1 inneholdt de uavhengige variablene «Kjønn», «Rekkefølge», «Karakter» og «Markering». Modell 1 var signifikant bedre enn modell 0 til å predikere valg av oppgaven. Modellen var godt tilpasset data, ettersom at Hosmer-Lemeshow signifikanstest indikerte at p-verdien var over 5 %.

Wald-testen viste at elevens terminkarakter var signifikans på 0,001 nivå. Fortegnet på logit-verdien var positiv. Det tolkes som at elever med høyere karakterer hadde større sannsynlighet til å få riktig svar på oppgaven, sammenlignet med elever med lavere karakterer. I tillegg viste signifikanstesten at elevens kjønn var signifikans på 0,05 nivå. Fortegnet på logit-verdien var negativt, og det indikeres derfor at det var større sjans for at

guttene fikk riktig svar på oppgaven sammenlignet med jentene.

Modell 2

Modell 2 testet samspillet mellom elevens karakter og oppgavens nivåmarkering, med samme kontrollvariabler som i de andre analysene. En tolkning av -2LL-verdien tyder på at modell 2 hadde lavest varians, da -2LL verdien var minst, sammenlignet med modell 1 og 0. Videre viste signifikanstesten at modell 2 ikke var signifikant *bedre* enn modell 1 til å forklare prestasjon på oppgave F. Det betyr at det er usikkert om den predikerte verdien *forbedres* da den uavhengige samspillvariabelen tas med. En sammenligning i mellom modell 2 og modell 0 tilsier at den relative endringen i -2LL verdi var signifikant bedre, på 0,001 nivå, til å forklare den avhengige variabelen.

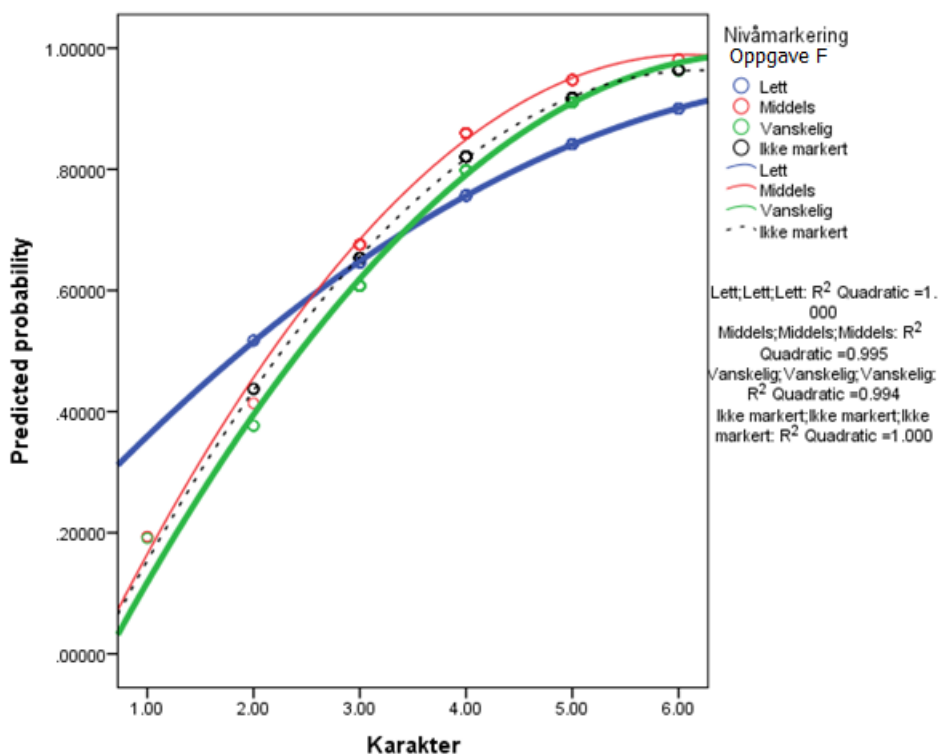
Hosmer-Lemeshows signifikanttest ga en reduksjon av Kji-kvadrat-verdien, og indikerte en signifikansverdi på 8,2 %. Av den grunn vurderes modell 2 som en god tilpasning til data (Skog, 2007, s. 384).

Som nevnt var Wald-testen den svakeste signifikanstesten, og ble derfor tolket med forsiktighet, da den er veldig følsom for små utvalgsstørrelser. Signifikanstesten indikerte at elevens terminkarakter var signifikant på 0,01 nivå. Fortegnet på logit-verdien var positiv. Det tolkes som om at elever med høyere karakter hadde større sannsynlighet til å få riktig svar på oppgaven, sammenlignet med elever med lavere karakterer. I tillegg viste signifikanstesten at elevens kjønn var signifikant på 0,05 nivå. Fortegnet på logit-verdien var negativt, og det tolkes derfor at gutter hadde størst sannsynlighet til å få riktig svar på oppgaven. En tolkning av fortegnene til samspillvariablene indikerte at det var større sannsynlighet for å få riktig svar på oppgaven når den markertes «lett» eller «middels», sammenlignet med når oppgaven ble markert «vanskelig», gitt elever med lave terminkarakterer. Dette da samspillvariablene ble sammenlignet med referansegruppen «Karakter*Vanskelig», som besto av de elevene som fikk oppgaven markert «vanskelig».

Oppsummert viste den første signifikanstesten at modell 2 var signifikant bedre, på 0,001 nivå, til å forklare den avhengige variabelen, sammenlignet med en modell uten uavhengige variabler. Wald-testen er svak for små utvalgsstørrelser, og det ble derfor også her foretatt en grafisk tolkning av samspilleffekten for å unngå en eventuell type II-feil.

Betinget effektplott av samspillvariabelen

Figur 7 illustrerer samspilleffekten mellom de uavhengige variablene «Karakter» og «Markering». En tolkning av grafene tilsier at når oppgaven markertes «lett», blå graf, var det større sannsynlighet for at elevene fikk riktig svar på oppgaven, sammenlignet med når oppgaven var markert «vanskelig», grønn graf. Dette gjaldt for elever med terminkarakter 1-3. For eksempel viser grafene at for elever med karakter 2 i faget var det ca. 50 % sjans for at elevene fikk riktig resultat når oppgaven markertes «lett», og 30 % sjans for at eleven presterte når oppgaven var markert «vanskelig».



Figur 7 Betinget effektplott som illustrerer samspilleffekten til prestasjon på oppgave F

Blå -og grønn grafs skjæringspunkt viste at det var like stor sannsynlighet for at elevene fikk riktig resultat på oppgaven. At grafene krysser hverandre i punktet (3.5, 0.70), betyr at det var 70 % sjans for at elever med 3,5 i terminkarakter fikk riktig resultat på oppgaven, gitt «lett» eller «vanskelig» nivåmarkering. Karakter 3,5 var ikke oppgitt som et svaralternativ på skjemaet, men var med i analysen da terminkarakteren betraktes på intervallnivå. Karakter 3,5 er gjennomsnittsverdien, bestående av like mange elever med karakterene 3 og 4 i faget.

En videre analyse tyder på at det var større sannsynlighet for at elevene med terminkarakter 4

-6 fikk riktig resultat på oppgave F da den markertes «vanskelig», sammenlignet med når oppgaven var markert «lett».

Signifikanstesten som ble utført i regresjonsanalysen indikerte at resultatene hadde en sammenheng med terminkarakteren. Effekten vurderes i figuren ved å betrakte grafenes stigningstall. At grafen stiger med høye terminkarakterer, tilsier at det er større sannsynlighet for at elevene fikk riktig svar på oppgaven for elever med høye karakterer i faget.

Som nevnt var ikke oppgavens markering signifikant på 0,05 nivå, gitt lik terminkarakter. En grafisk tolkning av resultatene tydet imidlertid på at oppgavens nivåmarkeringen, for elever med lik terminkarakter, hadde en effekt på om oppgaven besvartes riktig. Etter en grafisk vurdering kan det tenkes at en ikke-signifikant verdi skyldes en lav utvalgsstørrelse. Når det er sagt kan resultatene også her skyldes tilfeldigheter, og en kan ikke fastslå at liknende resultater vil forekomme med et annet utvalg.

4.2.4 Oppsummering av hovedfunn fra analyse 2

Det ble utført tre regresjonsanalyser. Hosmer-Lemeshow testene indikerte at datamaterialet som ble brukt passet godt til regresjonsmodellene. I SPSS ble det utført to signifikanstester, «Likelihood-ratio»- og Wald-testen. «Likelihood-ratio»-testene viste at samspilleffekten var statistisk signifikant, sammenlignet med en modell uten uavhengige variabler. For oppgave E ga Wald-testen støtte for at «middels» nivåmarkering og elevenes terminkarakterer hadde en effekt på sannsynligheten for å prestere på oppgaven. Som nevnt ga Wald-testen ikke grunnlag til å forkaste nullhypotesen fra oppgave D og F, noe som kan skyldes utvalgsstørrelsen. For å unngå en eventuell type 2 feil ble hypotesene vurdert ved å ta utgangspunktet i en grafisk fremstilling av samspillet mellom elevenes karakter og oppgavens nivåmarkering.

En sammenligning av grafene til hver av oppgavene D – F, antydte at «lett»- og «vanskelig» nivåmarkering hadde en tilsvarende effekt på elevenes resultater. Denne observasjonen medfører en antakelse om at utvalgsstørrelsen kan ha bidratt til ikke-signifikante verdier, på 0,05 nivå, for «lett» og «vanskelig» nivåmarkering. Igjen påpekes det at resultatene kan skyldes tilfeldigheter.

Ettersom de andre grafene varierte mer fra hver av effektplottene, ble analysen avgrenset til å fokusere på effekten ved å tilføre oppgavene «lett» og «vanskelig» nivåmarkering. Denne avgrensningene ble gjort for å gi et sterkere svar på delproblemstilling: *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes prestasjon på oppgaven?*

En samlet vurdering av regresjonsanalysene ga en indikasjon på at oppgavens nivåmarkering *har* en betydning for om elevene fikk riktig svar på oppgaven. Grafene viste at det er størst sannsynlighet for at elevene med terminkarakterer 1 – 4 fikk riktig svar på oppgaven når den var markert «lett», sammenlignet med når den *samme* oppgaven var markert «vanskelig». For elever med terminkarakter 5 eller 6 viste resultatene at det er størst sannsynlighet for å få riktig svar på oppgavene når den er markert «vanskelig», sammenlignet med når samme oppgave markertes «lett». Resultatene fra analysen viste også at elevenes terminkarakter hadde en effekt på prestasjonene. Elevene med høye terminkarakterer hadde størst sannsynlighet til å få riktig svar på oppgaven, sammenlignet med elevene med lave terminkarakterer.

Et annet funn fra analysen var at guttene hadde størst sjanse til å få riktig svar på oppgave F sammenlignet med jentene. Dette funnet blir ikke drøftet videre i oppgaven, ettersom det bare var signifikante sammenhenger i en av oppgavene, og kan derfor skyldes tilfeldigheter. I oppgavens avslutning blir det imidlertid gitt forslag til videre forskning på området.

Etter å ha sammenlignet «lett» og «vanskelig» nivåmarkering tyder resultatene på at nivåmarkeringen kan ha påvirket elevenes svar på oppgavene, og at effektstyrken varierte med elevenes terminkarakter. Hvordan denne effekten kan ha betydning for matematikkundervisning vil bli drøftet i oppgavens diskusjonsdel.

5 Diskusjon

I dette kapitlet gis det først en kort sammenfatning av masteroppgaven. Deretter foretas en kritisk refleksjon rundt bruk av nivådifferensierte oppgaver i matematikkundervisningen. Avslutningsvis gis det forslag til videre forskning som omhandler nivådifferensiering og mestringsforventning i matematikkfaget.

5.1 Oppsummering

Formålet med oppgaven var å undersøke om ulike nivåmarkeringer har en effekt på elevenes valg av matematikkoppgave, samt om markeringen påvirker elevenes resultater. Utvalget besto av 289 elever fra 9. trinn ved praksisskoler i Oslo, som bruker nivådifferensierte lærebøker. Det ble utført en kvantitativ undersøkelse som ga et øyeblikksbilde av situasjonen. Skjemaet var tredelt og ble besvart av en eksperimentgruppe og en kontrollgruppe.

I del 1 ble elevene bedt om å regne tre matematikkoppgaver med tilnærmet lik vanskelighetsgrad. Eksperimentgruppens oppgaver ble imidlertid påført forskjellige nivåmarkeringer til tross for at oppgavene altså hadde samme vanskelighetsgrad. Markeringene var tilfeldig fordelt på oppgavene og varierte innad i gruppen, hvilket betyr at noen elever fikk oppgave A markert som «lett», mens andre fikk den *samme* oppgaven markert som «middels», eller som «vanskelig». Kontrollgruppen fikk oppgavene utdelt uten markert vanskelighetsgrad. Hensikten var å måle om nivåmarkeringen hadde en påvirkning på om elevene fikk riktig eller galt svar på oppgavene.

I del 2 var formålet å registrere elevenes valg av oppgave. Oppgavene det kunne velges mellom hadde lik vanskelighetsgrad. Eksperimentgruppen fikk beskjed om å velge én av tre tilsynelatende nivådifferensierte matematikkoppgaver. Kontrollgruppen fikk lik beskjed, men deres oppgaver var uten markert vanskelighetsnivå. Hensikten var å sammenligne gruppene for å måle om nivåmarkeringen hadde en eksperimentell påvirkning på elevenes oppgavevalg.

I del 3 svarte elevene på forskjellige spørsmål hvor variablene «kjønn» og «terminkarakter» ble brukt senere i analysen. Karakterskalaen strekker seg fra 1 til 6, hvor 1 er laveste og 6 er høyeste grad av fagkompetanse.

Problemstillingen var todelt, og ble analysert hver for seg. Hensikten med den første analysen («analyse 1»), var å undersøke; *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes valg av oppgave?* «Binomisk logistisk regresjon» ble anvendt som analysemetode. Den samme metoden ble brukt i den andre analysen («analyse 2»), for å besvare følgende; *I hvilken grad kan matematikkoppgavens markerte vanskelighetsnivå ha en effekt på elevenes prestasjon på oppgaven?*

5.2 Hovedfunn

Resultatene fra den første analysen viste at oppgavens markerte nivå hadde en prediktiv effekt på elevenes valg av oppgave. Oppgavevalget hadde en sammenheng både med (1) elevens terminkarakter og (2) oppgavens nivåmarkering;

(1) En sammenligning mellom elevenes terminkarakterer og oppgavevalget tydet på at det var størst sannsynlighet for at oppgaven med «lett» nivåmarkering ble valgt for elevene med lave terminkarakterer, og minst sannsynlighet for at den samme oppgaven ble valgt for elevene med høye terminkarakterer. Elevene med høye terminkarakterer tenderte i større grad til å velge oppgaven som tilfeldigvis var markert «vanskelig». Andre funn fra analysen var at det ikke var like stor sammenheng mellom oppgavevalget og terminkarakteren for kontrollgruppen. Det samme viste seg for elevene i eksperimentgruppen som fikk oppgaven markert som «middels».

(2) Et annet funn fra analysen var at oppgavens nivåmerkingen, gitt elever med *like* terminkarakterer, hadde en effekt på oppgavevalget. En sammenligning mellom eksperimentgruppen og kontrollgruppen viste at nivåmerkingene «lett» og «vanskelig» avvakte mest fra kontrollgruppen, og ble derfor vurdert til å ha størst effekt på om oppgaven ble valgt. Det betyr at selv om elevene hadde like terminkarakterer var oppgavens nivåmarkering en avgjørende faktor for om oppgaven ble valgt. Valg av «middels» nivåmarkering skilte seg minst fra kontrollgruppen, og ble derfor vurdert til å ha lavest effekt på oppgavevalget.

Resultatene fra den andre analysen antydte at forskjellige nivåmarkeringer av identiske matematikkoppgaver hadde betydning for elevenes evne til å løse oppgaven riktig. Elevenes resultater hadde en sammenheng både med (1) elevens terminkarakter og (2) oppgavens nivåmarkering;

(1) Et funn, som ikke var overraskende, var at elevenes terminkarakter hadde en effekt på prestasjonene. Elevene med høye terminkarakterer hadde størst sannsynlighet for å få riktig svar på oppgaven, sammenlignet med elevene med lave terminkarakterer.

(2) En tolkning av grafene kan tyde på at *selv* om elevene hadde *like* terminkarakterer, karakterene 1, 2, 3 eller 4, hadde elevene som fikk oppgaven markert «lett» størst sannsynlighet til å få riktig svar på oppgaven, sammenlignet med de som fikk den *samme* oppgaven markert «vanskelig». For elevene med terminkarakterene 5 eller 6 viste en tolkning av grafene at nivåmarkeringen hadde en *motsatt* effekt. Det betyr at *selv* om elevene hadde *like* terminkarakterer var det flere som fikk riktig resultat når oppgaven var markert «vanskelig», i motsetning til de elevene som fikk den *samme* oppgaven markert «lett».

5.3 Konklusjon

Innledningsvis ble det satt spørsmålstegn til intensjonen bak nivådifferensiering; Kan det tenkes at nivådifferensierte oppgaver bidrar til at det er *elevene* som tilpasser seg oppgavenivået, og ikke at *oppgavenivået* tilpasses eleven, slik det er intendert at det skal gjøre? En samlet tolkning av analysen peker i denne retningen.

En vurdering av analyseresultatene kan tyde på at elevene lot nivåmarkeringen være bestemmende for valg av oppgave. En intensjonal forklaring av oppgavevalget, for elevene med lave terminkarakterer, kan være at de ønsket å regne lette matematikkoppgaver. Videre trodde elevene at oppgavene på skjemaet hadde forskjellige vanskelighetsnivåer, og resultatet ble at de valgte den oppgaven som var markert lettest. En motsatt forklaring kan gjelde for elevene med høye terminkarakterer. En slik forklaring *kan* tyde på at elevene ikke vurderte oppgavenivået i seg selv, men valgte det oppgavenivået som de trodde samsvarte best til egne ferdigheter.

Analyseresultatene antydte også at elevenes evne til å besvare oppgavene riktig kan en sammenheng med nivåmarkeringen. En intensjonal forklaring på elevenes oppgaveprestasjon kan være at de faglig svake elevene ønsket å få riktig svar på den «enkle» oppgaven, elevene trodde oppgaven var lett, og resultatet ble at de klarte oppgaven. Motsatt forklaring kan gjelde for de faglig sterke elevene. Dette *kan* bety at oppgavens faktiske vanskelighetsgrad ikke var avgjørende for om oppgaven ble besvart riktig, men heller at det var elevenes forestilling om

vanskelighetsgrad, ut fra nivåmarkeringen, som var utslagsgivende for resultatene.

Nedenfor gis en kritisk refleksjon over hvilken betydning forskningsresultatene kan ha for læring og undervisning av matematikkfaget.

5.4 Kritisk refleksjon rundt nivådifferentierte oppgaver

Selv om undervisningen skal tilpasses slik at elevene får medvirkning til opplæringen, stilles det med denne oppgaven kritiske spørsmål til at elevene på 9. trinn selv får ansvar for å velge vanskelighetsgrad på oppgaven. Dette begrunnes med funn fra analysen som tyder på at elevene velger oppgave etter en vurdering av oppgavens nivåmarkering.

Som nevnt viste resultatene at det var størst sannsynlighet for at elevene med lavest terminkarakter valgte oppgaven som var markert «lett», sammenlignet med når den samme oppgaven var markert som «vanskelig». En mulig forklaring kan være at eleven hadde en lav mestringsforventning. Dette støttes på mange måter av Bandura (1997) som hevder at den selektive prosessen påvirkes av mestringsforventningen, og at elevene med lav forventning om mestring velger mindre krevende oppgaver. En lav mestringsforventning kan skyldes de tidligere tilbakemeldingene elevene har fått under oppgaveløsning. Det kan tenkes at elevene som valgte den «lette» oppgaven har hørt fra venner, familie og/eller lærere at de bør regne den laveste nivåmarkeringen. Slike kommentarer kan være med på å redusere mestringsforventning i faget, noe som kan resultere i at elevene ikke våger å prøve seg på en høyere nivåmarkering i fremtiden. Bruk av nivådifferentierte oppgaver i matematikkundervisningen kan på bakgrunn av dette få konsekvenser hvis elevene velger oppgavenivået selv.

En negativ følge av nivåmarkeringen kan være at elever som *egentlig* hadde klart oppgaver på et vanskeligere nivå, velger for lette oppgaver fordi de har en lav mestringsforventning. Hvis så er tilfelle, kan dette også få konsekvenser for deres medelever; andres mestringsopplevelser er også med på å forme egne mestringsforventninger. Dersom en elev sammenligner seg med en medelev som velger et for lett oppgavenivå, kan dette altså medføre at denne eleven også velger en for lett oppgave på grunnlag av nivåmarkering. Dette kan få betydning for eleven,

siden arbeid med for enkle oppgaver kan resultere i kjedsomhet og et lavt læringsutbytte (Csikszentmihalyi, 2005). Bruk av nivå-differensierte lærebøker kan på bakgrunn av dette få negative konsekvenser for eleven hvis nivå-markeringen «skremmer» elevene til å ikke velge andre oppgaver enn det de tror de vil klare. Dersom oppgavens markering bidrar til at elevene blir «stående fast» på et nivå, kan i verstefall bruk av nivå-markerte oppgaver resultere i at elevene kategoriseres som tapere og vinnere i faget.

Signifikanstestene fra analysen viste også at det var størst sannsynlighet for at elevene med høye terminkarakterer valgte oppgaven med «vanskelig» nivå-markering. Dette samsvarer med tidligere forskning som viser at faglig sterke elever ønsker å velge oppgaver som er utfordrende (Bandura, 1997). Ettersom verbal overtalelse er en viktig kilde som konstruerer mestringsforventningen, kan en mulig forklaring være at elevene som valgte den «vanskelige» oppgaven har blitt *oppmuntret* til å velge oppgavene fra læreboken med høyest nivå-markering. Å markere oppgavens vanskelighetsgrad kan være til fordel for disse elevene, siden de på egen hånd får muligheten til å unngå de letteste oppgavene til fordel for å velge de oppgavene som er markert vanskeligere. Dette kan øke sjansen for at eleven havner innenfor sin flytsone, og dermed oppnår mestring og optimal læringseffekt. Når det er sagt kan nivå-markerte oppgaver forårsake at elever velger et for høyt oppgavenivå. Resultatet kan dermed bli at elevene «havner» ovenfor flytsonen, hvilket kan føre til prestasjonsangst og frustrasjon, samt lite faglig utbytte (Csikszentmihalyi, 2005). Dersom dette i tillegg medfører en lavere mestringsforventning kan det hende at elevene ikke våger å prøve seg på liknende oppgaver i fremtiden. Dette er i tråd med Bandura (1997) som hevder at tidligere mestrings erfaringer påvirker mestringsforventningene, og at nederlag over tid kan medføre unngåelse av krevende oppgaver.

Resultatene fra undersøkelsen viste også at nivå-markeringen kan påvirke elevenes evne til å få riktig eller galt svar på oppgaven. Analyseresultatene antydde at selv om elevene hadde like terminkarakterer (karakterene 1, 2, 3 eller 4), var det mindre sannsynlig å få riktig svar på oppgaven for elevene som fikk oppgaven markert «vanskelig», sammenlignet med «lett» nivå-markering. En tolkning kan være at de affektive prosessene påvirket elevenes resultater. Kanskje var det noen elever som ble stresset og/eller engstelige når de fikk opplyst at oppgavens nivå-markering var «vanskelig». Hvis dette var tilfelle er det en enda større grunn til å anta at elevenes mestringsforventning var en mellomforliggende variabel som påvirket

resultatene i undersøkelsen. Dette støttes av Collins (1982) som viser til en undersøkelse der elevenes mestringsforventning påvirket oppgaveresultatene, og at lav mestringsforventning reduserer sjansen for å oppnå gode resultater (referert i, Bandura, 1993). I tillegg viser forskning at mestringsforventning har betydning for utholdenhet og innsats (Bandura, 1997). Det kan derfor spekuleres i om elevene viste en lavere innsats og utholdenhet når oppgaven var markert «vanskelig», sammenlignet med når oppgaven var markert «lett». En tolkning kan være at disse elevene ikke gjorde et ordentlig forsøk på den «krevende» oppgaven fordi de *trodde* de ikke ville få riktig svar på oppgaven. Bruk av nivåmarkerte oppgaver kan på bakgrunn av dette være problematisk, da det er arbeid med *krevende* oppgaver som bidrar til et større læringsutbytte (Botten, Daland, & Dalvang, 2008). Hvis elevene ikke våger å prøve på vanskeligere oppgaver kan det tenkes at de forblir på samme oppgavenivå gjennom hele ungdomsskolen.

En motsatt effekt av nivåmarkering syntes å gjelde for elever med terminkarakterer 5 eller 6. Empirisk forskning viser at elevens karakter ofte korrelerer med mestringsforventningen, og at elever med høy mestringsforventning ofte viser større innsats og utholdenhet under oppgaver som er krevende (Bandura, 1997). På bakgrunn av dette kan en mulig forklaring være at elevene med høye terminkarakterer hadde en høy mestringsforventning, noe som resulterte i økt innsats og utholdenhet når oppgaven var markert «vanskelig». Feil svar på oppgaven markert «lett» kan på den andre siden skyldes slurv og mindre tidsbruk på oppgaven. Dersom den «vanskelige» markeringen bidro til en økt mestringsforventning, og dermed resulterte i at elevene fikk større innsats og utholdenhet under oppgaveløsingen, kan det være fordelaktig for disse elevene at oppgavene markeres med en høy vanskelighetsgrad.

Det er viktig å presisere at nivådifferensierte oppgaver er et viktig tiltak som brukes for å tilpasse opplæringen, spesielt med tanke på at Norge ønsker en enhetsskole (Solvang, 1992, s. 7). I denne oppgaven er det imidlertid nivådifferensiering satt i system, og de pedagogiske tiltakene som benyttes, det rettes et kritisk søkelys på.

5.5 Ytre validitet

Utvalgets representativitet i forhold til populasjonen må vurderes for å kunne konstituere den ytre validiteten, samt undersøkelsens kontekst.

At undersøkelsen foregikk i elevenes klasserom styrker antagelsen om en naturlig situasjon, og konteksten er derfor med på å styrke den ytre validiteten (Kleven, 2011e, s. 136).

Situasjonen *kan* likevel oppfattes som kunstig, fordi matematikkoppgavene opphav var på et skjema og ikke fra elevenes lærebok. At alle elevene besvarte skjemaet i klasserommet, vurderes som at utvalget hadde nokså lik kontekst. I tillegg styrkes den ytre validiteten ettersom undersøkelsen tilsammen ble utført i 12 klasserom. Antakelsen er at en test utført i mange klasserom, gir en større sannsynlighet for at resultatene også er gyldige i andre klasserom (Kleven, 2011e, s. 135-138).

Som nevnt ble utvalget inndelt etter oppgavens nivåmarkering og elevens terminkarakter, hvilket gjorde at hver utvalgsgruppe ble forholdsvis liten. Denne inndelingen medfører at den ytre validiteten svekkes. I tillegg svekkes den ytre validiteten fordi det er foretatt et bekvemmelighetsutvalg (Kleven, 2011e, s.123-134). Selv om det derfor ikke gis grunnlag for en statistisk generalisering, reiser resultatene likevel viktige problemstillinger som bør undersøkes nærmere (se neste punkt).

Resultatene fra undersøkelsen gir grunn til å anta at funnene kan være generelt gyldige for elever som bruker denne formen for nivåmarkerte oppgaver. Når det er sagt vil trolig lærerens bruk av slike oppgaver ha betydning for elevenes erfaringer med nivådifferensiering. For å styrke antakelsen om at resultatene er gyldige bør studiet gjentas i flere utvalg. Hvis resultatene fra analysen viser seg å gjelde i en bredere skala, kan det være behov for å revurdere de didaktiske tiltakene som brukes i dag. Dette fordi et «feil» valg av vanskelighetsnivå kan få negativ betydning for elevenes motivasjon, samt læringsutbytte (Csikszentmihalyi, 2005). På bakgrunn av dette vil neste del inneholde forslag til videre forskning rundt bruk av nivåmarkerte oppgaver i matematikkfaget.

5.6 Forslag til videre forskning

Ettersom det ble utført en tverrsnittstudie er forskningsdesignet begrenset. Det er derfor uvisst om effekten av nivåmarkerte oppgaver reduseres, økes eller holdes konstant over tid. Et forbedringsforslag er å foreta et panelstudie som undersøker de samme observasjonsheter over en lengre periode. Det kan gi informasjon om hvordan nivådifferensierte matematikkoppgaver påvirker matematikkundervisningen over tid. Det er mulig ny forskning

gir resultater som tilsier at effekten av å nivåmarkerte oppgaver øker med tiden. Et slikt funn kan gi viktig informasjon til matematikkdiraktikken og bidra til drøfting om hvorvidt nivåmarkerte oppgaver er det beste for eleven.

Et lite utvalg i denne testen kan bidra til type 2-feil ved signifikanstesting, hvilket tilsier at testen bør utføres i en større skala. Størrelsen på utvalget gjorde det vanskelig å vurdere om elevenes kjønn og oppgavens rekkefølge hadde en effekt på den avhengige variabelen. På bakgrunn av teori utelukkes det ikke at oppgavens rekkefølge hadde betydning for resultatene, og at effekten av nivåmarkeringen hadde en sammenheng for ulike kjønn. Det samme gjelder for vurderingen av hvorvidt nivåmarkeringen hadde en effekt på om elevene fikk riktig eller galt svar på oppgavene. Dette styrkes nødvendigheten av å utføre testen med et større utvalg. På denne måten kan resultatene bli mer presise, hvilket gjør det lettere å vurdere effektstyrken. I tillegg kan et større utvalg, der utvelgelsen skjer tilfeldig, bidra til at resultatene i større grad blir generaliserbare (Kleven, 2011d).

Datamaterialet i denne oppgaven gir ikke nok informasjon til å forklare hvorfor markering hadde en effekt på oppgavevalget, og hvorfor elevene fikk riktig eller galt svar på oppgavene. Selv om mestringsforventning ikke ble målt i undersøkelsen ble det tatt for gitt at den var en mellomliggende variabel. Et forbedringspotensial kan derfor være å måle elevenes mestringsforventning både før og under oppgaveløsningen. For å få informasjon om de kognitive prosessene vil det være behov for å konstruere et mer avansert skjema. En god idé kan derfor være å bruke «mixed-method» for å få et rikere datamateriell. For eksempel kan data fra intervju bidra til kunnskap om mulige forklaringsvariabler, som senere kan anvendes under utformingen av et mer omfattende skjema. Dette støttes av Hannula, som påpeker viktigheten av å bruke flere forskningsmetoder for å måle psykologiske begreper i matematikk (2006).

Et annet forslag er å utføre oppgaveskjemaet elektronisk hvor datamaskinen registrerer elevene tidsbruk per oppgave. Dette kan gi informasjon om elevenes utholdenhet og innsats under arbeid med forskjellige nivåmarkerte oppgaver. I tillegg kan et dataprogram som registrerer elevenes øyebevegelser benyttes til å beskrive hvor på skjermen elevene leser. Dette kan utføres ved å bruke et infrarødt kamera, for eksempel kameraet RED (remote eye tracking device). Denne metoden kan gi et materiell som beskriver elevenes oppmerksomhetsfokus under oppgavevalget.

Litteraturliste

- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), s. 117-148.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.
- Bandura, A., & Locke, E. A. (2003). Negative Self-Efficacy and Goal Effects Revisited. *Journal of Applied Psychology*, 88(1), s. 87-99.
- Botten, G., Daland, E., & Dalvang, T. (2008). Tilpasset matematikkopplæring i en inkluderende skole. *Tangenten 2/2008*, s. 23-27.
- Christophersen, K.-A. (2012). *IBM SPSS/ AMOS: - Databehandling og statistisk analyse*. Oslo: Akademika forlag.
- Christophersen, K.-A. (2013). *Introduksjon til statistisk analyse: Regresjonsbaserte metoder og anvendelse*. Oslo: Gyldendal Akademisk.
- Cohen, L., Manion, L., & Morrison, K. (2011). Research Methods in Education. I *Surveys, longitudinal, cross-sectional and trend studies* (s. 256-287). London: Routledge.
- Csikszentmihalyi, M. (2005). *Flow og engagement i hverdagen*. Virum: Dansk psykologisk Forlag.
- Dale, E. L., & Wærness, J. I. (2003). *Differensiering og tilpassning i grunnsopplæringen: Rom for alle - blick for den enkelte*. Oslo: Cappelen akademisk forlag.
- Eikemo, T. A., & Clausen, T. H. (2012). *Kvantitativ analyse med SPSS: - En praktisk innføring i kvantitative analyseteknikker*. Trondheim: Tapir Akademisk Forlag.
- Evans, T. D. (1997). Utvikling i barne- og ungdomsårene. I G. Høstmark, & K. Raaheim (Red.), *En innføringsbok i psykologi for universiteter og høyskoler* (s. 319-361). Oslo: Cappelen Akademisk Forlag.
- Everitt, B. S. (2002). *The Cambridge dictionary of statistics*. Cambridge: Cambridge University Press.

- Grunntall. (u.d.). *Hvordan læreverket er bygd opp og tankene bak*. Hentet April 24, 2014 fra Grunntall: <http://grunntall.no/images/pdf/Produkter/OmLaereverket.pdf>
- Hannula, M. S. (2006). Affect in Mathematical Thinking and Learning: Towards Integration of Emotion, Motivation, and Cognition. I J. Maasz, & W. Schloeglmann (Red.), *New Mathematics Education Research and Practice* (s. 209-232). Rotterdam: Sense Publishers.
- Imsen, G. (1997). *Lærerens verden: innføring i generell didaktikk*. Oslo: Tano Aschehoug.
- Imsen, G. (2008). *Elevenes verden: innføring i pedagogisk psykologi*. Oslo: Universitetsforlaget.
- Kinnear, P. R., & Gray, C. D. (2004). *SPSS 12 Made Simple*. New York: Psychology Press.
- Kleven, T. A. (2011a). Forskning og forskningsresultater. I T. A. Kleven (Red.), F. Hjordemaal, & K. Tveit, *Innføring i pedagogisk forskningsmetode: En hjelp til kritisk tolkning og vurdering* (s. 9-26). Oslo: Unipub.
- Kleven, T. A. (2011b). Data og datainnsamlingsmetoder. I T. A. Kleven (Red.), F. Hjordemaal, & K. Tveit, *Innføring i pedagogisk forskningsmetode: En hjelp til kritisk tolkning og vurdering* (s. 27-48). Oslo: Unipub.
- Kleven, T. A. (2011c). Hvordan er begrepene operasjonalisert? Spørsmål om begrepsvaliditet. I T. A. Kleven (Red.), F. Hjordemaal, & K. Tveit, *Innføring i pedagogisk forskningsmetode: En hjelp til kritisk tolkning og vurdering* (s. 85-102). Oslo: Unipub.
- Kleven, T. A. (2011d). Hvilke alternative forklaringer er mulige? Spørsmål om indre validitet. I T. A. Kleven (Red.), F. Hjordemaal, & K. Tveit, *Innføring i pedagogisk forskningsmetode: En hjelp til kritisk tolkning og vurdering* (s. 103-121). Oslo: Unipub.
- Kleven, T. A. (2011e). Hvilken kontekst er resultatene gyldige i? Spørsmålet om ytre validitet. I T. A. Kleven (Red.), F. Hjordemaal, & K. Tveit, *Innføring i pedagogisk forskningsmetode: En hjelp til kritisk tolkning og vurdering* (s. 123-138). Oslo: Unipub.

- Kunnskapsdepartementet. (2011). *Motivasjon – Mestring – Muligheter. Ungdomstrinnet* (St. meld nr. 22 (2010-2011)). Oslo: Kunnskapsdepartementet. Hentet fra <http://www.regjeringen.no>
- Mathiassen, K. (2009). Lektor – adjunkt – lærer: Artikler for studiet i praktisk-pedagogisk utdanning. I R. Mikkelsen, & H. Flademoe (Red.), *Differensiert undervisning* (s. 123-136). Oslo: Universitetsforlaget.
- Meyer, W.-U. (1987). Perceived ability and achievement-related behavior. I F. Halisch, & J. Kuhl (Red), *Motivation, intention, and volition* (ss. 73-86). Berlin: Springer-Verlag.
- Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology*, 20, s. 426-443.
- Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, 86(2), s. 193-203.
- Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of Counseling Psychology*, 42(2), s. 190-198.
- Skaalvik, E. M., & Skaalvik, S. (2005). *Skolen som læringsarena: Selvoppfatning, motivasjon og læring*. Oslo: Universitetsforlaget.
- Skog, O.-J. (2007). *Å forklare sosiale fenomener: En regresjonsbasert tilnærming*. Oslo: Gyldendal Akademisk.
- Solvang, R. (1992). *Matematikdidaktikk*. Oslo: NKI Forlaget.
- Teigen, K. H. (1997). Motivasjon. I G. H. Nielsen, & K. Raaheim (Red.), *En innføringsbok i psykologi for universiteter og høyskoler* (s. 195-209). Oslo: Cappelen Akademiske Forlag.
- Thronsen, I. S. (2005). *Selvregulert læring av matematikkferdigheter: en studie av elever fra begynnertrinnet*. Oslo: Universitas.

- Utdanningsdirektoratet. (2012, Januar 23). *Prinsipp for opplæringen*. Hentet Januar 21, 2014 fra Utdanningsdirektoratet: <http://www.udir.no/Lareplaner/Kunnskapsloftet/Prinsipp-for-opplaringa/>
- Utdanningsdirektoratet. (u.d.a). *Veiledning i lokalt arbeid med læreplaner*. Hentet Januar 21, 2014 fra Utdanningsdirektoratet: <http://www.udir.no/Lareplaner/Veiledninger-til-LK06/Veiledning-i-lokalt-arbeid-med-lareplaner/tilpasset-opplaring/>
- Utdanningsdirektoratet. (u.d.b). *Den generelle delen av læreplanen: Det arbeidende mennesket*. Hentet April 26, 2014 fra Utdanningsdirektoratet: <http://www.udir.no/Lareplaner/Kunnskapsloftet/Generell-del-av-lareplanen/Det-arbeidende-mennesket/>
- Utdanningsdirektoratet. (u.d.c). *Læreplan i matematikk fellesfag - Kompetansemål*. Hentet Mai 8, 2014 fra Utdanningsdirektoratet: <http://www.udir.no/kl06/MAT1-03/Kompetansemaal/?arst=98844765&kmsn=334280449>
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational psychology*, 82 (1), s. 51-59.

Vedlegg

Vedlegg 1: Kalibrering av matematikkoppgavene

Vedlegg 2: Et eksempel på et oppgaveskjema

Vedlegg 3: Følgebrev

Vedlegg 1: Kalibrering av matematikkoppgavene

Sannsynlighetsoppgaver

Nedenfor vises oppgavene som ble kalibrert i Hamar (kodes 1 og 2) og Lørenskog (kodes 3 og 4). For eksempel betyr «1: 38 %», under oppgave 1, at det var 38 % av elevene fra Hamar som fikk riktig svar på oppgaven.

Oppgave 1 – Valgt

Per er 23 år og bor i Oslo. Hvor stor er sannsynligheten for at bursdagen til Per er på en søndag i år?

1: 38 % 2: 60 % 3: 62 % 4: 67 % **Snitt: 57 %**

Oppgave 2

I et lotteri der det selges 100 000 lodd, er det 200 gevinster. Hvor stor er sannsynligheten for å vinne en av gevinstene hvis du kjøper ett lodd?

1: 77 % 2: 72 % 3: 62 % 4: 70 % **Snitt: 70 %**

Oppgave 3 – Valgt

I klasse 9 G er det til sammen 27 elever, 11 jenter og 16 gutter. En av elevene er ordenselev. Hvor stor er sannsynligheten for at Ida er ordenselev?

1: 65 % 2: 72 % 3: 73 % 4: 63 % **Snitt: 68 %**

Oppgave 4

Foran hver kamp trekker spillerne på fotballaget plassen de skal spille på. Lars er en av de elleve spillerne. Hvor stor er sannsynligheten for at Lars blir trukket ut til å være målvakt i dagens kamp?

1: 77 % 2: 100 % **Snitt: 88 %**

Oppgave 4

Du og vennene dine spiller et spill. Det er din tur å kaste en vanlig terning. Hva er sannsynligheten for å få et oddetall når du kaster terningen?

3: 0,62 4: 0,78 **Snitt: 70 %**

Oppgave 5

Silje kjøpte en kurv jordbær. Seks av de 30 bærene var råtne. Hvor stor var sannsynligheten for at Silje ikke tok ett råttent jordbær, dersom hun ikke så på bærene i kurven da hun smakte på et jordbær?

1: 15 % 2: 24 % **Snitt: 20 %**

Oppgave 5

Det er 24 elever i en gruppe, 14 jenter og 10 gutter. Læreren vil høre en tilfeldig elev i lekse. Hva er sannsynligheten for at en gutt blir spurt?

3: 54 % 4: 74 % **Snitt: 64 %**

Oppgave 6

I en fyrstikkeske ligger det 4 brukte og 16 ubrukte fyrstikker. Kalid tar fram en fyrstikk for å tenne et lys i mørket. Hvor stor er sannsynligheten for at han tar en fyrstikk som ikke er brukt?

1: 27 % 2: 28 % **Snitt: 27 %**

Oppgave 6 – Valgt

I en skål ligger det fem kuler, 3 blå og 2 røde. Hanna har bind foran øynene, og trekker en kule opp fra skålen. Hva er sannsynligheten for at Hanna trekker en blå kule?

3: 62 % 4: 70 % **Snitt: 66 %**

Oppgave 7

Alle elevene skal opp til skriftlig eksamen i 10. klasse. Mulige prøvefag er norsk, engelsk og matematikk. Prøvefaget trekkes, og det er like stor sannsynlighet for alle tre fagene. Hvor sannsynlig er det at matematikk trekkes ut?

1: 73 % 2: 96 % 3: 89 % 4: 100 % **Snitt: 89 %**

Problemløsningsoppgaver

Nedenfor vises oppgavene som ble kalibrert i Hamar (kodes 1 og 2) og Lørenskog (kodes 3 og 4). For eksempel betyr «3: 50 %», under oppgave 1, at det var 50 % av elevene fra Lørenskog som fikk riktig svar på oppgaven.

Oppgave 1 – Valgt

Julie og Torstein er to søsken som bor i Tromsø. Torstein er eldst, og er fire år eldre enn Julie. Til sammen er de 28 år. Hvor gammel er Julie?

1: 74 % 2: 60 % 3: 50 % 4: 70 % **Snitt: 66 %**

Oppgave 2

En dag Per og Pål var i svømmehallen, svømte de til sammen 860 m. Pål svømte 120 m lengre enn Per. Hvor langt svømte Per?

1: 52 % 2: 36 % 3: 35 % 4: 33 % **Snitt: 39 %**

Oppgave 3

Faren til Margit er tre ganger så gammel som Margit. Til sammen er de 68 år. Hvor gammel er Margit?

1: 22 % 2: 32 % **Snitt: 27 %**

Oppgave 3

Hanna og Herman vil dele en pose med 47 karameller slik at Hanna får 11 karameller mer enn Herman. Hvor mange karameller får de hver?

3: 46 % 4: 30 % **Snitt: 38 %**

Oppgave 4

Martin har 50 kr mer enn Lotte. Simen har 20 kr mindre enn Lotte. De har 690 kr til sammen. Hvor mange kroner har hver av dem?

1: 39 % 2: 40 % 3: 23 % 4: 11 % **Snitt: 28 %**

Oppgave 5

Siri er halvparten så gammel som Sander. Sander er tre år eldre enn Knut. Til sammen er de 53 år gamle. Hvor gammel er Siri?

1: 0 % 2: 4 % **Snitt: 2 %**

Oppgave 5

Sara kjøper 5 pizzaer og 10 brus til en klassefest. Det koster til sammen 600 kr. Hvor mye koster en pizza dersom brusen koster 20 kr per flaske?

3: 42 % 4: 37 % **Snitt: 40 %**

Oppgave 6 – Valgt

Lise, Lene og Lotte har vært på bærtur. Lise plukket 2 kg mer enn Lene. Lotte plukket dobbelt så mange kg bær som Lise. Lene plukket 3 kg blåbær og 2 kg tyttebær. Hvor mange kg bær plukket Lise?

1: 65 % 2: 96 % 3: 69 % 4: 33 % **Snitt: 65 %**

Oppgave 7

Ask og Birk skal ta samme tog. Birk går om bord på vogn nr. 12 telt bakfra, mens Ask går om bord på vogn nr. 34 telt forfra. Begge ender opp i samme vogn. Hvor mange vogner har toget?

1: 17 % 2: 16 % **Snitt: 16 %**

Oppgave 7 – Valgt

Anna fra Bergen er 3 år eldre enn søsteren sin. De er 29 år til sammen. Hvor gammel er søsteren til Anna?

3: 54 % 4: 48 % **Snitt: 51 %**

Resultat 1 – Hamar, N = 50 elever

Tabell 11 Resultater fra Hamar, sannsynlighetsoppgavene, N = 50

Hamar	Oppg. 1	Oppg. 2	Oppg. 3	Oppg. 4	Oppg. 5	Oppg. 6	Oppg. 7
9 A	38 %	77 %	65 %	77 %	15 %	27 %	73 %
9 B	60 %	72 %	72 %	100 %	24 %	28 %	96 %
Totalt	49 %	75 %	69 %	88 %	20 %	27 %	84 %

Tabellen viser resultatene etter kalibreringen av sannsynlighetsoppgavene fra Hamar kommune.

Tabell 12 Resultater fra Hamar, problemløsningsoppgavene, N = 50

Hamar	Oppg. 1	Oppg. 2	Oppg. 3	Oppg. 4	Oppg. 5	Oppg. 6	Oppg. 7
9 A	74 %	52 %	22 %	39 %	0 %	65 %	17 %
9 B	60 %	36 %	32 %	40 %	4 %	96 %	16 %
Totalt	65 %	43 %	27 %	39 %	2 %	80 %	16 %

Tabellen viser resultatene etter kalibreringen av problemløsningsoppgavene fra Hamar kommune.

Før oppgavene ble testet videre i Lørenskog ble noen av oppgavene byttet ut for å unngå «tak-» og/eller «bunneffekt». Oppgavene som ble erstattet var nr. 4, 5 og 6 fra sannsynlighetsoppgavene og nr. 3, 5 og 7 fra problemløsningsoppgavene.

Resultat 2 – Lørenskog, N = 53 elever

Tabell 13 Resultater fra Lørenskog, sannsynlighetsoppgavene, N = 53

Lørenskog	Oppg. 7	Oppg. 6 (ny)	Oppg. 5 (ny)	Oppg. 4 (ny)	Oppg. 3	Oppg. 2	Oppg. 1
9 A	89 %	62 %	54 %	62 %	73 %	62 %	62 %
9 B	100 %	70 %	78 %	70 %	63 %	70 %	67 %
Totalt	94 %	66 %	70 %	66 %	68 %	66 %	64 %

Tabellen viser resultatene etter kalibreringen av sannsynlighetsoppgavene fra Lørenskog kommune.

Tabell 14 Resultater fra Lørenskog, problemløsningsoppgavene, N = 53

Lørenskog	Oppg. 7 (ny)	Oppg.6	Oppg. 5 (ny)	Oppg. 4	Oppg. 3 (ny)	Oppg. 2	Oppg. 1
9 A	54 %	69 %	42 %	23 %	46 %	35 %	50 %
9 B	48 %	33 %	37 %	11 %	30 %	33 %	70 %
Totalt	51 %	51%	40 %	17 %	38 %	34 %	60 %

Tabellen viser resultatene etter kalibreringen av problemløsningsoppgavene fra Lørenskog kommune.

Samlet resultat – Hamar og Lørenskog

Tabell 15 Samlet resultat fra Hamar og Lørenskog, sannsynlighetsoppgaver, N = 104 og problemløsningsoppgaver, N = 101

Hamar og Lørenskog	Oppg. 1	Oppg. 2	Oppg. 3	Oppg. 4	Oppg. 5	Oppg. 6	Oppg. 7
Sannsynlighetsoppgaver	57 %	70 %	68 %				89 %
Problemløsningsoppgaver	66 %	39 %		28 %		65 %	

Tabellen viser samlet resultat etter kalibreringen av matematikkoppgavene. De blanke rutene er oppgavene som ble fjernet etter den første kalibreringen (Hamar kommune), og de grønne rutene viser hvilke oppgaver som ble tatt ut for bruk i undersøkelsen.

Tabell 16 Resultater fra sannsynlighetsoppgavene som bare ble brukt i Lørenskog, N = 53

	Oppg. 4 (Lørens.)	Oppg. 5 (Lørens.)	Oppg. 6 (Lørens.)
Sannsynlighetsoppgaver	70 %	39 %	66 %

Tabellen viser resultatene fra de ny oppgavene som ble kalibrert i Lørenskog kommune. De grønne rutene viser hvilke oppgaver som ble tatt ut for bruk i undersøkelsen.

Tabell 17 Resultater fra problemløsningsoppgavene som bare ble brukt i Lørenskog, N = 53

	Oppg. 3 (Lørens.)	Oppg. 5 (Lørens.)	Oppg. 7 (Lørens.)
Problemløsningsoppgaver	38 %	40 %	51 %

Tabellen viser resultatene fra de oppgavene som ble erstattet, og kalibrert i Lørenskog

kommune. De grønne rutene viser hvilke oppgaver som ble tatt ut for bruk i undersøkelsen

Etter en vurdering av resultatene ble følgende oppgaver valgt:

Sannsynlighetsoppgaver:

Oppgave 1: 57 % av elevene fikk riktig resultat (N = 104)

Oppgave 3: 68 % av elevene fikk riktig resultat (N = 104)

Oppgave 6: 66 % av elevene fikk riktig resultat (kun elever fra Lørenskog, N= 53)

Problemløsningsoppgaver:

Oppgave 1: 66 % av elevene fikk riktig resultat (N = 101)

Oppgave 6: 65 % av elevene fikk riktig resultat (N = 101)

Oppgave 7: 51 % av elevene fikk riktig resultat (kun elever fra Lørenskog, N= 53)

Vedlegg 2: Et eksempel på et skjema

Del 1

I del 1 skal du løse alle oppgavene

Oppgave 1 - Lett vanskelighetsgrad • (Blå oppgave)

Julie og Torstein er to søsken som bor i Tromsø. Torstein er eldst, og er fire år eldre enn Julie. Til sammen er de 28 år. Hvor gammel er Julie?

Oppgave 2 - Middels vanskelighetsgrad ▲ (Rød oppgave)

Anna fra Bergen er tre år eldre enn søsteren sin. De er 29 år til sammen. Hvor gammel er søsteren til Anna?

Oppgave 3 - Vanskelig vanskelighetsgrad ■ (Grønn oppgave)

Lise, Lene og Lotte har vært på bærtur. Lise plukket 2 kg mer enn Lene. Lotte plukket dobbelt så mange kg bær som Lise. Lene plukket 3 kg blåbær og 2 kg tyttebær. Hvor mange kg bær plukket Lise?

Del 2

I del 2 skal du velge **én** oppgave, og deretter løse den. For å løse oppgaven kan du bruke denne formelen:

$$\text{Sannsynlighet} = \frac{\text{Antall gunstige utfall}}{\text{Antall mulige utfall}}$$

Eksempel:

- Hva er sannsynligheten for å få en firer når du kaster en vanlig terning?

Løsning:

Det finnes bare én firer på terningen. *Antall gunstige utfall* blir derfor: **1**

Når du kaster terningen har du seks muligheter. Du kan enten få 1, 2, 3, 4, 5 eller 6 på terningen. *Antall mulige utfall* blir derfor: **6**

$$\text{Sannsynlighet} = \frac{\text{Antall gunstige utfall}}{\text{Antall mulige utfall}} = \frac{1}{6}$$

Sannsynligheten er $\frac{1}{6}$ for at du får en firer når du kaster en vanlig terning.

Velg én oppgave:

I **Lett vanskelighetsgrad • (Blå oppgave)**

Per er 23 år og bor i Oslo. Hvor stor er sannsynligheten for at bursdagen til Per er på en søndag i år?

II **Middels vanskelighetsgrad ▲ (Rød oppgave)**

I en skål ligger det fem kuler, 3 blå og 2 røde. Hanna har bind foran øynene, og trekker en kule opp fra skålen. Hva er sannsynligheten for at Hanna trekker en blå kule?

III **Vanskelig vanskelighetsgrad ■ (Grønn oppgave)**

I klasse 9 G er det til sammen 27 elever, 11 jenter og 16 gutter. En av elevene er ordenselev. Hvor stor er sannsynligheten for at Ida er ordenselev?

Skriv utregning og svaret på den oppgaven du valgte:

Del 3

For hvert spørsmål nedenfor skal du markere **ett** kryss på det svaralternativet som du mener passer best.

1. Hvilket kjønn er du?

- Jente
- Gutt

2. Hvor ofte regner du de blå oppgavene fra matematikkboken? (Sett ett kryss)

- Alltid
- Ofte
- Noen ganger
- Sjelden
- Aldri
- Vet ikke

3. Hvor ofte regner du de røde oppgavene fra matematikkboken? (Sett ett kryss)

- Alltid
- Ofte
- Noen ganger
- Sjelden
- Aldri
- Vet ikke

4. Hvor ofte regner du de grønne oppgavene fra matematikkboken? (Sett ett kryss)

- Alltid
- Ofte
- Noen ganger
- Sjelden
- Aldri
- Vet ikke

5. Hvilken oppgavefarge regner du oftest? (Sett ett kryss)

- Blå
- Rød
- Grønn
- Like mye av blå og rød
- Like mye av rød og grønn
- Like mye av alle fargene
- Annet
- Vet ikke

6. Hvem bestemmer oftest hvilken oppgavefarge du regner? (Sett ett kryss)

- Mest meg selv
- Mest læreren
- Mest familien
- Mest venner
- Like mye meg selv og læreren
- Like mye meg selv og familien
- Like mye meg selv, læreren og familien
- Annet
- Vet ikke

7. Hvor ofte får du riktig svar på de blå oppgavene fra matematikkboken? (Sett ett kryss)

- Alltid
- Ofte
- Noen ganger
- Sjelden
- Aldri
- Vet ikke

8. Hvor ofte får du riktig svar på de røde oppgavene fra matematikkboken? (Sett ett kryss)

- Alltid
- Ofte
- Noen ganger
- Sjelden
- Aldri
- Vet ikke

9. Hvor ofte får du riktig svar på de grønne oppgavene fra matematikkboken? (Sett ett kryss)

- Alltid
- Ofte
- Noen ganger
- Sjelden
- Aldri
- Vet ikke

10. Hvilken matematikkarakter fikk du på terminkortet? (Sett ett kryss)

- Karakter 1
- Karakter 2
- Karakter 3
- Karakter 4
- Karakter 5
- Karakter 6
- Vet ikke

Vedlegg 3: Følgebrev

Informasjon til eleven

Hensikten med spørreundersøkelsen er å få informasjon om hvordan elever fra 9. klasse arbeider med forskjellige matematikkoppgaver. Oppgavene er hentet fra ulike lærebøker som blir brukt på ungdomsskoler.

Spørreundersøkelsen skal brukes i tilknytning til arbeidet med min masteroppgave i realfagsdidaktikk. Masteroppgaven vil bli levert våren 2014, ved Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.

Det er frivillig å delta på spørreundersøkelsen. Ved å levere spørreskjemaet ferdig utfyllt regnes det som din godkjenning for at informasjonen brukes i arbeid med masteroppgaven. Alle elever og skoler som deltar på spørreundersøkelsen vil bli behandlet anonymt. Du skal derfor ikke skrive navnet ditt på spørreskjemaet.

Har du spørsmål må du gjerne ta kontakt på telefon eller e-post.

Mvh.
Maria Herset

Tlf.: 416 45 311

E-post: mkherset@gmail.com