



# Målrettet og tilrettelagt matematikkundervisning for elever som presterer svakt i matematikk

Resultater fra et randomisert kontrollert forsøk på 8. trinn og videregående skole i Oslo

TALL

SOM  
FORTELLER

Mathilde Bjørnset (Fafo), Trude Gunnes (SSB), Marianne Takvam Kindt (Fafo), Lars J. Kirkebøen (SSB), Lena Lindenskov (DPU), Jon Rogstad (Fafo) og Marte Rønning (SSB)

RAPPORTER / REPORTS

2020 / 7



*Mathilde Bjørnset (Fafo), Trude Gunnes (SSB),  
Marianne Takvam Kindt (Fafo), Lars J. Kirkebøen  
(SSB), Lena Lindenskov (DPU), Jon Rogstad (Fafo)  
og Marte Rønning (SSB)*

**Målrettet og tilrettelagt matematikkundervisning  
for elever som presterer svakt i matematikk**

Resultater fra et randomisert kontrollert forsøk på 8.  
trinn og videregående skole i Oslo

I serien Rapporter publiseres analyser og kommenterte statistiske resultater fra ulike undersøkelser. Undersøkelser inkluderer både utvalgsundersøkelser, tellinger og registerbaserte undersøkelser.

© Statistisk sentralbyrå  
Ved bruk av materiale fra denne publikasjonen skal Statistisk sentralbyrå oppgis som kilde.

Publisert 11. februar 2020

ISBN 978-82-587-1068-1 (trykt)  
ISBN 978-82-587-1069-8 (elektronisk)  
ISSN 0806-2056

<b>Standardtegn i tabeller</b>	<b>Symbol</b>
Tall kan ikke forekomme	.
Oppgave mangler	..
Oppgave mangler foreløpig	...
Tall kan ikke offentliggjøres	:
Null	-
Mindre enn 0,5 av den brukte enheten	0
Mindre enn 0,05 av den brukte enheten	0,0
Foreløpig tall	*
Brudd i den loddrette serien	—
Brudd i den vannrette serien	
Desimaltegn	,

## Forord

Denne rapporten beskriver gjennomføring og evaluering av et prosjekt med målsetning om å øke fullføringen av videregående opplæring gjennom å forbedre matematikkprestasjonene til svakt presterende elever på 8. trinn og Vg1.

Prosjektet er en del av et større program - Program for bedre gjennomføring (PBG) finansiert av Kunnskapsdepartementet. I PBG samarbeider fylkeskommuner og forskningsmiljøer om å utvikle tiltak for å redusere frafall i videregående opplæring, samt gjennomføre og evaluere tiltakene.

Prosjektet har vært et samarbeid mellom SSB, Fafo, Danmarks institutt for Pædagogik og Uddannelse (DPU) og Utdanningsetaten i Oslo kommune (UDE). Tiltaket er gjennomført av UDE, i samarbeid og med faglig veiledning fra DPU. SSB har vært prosjektleder og ansvarlig for kvantitativ effektevaluering. Fafo har vært ansvarlig for kvalitativ prosessevaluering.

Det er tidligere publisert et notat fra dette prosjektet (Kirkebøen m.fl., 2018), som beskriver gjennomføring og foreløpige resultater fra evalueringen. I denne rapporten studeres i første omgang hvordan tiltaket har påvirket elevprestasjoner. DPU hatt et særlig ansvar for denne rapportens kapittel 2, SSB for kapitlene 4-6 og Fafo for kapittel 7. UDE har beskrevet gjennomføring i tiltaket i kapittel 3. SSB, Fafo og DPU står felles ansvarlig for innledning og oppsummering.

Vi vil takke UDE generelt for bidraget til denne rapporten, og spesielt Christine Nybø Brattenborg og Terje Engh Wiig. Videre takker vi kollega Simon Bensnes og Ola Vestad samt forskningsdirektør Brita Bye for konstruktive kommentarer og tilbakemeldinger.

Statistisk sentralbyrå, 01.02.2020

Brita Bye

## Sammendrag

Formålet med denne rapporten er å beskrive og evaluere et tiltak gjennomført i Oslo-skolen. Tiltaket var rettet mot elever med svake matematikkferdigheter på 8. trinn og 1. trinn på videregående (Vg1). Målsetningen for tiltaket var å sette flere elever i stand til å fullføre videregående opplæring (VGO) ved å bedre deres ferdigheter i matematikk. Tiltaket inneholdt to elementer: 1) kursing av lærere i didaktikk egnet for denne målgruppen av elever; 2) gi elevene tilrettelagt undervisning (i stedet for ordinær undervisning) i to perioder á 4-6 uker høst og vårr. På grunnskolen foregikk disse øktene i små grupper, mens de på videregående ble gjennomført i vanlige grupper.

Tiltaket ble gjennomført som et eksperiment, der 24 av 48 ungdomsskoler i Oslo og 9 av 17 aktuelle videregående skoler ble tilfeldig valgt ut til å delta. Vi evaluerer tiltaket ved å sammenligne resultater i tiltaks- og kontrollskoler. Vi kan ennå ikke studere fullføring av VGO, og fokuserer derfor på resultater fra nasjonal prøve i regning på 9. trinn (NP9) og Oslo-prøven i matematikk på Vg1.

I ungdomsskolen finner vi at gjennomsnittlig NP9 blant målgruppeelevene øker med 0,2-0,4 skalapoeng, men vi kan ikke konkludere med at dette er en effekt av tiltaket. Til tross for tilfeldig tilordning av skoler til tiltak og kontroll finner vi resultatforskjeller forut for tiltaket. Ulike tilnærminger for å ta hensyn til forskjeller i forutsetninger gir forskjellige resultater, noe som gjør tolkningen av resultatene krevende. De fleste resultatene er ikke statistisk signifikante, men vi finner at andelen målgruppeelever på laveste mestringsnivå reduseres med 2 prosentpoeng, fra et utgangsnivå på omtrent 10 prosent. Denne forskjellen er signifikant på 10 prosent nivå. Vi finner ikke effekter på elever utenfor målgruppen. Basert på effektestimater og tidligere forskning gjør vi anslag på økonomisk nytte av tiltaket. Vi kan ikke konkludere med at tiltaket er samfunnsøkonomisk lønnsomt. Samtidig er det stor usikkerhet i estimatene, slik at heller ikke kan utelukke effekter som gir samfunnsøkonomisk lønnsomhet.

I videregående skole finner vi ingen effekt på elevenes ferdigheter målt ved Oslo-prøven. Vi finner imidlertid økt fullføring av Vg1 for de to første årskullene (hvor vi har data for fullføring av Vg1). Vi kan ennå ikke konkludere om økt fullføring er en effekt av tiltaket. Dersom det på sikt blir en tilsvarende effekt på fullføring av VGO vil tiltaket i videregående ha vært samfunnsøkonomisk lønnsomt.

Vi har også gjennomført en kvalitativ evaluering, basert på intervjuer og spørreundersøkelser. Denne viser en endring over tid i holdningen til tiltaket. Deltakende lærere var i utgangspunktet positive til innholdet i tiltaket, men dels skeptiske til gjennomføringen. Andre år endret dette seg til en mer generell positiv holdning, mens tiltaket siste år i større grad ble en del av den ordinære undervisningen, og skolens fokus flyttet til nye prosjekter. Lærernes svar om egen undervisning samt observasjon av undervisningsøkter tyder på at tiltaket har påvirket undervisningen. Tiltaksundervisningen er likevel ikke fullt i tråd med de didaktiske prinsippene for tiltaket, og disse prinsippene brukes også i noen grad ved andre skoler. Organisatoriske forhold kan ha redusert effekten av tiltaket. Det har vært utfordringer knyttet til jobbskifter blant tiltakslærere, og strenge rammer for utvelgelse av elever kan ha gjort det vanskelig å velge de elevene som har størst nytte av tiltaket.

Oppsummert har tiltaket i stor grad lyktes i å nå fram til tiltakslærerne, og oppleves som relevant for disse. Vi kan derimot ikke konkludere klart med at tiltaket har lyktes i å bedre elevenes resultater, selv om det er tegn til dette. På lengre sikt vil det være mulig å studere eventuelle effekter på faktisk fullføring av VGO.

## Abstract

The purpose of this report is to describe and evaluate an intervention implemented in public lower and upper secondary schools in Oslo. The intervention targeted students with poor mathematics skills in 8th grade and the first year of upper secondary education (Vg1) and aimed to enable more students to complete upper secondary education (VGO) by improving their math skills. The intervention consisted of: 1) didactics training of teachers and 2) adapted instruction instead of regular instruction for students during two periods of 4-6 weeks (one in autumn and one in spring). In lower secondary school, adapted instruction took place in small groups, in upper secondary schools in regular groups.

The intervention was implemented as a randomized experiment, in which 24 of 48 lower secondary schools in Oslo and 9 of 17 relevant upper secondary schools were randomly selected to participate. We evaluate the intervention by comparing results in treatment and control schools. We are not yet able to study the completion of VGO, and therefore focus on the results on the standardized national test in numeracy in 9th grade (NP9) and the Oslo test in mathematics at Vg1.

We find that average NP9 in the target group in lower secondary school increases by 0.02-0.04 percent of a standard deviation but cannot conclude that this is an effect of the intervention. Most effect estimates are not statistically significant. However, we do find that the share of target students at the lowest mastery level is reduced by 2 percentage points, from an initial level of about 10 percent. This difference is significant at the 10 percent level. We do not find any signs of effects on students outside the target group. Based on the effect estimates and previous research, we estimate the economic benefit of the intervention. We cannot conclude that the intervention is cost-effective. At the same time, there is substantial uncertainty in the estimates, and we cannot exclude effects that make the intervention cost-effective.

In upper secondary, we find no effect on the students' mathematics skills as measured by the Oslo test. However, we find evidence of increased completion of Vg1 for the first two cohorts (where we have data for completion of Vg1). We cannot yet conclude whether increased completion is an effect of the intervention. If in the long term, there is a corresponding effect on the completion of VGO, the intervention in upper secondary will have been cost-effective.

We also conducted a qualitative evaluation, based on interviews and surveys. This evaluation shows a change over time in the attitude towards the intervention. Participating teachers were initially positive to the content of the intervention but sceptical about features of the implementation. This changed to a more generally positive attitude. In the final year, the intervention became a part of regular teaching, and the schools' focus shifted to new projects. Teachers' responses about their teaching and observations of lessons indicate that the intervention had an impact on teaching. However, teaching during the intervention is not fully in line with the didactic principles of the intervention, nor are these principles exclusive to the treatment schools. Organizational conditions may have limited the effect of the intervention. There were challenges with turnover among teachers, and possibly with selecting the students who would benefit most from the intervention.

In summary, the intervention has largely succeeded in connecting with and motivating the treatment school teachers and is seen as relevant by these. On the other hand, we cannot conclude that the intervention has succeeded in improving the students' results, although there are some signs of this. In the longer term, it will be possible to study any effects on actual completion of VGO.

# Innhold

<b>Forord</b> .....	<b>3</b>
<b>Sammendrag</b> .....	<b>4</b>
<b>Abstract</b> .....	<b>5</b>
<b>1. Innledning</b> .....	<b>7</b>
<b>2. Beskrivelse av tiltaket</b> .....	<b>12</b>
2.1. Tiltakets faglig/didaktiske innhold.....	13
2.2. Etterutdanning av lærere og utforming av undervisningsopplegg .....	16
2.3. Organisering av undervisningen.....	20
<b>3. Organisering av tiltaket</b> .....	<b>21</b>
3.1. Faglig innhold.....	21
3.2. Forankring.....	22
3.3. Implementering .....	25
3.4. Evaluering .....	26
3.5. Avsluttende kommentar .....	26
<b>4. Effektevaluering og utforming av tiltaket</b> .....	<b>27</b>
4.1. Identifisering av effekter .....	27
4.2. Forventede resultater, virkningsmekanismer og styrke- beregning .....	32
<b>5. Effektestimater – grunnskolen</b> .....	<b>34</b>
5.1. Datagrunnlag.....	34
5.2. Balansering av tiltak- og kontrollskoler .....	36
5.3. Effekter.....	39
5.3.1. Effekter for alle elever (målgruppeelever + øvrig elever) .....	46
5.4. Oppsummering og kost/nytte – grunnskole .....	46
<b>6. Effektestimater – videregående</b> .....	<b>49</b>
6.1. Datagrunnlag.....	49
6.1.1. Andre bakgrunnskjennetegn ved elevene .....	51
6.2. Balansering av tiltak- og kontrollskoler .....	51
6.3. Effektestimat – videregående.....	52
6.4. Oppsummering og kost/nytte – videregående.....	53
<b>7. Kvalitativ analyse</b> .....	<b>55</b>
7.1. Analytisk rammeverk: Implementering og endring .....	55
7.2. Data og metode.....	56
7.3. Resultater: Normalisering - fra Singapore til Oslo-skolen.....	58
7.4. Rektorene – implementeringsansvarlig på skolen.....	59
7.4.1. Lærerne: Fra instruks til oppslutning .....	63
7.4.2. Elevenes opplevelse .....	69
7.5. Hvordan nyttiggjøre erfaringene fra tiltaket i framtiden?.....	72
7.6. Avslutning .....	73
<b>8. Sammenstilling av resultater</b> .....	<b>75</b>
<b>Referanser</b> .....	<b>77</b>
<b>Figurregister</b> .....	<b>79</b>
<b>Tabellregister</b> .....	<b>80</b>



## 1. Innledning

Omtrent 25 prosent av elevene som starter videregående opplæring fullfører ikke i løpet av 5 år (SSB). De individuelle og samfunnsøkonomiske kostnadene ved frafall er betydelig (Lillejord m.fl., 2015; Falch m.fl., 2009). Å øke gjennomstrømningen i videregående opplæring (heretter VGO) står derfor høyt på den politiske dagsorden.

Fullføring av VGO er vesentlig dårligere blant elever med svake faglige resultater fra grunnskolen (Falch m. fl., 2014b; Utdanningsdirektoratet, 2015a) og ferdigheter i matematikk viser seg å være spesielt viktig for senere skoleprestasjoner og fullføring. Hægeland m.fl. (2010) finner at sammenhengen mellom fullføring av Vg1 og matematikkarakter fra grunnskolen er sterkere enn sammenhengen mellom grunnskolekarakterene i norsk og engelsk og fullføring av Vg1. Falch m.fl. (2011) finner at matematikk er et av fire fag der resultatene fra grunnskolen er særlig assosiert med fullføring av VGO. Matematikk er også faget der flest elever ikke består, noe som direkte bidrar til at elever ikke fullfører VGO med bestått i alle fag.

Formålet med denne rapporten er å utforme og evaluere et tiltak rettet mot elever som presterer svakt i matematikk, hvis hensikt er å øke sannsynligheten for at de fullfører videregående opplæring.

### Tiltak

Tiltaket som vi både utformer og evaluerer i denne rapporten er utført i Oslo og har bestått av to deler. Den første delen gikk ut på å gi didaktiske verktøy av ulik konkretiseringsgrad til matematikklærere ved enkelte tilfeldig utvalgte Oslo-skoler gjennom en målrettet etterutdanning. I den andre delen av tiltaket har omtrent 3000 elever ved disse tilfeldig utvalgte Oslo-skolene (årlig omtrent 500 på 8. trinn og 500 på Vg1 over tre år) som presterte svakt i matematikk (såkalt målgruppeelever) mottatt strukturert og tilrettelagt undervisning av disse kursede lærerne.

På 8. trinn har målgruppeelevene på tiltaksskolene mottatt målrettet undervisning i løpet av to avgrensede perioder på 4-6 uker høst og vår. Et flertall av målgruppeelevene har fått tilrettelagt undervisning i mindre grupper (maksimalt 6-8 elever) sammensatt for formålet. Intensivundervisningen erstattet all ordinær matematikkundervisning for elevene i tiltaksperiodene. På Vg1 har undervisningen foregått i grupper av vanlig størrelse bestående av de elevene som tilhører klassen. I praksis betyr dette at det er et mindre klart skille mellom tiltaksperioden og øvrig undervisning på Vg1 enn på 8. trinn.

Utdanningsetaten (UDE) i Oslo kommune har i nært samarbeid med, og under faglig rettleiding av Danmarks institutt for Pædagogik og Uddannelse (DPU), vært ansvarlig for å utforme det faglige innholdet i tiltaket. UDE har i tillegg vært tiltakseier og ansvarlig for forankring, implementering og oppfølging av tiltaket. Tiltaket og intensivundervisningen har vært rettet mot grunnleggende matematikkferdigheter.

### Tidligere forskning

Både forskning og erfaring peker på at ekstra innsats og forbedring i matematikkundervisningen kan bidra til økt gjennomføring av videregående opplæring.

Det finnes internasjonal forskning som tyder på at tiltak rettet mot bedrede matematikkferdigheter kan bidra til å redusere frafall fra VGO. Cortes m.fl. (2014) studerer effekten av en dobling av timetallet i matematikk i det første året av videregående i Chicago. Studien finner at blant elevene som fikk ekstraundervisning i matematikk var det omtrent 10 prosentpoeng flere som fullførte enn

blant andre tilsvarende elever (fullføring økte fra ca. 60 til 70 prosent). Cook m.fl. (2015) studerer et randomisert kontrollert eksperiment der en veileder gir intensiv matematikkopplæring. Forfatterne finner at deltakelse i veiledningen har en sterk effekt på skoleprestasjoner. Matematikkresultatene økte med 0,2-0,3 standardavvik (i en norsk karakterfordeling svarer det til omtrent like mange karakterpoeng), og andelen som strøk i matematikk ble halvert.

En norsk studie, Falch m.fl. (2014a), sammenligner elever som kommer opp i skriftlig eksamen i forskjellige fag ved avsluttet grunnskole. Studien finner at det å bli tilfeldig trukket ut til matematikk eksamen på ungdomsskole reduserer frafall i videregående med omtrent 0,7 prosentpoeng, og tolker dette som en effekt av at elevene arbeider intensivt med matematikk i eksamensforberedelsene.

Fryer (2017) finner i en omfattende metaanalyse at såkalt «high-dosage tutoring», som vi kan oversette med intensiv smågruppeundervisning, er et av få typer tiltak som ofte har stor effekt på elevenes skoleprestasjoner. I følge Fryer krever dette en gruppestørrelse på maks seks elever og et omfang tilsvarende 50 timer i løpet av 36 uker. Han finner at øvrig smågruppeundervisning, med større grupper og/eller mindre omfang, gjennomgående har liten effekt.

Forskning viser at bruk av målrettet undervisning kan ha god effekt i barneskolen (Torgersen m.fl., 2011; Clements m.fl., 2013). Slike måter å undervise på er imidlertid mindre utbredt i høyere trinn. En av grunnen til det er forskning som viser at tidlig innsats er gunstig (Carneiro og Heckman, 2003). Samtidig sliter mange elever på ungdomsskolen med 3. og 4. klasse pensum. Dette er hovedargumentet for at det er valgt å tilby eldre elever i 8. trinn og Vg1 målrettet undervisning i grunnleggende matematikk.

## Evaluerings

Evaluerings av tiltaket baserer seg på forskjellige typer analyser: Effekt-evaluering, kvalitativ prosessevaluering, samt analyser av undervisningen. De to første er diskutert i denne rapporten mens analyser av undervisningen, basert på klasseromsobservasjoner ble presentert i undervisningsrapporten (Kirkebøen m.fl., 2018).

Effektevaluering baserer seg på å sammenligne resultatene til elever i tiltaksskolene med elever i kontrollskoler. 24 ungdomsskoler og 9 videregående skoler er tilfeldig valgt ut som tiltaksskoler, fra totalt 48 tilgjengelig ungdomsskoler i Oslo og 17 aktuelle videregående skoler (Oslo-skoler med et tilstrekkelig antall målgruppeelever med matematikk 1P<sup>1</sup>). På den måten sikrer vi (forventningsmessig) likhet mellom tiltak- og kontrollgruppen slik at forskjeller i resultat mellom de to gruppene kan gis en kausal tolkning, dvs. tolkes som effekter av tiltaket.

I denne rapporten presenterer vi resultater for elever som gikk på 8. trinn i de siste to tiltaksårene, 2017/18 og 2018/19, og elever fra alle tiltaksårene (2016/17, 2017/18 og 2018/19) i Vg1. I 2016/17 var tiltaket organisert på en annen måte på 8. trinn, der et flertall av skolene kun fikk finansiering av smågrupper, og ikke kursing av lærere. Resultater fra dette året er omtalt i mer detalj i Kirkebøen m.fl. (2018).

Den langsiktige målsetningen er å øke elevenes fullføring av VGO. Det er imidlertid for tidlig for å studere dette, og vi vil i denne rapporten fokusere på ferdigheter i regning/matematikk, målt ved nasjonal prøve i regning på 9. trinn

---

<sup>1</sup> På Vg1 kan elevene velge mellom 1P og 1T der P står for praktisk matematikk og T for teoretisk matematikk.

(heteretter NP9, for tiltakselever i 8. trinn) og Oslo-prøven i matematikk (for tiltakselever i Vg1). For de to første årskullene med tiltak i Vg1 har vi også data for fullføring av Vg1.

I de kvalitative analysene settes søkelys på *subjektive erfaringer*, hvilket innebærer at aktørens egen forståelse legges til grunn i analysene.<sup>2</sup> Disse to tilnærmingene kan supplere hverandre. På den ene siden er det relevant å tallfeste effektene av et tiltak. På den andre siden er det også av betydning å lære av aktørens erfaringer, som ikke alltid lar seg måle med ferdig definerte kategorier. Med bruk av en eksplorerende tilnærming kan vi også få svar på spørsmål vi ikke har kunnskap nok til å formulere på forhånd. Dessuten er det rimelig å anta at de involvertes subjektive oppfatninger vil virke inn på det som kan måles statistisk. Blant annet vil oppfatninger om tiltaket virke inn på forhold som lojalitet, og derigjennom gjennomføring av tiltaket.

### Resultater effektevaluering grunnskole

Fra effektevalueringen finner vi ikke noe klart grunnlag for å konkludere med at tiltaket har ført til økte gjennomsnittlige faglige ferdigheter, verken i målgruppen eller blant øvrige elever.

I målgruppen finner vi at gjennomsnittlig resultat på NP9 er omtrent 0,2 skalapoeng høyere i tiltaksskolene i vår hovedanalyse. Denne forskjellen er imidlertid ikke statistisk signifikant, og kan derfor også skyldes tilfeldigheter.

Til tross for at skolene er tilfeldig fordelt til tiltak og kontroll finner vi en forskjell i elevenes familiebakgrunn og resultater på nasjonal prøve i regning på 8. trinn (NP8), dvs. før tiltaket i favør tiltaksskolene. En slik forskjell kan oppstå tilfeldig fordi vi har et begrenset antall skoler å fordele til tiltak og kontroll (vi diskuterte dette mer i Kirkebøen m.fl, 2018). I effektanalysene må vi ta hensyn til denne ubalansen i tiltaks- og kontrollskolene ved å kontrollere for elevenes forutsetninger.

I hovedanalysene bruker vi NP8 som mål på elevenes forutsetninger. I alternative analyser bruker vi enten nasjonal prøve i regning på 5. trinn (NP5) og/eller bare elevenes familiebakgrunn. Dette gir høyere effektestimater, omtrent 0,4 skalapoeng. Med kontroll for NP5 er dette statistisk signifikant på 10 prosent nivå, men ikke på 5 prosent nivå.

Forskjellene i målte forutsetninger mellom tiltaks- og kontrollgruppene, og de litt avvikende resultatene vi får når vi på forskjellige måter tar hensyn til disse forskjellene, gjør det krevende å tolke resultatene. NP8 er i utgangspunktet vårt foretrukne mål på elevens forutsetninger, men det er ikke åpenbart at dette er det beste målet. Vi finner også en liten positiv forskjell i andelen elever som i det hele tatt gjennomfører NP9, som bidrar til å trekke resultatene i tiltaksskolene ned.

I sum mener vi det er sannsynlig, om enn ikke sikkert, at snittresultatet på NP9 til elevene i målgruppen er bedret med 0,2-0,4 skalapoeng. Dette er en liten effekt sammenlignet med de tidligere refererte studiene, og med målsetningen for tiltaket, som var en bedring på 1 skalapoeng eller mer.

Vi finner at andelen elever på laveste mestringsnivå i målgruppen er redusert med omtrent 2 prosentpoeng. Dette resultatet er statistisk signifikant på 10 prosent nivå.

---

<sup>2</sup> Denne tilnærmingen åpner for et potensielt avvik mellom det som framkommer når aktørene selv forteller og resultater målt med aggregerte data. Et eksempel på denne typen avvik kan hentes fra en annen studie Fafo har gjort i samarbeid med SSB og Frisch-senteret (Kirkebøen mfl. 2017). Da var temaet evaluering av lærertetthet. Mens elever og lærere ga uttrykk for at økt lærertetthet var positivt for trivsel og skoleprestasjoner, fant de statistiske analysene ikke noen sammenheng.

Vi har dermed noe grunnlag for å konkludere, selv om det er svakere enn det ville vært om effekten var signifikant på 5 prosent nivå.

For øvrige elever finner vi effektestimater som er svært nær null for gjennomsnittlige resultater på NP9. Vi finner litt økte andeler på lave mestringsnivåer, men forskjellen er mindre enn forskjellen i andel elever som gjennomfører prøven. I sum mener vi ikke det er tegn til effekter på øvrige elever, verken positive eller negative.

Basert på effektestimaterne, observert sammenheng mellom NP9 og fullføring av VGO samt økonomisk verdsetting av fullføring av VGO fra tidligere forskning, gjør vi anslag på økonomisk nytte av tiltaket. Anslagene tilsier en økning i fullføring av VGO på omtrent 0,2-0,5 prosentpoeng. Med 1000 målgruppeelever i tiltaksskolene (fordelt på de siste to årene) svarer det til to til fem ekstra elever som fullfører VGO, som vi verdsetter til 2,5-6,25 millioner. Dette er mindre enn kostnaden av tiltaket, og trekker i retning av at tiltaket ikke er samfunnsøkonomisk lønnsomt. Disse anslagene er derimot usikre. Med usikkerheten som er i estimatene kan vi ikke utelukke effekter som gir samfunnsøkonomisk lønnsomhet, slik vi heller ikke kan utelukke ingen effekt.

### **Resultater effektevaluering videregående**

I videregående finner vi ingen effekt på elevenes ferdigheter målt ved Oslo-prøven. Derimot finner vi tegn til økt fullføring av Vg1 for de to første årskullene, som har data for fullføring av Vg1. Fullføring av Vg1 øker med omtrent 3-5 prosentpoeng, mest i målgruppen. Disse forskjellene er signifikant på 10 prosent nivå, men ikke på 5 prosent nivå. Vi mener det er for tidlig å konkludere klart med om økt fullføring er en effekt av tiltaket.

Vi har også vurdert økonomisk verdsetting av eventuelle effekter i Vg1, tilsvarende hva vi gjorde på ungdomstrinnet. 3-5 prosentpoeng økt fullføring av Vg1 svarer til omtrent 1 prosentpoeng økt fullføring av VGO, som vi verdsetter til 12 500 kr per elev som deltar i tiltaket. Dette er mer enn kostnaden av tiltaket per elev. Dersom tiltaket i Vg1 faktisk øker fullføring av Vg1 med 1 prosentpoeng vil det være samfunnsøkonomisk lønnsomt.

### **Resultater implementeringsstudie**

Den kvalitative implementeringsstudien viser en endring i synet på tiltaket underveis. I grunnskolen var lærerne i utgangspunktet positive til innholdet i tiltaket, selv om skolene og lærerne ikke selv hadde valgt å delta i tiltaket. Praktiske problemer knyttet til gjennomføringen skapte imidlertid utfordringer og frustrasjon høsten første tiltaksår. Det var også uklart for skoleledere hvordan forskningen la føringer for gjennomføringen av tiltaket på skolene (bl.a. utvelgelse av elever, tiltaksperioden og innhold), til dels ble føringene oppfattet som rigide, og en del respondenter i intervjuer og spørreundersøkelser uttrykte motstand mot tiltaket.

I lys av disse tilbakemeldingene gjorde prosjektet endringer i kommunikasjonen med lærerne. Gjennomføringen av prosjektet ble fortsatt styrt av Oslo kommune langs vanlige styringslinjer. Forskerne bidro imidlertid noe mer med å forklare bakgrunnen for prosjektet og begrunne de forskjellige føringene, først og fremst gjennom korte innledninger på samlinger for skoleledere og tiltakslærere. Samtidig ble de praktiske utfordringene redusert, ettersom materiell ble klart og de forskjellige aktørene fikk erfaring med relevante utfordringer, f.eks. skolens behov for timeplanlegging. I tråd med dette økte oppslutningen blant lærerne fra høsten 2016 til høsten 2017.

Mot slutten av prosjektet finner vi tegn til en normalisering: Tiltaket oppfattes i mindre grad som en spesiell satsning, men har i stedet, i større eller mindre grad, blitt integrert som en del av den ordinære undervisningen. Skolenes fokus på utviklingsarbeidet er flyttet videre, til nye og kommende prosjekter.

I videregående var det generelt sett mindre entusiasme for prosjektet, det var også flere som var skeptiske til didaktikken.

Mobilitet blant lærerne framstår som en utfordring for gjennomføringen av prosjektet. Det var krevende å finne tilstrekkelig lærere med riktig kvalifikasjoner, og det var krevende å bygge opp kompetanse i lærerkollegiet over tid.

Rapporten er bygd opp på følgende måte: kapittel 2 gir en grundig beskrivelse av tiltaket. I kapittel 3 redegjør Utdanningsetaten i Oslo kommune for den praktiske utformingen og gjennomføringen av tiltaket. Kapittel 4 forklarer effektevaluering mens kapittel 5 og 6 presenterer resultatene fra den kvantitative effektevalueringen for henholdsvis grunnskolen og videregående opplæring. Resultatene fra den kvalitative studien gis i kapittel 7. I kapittel 8 ser vi resultatene fra de forskjellige analysene i denne rapporten og i underveisnotatet (Kirkebøen m.fl., 2018) i sammenheng.

## 2. Beskrivelse av tiltaket

I dette kapittelet beskrives tiltaket og de erfaringer og forskningsresultater som tiltaket bygger på. Det gis også enkelte eksempler fra lærerkursingen og gjennomført undervisning.

Som nevnt innledningsvis viser forskning at svake matematikkferdigheter er negativt korrelert med senere skoleprestasjoner, som for eksempel fullføring av VGO.<sup>3</sup> Målgruppen for tiltaket er derfor elever som står i fare for ikke å fullføre VGO, altså de svakest presterende elevene i matematikk fra henholdsvis 8. trinn og 1P (praktisk matematikk) i videregående skole.<sup>4</sup> Mer konkret er målgruppen for tiltaket elever som presterer under en klart definert grense (mestringsnivå $\leq$ 2) på nasjonal prøve i regning på 8. trinn og 1P-elever som presterer under en tilsvarende grense (mestringsnivå $\leq$ 2) på læringsstøttene prøve i Vg1.

Generelt, det som kjennetegner disse elevene er at de ikke har tilegnet seg solid tallforståelse på småtrinnet, har svake ferdigheter innenfor de fire regningsartene, har ineffektive regnestrategier som gjør matematikk vanskelig og de har liten læreeffekt av ordinær undervisning som omhandler desimaltall og prosent (Borge m.fl., 2014). De har heller ikke tilegnet seg solid forståelse av brøk og algebra, herunder ligninger. Både Berge m.fl. (2014) og Borge m.fl. (2014) trekker frem overgangene mellom de ulike trinnene i utdanningsløpet som utfordrende. Uten kunnskap og ferdigheter fra forrige skoleår har disse elevene svært dårlige forutsetninger for å mestre matematikken videre, og utfordringene forplanter seg i det videre skoleløpet. Videre trekkes det frem at selv om læreplanen i praktisk matematikk i videregående skole i stor grad er en repetisjon av læreplanmål for ungdomstrinnet, settes det ikke av tid til å reparere grunnleggende faglige mangler.

Tiltaket i dette prosjektet retter seg mot grunnleggende matematikkferdigheter og består av to hovedelementer:

- Å gi **lærere** et sett med matematikdidaktiske prinsipper og verktøy rettet mot å undervise elever som presterer svakt i matematikk, og som står i fare for å ikke fullføre VGO.
- Å gi **elever** som presterer svakt i matematikk tilrettelagt (og strukturert) undervisning i to avgrensede perioder (4-6 uker vår og høst), med den hensikt at de skal ha økt utbytte av ordinær undervisning etter dette.

Den tilrettelagte undervisningen har blitt gjennomført i grupper sammensatt for formålet og bestått utelukkende av elever som presterer svakt i matematikk. På 8. trinn har disse gruppene hatt mindre spredning i nivået på elevenes matematikkferdigheter enn i den ordinære undervisningen. På Vg1 har derimot undervisningen vært i grupper av vanlig størrelse. Det vil si at hele klasser har blitt undervist med matematikdidaktiske prinsipper og verktøy som lærerne har blitt etterutdannet i. Det er flere grunner til dette. I VGO har elevene allerede fordelt seg til studieretninger og fag, slik at elevmassen i et gitt fag er mer ensartet. Det er større skoler,

<sup>3</sup> Av de 20 prosent svakest presenterende elevene ved nasjonal prøve i regning ved 8. trinn er omtrent halvparten blant de 20 prosent svakest presterende ved fullført grunnskole.

<sup>4</sup> To forhold er verdt å kommentere: Først, på yrkesfaglige utdanningsprogram har Oslo iverksatt mange tiltak for å bedre elevenes matematikkferdigheter og gjennomføring. Vi har derfor ikke rettet noen nye tiltak mot denne målgruppen. For det andre, mens læringsstøttene prøve på Vg1 måler matematikkferdighet måler nasjonal prøve på 8. trinn ferdigheter i regning. Denne distinksjonen har imidlertid mindre betydning når det gjelder å velge ut elever. Rapporter etter de nasjonale prøvene peker på at elever som strever med grunnleggende tallforståelse og matematisk kompetanse presterer dårlig på de nasjonale prøvene i regning (Ravlo og Johansen, 2013). Dette er i tråd med den sterke sammenhengen mellom nasjonal prøve i regning og resultat i matematikk ved fullført grunnskole, og tilsier at vi ville valgt omtrent de samme elevene om vi skulle valgt elever ut fra matematikkferdigheter i 8. trinn.

som har bedre muligheter til å justere gruppesammensetningen utfra elevenes forutsetninger, samt at det er færre skoler, noe som gjør det urealistisk å gjøre en presis sammenligning av flere forskjellige tiltak.<sup>5</sup>

## 2.1. Tiltakets faglig/didaktiske innhold

Utdanningssetaten (UDE) i Oslo kommune har vært ansvarlig for å utforme det faglige innholdet i tiltaket. Innholdet har blitt utformet i nært samarbeid med og under faglig rettleiding av Danmarks institutt for Pædagogik og Uddannelse (DPU). DPU ble involvert i arbeidet etter det overordnede designet av tiltaket var etablert. DPU fant prosjektets overordnede problemstilling om frafall og svak matematikkompetanse relevant ut i fra et utdanningspolitisk- og matematikk-didaktisk ståsted.

DPU har kjennskap til UDEs mangeårige arbeid med å styrke matematikkundervisningen, herunder Oslo-prøvene, grundige analyser av elevsvar på prøve-spørsmål samt UDEs prioritering av intervensjoner som foregår gjentagende og over lengre tid. DPUs bidrag til prosjektet er fundert på en bred viten om internasjonale undersøkelser og innsatts rettet mot elever med svak matematikkompetanse, og på danske erfaringer med tiltak rettet mot frafall og svak matematikkompetanse. DPU har adaptert og utviklet denne kunnskapen og erfaringen til forhold i Oslo-konteksten i et reelt samarbeid med UDE.

Lillejordet m.fl. (2015) trekker fram betydningen av en grundig analyse av hva elevenes problemer består i før tiltak iverksettes. Nettopp derfor har UDE analysert konkrete mangler og misoppfatninger i Oslo-elevenes kunnskaper på hhv. 5-10. trinn (vha. analyseverktøyet Alle Teller) og ved fullført Vg1 (vha. Osloprøve i matematikk). Disse analysene og annen kjennskap til elevenes svakheter innenfor faget, har dannet grunnlaget for utformingen av tiltaket.

Analysene viser at en betydelig andel av elevene har for dårlige kunnskaper innenfor grunnleggende tallforståelse; særlig i å forstå plassverdisystemet, brøk og desimaltall, men også prosent, bruk av måleenheter og målestokk. Innenfor disse emnene viser elevene manglende dybdelæring og forskjellige misoppfatninger fremkommer. Misoppfatningene har vist seg å oppstå tidlig i opplæringsløpet og er stabilt til stede gjennom klassetrinnene, selv om de er knyttet til læreplanmål for langt lavere trinn. Både læreplanmålene og de identifiserte manglene og misoppfatningene har dannet utgangspunktet for målrettingen av tiltaket. Tiltaket er plassert i kategorien «recovery»: det fanger opp elever i overgangen til ungdomsskole og VGO, og gjennom en spesielt tilrettelagt innsats gir et grunnlag for bedre læreeffekt av videre undervisning.

Det er et mål at elevene gjennom tiltaket skal ha oppnådd en bedre tallforståelse som grunnlag for å beherske matematikken videre. Valenta (2015) viser at utvikling av tallforståelse i mange studier framheves som svært viktig for elevenes læring av matematikk. Definisjonen av tallforståelse som hun bygger på, beskriver matematisk kompetanse som sammensatt av fem komponenter; begrepsforståelse, beregning, anvendelse (strategisk tankegang), resonnering og engasjement. Tiltaket vårt bygger på denne modellen ved at lærerne bruker disse komponentene som læringsmål. Elevene arbeider kognitivt med oppgaver som fremmer resonnering og forståelse, og diskuterer ulike løsningsstrategier med hverandre. Med andre ord,

---

<sup>5</sup> Det begrensede antall videregående skoler gjør det vanskelig å teste ut undervisning både i små og store grupper. Ettersom de videregående skolene allerede har noen ressurser de prioriterer til matematikkundervisning, har vi ønsket en organiseringsform som er praktisk mulig å gjennomføre uten økte kostnader. Vi har derfor prioritert undervisning i vanlige grupper i VGO. Vi tror også at spesifiserte undervisningsopplegg oppleves som mer fremmed i VGO enn i grunnskolen.

tiltaket bidrar til at elevene utvikler dybdekunnskap og metakognitiv kompetanse, og oppgavene elevene arbeider med gir mulighet til dette.<sup>6</sup>

Det didaktiske innholdet i tiltaket bygger på eksisterende forskning og materiale. Resultatene fra de seneste årenes TIMSS-undersøkelser viser at høy andel elever i Singapore presterer på et avansert nivå i matematikk, samtidig som de har en meget lav andel elever på det laveste nivået. UDE sine erfaringer med å nyttiggjøre matematikdidaktiske verktøy utviklet i Singapore er anvendt i tiltaket.

### Fra Singapore til Oslo

Utdanningsdepartementet i Singapore tok en rekke initiativ som følge av deres elevers relativt svake prestasjoner i internasjonale sammenlikninger i 1980-årene. Utdanningsdepartementet besluttet å undersøke teorien fra Jerome Bruner, Richard Skemp, Jean Piaget, Lev Vygotsky og Zoltan Dienes samt forskningsresultater om «best practices» i vestlige land med henblikk på å kunne sammenfatte dette som en «amalgamation of global ideas» og anvende denne fusjonen i matematikkundervisning i Singapore under overskriften «Singapore-matematikk» (Kaur et al. 2015). Utdanningsdepartementet anser resultater fra etterfølgende internasjonale sammenlikninger av elevers prestasjoner som tegn på at «Singapore-matematikk» har en positiv effekt i Singapore.<sup>7</sup> Nå går kunnskapsutvekslingen den andre veien - Singapore-didaktikere formidler og underviser «Singapore-matematikk» i andre land:

*A global maths teaching revolution: From Northern Europe to Kathmandu, we have introduced Singapore maths teaching to schools across the globe.*

Selv om det er en del kulturelle forskjeller mellom Norge og Singapore, blant annet statusen til faget matematikk, har Model-metoden<sup>8</sup>, som benyttes for de laveste klassetrinnene i Singapore, vist seg å fremme refleksjon, dybdeløring og god tallforståelse hos norske elever. I utformingen av tiltaket har vi derfor sett nærmere på Model-metoden og integrert denne som et verktøy i undervisningsmetoden i tiltaket, så vel ved lærerkursingen i planbasert opplegg. Videre er innholdet i tiltaket basert på læreplanmålene, og er dermed rettet mot grunnleggende matematikk.<sup>9</sup>

Ettersom matematikk er faget der flest elever i VGO stryker og dermed er en viktig bidragsyter til at elever ikke fullfører med alle fag bestått, har vi målrettet tiltaket mot ferdighetene elevene trenger for å bestå matematikk i VGO. I utformingen av det strukturerte tilrettelagte undervisningsopplegget (planbasert opplegg) til de små grupper på 8. trinn, er det dratt nytte av programmet og forskningen rundt Numbers Count.<sup>10</sup> Dette programmet retter seg mot barn i de laveste klassetrinnene. Her er det funnet betydelige effekter på matematikkferdigheter (Torgerson m.fl., 2011). For øvrig finnes det foreløpig lite forskningsmessig belegg for effekter av slike programmer på eldre elever. Som vist over har imidlertid svakt presterende elever

<sup>6</sup> Klasseromstudier knyttet til realfag viser at det benyttes mye tid både til helklasseundervisning og individuell oppgaveløsning. Det brukes mindre tid til å arbeide med kognitivt utfordrende og sammensatte problemstillinger (Bergem m.fl., 2014).

<sup>7</sup> I PISA 2018 i matematikk rangerer 15-åringer i Singapore med det nest beste gjennomsnittet av alle, 569. De nordiske landenes gjennomsnitt er for Danmark 509, for Finland 507, for Sverige 502, for Norge 501 og for Island 495. OECD-gjennomsnittet er 489.

<sup>8</sup> Model-metoden: <http://www.teach-kids-math-by-model-method.com/>

<sup>9</sup> I rammeverket for grunnleggende ferdigheter er den overordnede definisjonen av regning det å bruke matematikk på ulike måter (Utdanningsdirektoratet, 2012). For å kunne anvende matematikk til å løse problemer i dagligliv og yrkesliv, forutsettes det at elevene har en basis av fakta, ferdigheter og begrepsforståelse innen tall og tallregning.

<sup>10</sup> Se også <https://everychildcounts.edgell.ac.uk/wp-content/uploads/2015/11/Extract-from-the-NC-Handbook.pdf>



på 8. trinn og Vg1 utfordringer knyttet til læreplanmål for mye lavere trinn. Dette tilsier at det også kan være relevant å tilby eldre elever en didaktisk tilpasset lave ferdigheter. UDE har god erfaring med forsøk med tilsvarende programmer på de laveste klassetrinnene,<sup>11</sup> og også med å bruke småskoledidaktikere til å kurse lærere som skal undervise svakt presterende elever på høyere trinn. Derfor kan blant annet Numbers Count's matematikdidaktiske prinsipper og verktøyer anvendes på mange trinn, forutsatt at den konkrete gjennomførelsen tilpasses elevenes alder, spesifikke forutsetninger og motivasjonsstruktur.

I det følgende listes de seks utvalgte prinsippene og de fire utvalgte verktøyene i tiltaket. For prinsipp 1, 2 og 5 gis en kortfattet skissering av deres forskningsmessige bakgrunn.

Seks prinsipper på lærerkursingen. Læreren oppfordres til å:

1. Skape sammenheng mellom økter
2. Bruke aktiviteter med lavt gulv og høyt tak
3. Igangsette MO(tivation) - (Per)FORMANCE
4. Samtale og igangsette samtaler, som støtter forståelse
5. Sette realistiske, høye forventninger
6. Bruke journal

Disse prinsippene er konkretisert i fire metoder og verktøy:

- "Thinking Blocks" – blokkmetoden – Singapore modellmetoden
- Læringspartner
- IGP: Individuell, Gruppe, Plenum
- Samtaletrekk

Bakgrunnen for (1) er å oppfordre lærere til å skape eller medvirke til at elevene erfarer at det er sammenheng mellom økter. Dette er basert på internasjonal sammenlignende klasseromforskning i matematikkundervisning. Denne forskningen ble initiert av vestlige utdanningspolitikere og matematikdidaktikere som etter de første «International Association for the Evaluation of Educational Achievement» (IEA)-undersøkelsene av matematikkprestasjoner ble overraskede over at elever i 4. klasse og 8. klasse i asiatiske land presterte bedre enn elever i vestlige land. De etterfølgende studiene avdekte at organiseringen av den enkelte læringsøkt i klasserom i Japan skilte seg fra klasserom i USA og Vest-Tyskland med hensyn til fordeling mellom lærer- og elev-taletid, øktenes faseoppdeling, fokusering på viktige matematiske idéer og operasjoner, antall oppgaver samt nettopp *sammenheng mellom økter*. Japanske klasserom utmerket seg ved at læreren tydeliggjorde for elevene hvordan denne øktens innhold og arbeidsmetoder relaterte til foregående økter og eventuelt også til fremtidige økter (Stigler, 1993) (Okazaki, Okamoto, Morozumi, 2019).

Dette tiltakets vektlegging av (1) som matematikdidaktisk prinsipp bygger på hypotesen om at sammenheng-skapende elementer kan støtte konsentrasjon, opplevelse av meningsfullhet og aktivisering av relevante forforståelser hos målgruppeelever.

Bakgrunnen for å oppfordre lærere til å bruke aktiviteter med lavt gulv og høyt tak (2) er basert på forskningsresultater om nivådelt matematikkundervisning. Boaler (2011) har offentliggjort resultater fra nivådelte videregående skoler som senere ble ikke-nivådelte. Resultatene viser at alle studenter hadde godt av ikke-nivådeling, enten de presterte på lavt, mellomhøyt eller høyt nivå. Analysene viser at grunnen til fordelene var at undervisningen til alle studenter omfattet arbeid med rike

<sup>11</sup> Bl.a. programmet 1st class@numbers, <https://everychildcounts.edgehill.ac.uk/mathematics/what-is-1stclassnumber>

oppgaver, formativ vurdering og høye forventninger til studentene. I tillegg viser forskning at nivådelt undervisning gir ulemper for studenter som er svakt presterende. Forgasz (2010) undersøkte nivådeling for matematikk i Australia. Hun fant fordel for studenter på et høyt nivå, men ulemper for studenter som presterte svakt. Dette sannsynligvis fordi den nivådelte undervisningen gjorde det mulig for høypresterende å engasjere seg i matematiske utfordringer, men kun tilbød "drill og praksis" – metoder og kun meget enkel matematikk for svakt presterende elever, noe som kan ha innsnevret disse elevenes muligheter for å lære.

Å anvende riktige aktiviteter med lavt gulv og høyt tak er et prinsipp i undervisningsdifferensiering i tiltaket for målgruppeelever på forskjellig nivå, hvor alle elever har tilstrekkelige forutsetninger for å gå i gang (lav inngangsterskel), og hvor elever kan fortsette aktiviteten i mer og mer komplekse varianter så langt som deres forutsetninger gir dem muligheter for (høyt tak). Å anvende rike aktiviteter med lavt gulv og høyt tak er også et prinsipp i tiltaket for gruppe med elever der alle er målgruppeelever, for også å gi disse elevene mulighet for å engasjere seg i ekte matematiske utfordringer ut i fra et trygt startpunkt som de erfarer at de kan mestre.

Bakgrunnen for å oppfordre lærerne til (5), det vil si å stille realistiske, høye forventninger til elevene, er dels basert på forskning om nivådelt og ikke-nivådelt matematikkundervisning (se ovenfor), og dels forskning om læreroppfattelser av elever med matematikkvansker. Scherer et al. (2017) diskuterer teorier om årsaker til matematikkvansker, hvor det ene ekstreme er at vanskelighetene betraktes som feil ved individet (neurologisk eller psykologisk) og det andre ekstreme er vanskeligheter som fiaskoer i utdanningssystemet (didaktisk) eller andre samfunnsmessige trekk (sosiologisk). De hevder at lærere som bygger på neurologiske og psykologiske teorier får et syn på elevene som om de er i underskudd, besværlige og uten mulighet for å lære, mens lærere med mer relasjonelle synspunkter på matematikkvanskeligheter bedre kan se at alle studenter kan ha potensiale og motivasjon til matematikklæring. Faragher, Brady, Clarke og Gervasoni (2008) og Faragher (2014) angir muligheter for alle til å lære matematikk ved å vise til elever med Downs syndrom.

*Observasjoner i tiltaks-undervisningen viser at der av og til stilles lave kognitive forventninger til de studenter, hvilket innsnevrer muligheten for å lære. Dette er gjennomgående i alle nivåer og grupper av design, men mest utpreget på Vgl.*

## 2.2. Etterutdanning av lærere og utforming av undervisningsopplegg

En viktig del av tiltaket har vært å overføre det faglige innholdet i tiltaket til lærerne. Med andre ord å gi lærerne målrettede didaktiske verktøy gjennom etterutdanning og gjøre dem i stand til å bruke disse i undervisningen på en slik måte at elevene hever sine grunnleggende kunnskaper i matematikkfaget.<sup>12</sup> Innholdet i etterutdanningen er basert på problembeskrivelsen, teorien og didaktikken beskrevet over, og har til hensikt å gi lærerne både en forståelse av den teoretiske forankringen og en opplæring i klasseromsnær praksis.

Det er på ingen måte trivielt å forsøke å gi lærere målrettede didaktiske verktøy gjennom etterutdanning og gjøre dem i stand til å bruke disse. Det er vanskelig for mange barn, unge og voksne å lære matematikk, men forskning peker på at det er

<sup>12</sup> Det er gitt målrettet etterutdanning som ikke gir studiepoeng, altså ikke studiepoenggivende videreutdanning. Vi tar utgangspunkt i at tiltakslærere minimum oppfylle lovkravet til undervisningskompetanse. I Oslo kommune har en stor del av lærerstaben allerede minimum 60 studiepoeng når de underviser i matematikk på hhv ungdomstrinnene og ved videregående skole.

enda vanskeligere å undervise elever med matematikkvansker. Forskingen peker også på at læring av nye og komplekse instruksjonspraksiser i kursing og profesjonell utvikling er en ”rocky path” (Heyd-Metzuyanım et al., 2019).

Klasseromobservasjoner av intensivopplæringen i Ny Giv (Rønning, Hodgson, Tomlinson, 2013) viser en fallende tiltro etter hvert som lærerne får erfaring med intensivopplæringen, hvor lærerne mindre bruker ‘Ny Giv metoder’ og vender tilbake til metoder som de brukte før NY Giv prosjektet. Bruken av spill i regneundervisningen falt fra 37% i øktene til 22 %, 8% og til 0% gjennom prosjektets faser. Hva Rønning et al. betegner som «mer tradisjonell oppgaveløsning» økte fra 6% til 17%. Derimot, bruken av verktøy i matematikk økte svakt med erfaring.

Det er opplagt å spørre hvorfor lærernes profesjonelle utvikling og etterutdanning er en «rocky path» og videre hvorfor den påvirker deres undervisning så lite, som det er typisk å se i de prosjekter som undersøker dedikasjon. Likeledes er det opplagt å spørre, hvorfor tiltro og dedikasjon kan falle gjennom et prosjekts faser: Finnes generelle svar og matematikdidaktiske svar?

Generelle svar kan man få fra Michael Fullan (2001, s. 12), som viser at reformer i undervisning er komplekse og tar tid fra initiering til institusjonalisering. Det kan kreve 3- 8 år med intensiv innsats å skape forbedringer, og forbedringer er i tillegg sårbare: kanskje forlater nøkkelpersoner skolen, kanskje er skolens infrastruktur svak og skolekulturen negativ og kanskje er støtten fra kommunens skoleforvaltning svak. Hargreaves og Fullan (2012) påpeker at lærernes profesjonalisme (profesjonelle kapital) er et kjerneelement i reformer. I matematikdidaktisk forskning kan man finne to svar, som utdypet av Hargreaves og Fullan.

Den ene forklaringen er at intuitiv matematisk forståelse har stor innflytelse på lærerpraksis i all slags undervisning, og kanskje gjelder dette især i undervisning av elever med matematikkvansker. Lærernes pedagogiske kunnskap («pedagogical content knowledge») som de tilegnes i lærerutdanningen inneholder kunnskap om elevers problemløsningsstrategier i aritmetikk og om hva som er vanskelig, men denne kunnskapen kan bli overskygget av lærerens intuitive matematiske forståelse (Gvozdic, Sander, 2018). Det kan også være vanskelig for voksne som mestrer aritmetisk problemløsning – selv om de er lærerutdannede – å forstå hva elever opplever som vanskelig, og hvordan de opplever det. Elementær aritmetikk kan synes innlysende og banal for voksne som selv mestrer det.

Den andre forklaringen fra matematikdidaktisk forskning omhandler affektive aspekter av matematikkundervisningen. Som påpekt av Kazemi og Wæge (2015) er undervisningsarbeid verdiladet. Undervisning krever både intellektuelle og følelsesmessige investeringer og derfor må lærerkursingen i matematikk forsøke å kultivere holdninger og identitet hos lærerne, og lærernes reaksjon på å forsøke å kultivere holdninger og identitet har avgjørende betydning for hva lærerne lærer på kurs. Lærerutdannere skal derfor ifølge Kazemi og Wæge være oppmerksomme på både hvordan man respektfullt kan forsøke å kultivere lærernes identitet og hvordan man samtidig arbeider med lærernes tekniske dyktighet og kunnskap om undervisning. Det er ikke tilstrekkelig at deltagerne presenteres for didaktiske prinsipper og verktøyer, for *deltagerne må også innse hvorfor prinsipper og verktøyer er nyttige*. Likeledes har lærernes kunnskap om undervisningsinnhold og *kunnskap om elever* betydning for lærernes produktive utførelse av prinsipper og verktøyer.

I tillegg til Kazemi og Wæges betraktninger om læreridentitet kan en kulturell og eventuelt en nasjonal dimensjon tilføyes læreridentiteter. Det kan tenkes at norske matematikklæreridentiteter har en nasjonal forankring. Dette er ikke undersøkt, men

man kan derimot anta en nasjonal forankring i dansk matematikundervisningskultur blant danske matematikklærere med skepsis over å innta undervisningsprinsipper fra andre land, især fra ikke-demokratiske land. I tiltaket legges det i lærerkursingen vekt på å presentere begrunnelse og hypoteser for prinsippers og verktøyers effektivitet i forhold til målgruppen, som anbefalt av Kazemi og Wæge. En del av begrunnelsene stammer fra internasjonal forskning og er ikke utviklet ut fra norsk kontekst. Prinsippet om sammenheng mellom økter (med japansk inspirasjon) og prinsippet om at å igangsette MO-FORMANCE (med dansk inspirasjon) er lite utprøvd i Norge. Noen lærere har direkte uttrykt skepsis overfor internasjonale referanser, og denne skepsisen er kanskje blitt forsterket av at DPU og ikke et norsk universitet er partner i tiltaket. DPU antar at en del lærere ønsker at tiltaket hadde hatt et mer deltagerstyrt og mer «norsk-født»-design enn tilfellet var. Kursingen har derfor inneholdt en aktivitet hvor den enkelte lærer tydeliggjør sine egne erfaringer med målgruppeelever og med undervisning av målgruppeelever. Dette ble organisert som individuelt spørreskjema, etterfulgt av DPUs analyse og presentasjon av svar i plenum-refleksjon etter kursingen. Disse sammenfattede lærererfaringene har DPU deretter relatert til utvalgte prinsipper og verktøyer.

Lærerfaringer – fra enkelte lærere og sammenfattede – er løpende tatt opp i kursingen og foredragsholdere har bidratt med egne erfaringer og forskningsresultater om elever med matematikkvansker og om undervisning av disse elevene. I mange tilfeller har foredragsholdere formidlet egne erfaringer og kunnskap i form av elev-caser. Bland annet har DPU bidratt med elev-caser som kunne ha vært målgruppeelever og som har vist potensiale for matematisk tenkning som overrasker deres lærere positivt. Synliggjørelse av potensialer er betinget på lærerens interesse for elevene og evne til at inngå i matematiske samtaler med elevene.

Tiltakslærere har fått en etterutdanning bestående av begrunnelser for og øvelse i å ta i bruk grunnleggende matematikdidaktiske faglige prinsipper og verktøyer som beskrevet ovenfor. Formålet har vært å tilby et sett av «grep» i form av prinsipper og verktøy samt å gi en tydelig forankring av undervisningsmetode. Følgende omtaler vi dette som «veiledet» opplegg V.

I tillegg har lærere på enkelte skoler fått kursing i å gjennomføre detaljerte undervisningsøkter. Disse er nøye beskrevet etter en detaljert standard plan (heretter «planbasert» opplegg P). Dette er ment som en ytterligere støtte for læreren, og som en hjelp til å endre praksis i klasserommet på ønsket måte.

Vi kan gi følgende eksempler som illustrasjon på hvordan tilbud om et «grep» (her prinsippet om rike oppgaver med lavt gulv og høyt loft) og tydelig forankring av undervisningsmetode er foregått i kursingen:

### **Veiledning Vg1 lærerkurs**

- Lærerne presenteres for og utarbeider nye utgaver basert på tre oppgaver fra Osloprøven Vg1 2015. Lærerne diskuterer også følgende spørsmål: Hvordan kan man gjøre tradisjonelle oppgaver rike med lavt gulv og høyt tak?
- Overvei følgende:
- Hvorfor undervise med rike oppgaver?
- Hvordan endres tradisjonelle oppgaver til rike oppgaver?
- Hva kreves det av deg å motivere elevene til rike oppgaver?
- Hva kreves det av deg at du selv omskriver tradisjonelle oppgaver til rike oppgaver?

Vi kan gi følgende eksempler som illustrasjon på hvordan de planbaserte undervisningsøktene er beskrevet i det skriftlige lærermateriale i lærerkursingen for lærerne i 8. klasse, små grupper:

#### **Planbasert materiale – avsnitt algebra**

- 1 side Begrunnelse og hva læreren skal se etter.
- 3 sider *Oppstart* med aktivitet og fire oppgaver med eksempler på løsning og læreroppsummering.
- 2 sider *Algebra ved blokkmetoden* med to oppgaver og med lærerspørsmål og mulige elevvanskeligheter.
- 1 side *Aktivitet* med hint til hvordan læreren kan guide elever gjennom de algebraiske uttrykkene
- 1 side *Utfordringer og avslutning* hvor også forståelse av plassverdisystemet
- 2 sider *Ekstraoppgaver* med seks oppgaver

Det er helt sentralt for utfallet av tiltaket at lærerne faktisk tar i bruk innholdet i etterutdanningspakken i undervisningen sin. Erfaring fra en rekke innsatser (Ny GIV, 1st class@numbers samt lignende innsatser i lesing og språk) tyder på at resultatene i stor grad avhenger av om læreren følger anbefalte eller pålagte framgangsmåter i undervisningen. Det kan være mange årsaker til at lærere velger å avvike fra et anbefalt opplegg, for eksempel lærerens kompetanse og elevenes forutsetninger. Ved å prøve ut både et detaljert «planbasert» opplegg (P) (små grupper 8.klasse) og et "veiledet" opplegg (V) (stor gruppe 8 klasse og VG1), kan vi undersøke om systematiske forskjeller i implementering og effekter avhenger av hvor detaljert undervisningsøktene blir beskrevet.

Kursene har blitt gjennomført av UDE ved Avdeling for pedagogisk utvikling og kvalitet, som har kompetanse internt på å legge til rette for gode prosesser på lærersamlinger og lang erfaring med å tilby kompetanseheving og kurs til skoler. DPU har bistått i utarbeidelsen av kursene, og i gjennomførelsen av kursdagene.

Etterutdanningen har blitt strukturert i tre hoveddeler:

1. To generelle samlinger for både lærere og skoleledelse før oppstart av strukturert undervisning, der bakgrunn, målgruppe og problembeskrivelse har blitt presentert.
2. Et antall samlinger om høsten med separat etterutdanning for de planbaserte smågruppene i 8., store grupper i 8. samt Vg1. Disse har blitt lagt i forkant av og parallelt med første periode av den strukturerte undervisningen. Innholdet har vært nært knyttet til undervisningen lærerne gjennomfører for å bidra til endring i faktisk undervisningspraksis. Samlingene har også lagt opp til deling av erfaringer fra undervisningen.
3. Et mindre antall samlinger om våren – separat som om høsten - ifm. den strukturerte undervisningen. Her har det også vært fokus på praksisnær undervisning og deling av erfaringer fra undervisningen.

Flere lærere enn de som var ment å gjennomføre tiltakene har deltatt på kursene. Det er to grunner til dette: 1) Unngå at uforutsette hendelser hindrer gjennomføring, f.eks. lærere trer inn som vikarer om deltakende lærer ble syk. 2) For at kompetansen kan spres til skolens øvrige matematikklærere og bidra til å utvikle den ordinære undervisningen har deltakelse også kommet fra flere lærere ved hver skole, også fra skoleledelsen. Etterutdanningen har på denne måten karakter av å være et skoleutviklingsprosjekt.

### 2.3. Organisering av undervisningen

- De strukturerte undervisningsløpene for tiltakselevne har blitt gjennomført i to perioder både på 8. trinn og Vg1. En periode på høsten og en oppfriskingsperiode i april i skoleårene 2016-17, 2017-18 og 2018-19. Hver periode varte i 4-6 uker. I denne perioden ble elevene undervist i egne grupper og utfra de didaktiske prinsippene rundt grunnleggende matematikkferdigheter forklart over, for å avlære misoppfatninger og for å bli i stand til å ha utbytte av ordinær undervisning. Merk at undervisningstimene kom i stedet for ordinær undervisning i perioden, og ble lagt parallelt med den ordinære matematikkundervisningen.
- På 8. trinn har som allerede nevnt skoler gjennomført undervisningen i små og relativt homogene grupper (4-8 elever). I VGO har all undervisning vært i grupper av vanlig størrelse.
- I tillegg til etterutdanning av tiltakslærere, har skoleledelsen og øvrige matematikklærere ved skolene blitt godt informert om tiltaket. Dette for å øke kvaliteten på implementeringen. Mer om dette og implementeringen av tiltaket i neste kapittel.
- Vi understreker at lærere ved kontrollskoler kan ha kjennskap til metodene brukt i den strukturerte undervisningen, og de kan derfor ha anvendt disse metodene aktivt, men de har ikke mottatt etterutdanningen knyttet til tiltaket (forutsatt at de ikke har byttet skole underveis).
- Tiltaket er basert på et intensivkurs i matematikk. Det kan stille spørsmålsteget ved om eksponeringen av tiltaket varte lenge nok for elevenes del. Flere artikler derimot ser på virkninger av intensivkurs i matematikk og finner positiv effekt på elevprestasjoner (Rose and Betts, 2004; Joensen and Nielsen, 2009; Aughinbaugh, 2012; Goodman, 2019). Dog, ikke alle disse er rettet mot svakt presterende elever.

### 3. Organisering av tiltaket

Utdanningssetaten i Oslo har vært tiltakseier og hovedansvarlig for tiltakets:

- faglige innhold
- forankring
- implementering og oppfølging

I denne delen av rapporten redegjør UDE for prosjektet på de områdene de har vært hovedansvarlig for.

#### 3.1. Faglig innhold

##### Utvikling av undervisningsopplegg

Utdanningssetaten påbegynte arbeidet med å utforme innholdet i tiltaket med en gang meddelelsen om resultatet av konkurransen forelå fra Kunnskapsdepartementet. Det skulle utformes egne øktplaner for undervisningen av små grupper på 8. trinn samt prinsipper/matematikkdidaktiske verktøy for undervisningen av tiltakselever i store grupper på ungdomstrinnet og i videregående.

Som angitt i forrige kapittel ble innholdet utformet i nært samarbeid med og under faglig rettleiding av Danmarks institutt for Pædagogik og Uddannelse (DPU).

Det var avsatt relativt kort tid til utforming av det faglige innholdet i tiltaket, og det var ikke lagt opp til muligheten for å pilotere undervisningsplanene. Med tilbakemelding fra de ulike forskningsmiljøene og fra skoler/lærerne, ble undervisningsoppleggene revidert og forbedret fra 1. til 2. prosjektår. Etter 2. prosjektår var opplegget ferdig utviklet, og ble bare utført små endringer før siste prosjektår.

##### Etterutdanning

Med utgangspunkt i *Skisse for etterutdanning av lærere* ble plan for etterutdanningen/kursrekken for lærere utformet.

Planen har hatt følgende hovedtema:

- Motivasjon og mestring – relasjoner og elevsamtalen
- Misoppfatninger og kartlegging – typiske misoppfatninger i matematikk basert på lang erfaring med kartlegging i Osloskolen
- Dialogbasert undervisning – rike oppgaver og berikede oppgaver, undervisningsprinsipper med bruk av blant annet læringspartner, individuell-gruppe-plenum (IGP) og bruk av samtaletrekk
- Metoder og strategier – illustrasjoner for visualisering av elevenes tanker som f.eks. blokkmetoden, arealmetoden og tom tallinje.

Kursene ble delt i tre "løp", for henholdsvis videregående skole (undervisningsprinsipper), små grupper 8.trinn (øktplan) og store grupper 8.trinn (undervisningsprinsipper).

Hver samling blir evaluert, bl.a. på bakgrunn av deltakernes strukturerte tilbakemelding om innhold og nytteverdi av samlingen for egen undervisning. Basert på tilbakemeldingene, UDEs vurderinger og tilbakemeldinger gitt av forskningsmiljøene, er kursene kontinuerlig forbedret. De største endringene som ble gjennomført i løpet av første prosjektår var økt grad av workshop på samlingene for å gi lærere tid til å trene på bruk av prinsippene.

Det ble også i 2. prosjektår gjort noen endringer i etterutdanningsplanen. Opprinnelig var det lagt opp til at halvparten av kursene skulle gjennomføres før

tiltaksperioden og halvparten underveis i tiltaksperioden. Basert på tilbakemeldingene har det vist seg mer hensiktsmessig å legge så mange av kursene i forkant av tiltaksperiodene som mulig. Utdanningsetaten vurderte kontinuerlig kursenes innhold, form og tidspunkt for å tilpasse dem målgruppen, og for å sette lærerne bedre i stand til å gjennomføre en tilpasset undervisning som møter elevenes behov.

Etter 2. prosjektår så man at det fortsatt var behov for endringer. Dette viste også underveistrapporten "*Foreløpig beskrivelse av utforming og gjennomføring av tiltak*" av SSB notat 2018/15 (Kirkebøen et al.). Her fremgikk det at det var stor variasjon i hvilken grad lærerne brukte prinsippene og verktøyene i undervisningen, og det var forskjell mellom implementeringsgraden til de seks prinsippene som lå til grunn for prosjektet. Videre fremgikk det at lærernes motivasjon for å delta i prosjektet hadde økt gjennom tiltaksperioden, og at flere lærere ønsket å kunne bruke de anviste prinsipper, metoder og verktøy, men ikke synes de hadde forutsetninger og tilstrekkelig trening/øving til å kunne gjennomføre dem i praksis i klasserommet.

Styringsgruppen vedtok derfor å endre kursrekken slik at de to siste kursene før 2. tiltaksperiode ble ett heldags planleggingsseminar med veiledere fra Nasjonalt senter for Matematikk i opplæringen (Matematikksenteret).

I første del av seminaret ble planleggingens betydning gjennomgått, og selve planleggingen av økten gjennomført. I siste del gjennomførte en gruppe lærerne i samarbeid en undervisningstime med kyndig veiledning fra Matematikksenteret i et klasserom på en av tiltaksskolene. I etterkant var det tilbakemeldinger og diskusjoner rundt læringsprosessen til både elever og lærere.

Endringen innebar en forsterkning av et element som allerede var i kursrekken, men nå med veiledning og utprøving med elever i klasserommet. Lærerne har tidligere også fått i oppdrag å planlegge og gjennomføre undervisningsopplegg, men erfaringene har vist at dette ikke alltid har fungert etter intensjonen. Dette kunne ha vært løst ved tettere oppfølging av skoleledelsen, men da hadde lærerne kanskje ikke fått den matematikkdiraktiske veiledningen de etterspurte.

### 3.2. Forankring

I tillegg til tidlig forankring av prosjektet hos etatsledelsen og politisk ledelse, har forankringsarbeidet i øvrige ledd i organisasjonen vært sentralt.

#### **Forankring i Utdanningsetatens medbestemmelsesutvalg (MBU)**

Prosjektet ble behandlet i etatens MBU-møte 21.08.2016. Planene og organiseringen ble gjennomgått, og følgende begrunnelse for prosjektet ble gitt:

- Tiltaket adresserer to aktuelle problemer, frafall og svak matematikkompetanse, basert på et grundig vitenskapelig og erfaringsbasert grunnlag både hva gjelder forståelse av problemene og tiltakets potensial.
- Tiltaket formuleres eksplisitt (bl.a. i form av konkrete etterutdannings- og undervisningsopplegg, samt beskrivelse av undervisningen) og gjøres dermed reproduserbart, slik at det evt. vil kunne implementeres av andre skoleiere.
- Prosjektet legger opp til en solid effektevaluering inkl. en kost-nytte analyse, dette gir et godt grunnlag for å vurdere effekt og nytte ved evt. implementering av andre skoleiere samt for å motivere implementering dersom tiltaket viser seg å være effektivt.
- Observasjon av undervisningen og den kvalitative implementeringsstudien gir grunnlag for en bedre forståelse av effekter og for å identifisere kritiske



punkter og forbedringspunkter, dette legger til rette både for vellykket kopiering og for videreutvikling.

- Planlagte studier av forskjellige tiltaksvarianter (forskjellige organisering av strukturert undervisning og etterutdanning) og av heterogene effekter (forsøk både på 8. trinn og i Vg1, samt mulighet til å studere effekter på bestemte elevgrupper og skoler) gir grunnlag for forbedring/videreutvikling av prosjektet, samt for målretting mot spesielle grupper.

Medbestemmelsesutvalget tok saken til orientering.

### **Forankring gjennom referansegruppa**

Utdanningsetaten (UDE) opprettet en referansegruppe som har hatt 4-5 møter i året. Representantene i referansegruppen bestod av følgende:

- Prosjektgruppa (Assisterende direktør i Utdanningsetaten, assisterende direktør i avd. for pedagogisk utvikling og kvalitet (PED) og Prosjektleder UDE)
- To områdedirektører (OMD) fra grunnskolen
- Områdedirektør videregående skole
- To rektorer fra grunnskolen
- En rektor fra videregående
- Tre representanter fra fagforeningene (Utdanningsforbundet (UDF) vgs, UDF grunnskole og Skolelederforbundet)

Referansegruppens oppgaver var følgende:

- støtte prosjektet med råd, med bakgrunn i eget "ståsted"
- være diskusjonspartner og gi innspill på eget initiativ eller etter forespørsel
- bistå med forankring i hhv. administrasjonen og skolene

Samarbeidet i referansegruppen har fungert godt og deltakerne har meldt tilbake at de opplevde å bli hørt og at deres forslag til endringer har blitt tatt til etterretning.

### **Forankring i styringslinjen**

Osloskolene er delt inn i syv hovedområder med tilhørende områdedirektører. Områdedirektørene har ansvar for å følge opp skoler i eget område. Områdene A-F består av grunnskoler, og område G består av videregående skoler. Antall skoler i hver gruppe varierer fra 20 til 27 skoler. Områdedirektørene har ukentlig møter (OMD-møte) hvor hele toppledelsen i Utdanningsetaten deltar. Prosjektet har vært jevnlig tema i disse møtene.

For å sikre ansvarsfordelingen, ble det i et OMD-møte tidlig i prosjektet besluttet en rolle- og ansvarsmatrise, som ble kommunisert videre til involverte parter:

<b>Ansvarlig:</b>	<b>Ansvar og oppgaver:</b>	<b>Følges opp av:</b>
Prosjekteier i UDE:	<ul style="list-style-type: none"> <li>• Overordnet ansvar               <ul style="list-style-type: none"> <li>○ Leder styringsgruppa</li> <li>○ Forankrer prosjektet hos etatsledelsen</li> <li>○ Følger opp områdedirektører</li> <li>○ Setter PBG på agendaen på OMD-møter</li> </ul> </li> </ul>	
Prosjektleder:	<ul style="list-style-type: none"> <li>• Ledelse av prosjektet               <ul style="list-style-type: none"> <li>○ Koordinerer samarbeid med SSB, Fafo og DPU</li> <li>○ Sikrer at prosjektets leveranser har riktig kvalitet, leveres innen rett tid og i henhold til vedtatt budsjett</li> <li>○ Fatter beslutninger innenfor rammer gitt av styringsgruppa, herunder endringer som ikke påvirker prosjektets resultat</li> <li>○ Følger opp milepælsplan og rapporterer på eventuelle avvik</li> <li>○ Er styringsgruppens sekretær og rapporterer til styringsgruppa</li> <li>○ Koordinerer aktiviteter av tildelte ressursene på realfagsfeltet</li> <li>○ Sørger for at skolene får bestillinger og nødvendig informasjon knyttet til prosjektet</li> <li>○ Har overordnet ansvar for kursplan</li> </ul> </li> </ul>	Ledelsen i PED/Prosjekteier UDE
Ledelse i PED:	<ul style="list-style-type: none"> <li>• Lederoppfølging av prosjektleder</li> <li>• Overordnet faglig ansvar               <ul style="list-style-type: none"> <li>○ Setter inn ressurser på realfagsfeltet</li> </ul> </li> </ul>	
OMD:	<ul style="list-style-type: none"> <li>• Påser og følger opp regelmessig at den enkelte skole gjennomfører den besluttede organiseringen, at innhold og aktiviteter er i henhold til handlings-/tiltaksplan og er resultatansvarlig               <ul style="list-style-type: none"> <li>○ Månedlige samtaler med rektor når det gjelder PBGs fremdrift, mål og resultater</li> <li>○ Evaluering underveis med rektor og oppdatering av handlingsplanene i henhold til dette</li> <li>○ Sikrer at rektor sørger for nødvendig informasjon til personalet, elever og foresatte</li> <li>○ Sikrer eventuell videreføring etter avsluttet prosjekt</li> </ul> </li> </ul>	Prosjekteier UDE
Skoler:	<ul style="list-style-type: none"> <li>• Gjennomfører prosjektet i henhold til bestilling fra prosjektledelsen, skolens handlingsplan og styringsdialog</li> </ul>	OMD

### Forankring på skolenivå

Prosjektet ble forankret nedover i organisasjonen gjennom ulike informasjonsmøter med skoleledere i forkant av oppstarten, både på sentrale skoleledermøter og egne møter for ledere ved hhv ungdomsskoler og videregående skoler. På disse møtene deltok også de andre partene i forskningsprosjektet (SSB og Fafo). I tillegg mottok skolene skriftlig informasjon om prosjektet.

For å forenkle arbeidet videre til skolenes personale og elever/foresatte, ble det utarbeidet informasjonsmaterieell skolene kunne benytte:

- Skriv til foresatte
- PP-presentasjon til bruk på personalmøte og/eller foreldremøte

I forkant av 2. prosjektår ble informasjonsmateriellet til skolene revidert og distribuert på nytt, og det ble avholdt nye oppstartsmøter med alle tiltaksskolene. Dette ble også gjort i forkant av 3. prosjektår, men med kun få justeringer i materiellet.

### 3.3. Implementering

Utarbeidelsen av rolle- og ansvarsmatrisen (se under punkt 2.3) var et viktig ledd i å sikre implementeringen av tiltaket på skolene. I tillegg ble det utarbeidet egne sjekklister for rektor i oppfølging på skolen, som støttet rektor i å sikre en korrekt implementering av intervensjonen.

At lærerne gjennomførte tiltaket mot elevene slik det var ment å gjennomføres, var et sentralt punkt i implementeringsarbeidet. Det var derfor viktig at involverte lærerne hadde forståelse for hva de var en del av gjennom deltakelse i dette forskningsprosjektet og viktigheten av å bidra til å øke svakt presterende elevers ferdigheter i matematikk.

For å sikre en implementering i tråd med dette, gjennomførte UDE fortløpende møter med rektorer og områdedirektører. I tillegg har UDE også hatt møter med lærere og observasjon av undervisning.

#### Oppfølging

Det ble ansatt en prosjektleder for prosjektet samt en prosjektmedarbeider i UDE. Prosjektmedarbeider begynte 01.08.2017. I første prosjektår ble ressurser tilsvarende én stilling benyttet til å avsette tid hos fagkonsulenter og ledelsen i PED for å sikre utforming av og fremdrift i prosjektet.

Prosjektleder var sentral i oppfølgingen av prosjektet både internt i Utdanningsadministrasjonen, ut mot skolene og mot samarbeidspartnerne i prosjektet. Internt i Utdanningsadministrasjonen ble prosjektet fulgt tett av ledelsen og det ble gjennomført jevnlig statusmøter mellom prosjektleder, assisterende direktør i UDE og assisterende avdelingsdirektør i avdeling for pedagogisk utvikling og kvalitet (PED). Prosjektet var også et tema på møter med områdedirektørene minst fire ganger årlig.

I løpet av første prosjektår så vi behovet for at skolene utarbeidet lokale handlings/tiltaksplaner for å sikre oppfølgingen av prosjektet ved egen skole. Det ble utarbeidet en mal til skolene som støtte i dette arbeidet. Handlings- og tiltaksplanene på skolene har hatt følgende hovedpunkter:

- Elevenes utvikling
- Informasjon
- Implementering

Planen var sentral i oppfølgingsarbeidet og i oppfølgingsmøtene på skolene.

Prosjektleder og prosjektmedarbeideren var i jevnlig dialog med skolene og de prosjektansvarlige ute på skolene. Prosjektleder og prosjektmedarbeider har i hovedsak hjulpet med å besvare spørsmål knyttet til forskningsdesignet, til rettelegginger slik at skolene tilfredsstiller kravene til forskningsdesignet og hjulpet til

hvis skolen har møtt på problemstillinger knyttet til arbeidet med å endre matematikkundervisningen på skolene.

I tillegg har lærerne blitt fulgt opp; primært på samlingene, men også ved besøk på skolene og gjennom observasjoner i klasserommene (i samarbeid med DPU). Første og andre prosjektår har fagkonsulenter fra Pedagogisk avdeling i UDE vært med på observasjonene, i tillegg til representanter fra DPU og SSB. Dette for at de skulle få et bilde av hvordan øktplanene og det pedagogiske opplegget fungerte. Tilbakemeldinger fra skoler og lærere har også bidratt til justeringer og forbedringer av tiltakene og oppfølgingsarbeidet ut mot skolene. Da det ikke var behov for ytterligere justeringer siste prosjektår deltok ikke UDE i observasjoner dette året, og bisto kun med organiseringen av observasjonene

### 3.4. Evaluering

UDE har evaluert prosjektet på flere nivåer. Prosjektet ble evaluert fortløpende internt i UDEs prosjektgruppe, evaluering har inngått som et fast punkt i oppfølgingsmøtene med OMD, prosjektleder og ledelsen på skolene, det har blitt gjennomført evalueringer av lærerkursene, UDE har gitt organisasjonene og rektorene i referansegruppa evalueringssoppdrag og UDE har sendt ut en egen questback til alle interne prosjektledere og rektorer vår 2019. Tilbakemeldingene fra både ledelse og lærere har gått på at designet har vært noe rigid, men at de er svært fornøyd med innholdet i prosjektet. Med unntak av første prosjektår, da det ble en del forsinkelser i informasjonsflyten, har de også vært fornøyd med oppfølgingen fra UDE.

### 3.5. Avsluttende kommentar

Utdanningssetaten anser det som viktig å få økt forskningsbasert kunnskap om tiltak som kan bidra til å øke svakt presterende elevers ferdigheter i matematikk og om tiltak som kan bidra til å redusere frafall i videregående skole. Siden dette prosjektet er følgeforskning, har partene fått systematiske tilbakemeldinger underveis i prosjektet. Dette har muliggjort justeringer og forbedringer av tiltaket. Utdanningssetaten anser samarbeidet mellom de involverte partene i dette prosjektet som meget godt.

Forskning som skal måle effekt, stiller krav til forskningsdesign som også får implikasjoner på hvordan tiltaket skal gjennomføres. Dette gjelder i særlig grad områder der skolen normalt har stor råderett over organisering og innhold. Prosjektet har derfor prioritert å støtte skolene når det gjelder disse utfordringene, bl.a. med timeplanlegging og gruppeinndeling, informasjon til elever og foresatte, sikring av lik implementering i alle klasserom.

I løpet av prosjektet har Utdanningssetaten erfart hvor viktig det er at et prosjekt følges opp fra skoleeier. I tillegg kan det se ut som at det å ha en dedikert person i ledelsen som følger opp prosjektet lokalt på skolen, både med informasjon til elever og foresatte, samt å følge opp lærernes læring i kollegiet, har vært gunstig for en god gjennomføring. Det kan se ut som at på de skolene som har tenkt skoleutvikling og involvert hele matematikkseksjonen for å få til et felles løft, har fått mest utbytte av tiltaket.

Gjennom arbeidet med de ulike aktørene i prosjektet har Utdanningssetaten også erfart at det er viktig å lytte til deltakerne i prosjektet, og her har referansegruppen hatt en stor betydning. Gjennom regelmessige møter og kontakt med fagforeningene, rektorer og lærere, er det gjort justeringer underveis, slik at prosjektet oppleves som nyttig og bra.

## 4. Effektevaluering og utforming av tiltaket

For å kunne vite i hvilken grad et tiltak har virket etter hensikten, eller mer generelt ført til endringer overhodet, er vi nødt til å beregne den kausale effekten av tiltaket. Den kausale effekten er forskjellen mellom det faktiske og det kontrafaktiske resultatet – altså det som ville ha skjedd dersom tiltaket ikke hadde blitt implementert. Vi vet hvilke resultater elevene faktisk får, men kan ikke observere hvilke resultater de ville oppnådd i fravær av tiltaket - det kontrafaktiske resultatet. Dette må derfor estimeres. Generelt består utfordringen i å finne en god kontrollgruppe: En gruppe (i vårt tilfelle elever) som ikke deltar i tiltaket, men som ellers er lik elevgruppen som deltar. Faktisk resultat i kontrollgruppen vil da kunne være et anslag på hva tiltaksgruppen ville ha oppnådd i fravær av resultatet. Og forskjellen i resultatet mellom tiltak- og kontrollgruppen gir et anslag på hvordan tiltaket har virket, altså den kausale effekten.

### Randomisert kontrollert forsøk

I dette prosjektet gjennomfører vi et såkalt randomisert kontrollert eksperiment (randomized controlled trial – RCT) for å ha en gyldig kontrollgruppe. I forskningslitteraturen blir denne metoden ansett som gullstandarden for effektevaluering, da gruppen av deltagere som inngår i prosjektet tilfeldig deles inn i tiltak- og kontrollgrupper ved hjelp av loddtrekning. Ettersom det ikke er noen systematikk i utvelgelsen av tiltaks- og kontrollgrupper forventer vi at de to gruppene i utgangspunktet er like, og dermed at evt. forskjeller etter tiltaket må skyldes nettopp tiltaket.

Tiltaket har blitt gjennomført på skolenivå, noe som innebærer følgende: 1) alle elever i målgruppen på tiltaksskoler har fått målrettet undervisning; 2) vi gjør ingen analyse av klasserom og enkeltlærere som gjennomfører de tilrettelagte øktene. Som nevnt ovenfor, kunne alle lærere på tiltaksskoler med kompetanse i å undervise i matematikk delta på kursene som fulgte med tiltaket. Hvordan de ulike skolene har valgt å fordele disse lærerne til undervisningsøkter har vært opp til den enkelte skole og er utenfor vår kontroll. Dette vil derfor ikke (foreløpig) bli studert i den kvantitative effektevalueringen.

Randomiseringen av videregående skoler til tiltak- og kontrollgruppe fant sted i april 2016, og trekkingen av ungdomsskoler ble gjort i mai 2016. For effektevalueringen på 8. trinn har vi fordelt 48 ungdomsskoler tilfeldig til tiltaks- eller kontroll gruppen. Ettersom tiltaket retter seg mot svake elever har vi sortert ungdomsskolene etter antall og andel elever i målgruppen (basert på nasjonal prøve i regning på 8. trinn fra skoleåret 2015/16) og delt skolene i 24 ulike grupper, eller strata (hver gruppe bestående av to skoler, som dermed vil ha lignende antall og andel målgruppeelever). Fra hver gruppe har vi tilfeldig trukket én av de to skolene til tiltaksgruppen, mens gjenværende skole tilhører kontrollgruppen. På den måten øker vi sannsynligheten for at tiltaks- og kontrollskolene er så like som mulig.

Trekkingen på videregående foregikk på samme måte som for grunnskolen. Vi sorterte de 17 videregående skolene i Oslo etter antall og andel elever med standpunkt karakter i matematikk 1P lavere enn fire (fra skoleåret 2014/15) og delte inn i 8 ulike strata/grupper (på to skoler, en gruppe med tre skoler). Én av de to skolene ble tilfeldig valgt ut til å tilhøre tiltaksgruppen.

### 4.1. Identifisering av effekter

Elever som har resultat på mestringsnivå en eller to på nasjonal prøve i regning 8. trinn (heretter NP8) utgjør målgruppen i ungdomsskolen. Tilsvarende utgjør elever som har resultat på mestringsnivå en eller to på læringsstøttende prøve (se mer omtale i kapittel 6) målgruppen i videregående skole. I tillegg er elever som mottar

spesialundervisning utelukket fra tiltaket, ettersom disse allerede har et tilrettelagt individuelt opplegg.

Elevene som tilhører målgruppen og går på en skole som blir trukket ut til å delta i et tiltak vil dermed delta. Disse omtaler vi som tiltakselevne. Elever i målgruppen på skoler som ikke blir trukket ut fungerer som kontrollgruppe. Ettersom målgruppen er klart definert fra data tilgjengelig for alle elever, kan vi identifisere følgende grupper av elever som vi senere benytter for å identifisere forskjellige, relevante effekter av tiltaket:

1. Tiltakselever: Er i målgruppen og går på en tiltaksskole, disse tilrettelagt og målrettet undervisning.
2. Øvrige elever i tiltaksskoler: Deltar ikke fordi de ikke er i målgruppen.
  - a. For 8. trinn er det et viktig poeng er at disse elevenes gruppestørrelse og sammensetning også kommer til å endre seg i tiltaksperioden ettersom tiltaket innebærer at tiltakselevne tas ut av ordinær undervisning.
  - b. Som allerede nevnt i kapittel 2 foregikk den tilrettelagte undervisningen på Vg1 i ordinære klasserom av vanlig størrelse. Dette innebærer at også øvrige elever i tiltaksskoler på VGO mottar tilrettelagt undervisning, dvs 1P-elever som presterer over mestringsnivå 2 på læringsstøttene prøve i Vg1.
3. Kontrollgruppen - elever i målgruppen i skoler som ikke blir trukket ut: Ville deltatt om de gikk ved en tiltaksskole.
4. Øvrige elever i kontrollskoler: Befinner seg ikke i målgruppen og ville derfor ikke deltatt om de hadde gått på en tiltaksskole.

Hovedfokuset vil være effekten på elevene som deltar i tiltaket. Denne kan vi finne ved å sammenligne gruppe 1 med gruppe 3. Den tilfeldige tilordningen til skoler sikrer at elevene i gruppe 1 og 3 forventningsmessig er like i utgangspunktet og har likt forventet resultat i fravær av tiltak. Dermed gir denne sammenligningen et gyldig effektestimat. Dette effektestimatet skiller for øvrig ikke mellom hva som kan forklares av redusert gruppestørrelse og hva som kan forklares av målrettet didaktikk. Det er viktig å huske på når man tolker resultatene senere. Det kan tenkes at tiltaket også har en effekt på andre elever (som ikke tilhører målgruppen) ved tiltaksskolene (såkalte ringvirkninger). Slike ringvirkninger kan vi finne ved å sammenligne gruppe 2 og 4. Vi diskuterer disse i mer detalj nedenfor.

### Utforming og effektevaluering i ungdomsskolen

Den målrettede undervisningen i 8. trinn har blitt gjennomført en periode på høsten (5-6 uker, oppstart etter høstferien) og en periode på våren (4-6 uker) i skoleårene 2016/17, 2017/18 og 2018/19. Det første året fungerte som en innfasing og i praksis en pilot. 8 tilfeldig valgte skoler av de 24 tiltaksskolene gjennomførte tiltaket fullt ut det første året. Dvs., de fikk kursing av lærere og gjennomførte målrettet matematikkundervisning i små grupper. De resterende 16 tiltaksskolene ble kun tildelt ressurser og gitt et vagere formulert oppdrag om å tilby smågruppeundervisning til målgruppen. Fra og med skoleåret 2017/2018 skulle alle tiltaksskoler implementere tiltaket fullt ut.

Smågruppene hadde inntil åtte elever første år. Det maksimale antallet ble redusert til seks elever i andre og tredje år, fordi mindre grupper ble opplevd som mer vellykket. I så måte tilfredsstillende vår gruppestørrelse den størrelsen som ble anbefalt av Fryer (2017). Antall grupper varierte fra en til tre per skole, avhengig av antall elever i målgruppen. Antall målgruppeelever i 2015/16, dvs. før oppstart av tiltaket varierte mellom 6 og 64. Skoler som hadde 18 eller færre elever i målgruppen (24 elever første tiltaksår) fikk finansiering for det nødvendige antall smågrupper. Skoler som hadde flere enn 18 elever i målgruppen (24 første år) fikk

finansiering for to smågrupper. Målgruppeelevne med de svakeste resultatene på NP8 skulle få smågruppeundervisning. I tillegg fikk de en mindre finansiering for tilrettelegging for undervisning av øvrige målgruppeelever (med bedre resultater på NP8) i større grupper. Smågrupper krever ledige rom og må passes inn i læreres og elevers timeplaner. Det ble vurdert som for praktisk krevende å passe inn mer enn tre smågrupper, eventuelt to smågrupper dersom de samtidig skulle tilrettelegge for en annet tilbud for øvrige målgruppeelever.

I hele tiltaksperioden, ble det planlagt for totalt 51 smågrupper, dvs. i snitt 2,1 per tiltaksskole. Med åtte elever per gruppe første år ble det planlagt for at omtrent 370 elever av et antatt totalt antall på omtrent 600 målgruppeelever skulle få tilbud i smågrupper. Øvrige målgruppeelever, på forhånd beregnet til omtrent 230 elever ved 10 skoler, skulle få tilbud i store grupper. Da nasjonal prøve i regning ble gjennomført i tiltaksskolene årene 2016-2018 viste det seg at antall elever på mestringsnivå 1 eller 2, og dermed i målgruppen, ble omtrent 500 per år, og ikke 600. Antall smågrupper lå fast, slik at omtrent 130 elever fikk tiltaksundervisning i store grupper første tiltaksår. Med reduksjon i maksimal gruppestørrelse fra åtte til seks elever fra tiltaksår to ble det rom for omtrent 300 elever i smågrupper, og dermed omtrent 200 elever som fikk tilbud i store grupper.

De tilrettelagte øktene erstattet de ordinære undervisningsøktene, noe som innebar at målgruppeelevne på ungdomsskolen i tiltaksskolene ble tatt ut av ordinær matematikkundervisning de ukene tiltaket pågikk. Dette medførte at øvrige elever på tiltaksskoler (som ikke tilhørte målgruppen) også fikk endret gruppestørrelse og sammensetning under tiltaksperioden.

For å estimere effekten av tiltaket estimerer vi følgende ligning for ungdomsskoler.

$$(1) \quad y_{ist} = \beta T_s + \alpha D_t + \delta G_s + \eta X_{ist} + \varepsilon_{ist}$$

$y_{ist}$  er resultatet til en elev  $i$  ved skole  $s$  i år  $t$ . I effektanalysene vil dette i hovedsak være resultater fra nasjonal prøve i regning på 9. trinn (NP9), enten skalapoeng eller indikatorer for mestringsnivå.

$T_s$  er lik 1 når skole  $s$  er en tiltaksskole og 0 ellers.  $\beta$  viser dermed forskjellen mellom tiltaks- og kontrollskoler. Dette er effektestimaten, eller mål på forskjeller i kjennetegn mellom skolene som vi vil vise i resultattabellene. Estimering av (1) gir også et mål på usikkerheten i estimatene, i form av beregnede standardfeil,  $\varepsilon_{ist}$ . Disse kan brukes til å angi et intervall som vi forventer en «sann» effekt vil være innenfor med en gitt, høy, sannsynlighet, gjerne i form av et 95 prosent konfidensintervall.<sup>13</sup> Et særlig spørsmål vil være om resultatene støtter en hypotese om ingen effekt (dvs.  $\beta=0$ ). Dette vurderes gjerne ved å beregne om estimerer er signifikante på 5 prosent nivå (dvs. 0 ikke inneholdt i et 95 prosent intervall) eller på 10 prosent nivå (som betyr en større sannsynlighet for å konkludere med at det er effekter som ikke faktisk finnes, og som dermed gir en svakere konklusjon). Det er likevel verdt å merke seg at estimaten  $\beta$  er det beste anslaget på en forskjell eller effekt, selv om manglende statistisk signifikans ikke gir grunnlag for klart å konkludere med at det er en effekt.

I tillegg til deltakelse i tiltaket kontrollerer vi for noen andre forskjeller mellom elevene, som er mindre sentrale for tolkningen av resultatene, og som vi derfor ikke kommer til å vise detaljerte resultater for:  $D_t$  er et sett av indikatorvariable for hvert elevkull, slik at  $\alpha$  gir forskjellene i snittverdi mellom årene.  $G_s$  er indikatorvariable for hvilken gruppe (eller strata, jf. avsnitt 4 på side 26) skole  $s$  er plassert ifm randomiseringen, og  $X$  er kjennetegn ved elevene. Dette kan omfatte kjønn og

<sup>13</sup> Et 95 prosent konfidensintervall vil, dersom vi gjør mange uavhengige analyser, i 95 % av tilfellene inneholde den sanne verdien av  $\beta$ .

familiebakgrunn (foreldres utdanningsnivå, innvandringsbakgrunn), og i en del analyser også resultater tidligere i skoleløpet, enten fra NP8 eller nasjonal prøve i regning 5. trinn (NP5). I estimeringen tar vi hensyn til at det kan være sammenheng i resultatene innad i skoler ved å tillate restleddet  $\varepsilon$  å være korrelert over tid for en gitt skole (ved å bruke såkalt klynge- eller «cluster-robuste» standardfeil).

På grunn av at vi tilfeldig har fordelt skoler til tiltak- og kontrollgrupper, forventer vi at vi kan gi  $\beta$  en kausal tolkning, nemlig effekten på nasjonale prøver i 9. klasse av å motta målrettet og tilrettelagt matematikkundervisning i små grupper. Det er verdt å merke seg at dette er en forventet egenskap ved effektestimaterne, og ikke garantert. Dersom vi tilfeldig hadde trukket et stort antall tiltaks- og kontrollskoler kunne vi vært tryggere på at gruppene var like i utgangspunktet. Når vi trekker et begrenset antall skoler vil det kunne bli tilfeldige forskjeller mellom skolene og elevgruppene. Stratifisert trekning, dvs. gruppering av skoler og trekking innen gruppene, som beskrevet over, er en strategi for å begrense mulighetene for slike tilfeldige forskjeller.

Etablering av flere små grupper på 8. trinn innebærer en økning i antall lærertimer. Dette betyr at lærere må omfordeles sammenlignet med hvilke grupper og fag de ellers ville undervist og/eller nye lærere må ansettes. Begge deler kan ha en betydning for resultatene til elevene, og vil evt. være en effekt av tiltaket. Dette er en relevant effekt, eventuelt andre skoleeiere som ønsker å kopiere tiltaket vil ha tilsvarende behov for omfordeling og/eller nyansettelser av lærere.

For effektevalueringen sin del kunne det vært interessant å styre fordelingen av lærere, for å studere effekten av dette. Imidlertid innser vi at dette ville vært en utfordring for skolene, og vanskelig gjennomførbart. I tiltaket ligger det en føring om at lærerne som underviser skal tilfredsstille kompetansekravene. Dette medfører at skolene må forholde seg til en (mer eller mindre) gitt gruppe lærere samtidig som at de må dekke opp annen undervisning. Det vil derfor ikke være mulig å legge sterke ytterligere føringer på fordelingen av lærere til grupper.

### **Utforming og effektevaluering i videregående skole**

Også lærerne i de videregående tiltaksskolene fikk kursing i å tilrettelegge matematikkundervisningen for elever med svak matematikkompetanse, jf. de didaktiske elementene beskrevet tidligere i rapporten. Skolene fikk imidlertid ikke ressurser til smågruppeundervisning, kun begrensede midler for vikarer i for bindelse med at lærerne skulle på kurs. Det vil si at undervisningen foregikk i ordinære grupper, eventuelt i grupper som ble opprettet ved å omprioritere skolens øvrige ressurser. I praksis er det vårt inntrykk at den målrettede undervisningen i hovedsak ble gjennomført i ordinære 1P klasserom. Dette innebærer at selv om undervisningen var rettet mot de svakeste elevene på 1P, var også øvrige elever til stede i klasserommene når undervisningen foregikk.

Bakgrunnen for at videregående skoler ikke fikk ressurser til grupper er todelt. For det første, har elevene på videregående skole allerede selektert seg selv til mer homogene grupper. For eksempel, de elevene som i utgangspunktet har valgt 1P matematikk presterer normalt dårligere i matematikk enn elever på 1T<sup>14</sup>. Andelen målgruppeelever (blant 1P elevene) på videregående er dermed høyere sammenliknet med grunnskolen, og det er færre elever med veldig forskjellige matematikkferdigheter. For det andre vurderte vi at de videregående skolene er større, med flere aktuelle lærere og tilstrekkelig mulighet til å tilrettelegge for undervisningen innenfor eksisterende rammer.

---

<sup>14</sup> Se fotnote 1.



For å finne effekter i videregående estimerer vi følgende ligning:

$$(2) \quad y_{ist} = \beta T_s + \alpha D_t + \delta G_s + \eta X_{ist} + \varepsilon_{ist}$$

Ligningen er helt tilsvarende ligningen for grunnskolen.  $y_{ist}$  er resultatet til en elev  $i$  ved skole  $s$  i år  $t$ . For videregående vil vi i hovedsak fokusere på resultater fra Oslo-prøven, samt tidlige mål på gjennomstrømning. Øvrige forskjeller mellom videregående og grunnskole er hvilke tidligere resultater vi tar hensyn til (dvs som i noen spesifikasjoner inngår i  $X$ ): i videregående bruker vi resultat på læringsstøttende prøve samt standpunktarakter fra grunnskolen. I videregående tar vi hensyn til korrelerte restledd innen skole og år, i motsetning til grunnskolen der vi tar hensyn til korrelasjoner innen skoler mellom år. Motivasjonen for dette er dels at det er færre videregående skoler, som kan gi utfordringer med å beregne korrelasjoner, samt at de videregående skolene er større og at mye interaksjon skjer innen trinnene. Vi finner imidlertid ikke store forskjeller om vi lar restledd være korrelerte innen skoler eller skoler og år.

### Ringvirkninger til elever som presterer bedre i matematikk

Som nevnt, kan det tenkes at tiltaket også påvirker øvrige elever (dvs. elever utenfor målgruppen). På ungdomsskolen er den mest åpenbare årsaken at ringvirkninger oppstår på grunn av at gruppestørrelse og -sammensetning også endres for elevene i gruppe 2. Ettersom elevene i gruppe 1 tas ut av ordinær undervisning vil også elever tilhørende gruppe 2 på tiltaksskoler motta matematikkundervisning i mindre grupper der det er mindre spredning i det faglige nivået. Dette åpner derfor muligheten til også å studere effekten av nivådeling («tracking» i internasjonal litteratur, f.eks. Duflo m.fl., 2011), både for elevene som deltar i tiltaket og de øvrige elevene. Å motta undervisning i nivådelte grupper kan tenkes å øke læringsutbyttet, enten fordi det i en homogen gruppe er lettere for lærerne å tilpasse seg elevenes behov, eller at man lærer best sammen med elever på samme nivå. På videregående kan øvrige elever påvirkes på grunn av at den målrettede undervisningen gis i ordinære klasserom hvor alle elevene er tilstede. Kursingen av lærere kan også ha karakter av generell skoleutvikling, og påvirke hvordan lærerne underviser også utenom tiltaksundervisningen.

### Heterogene effekter

Tiltakene kan ha ulik effekt på ulike elevgrupper. En fordel med å gjennomføre (samme eller lignende) tiltak både i ungdomsskoler og videregående skoler er at det vil være mulig å sammenligne effekter av tilsvarende tiltak ved forskjellige aldre (nivåer).

Vi har også gjort separate estimeringer for forskjellige elevgrupper, og dermed studert effektheterogenitet utfra kjønn, innvandringsbakgrunn, foreldres utdanning eller faglige prestasjoner før tiltaket. I tillegg har det vært mulig å dele tiltakselever på samme skole til ulike nivådelte grupper (for eksempel grupper bestående utelukkende av elever som scorer hhv. på nivå 1 og 2 på nasjonale prøver).

I grunnskolen har vi også sett på forskjellige grupperinger av skoler. De mest sentrale grupperingene går på antall og andel elever i målgruppen samt om skolene startet med fullt tiltak første år.

Når vi ser på heterogene effekter vil vi beregne effekter for grupper av elever eller skoler. Dette medfører at antallet elever (og evt. skoler) bak hvert effektestimert blir mindre, og estimatene mindre presise (mer usikre). Dette er særlig et problem for forskjeller mellom skoler, der det vil være vanskelig å vurdere om eventuelt forskjeller i effekter skyldes, eller bare samvarierer med, bestemte observerte kjennetegn ved skolene. Det er to spesielle utfordringer her: For det første er

kjennetegn ved skolene ikke tilfeldig fordelt, i motsetning til deltakelse i tiltaket, slik at vi må ha en relevant sammenligningsgruppe for hver gruppe av tiltaksskoler. For det andre, må vi ha en relevant sammenligningsgruppe for å finne gyldige effektestimater. Det vil normalt kreve at vi kan gruppere tiltaks- og kontrollskoler på tilsvarende måte, dvs. vi kan ikke lage grupper basert på informasjon vi bare har for tiltaksskolene.<sup>15</sup>

I sum gjør dette at heterogene effektestimater er mer usikre enn snitteffekter, og vi vil derfor i begrenset grad fokusere på disse. For videregående skole er det for få skoler til at vi kan ha forhåpninger om å skille mellom forskjellige effekter knyttet til forskjeller i skolekjennetegn.

## 4.2. Forventede resultater, virkningsmekanismer og styrkeberegning

Tiltaket er i første omgang rettet mot lærerne. Hensikten er at lærerne skal få kjennskap til nyttige prinsipper og didaktiske virkemidler for undervisning av matematikk til målgruppen. Disse delementene er ikke unike for tiltaket, men vår oppfatning er at de er forholdsvis lite kjent blant lærere på ungdomstrinn og i videregående, og at disse lærerne i begrenset grad kjenner til oppdatert forskning og erfaring med denne didaktikken. Videre forventer vi at større kjennskap til tiltaksdidaktikken påvirker lærernes undervisning, i det minste i tiltaksperioden, slik at denne i større grad inneholder de didaktiske elementene i tiltaket. Vi har mindre klare forventninger til undervisningen utenom tiltaksperiodene, men, basert på motivasjonen for og forskningen bak tiltaket, tror vi tiltaksdidaktikken er relevant også utenom tiltaksundervisningen. Derfor forventer vi også at undervisningen ellers påvirkes i noen grad.

For elevene forventer vi at deres matematikkferdigheter bedres gjennom tiltaket. På sikt arbeider Oslo kommune mot å kun ha 10 prosent elever på mestringsnivå 1 og 2 ved nasjonale prøver i regning på 9. trinn og at gjennomsnittskarakteren ved skriftlig eksamen i matematikk på 10. trinn øker til 3,8. Det er et mål at vi i tiltaksperioden kan se at vi nærmer oss målene. På kort sikt, i løpet av tiltaket, har vi vurdert det som realistisk med en effekt på 0,1-0,2 standardavvik på faglige ferdigheter for målgruppen, dvs. 1-2 skalapoeng på nasjonale prøver. Dette anslaget baserer seg på internasjonal forskning (som Cortes m.fl., 2014) og UDEs erfaringer med andre satsninger, bl.a. på matematikk på yrkesfag.

Sammenhengen mellom matematikkresultater og fullføring tilsier at det dermed også kan ha en effekt på fullføring. Måltallet Oslo kommune her opererer med er at 80% av elevene skal ha fullført og bestått videregående opplæring i løpet av fem år.<sup>16</sup>

Ettersom elevene i utgangspunktet er svaktpresterende vil økte ferdigheter kunne ha en betydelig effekt på andelen som består matematikk.<sup>17</sup> Dette gir en direkte effekt på elevenes muligheter til å fullføre og bestå VGO. Økt mulighet til å fullføre med bestått vil også kunne ha betydning for elevenes motivasjon for å fullføre i det hele tatt, eller for å bestå andre fag.

<sup>15</sup> Om vi ønsker å gjøre dette må vi anta at alle grupper av tiltaksskoler i utgangspunktet er tilsvarende, og dermed at hele gruppen av kontrollskoler er en relevant sammenligning for alle gruppene. Dette vil ofte være en sterk antakelse.

<sup>16</sup> Cortes m.fl. (2014) finner en effekt på fullføring som er større enn forventet effekt utfra effekten på matematikkferdigheter, og større enn effekten på bestått i matematikkfag.

<sup>17</sup> Fra 2013/14 til 2014/15 økte snittresultatet på 1P-Y i Oslo fra 2,88 til 2,93. Samtidig økte andelen bestått fra 85,3 prosent til 90,0.

Tiltaket kan også ha en mer indirekte effekt på fullføring av VGO. Hvis elevene får økte ferdigheter og resultater i matematikk vil dette kunne påvirke deres selvfølelse, opplevd mulighet til å fullføre VGO og trivsel og motivasjon for skolearbeidet. Dette kan påvirke elevenes videre innsats, læring og sannsynlighet for å fullføre VGO. Samlet sett, basert på forventede effekter på faglige ferdigheter og tidligere forskning som Cortes m.fl., forventer vi en effekt for målgruppen på 5-10 prosentpoeng på fullføring.

Som en del av arbeidet med prosjektbeskrivelsen ble det gjennomført styrkeberegninger. Dette er beregninger som viser i hvilken grad det vil være mulig eller sannsynlig å finne effekter av forskjellige størrelser. Styrkeberegninger for tiltaket i ungdomsskolen tilsier at vi sannsynligvis vil være i stand til å påvise en effekt på 0,05 standardavvik (0,5 skalapoeng) på nasjonal prøve, mens vi med omtrent 90 prosent sannsynlighet vil kunne påvise en effekt på 0,1 standardavvik. Tiltaket i videregående ble utformet med en beregnet styrke som tilsier at vi ville finne evt. effekter større enn 0,2 standardavvik. Fordi datagrunnlaget fra videregående var mer begrenset enn for ungdomsskolen ved utformingen av tiltaket var disse beregningene mer usikre.

## 5. Effektestimater – grunnskolen

I dette kapitlet presenterer vi de kvantitative analysene for grunnskolen. Vi starter med å presentere datagrunnlaget. Deretter undersøker vi hvorvidt tiltak- og kontrollgruppen er sammenliknbare, før vi diskuterer resultatene for de siste to årene, 2017/18 og 2018/19.

I Kirkebøen m.fl. (2018) presenterte vi de kvantitative resultatene for 2016-17. For grunnskolen gjennomførte vi to forskjellige varianter av tiltaket i 2016-17. Som vist i Kirkebøen m.fl. var det også særlig store utfordringer med balansering (tilvarende utgangspunkt i tiltaks- og kontrollskoler) for den av tiltaksvariantene som er videreført i de siste to årene. Første semester med kursing av lærerne i grunnskolen var også preget av en del oppstartsproblemer, som det sannsynligvis vil være mulig å unngå eller redusere med bedre tid til å forberede oppstart. Til slutt vil det begrensede antall elever i første tiltaksår ikke ha stor betydning for resultatene fra evalueringen av tiltaket i grunnskolen. Vi har derfor valgt å utelate diskusjonen av resultater fra 2016/17 fra denne rapporten og henviser til Kirkebøen m.fl. (2018) for ytterligere diskusjon.

Vi avslutter kapitlet med å oppsummere våre funn, vurdere effektstørrelser opp mot tidligere tiltak i ungdomsskolen, samt å vurdere kostnader av tiltaket og anslag på framtidige gevinster av evt. effekter på læringsutbytte.

### 5.1. Datagrunnlag

Datamaterialet baserer seg i hovedsak på data fra administrative registre, dels nasjonale registre og dels data fra Oslo kommune. I tillegg supplerer vi med noe data samlet inn i regi av prosjektet: Data fra spørreundersøkelser gjennomført av Fafo, samt data om tilordning av elever til grupper på 8. trinn. Fra nasjonal utdanningsdatabase (NUDB) finner vi elevenes tidligere resultater fra nasjonale prøver på 8. trinn (vi vil fokusere på regning, som vi heretter omtaler som NP8) samt nasjonal prøve i regning på 5. trinn (NP5). NUDB gir oss også informasjon om elevenes familiebakgrunn, dvs. elvenes fødselsår og kjønn, foreldrenes høyeste utdanningsnivå og elevenes og foreldrenes innvandringsbakgrunn. Fra Oslo kommune får vi data for spesialundervisning og nasjonal prøve i regning på 9. trinn (NP9).<sup>18</sup> ..

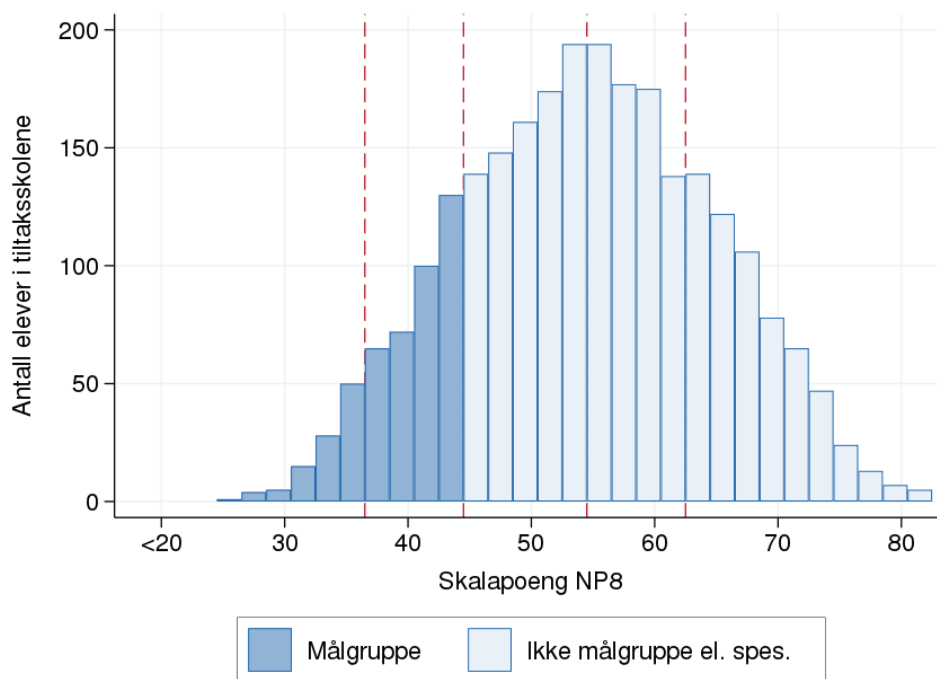
NP8 definerer elevpopulasjonen vår: Vår utgangspopulasjon består av alle elever som er registrert i dataene som inneholder NP8. For NP8 gjennomført i 2017 og 2018 utgjør det omlag 11000 elever. Analysesamplet, dvs. de elevene som vi vil undersøke i analysene, omfatter alle elever som er registrert med et faktisk resultat fra NP8 (dvs. ikke fritatt eller fraværende) og som ikke er registrert med spesialundervisning. Analysesamplet består av om lag 9900 elever, eller rundt 90 prosent av elevpopulasjonen (4,4 prosent mangler NP8 og 8,1 prosent er registrert med spesialundervisning, med noe overlapp mellom disse gruppene). Målgruppen er elever i analysepopulasjonen som har et resultat fra NP8 på mestringsnivå 1 eller 2. Totalt utgjør dette 1977 elever, som igjen er 20 prosent av analysesamplet.

Figur 5.1 viser fordelingen av resultater på NP8 (skalapoeng) for elevene høsten 2017. Skillene mellom mestringsnivåer er markert med stiplede vertikale linjer (fire linjer for skiller mellom fem mestringsnivåer fra 1-5), og målgruppen er

<sup>18</sup> De nasjonale dataene fra nasjonale prøver på 9. trinn i 2019 var ikke tilgjengelig da analysene ble gjennomført. Dette betyr at vi ikke får ved elever som har avlagt NP9 utenom kommunale skoler i Oslo. Vi har resultater fra NP9 for 94 prosent av analysepopulasjonen. For NP 2018 har vi hatt muligheten til å sammenligne resultater basert på nasjonale data og data fra Oslo kommune. Vi finner da at det er få ytterligere elever med resultater fra andre skoler/kommuner, og at resultatene ikke endres vesentlig om vi bruker nasjonale data eller data fra Oslo kommune. For å være konsistente i bruk av datakilde vil vi derfor bruke NP9-data fra Oslo kommune gjennomgående.

markert med en mørkere farge. Fra figuren ser vi at det er betydelig variasjon i resultatene innad i målgruppen, men at en stor andel av målgruppen ligger nær skillet mellom mestringsnivå 2 og 3, dvs mellom 40 og 44,5 skalapoeng. Til sammenligning er det færre elever som har resultater på mestringsnivå 1, og selv om laveste resultat er omtrent 25 skalapoeng er det svært få som har mindre enn omtrent 30.

Figur 5.1 Fordeling av skalapoeng (NP8) og mestringsnivå – høst 2017



Kilde: Statistisk sentralbyrå.

### Utfallsvariable og øvrige variable som måler elevprestasjoner på ungdomsskolen

Det overordnede formålet med tiltaket er å øke andelen som fullfører og består videregående opplæring. Ettersom tiltakselevne fortsatt er i grunnskolen eller første år på videregående skole (VG1) er det for tidlig å analysere dette. Effektevalueringen i dette kapitlet vil derfor fokusere på elevens ferdigheter i regning, målt ved NP9. Som omtalt tidligere, er målsetningen at tiltaket i første omgang skal bedre elevenes forståelse for og ferdigheter i regning. Vi forventer at bedringen er varig, slik at også ferdigheter og resultater i matematikk i videregående bedres av tiltaket i grunnskolen. Bedrede ferdigheter og resultater i matematikk antas å ha en positiv effekt på elevenes muligheter til å gjennomføre videregående opplæring. Sterkere resultater i matematikk kan 1) direkte føre til at flere elever består videregående (for resultater i videregående); og 2) gi en økt opplevelse av mestring som gir seg utslag i bedre karakterer generelt sett. På lengre sikt har vi til hensikt å studere elevenes resultater også ved fullført grunnskole (dvs. karakterer, særlig i matematikk), samt gjennomstrømning i videregående opplæring. Dette vil kunne studeres ved hjelp av data fra NUDB.

I Tabell 4.1 gir vi en deskriptiv oversikt over NP9 som er vår utfallsvariabel for ungdomsskoleelevene samt øvrige pre-determinerte variable som måler elevprestasjoner i forkant av tiltaket. Ikke overraskende presterer elevene i målgruppen dårligere enn øvrige elever. For eksempel er snittet på NP8 i målgruppen 39,2, mot 57,2 blant øvrige elever. Til sammenlikning er snittet i hele analysesamplet 53,6. Analysesamplet vårt er betinget på at elevene har gjennomført nasjonale prøver på 8. trinn, til sammen 9929 elever. Noen av elevene i

analysesamplet mangler resultat fra NP5 og/eller NP9, og dermed varierer antall observasjoner noe innen hver elevgruppe.

**Tabell 5.1 Elevprestasjoner på ungdomsskolen – analysesamplet<sup>1</sup> skoleårene 2017/18 og 2018/19**

	Analysesamplet		Målgruppen		Øvrige elever	
	Snitt	Ant. obs	Snitt	Ant. obs	Snitt	Ant. obs
<b>Utfallsvariable</b>						
NP9	56,887	9371	43.817	1775	59.941	7596
NP9=1	0,020	9371	0.103	1775	0.001	7596
NP9<=2	0,123	9371	0.540	1775	0.025	7596
<b>Pre-determinerte resultat</b>						
NP8	53,629	9929	39.179	1977	57.222	7952
NP5	53,683	9282	42.056	1599	56.103	7683

<sup>1</sup> Alle elever med registrert resultat fra NP8.

Kilde: Statistisk sentralbyrå.

### Andre bakgrunnskjenne tegn ved elevene

I Tabell 5.2 gir vi en oversikt over elevenes bakgrunn. Som i Tabell 5.1 deler vi opp i tre grupper: analysesamplet, målgruppen og øvrige elever. Det er flere gutter i målgruppen (enn jenter), andelen elever med lavt utdannede foreldre<sup>19</sup> er høyere i målgruppen enn blant øvrige elever, og det er flere elever med utenlandskfødte foreldre i målgruppen sammenliknet med gruppen av øvrige elever.

**Tabell 5.2 Deskriptiv oversikt over variable som måler familiebakgrunn på ungdomsskolen – årene 2017/18 og 2018/19**

	Analysesamplet		Målgruppen		Øvrige elever	
	Snitt	Ant. obs	Snitt	Ant. obs	Snitt	Ant. obs
<b>Kjønn</b>						
-Gutt	0,508	9927	0,592	1977	0,487	7952
-Jente	0,492	9927	0,408	1977	0,513	7952
<b>Foreldres utdanning</b>						
-Lav	0,355	9927	0,671	1977	0,276	7952
-Høy	0,645	9927	0,329	1977	0,724	7952
<b>Innvandrerbakgrunn</b>						
-Norskfødte foreldre	0,688	9927	0,424	1977	0,754	7952
-Utenlandskfødte foreldre	0,312	9927	0,576	1977	0,246	7952

Kilde: Statistisk sentralbyrå.

## 5.2. Balansering av tiltak- og kontrollskoler

Skolene er tilfeldig fordelt til tiltak og kontroll. Vi forventer derfor ikke systematiske forskjeller mellom tiltak- og kontrollskoler når det gjelder antall elever, elever med bestemte kjennetegn (for eksempel familiebakgrunn eller tidligere resultater), deltakelse i nasjonale prøver e.l. Ettersom vi har et begrenset antall skoler kan det likevel oppstå tilfeldige forskjeller.

I tabell 5.3 sammenligner vi tiltaks- og kontrollskolene for ungdomsskolen ytterligere for å undersøke om det er forskjeller mellom tiltak- og kontrollskoler. De rapporterte tallene kommer fra å estimere ligning (1) hvor avhengig variabel er de ulike variablene i rad to i tabell 5.3. Totalt rapporterer tabell 5.3 seks ulike regresjoner, hvor hver kolonne representerer en regresjon. Gjennomsnittsverdien til utfallsvariablen rapporteres i siste rad.

<sup>19</sup> Foreldres utdanning er definert til å være høy dersom minst en av foreldrene har høyere utdanning (minst 3 år med høyere utdanning), og ellers lav.

**Tabell 5.3 Balansering av tiltak- og kontrollskoler for grunnskolen, årene 2017/18 og 2018/19<sup>1</sup>**

	(1)	(2)	(3)	(4)	(5)	(6)
	Har resultat	Spesial- NP8 undervisning	Inngår i analyse- samplet	Indeks familie bakgrunn	Resultat NP8	Inngår i målgruppen
Tiltak	-0.002 (0.008)	-0.003 (0.008)	0.001 (0.011)	0.970** (0.355)	0.756* (0.447)	-0.022* (0.012)
Ant observasjoner	11106	11106	11106	9929	9929	9929
Snitt avhengig variabel	0.956	0.081	0.894	55.955	53.629	0.199

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (1) med forskjellige uavhengige variable ( $y$ ) i kolonner. Alle estimater kontrollerer for årskull og strata (gruppe i randomiseringen). Klyngerobuste beregnede standardfeil i parentes under hvert effektestimater. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

I de første tre kolonnene bruker vi hele datagrunnlaget, dvs. 11106 elever som er registrert på 8. trinn i Oslo i 2017 eller 2018. Siste rad i kolonne (1) viser at 95,6 prosent av disse har faktisk resultat fra nasjonal prøve i regning 8. trinn. De øvrige 4,5 prosentene var fritatt eller fraværende. Elever uten resultat fra nasjonale prøver har ofte svakere forutsetninger. «Tiltak» er beregnet forskjell mellom tiltak- og kontrollgruppene (dvs. estimert  $\beta$  fra ligning (1)), og som vi ser er det ingen tydelig forskjell mellom tiltak- og kontrollskolene når det gjelder gjennomføringen av NP8. Der er heller ingen tydelig forskjell i andelen som er registrert med spesialundervisning: både i tiltak- og kontrollskolene er denne omtrent 8 prosent (kolonne 2). Som allerede nevnt ovenfor, i analysesamplet for ungdomsskolen inkluderer vi elever som har resultater på nasjonale prøver på 8. trinn og som ikke mottar spesialundervisning. Dette utgjør om lag 90 prosent av elevene i utgangspopulasjonen. I kolonne (3) undersøker vi om denne andelen er ulik på tvers av tiltak- og kontrollskolene. Som forventet er det ingen forskjell i andelen av elevene som inngår i analysesamplet.

I kolonnene (4) – (6) begrenser vi datamaterialet til elever som inngår i analyse-samplet vårt (totalt 9929 elever). Dette er elevene som senere inngår i effekt-analysene. I kolonne (4) ser vi at det er en forskjell i elevenes familiebakgrunn i favør tiltaksskolene. Familiebakgrunn er her målt ved en indeks som svarer til skalapoeng på nasjonale prøver.<sup>20</sup> Vi finner dermed en forskjell i familiebakgrunn som gjør at vi forventer at elevene i tiltaksskolene oppnår omtrent ett skalapoeng mer på nasjonal prøve i regning 8. trinn enn elevene i kontrollskolene. I kolonne (5) viser vi den faktiske forskjellen i NP8, denne er omtrent 0,76 skalapoeng. Som ventet utfra forskjellen i familiebakgrunn oppnår elevene i tiltaksskolene bedre resultater enn elevene i kontrollskolene. Forskjellen i faktiske resultater er imidlertid mindre enn forventet forskjell basert på forskjell i familiebakgrunn. Til slutt viser vi forskjellen i andelen elever i målgruppen, dvs. elever på mestringsnivå 1 og 2 på NP8. I tillegg til å score høyere på NP8 i snitt har også tiltaksskolene færre elever på de lave mestringsnivåene. I snitt er andelen i målgruppen omtrent 20 prosent, denne andelen er 2,2 prosentpoeng lavere i tiltaks- enn i kontrollskolene.

Forskjellene i NP8 og andel i målgruppen er på grensen til hva en skulle vente basert på tilfeldig variasjon (statistisk signifikante på 10 prosent nivå, men ikke på 5 prosent). Forskjellen i familiebakgrunn er større, og klart statistisk signifikant. Ettersom skolene er tilfeldig fordelt til tiltak og kontroll venter vi ikke denne typen signifikante forskjeller. På den andre siden, gjør det at vi har et såpass begrenset antall skoler (totalt 48, hvorav 24 trekkes ut som tiltaksskoler), og at skolene er ulike med tanke på elevtall og elevkjennetegn, at vi likevel kan få forskjeller

<sup>20</sup> Vi har gjort en regresjon av resultat på nasjonal regning 8. trinn som avhengig variabel og forskjellige familiekjennetegn (elvenes kjønn, foreldrenes utdanning, hvorvidt foreldrene/eleven er født i utlandet). Basert på resultatene for denne har vi så beregnet familiebakgrunnsindeksen som den forventede verdien for hver elev. Denne indeksen er dermed et vektet gjennomsnitt av kjennetegnene som inngår i regresjonen, med vekt bestemt av hvor sterk sammenheng det er mellom kjennetegnet og resultat fra nasjonale prøver.

mellom gruppene. I Kirkebøen m.fl. (2018) diskuterer vi, basert på simuleringer, hvor sannsynlig det er at det oppstår vesentlige forskjeller mellom tiltaks- og kontrollskolene.

### Målgruppen

Vi fortsetter med å se på balansering i målgruppen, dvs. elever med resultat fra NP8 på mestringsnivå 1 eller 2. Totalt er det 1977 elever i målgruppen, fordelt på tiltaks- og kontrollskoler og på skoleårene 2017/18 og 2018/19. Disse resultatene er rapportert i Tabell 5.4, og hver kolonne representerer en spesifikkasjon.

**Tabell 5.4 Balansering av målgruppen i tiltak- og kontrollskoler for grunnskolen, årene 2017/18 og 2018/19<sup>1</sup>**

	(1)	(2)	(3)	(4)
	Indeks familie bakgrunn	Resultat NP8	Resultat NP5	Har resultat NP5
Tiltak	0.653** (0.268)	0.113 (0.154)	-0.203 (0.399)	-0.019* (0.011)
Indeks for familiebakgrunn		x	x	x
Ant observasjoner	1977	1977	1599	1977
Snitt avhengig variabel	51.684	39.179	42.056	0.809

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (1) med forskjellige uavhengige variable ( $y$ ) i kolonner. Alle estimater kontrollerer for årskull og strata (gruppe i randomiseringen). Kontroll for indeks av familiebakgrunnsvariable i alle kolonner dette er angitt. Klyngerobuste beregnede standardfeil i parentes under hvert effektestimater. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

Kolonne (1) i tabell 4.4 viser at også målgruppeelevene har forskjellig familiebakgrunn i tiltaks- og kontrollskolene. Dette er som ventet, basert på forskjellene vi fant for alle elever i forrige avsnitt. Forskjellene er imidlertid mindre enn for alle elever under ett, tilsvarende en forventet forskjell på omtrent 0,65 skalapoeng. I kolonne (2) ser vi på forskjeller i resultatet på NP8 betinget på familiebakgrunnsindeksen. Målgruppeelevene i tiltaksskolene scorer 0,11 skalapoeng høyere på NP8 enn målgruppeelevene i kontrollskolene, utover hva vi kan forklare med familiebakgrunn. Denne forskjellen er imidlertid godt innenfor hva vi kan vente fra tilfeldig variasjon. I tillegg til å se på balansering i NP8, ser vi også på hvorvidt NP5 er balansert på tvers av tiltaks- og kontrollskolene (kolonne (3) og (4)). Også i analysene for NP5 tar vi hensyn til forskjeller i familiebakgrunn. Målgruppeelevene i tiltaksskolene har 0,20 skalapoeng lavere resultat på NP5, dette er også innenfor hva som kan skyldes tilfeldige forskjeller. Tiltaksskolene har også flere elever som ikke har prøveresultat fra NP5. Elever som mangler en prøve har i gjennomsnitt svakere resultater på øvrige prøver. Dette kan tyde på at forskjellen i forutsetninger basert på NP5 er større enn 0,20 skalapoeng.

Med tanke på nasjonale prøver (NP8 og NP5) finner vi altså et sammensatt bilde hva gjelder forskjeller mellom målgruppeelevene. Derimot er det en forholdsvis tydelig forskjell i elevsammensetningen, der elevene i tiltaksskolene har bedre forutsetninger (målt ved familiebakgrunnsindeks) enn elevene i kontrollskolene. Når vi tar hensyn til denne forskjellen er det en forskjell i NP8 i favør tiltaksskolene, og en forskjell i NP5 i favør kontrollskolene. Disse forskjellene er imidlertid mindre tydelige.

Oppsummert, ettersom skoler er tilfeldig fordelt til tiltak og kontroll forventer vi ikke systematiske forskjeller uavhengig av tiltaket. Analysene av grunnskolen viser hvordan det likevel kan oppstå tilfeldige forskjeller. Forskjellene i elevsammensetning kan ventes å skape forskjeller mellom tiltaks- og kontrollgruppen utover evt. forskjeller som skyldes tiltaket, og er dermed noe vi må ta hensyn til i analysene. Vi gjør det ved først å fremst å kontrollere for familiebakgrunn og NP8, men vil også kontrollere for bare familiebakgrunn og for familiebakgrunn og NP5.



NP8 gjennomføres i forkant av tiltaket, dvs. før elevene fordeles til målgrupper og ikke målgrupper, og før målgruppeelevene gis undervisning i smågrupper. Ettersom NP8 ligger tett opp mot oppstart av tiltaksundervisningen og er helt tilsvarende NP9, som er det første resultatet vi måler etter tiltaket, vil NP8 gi et godt og presist mål på elevenes forutsetninger. Spesifikasjonene med kontroll for NP8 er derfor våre hovedspesifikasjoner, i tråd med planlagt evaluering slik den er beskrevet før oppstart av tiltaket.

NP8 gjennomføres imidlertid etter at skolene er fordelt til tiltak og kontroll (høsten 2016), og etter at lærerne er i gang med kurs. Ettersom vi her ser på NP8 fra 2017 og 2018 vil noen lærere ved tiltaksskolene ha gjennomført kurs allerede i 2016. Vi kan ikke utelukke at det påvirker gjennomføring av NP8 for senere kohorter, noe som igjen kan påvirke resultatene. Dette kan for eksempel være fordi prøvene på skolene som er trukket ut til tiltaksgruppen gjennomføres med en visshet om at resultatene på prøvene påvirker tilbudet til elevene. Vi har ikke indikasjoner på at dette har skjedd, verken i form av forskjellige andeler som gjennomfører prøver eller i form av rapporter fra skolene, men kan heller ikke helt utelukke at dette har skjedd. I så fall kan det være problematisk å bruke NP8 som kontroll. Resultater fra NP5 eller bare familiebakgrunn er slik sett bedre måter å kontrollere for skjevheten på, ettersom disse ikke kan være direkte påvirket av tiltaket. På den annen side er disse mindre presise kontroller for forskjeller i elevenes forutsetninger.

### 5.3. Effekter

I det følgende presenterer vi resultatene fra den kvantitative effektevalueringen for grunnskolen. Vi starter med å se på hvordan tiltaket påvirket elevene i målgruppen, altså de svakest presterende elevene som tiltaket er rettet mot. Deretter studerer vi øvrige elever ved tiltaks- og kontrollskolene.

Som vist ovenfor er ikke tiltaks- og kontrollgruppene perfekt balansert i observerbare kjennetegn ved elevene. Fra balanseringstestene vet vi blant annet at elevene i tiltaksskolene har bedre forutsetninger, målt ved familiebakgrunn. Det er derfor også forventet at disse elevene vil prestere bedre, også i fravær av tiltaket. I den kvantitative effektevalueringen tar vi hensyn til dette ved å kontrollere for familiebakgrunn og resultatet på nasjonale prøver i 8. og 5. klasse. Såfremt vi lykkes i å ta hensyn til slike forskjeller vil våre estimater være gyldige anslag på kasuale effekter. Vi vil diskutere dette videre i omtalen av resultatene.

#### Målgruppen

I Tabell 4.5 presenterer vi effektene av tiltaket på nasjonale prøver i 9. klasse. Dette er hovedgrunnet for å vurdere effekter av tiltaket på ungdomsskolen.

**Tabell 5.5 Effektestimater målgruppen i grunnskolen: Forskjeller i NP9 mellom målgruppeelever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19<sup>1</sup>**

	(1)	(2)	(3)	(4)	(5)
	Resultat NP9	Resultat NP9	Mestringsnivå NP9=1	Mestringsnivå NP9<=2	Har resultat NP9
Tiltak	0.498* (0.261)	0.363 (0.245)	-0.022* (0.011)	-0.019 (0.019)	0.012 (0.011)
Tiltak gitt NP8		0.222 (0.201)	-0.017* (0.010)	-0.009 (0.016)	0.011 (0.011)
Tiltak gitt NP5		0.418* (0.219)	-0.024** (0.011)	-0.023 (0.019)	0.014 (0.011)
Indeks fam.bakgr		x	x	x	X
Ant observasjoner	1775	1775	1775	1775	1977
Snitt avhengig variabel	43.8	43.8	1.103	0.540	0.898

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (1) med forskjellige uavhengige variable (y) i kolonner, og forskjellige sett av kontrollvariable (x) i rader. Alle estimerer kontrollerer for årskull og strata (gruppe i randomiseringen). Kontroll for indeks av familiebakgrunnsvariable i alle kolonner dette er angitt. Kontroll for (tredjegrads polynom) i NP8 eller NP5 der dette er angitt. Klyngerobuste beregnede standardfeil i parentes under hvert effektestimat. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

Tabellen inneholder forskjellige rader og kolonner, som viser resultatene fra analyser der vi på forskjellige måter tar hensyn til at elvenes forutsetninger ikke er perfekt balansert på tvers av tiltak- og kontrollgruppen. Dvs., hver celle i tabellen representerer en separat spesifisering. Nest nederste rad rapporterer gjennomsnittsverdien i avhengig variabel.

Første kolonne/rad i tabellen viser forskjell i gjennomsnittlig NP9 i målgruppen uten noen form for kontroll for forutsetninger. Vi ser at målgruppeelevene i tiltaksskolene i snitt presterer 0,50 skalapoeng bedre enn målgruppeelevene i kontrollskolene. Denne koeffisienten er også signifikant på 10 prosents nivå. Når vi kontrollerer for familiebakgrunn (andre kolonne/første rad) reduseres forskjellen, men elvene i tiltaksskolene presterer fortsatt bedre enn elevene i kontrollskolene. Når vi tar hensyn forskjellen i familiebakgrunn gjenstår det en forskjell på 0,36 skalapoeng. Denne forskjellen er ikke statistisk signifikant (med en t-verdi på 1,5).

I andre kolonne/rad kontrollerer vi ytterligere for NP8. Forskjellen i NP9 reduseres da ytterligere, til 0,22 skalapoeng. Som bemerket over er dette vårt hovedresultat, og fremste mål på effekten av tiltaket. Dette estimatet er ikke statistisk signifikant, og gir dermed ikke noe sterkt grunnlag for å konkludere med at tiltaket har hatt en effekt på faglige ferdigheter. Et 95 prosent konfidensintervall er (-0,17, 0,62). For å oppsummere er det beste anslaget, gitt NP8 som mål på elevenes forutsetninger, en effekt på 0,22 skalapoeng. Det er usannsynlig med en effekt lavere enn -0,17 skalapoeng, eller høyere enn 0,62 skalapoeng, og ikke grunnlag for noen klar konklusjon om noen effekt.

Jf. diskusjonen av NP8 og NP5 som kontrollvariable er det imidlertid også mulig å argumentere for at NP5 er en tryggere, om enn svakere, måte å ta hensyn til forskjeller i elevenes forutsetninger. Når vi kontrollerer for NP5 i stedet for NP8 finner vi en forskjell på 0,42 skalapoeng. Denne forskjellen er signifikant på 10 prosent nivå, men ikke på 5 prosent nivå. Dvs, forskjellen er så stor at vi langt på vei kan utelukke at den kun skyldes tilfeldige forskjeller/variasjon, og tyder på en effekt av tiltaket. Resultatet er ikke signifikant forskjellig fra resultatet med kontroll for NP8, dvs. forskjellene kan skyldes tilfeldige resultatforskjeller mellom NP5 og NP8. Det høyere estimatet med kontroll for NP5 kan likevel tyde på estimatet der vi kontrollerer for NP8 er (for) lavt.

Kolonnene (3) og (4) viser forskjeller i andel elever på henholdsvis mestringsnivå 1 (kolonne (3)) og mestringsnivå 1 eller 2 (kolonne (4)) på NP9. I alle spesifiseringene er det færre målgruppeelever i tiltaksskolene enn i kontrollskolene som

presterer lavt. Andelen som presterer på mestringsnivå 1 reduseres med omtrent 2 prosentpoeng (denne forskjellen kan sammenlignes med en gjennomsnittlig andel på 10 prosent for elevene i målgruppen, jf. nest nederste rad som rapporterer snittet i avhengig variabel) og andelen på mestringsnivå 1 eller 2 reduseres med 1-2 prosentpoeng (gjennomsnittlig andel 54 prosent). Forskjellen i andelen på mestringsnivå 1 er lik på tvers av spesifikasjoner, og lite følsom til hvordan vi kontrollerer for ubalansen på tvers av tiltak- og kontrollgruppen. Estimaten er også statistisk signifikant, hovedsakelig på 10 prosents nivå.

Forskjeller i snittresultater eller andel på lave nivåer kan i prinsippet skyldes ulik prøvedeltagelse i tiltak- og kontrollskolene. Gjennomføring av tiltak kan påvirke hvordan prøver gjennomføres, og det trenger ikke være tilfeldig hvilke elever som gjennomfører prøver eller ikke. Elever med svake forutsetninger er overrepresentert ved manglende prøvedata. Den siste kolonnen viser forskjellen i prøvegjennomføring. Omtrent 90 prosent av elevene har registrert prøveresultat på NP9.<sup>21</sup> Tabellen viser at andelen med resultat fra NP9 er 1 prosentpoeng høyere i tiltaksskolene, men at denne forskjellen ikke er større enn at den kan skyldes tilfeldig variasjon.

Oppsummert scorer elever i målgruppen på tiltaksskoler høyere på NP9 enn tilsvarende elever på kontrollskoler. Forskjellen på omtrent 0,5 skalapoeng kan imidlertid delvis forklares med forskjeller i elevenes forutsetninger, og når vi på forskjellige måter tar hensyn til forskjellene i forutsetninger reduseres forskjellen. Det er likevel en forskjell i favør tiltaksskolene på omtrent 0,2 til 0,4 skalapoeng. Denne forskjellen er i de fleste analysene, inkludert hovedresultatene, for liten til at vi kan konkludere at den skyldes tiltaket, og ikke tilfeldige forskjeller. Det er dermed indikasjoner på positiv effekt på gjennomsnittlige resultater, men ikke grunnlag for å trekke en sterk konklusjon.

Når vi ser på andelen elever på lave mestringsnivåer er denne redusert med omtrent 2 prosentpoeng. Forskjellen i elever på mestringsnivå 1 eller 2 er ikke større enn at den kan skyldes tilfeldige forskjeller. Forskjellen i andelen på mestringsnivå 1 er ikke signifikant på 5 prosent nivå, men er signifikant på 10 prosent nivå, og det er dermed indikasjoner på at tiltaket har redusert andelen elever på mestringsnivå 1.

### **Heterogene effekter – målgruppen**

Mange tiltak har sterkere effekter for noen grupper av elever enn andre, eller lykkes bedre i noen grupper av skoler enn andre. I tabell 5.6 ser vi på forskjeller i effektestimater knyttet til en del kjennetegn, både ved eleven, men også ved skolen. På individnivå beregner vi separate effekter for grupper definert ved følgende kjennetegn: kjønn, mestringsnivå, foreldres utdanning (minst en forelder med høyere utdanning), innvandrerbakgrunn og årskull, samt om elevene skal ha fått undervisning i en liten gruppe. På skoler med få elever i målgruppen vil alle målgruppeelever få undervisning i små grupper. På skoler med mange målgruppeelever vil bare de svakest presterende få undervisning i små grupper, jf beskrivelsen i kapittel 4. På skolenivå grupperer vi etter antall elever i målgruppen. Vi kontrollerer for variabelen vi grupperer etter samt familiebakgrunnsindeks i alle spesifikasjoner. Hvorvidt vi kontrollerer for NP8 og NP5 er angitt i de to nederste radene i tabellen.

I den øverste delen av tabellen rapporterer vi punktestimaten når vi kjører analyser for ulike grupper av elever.<sup>22</sup> Gjennomgående kan vi ikke konkludere klart med at

<sup>21</sup> Dette inkluderer ikke elever som har avlagt NP9 utenfor Oslo-skolen, og som derfor foreløpig ikke har resultater for, men som finnes i nasjonale NP9 data.

<sup>22</sup> Vi gjør en felles analyse, der vi lar effekten av tiltaket variere mellom grupper av elever eller skoler. Som beskrevet tar vi også hensyn til snittforskjeller i resultater mellom gruppene. Vi har ikke interaksjoner med gruppene av elever/skoler for øvrige kontrollvariable.

det er forskjeller mellom beregnede effekter for forskjellige grupper. I tillegg viser også balanseringstester at de ulike gruppene ikke er perfekt balansert på tvers av tiltak- og kontrollskoler. Dette er en tilsvarende forskjell som vi har sett i balanseringstester av målgruppen under ett, men for noen grupperinger av elever og skoler er forskjellen i større grad til stede for en av gruppene. Som i hovedanalysene av effekter gir det ytterligere usikkerhet i tolkning av estimatene, ettersom dette avhenger av at vi lykkes i å kontrollere for forskjeller i forutsetninger.

Vi er derfor forsiktige i tolkningen av forskjellene i effektestimater. Med disse forbeholdene ser vi imidlertid noen mønstre. Gitt forskjeller i NP8 ser det ut til at gutter i målgruppen i tiltaksskolene gjør det bedre enn tilsvarende gutter i kontrollskolene. For jenter er det liten forskjell. Det er også liten forskjell i beregnede effekter for elever på mestringsnivå 1 og 2 fra NP8 (men det er vesentlig færre elever, og dermed mindre presise resultater). Det er større beregnet effekt for elever med høyere utdannede foreldre, og lite forskjell knyttet til foreldrenes innvandringsbakgrunn. Effekten ser ut til å være knyttet særlig til elevkullet 2017-18, mens det er lite tegn til effekt for 2018-19.

Vi finner klart større effekt for målgruppeelever som (vi forventer) får undervisning i smågrupper. Denne forskjellen er imidlertid noe vanskelig å tolke, ettersom tilbud i små/store grupper samvarierer med andre forhold. For eksempel finnes store grupper kun på skoler med mange målgruppeelever, som også viser seg å være assosiert med høyere beregnet effekt. Skoler med få målgruppeelever har også relativt lav andel målgruppeelever og relativt høy gjennomsnittlig skår på NP8.

Oppsummert er det tendenser til høyere effektestimater for noen grupper elever og skoler. Estimaten er imidlertid moderate også for disse gruppene, inntil omtrent 0,5 skalapoeng når vi tar hensyn til NP8, og det er ikke signifikante forskjeller i effekter. Vi er derfor forsiktige med å trekke konklusjoner om tiltaket har hatt effekt for noen elever/skoler og ikke for andre.

**Tabell 5.6 Heterogene effekter på NP9 av tiltaket for målgruppeelever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19<sup>1</sup>**

	(1)	(2)	(3)
<b>ELEVKJENNETEGN</b>			
Kjønn			
-Jente	0.290 (0.286)	0.082 (0.253)	0.295 (0.279)
-Gutt	0.467 (0.467)	0.422 (0.447)	0.594 (0.436)
Mestring NP8			
-Nivå 1	0.315 (0.766)	0.281 (0.718)	0.241 (0.695)
-Nivå 2	0.190 (0.251)	0.201 (0.226)	0.280 (0.248)
Foreldres utdanning			
-Lav	0.131 (0.312)	-0.007 (0.289)	0.217 (0.285)
-Høy	0.802** (0.367)	0.657* (0.330)	0.799** (0.364)
Innvandrerbakgrunn			
-Norskfødte foreldre	0.490 (0.390)	0.187 (0.316)	0.527 (0.374)
-Utenlandskfødte foreldre	0.269 (0.294)	0.248 (0.280)	0.338 (0.259)
År (for NP8)			
-2017	0.584 (0.493)	0.441 (0.384)	0.681 (0.475)
-2018	0.164 (0.481)	0.026 (0.404)	0.182 (0.436)
Gruppestørrelse (forventet)			
-Stor	-0.045 (0.390)	-0.012 (0.306)	0.047 (0.353)
-Liten	1.208** (0.337)	0.366 (0.267)	1.122** (0.319)
<b>SKOLEKJENNETEGN</b>			
Antall elever i målgruppen			
-Få	1.015** (0.322)	0.560** (0.241)	1.119** (0.321)
-Mange	-0.473 (0.355)	-0.201 (0.418)	-0.487* (0.289)
NP8	Nei	Ja	Nei
NP5	Nei	Nei	Ja

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (1) med NP9 som uavhengig variable (y), forskjellige sett av kontrollvariable (x) i kolonner, og forskjellige grupperinger av skoler/elever i rader. Alle estimater kontrollerer for årskull, strata (gruppe i randomiseringen) og indeks for familiebakgrunnsvariabel. Kontroll for (tredjegrads polynom) i NP8 eller NP5 der dette er angitt. Klyngebuste beregnede standardfeil i parentes under hvert effektestimater. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

### Øvrig elever

I forrige avsnitt studerte vi effekter på målgruppen. Tiltaket kan også ha påvirket andre elever, enten fordi kursingen påvirker undervisningen til disse elevene, eller fordi det å gi noen elever undervisning i smågrupper påvirker undervisningen til de gjenværende. I dette avsnittet studerer vi effekter på elever i analysedataene som er utenfor målgruppen. De utgjør omlag 80 prosent av elevpopulasjonen på 8. trinn i skoleårene 2017/18 og 2018/19 og presterer på mestringsnivå tre til fem på NP8.

Tabell 5.7 viser balansering av øvrige elever mellom tiltak- og kontrollskolene. Vi ser at det er en enda sterkere ubalanse i elevenes familiebakgrunn enn hva vi så for målgruppeelevene. Når vi tar hensyn til denne forskjellen er det ikke ytterligere forskjeller i resultater fra NP5 eller NP8.

**Tabell 5.7 Balansering av øvrige elever i tiltak- og kontrollskoler for grunnskolen, årene 2017/18 og 2018/19<sup>1</sup>**

	(1)	(2)	(3)	(4)
	Indeks familie bakgrunn	Resultat NP8	Resultat NP5	Har resultat NP5
Tiltak	0.939** (0.354)	-0.004 (0.003)	-0.067 (0.268)	-0.026 (0.233)
Indeks for familiebakgrunn		X	X	X
Ant observasjoner	7952	7952	7683	7952
Snitt avhengig variabel	57.016	0.966	56.103	57.222

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (1) med forskjellige uavhengige variable (y) i kolonner. Alle estimater kontrollerer for årskull og strata (gruppe i randomiseringen). Kontroll for indeks av familiebakgrunnsvariable i alle kolonner dette er angitt. Klynge robuste beregnede standardfeil i parentes under hvert effektestimater. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

Også når vi studerer effektresultatene for øvrige elever må vi ta hensyn til at elevene i tiltaksskolene har en noe mer fordelaktig bakgrunn (målt ved familiebakgrunnsindeksen) sammenliknet med elevene i kontrollgruppen.

Resultatene er rapportert i Tabell 5.8. Tabellen er satt opp på samme måte som Tabell 5.5 som viser effektestimater for målgruppen. I kolonne (1) ser vi at øvrige elever i tiltaksskoler presterer bedre på nasjonale prøver i 9. klasse enn øvrige elever i kontrollskoler når vi ikke betinger på den initielle forskjellen i de to gruppene. Når vi kontrollerer for ulike forutsetninger (familieindeks, NP8 eller NP5) reduseres forskjellen i NP9 betraktelig (se kolonne (2)), og blir svært nær null. Estimaterne er langt unna å være statistisk signifikante, og det er heller ikke store forskjeller mellom estimatene avhengig av hvordan vi tar hensyn til forskjeller i elevenes bakgrunn. Konfidensintervallet for hovedresultatet, der vi kontrollerer for NP8, er nesten symmetrisk rundt null: (-0,27, 0,25). Det er altså ikke tegn til effekter, og vi kan langt på vei utelukke effekter større enn estimatene for målgruppen.

I kolonne (3) og (4) ser vi på om andelen lavt-presterende elever er endret. Det er ingen tendenser til endring i andelen som presterer på det laveste mestringsnivået, nivå 1 (kolonne (3)). Derimot ser vi at andelen som presterer på lavest og nest lavest nivå, nivå 1 eller 2 er høyere i tiltaksskolene enn i kontrollskolene (kolonne (4)). En forklaring på dette kan være at vi finner positive effekter av tiltaket på prøvedeltagelse (kolonne (5)). Elever som ikke deltar ved enkelte prøver har i snitt svakere forutsetninger enn gjennomsnittet. Når flere av disse elevene gjennomfører prøven vil snittet forventningsmessig trekkes ned samt at antallet med svake resultat vil økes.

**Tabell 5.8 Effektestimater øvrige elever i grunnskolen: Forskjeller i NP9 mellom øvrig elever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19<sup>1</sup>**

	(1)	(2)	(3)	(4)	(5)
	Resultat NP9	Resultat NP9	Mestrings- nivå NP9=1	Mestrings- nivå NP9<=2	Har resultat NP9
Tiltak	0.431 (0.351)	-0.064 (0.219)	0.001 (0.001)	0.006** (0.003)	0.010** (0.005)
Tiltak gitt NP8		-0.011 (0.131)	0.001 (0.001)	0.006* (0.003)	0.010** (0.005)
Tiltak gitt NP5		0.009 (0.177)	0.001 (0.001)	0.005 (0.003)	0.010** (0.004)
Indeks fam.bakgr		x	x	x	x
Ant observasjoner	7596	7596	7598	7598	7952
Snitt avhengig variabel	59.941	59.941	0.001	0.025	0.955

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (1) med forskjellige uavhengige variabel (y) i kolonner, og forskjellige sett av kontrollvariable (x) i rader. Alle estimater kontrollerer for årskull og strata (gruppe i randomiseringen). Kontroll for indeks av familiebakgrunnsvariable i alle kolonner dette er angitt. Kontroll for (tredjegrads polynom) i NP8 eller NP5 der dette er angitt. Klynge robuste beregnede standardfeil i parentes under hvert effektestimater. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

### Heterogene effekter – øvrige elever

I Tabell 5.9 ser vi på heterogene effekter for øvrige elever. Tabellen er satt opp på akkurat samme måte som Tabell 5.6 som så på heterogene effekter for målgruppen. Det første som er verdt å merke seg er at veldig få av punkttestimatene er signifikante. I tillegg viser balanseringstester at elevgruppene vi studerer ikke er balansert på tvers av tiltak- og kontrollskolene. Dette er viktig å huske på når man tolker resultatene.

**Tabell 5.9 Heterogene effekter på NP9 av tiltaket for øvrig elever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19<sup>1</sup>**

	(1)	(2)	(3)
<b>ELEVKJENNETEGN</b>			
Kjønn			
-Jente	-0.196 (0.295)	-0.142 (0.141)	-0.221 (0.227)
-Gutt	0.061 (0.312)	0.114 (0.213)	0.229 (0.217)
Mestring NP8			
-Nivå 3	-0.130 (0.183)	-0.096 (0.174)	-0.029 (0.169)
-Nivå 4	0.149 (0.186)	0.287* (0.162)	0.152 (0.198)
-Nivå 5	-0.169 (0.276)	-0.261 (0.230)	-0.173 (0.279)
Foreldres utdanning			
-Lav	-0.445 (0.350)	-0.190 (0.221)	0.093 (0.291)
-Høy	0.082 (0.271)	0.058 (0.152)	-0.023 (0.204)
Innvandrerbakgrunn			
-Norskfødte foreldre	0.054 (0.269)	0.092 (0.152)	0.046 (0.191)
-Utenlandskfødte foreldre	-0.418 (0.381)	-0.318 (0.193)	-0.102 (0.381)
År (for NP8)			
-2017	0.431 (0.315)	0.275 (0.258)	0.433 (0.283)
-2018	-0.562 (0.345)	-0.299 (0.194)	-0.419 (0.314)
<b>SKOLEKJENNETEGN</b>			
Skolestørrelse			
-Liten	-0.152 (0.344)	0.312 (0.230)	0.489 (0.449)
-Stor	-0.027 (0.289)	-0.150 (0.167)	-0.195 (0.180)
Antall elever i målgruppen			
-Få	0.104 (0.232)	0.180 (0.150)	0.201 (0.215)
-Mange	-0.535 (0.471)	-0.595** (0.250)	-0.601* (0.354)
Snitt NP8			
-Lav	-0.708** (0.306)	-0.283 (0.274)	-0.430 (0.267)
-Høy	0.031 (0.232)	0.085 (0.154)	0.067 (0.195)
NP8	Nei	Ja	Nei
NP5	Nei	Nei	Ja

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (1) med NP9 som uavhengig variable ( $y$ ), forskjellige sett av kontrollvariable ( $x$ ) i kolonner, og forskjellige grupperinger av skoler/elever i rader. Alle estimater kontrollerer for årskull, strata (gruppe i randomiseringen) og indeks for familiebakgrunnsvariabelt. Kontroll for (tredjegrads polynom) i NP8 eller NP5 der dette er angitt. Klyngerobuste beregnede standardfeil i parentes under hvert effekttestimat. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

Også for øvrige elever ser tiltaket ut til å ha en større virkning for gutter enn for jenter med tanke på resultater på NP9. Angående mestringsnivå er det tendenser til at elever på mestringsnivå 4 tjener mer på å tilhøre en tiltaksskole (selv om de ikke selv mottar tilrettelagt undervisning) enn elever på andre mestringsnivå. Derimot når det gjelder foreldres utdanning- og innvandringsbakgrunn er det få forskjeller på tvers grupper. Det er heller ikke mange tydelige forskjeller på skolenivå. Det eneste unntaket er at øvrige elever på tiltaksskoler med mange elever i målgruppen

muligens presterer noe svakere sammenliknet med tilsvarende elever på kontroll-skoler.

For øvrig, som for målgruppen er den generelle oppsummeringen at vi ikke finner vi tydelige, konsistente eller signifikante forskjeller og vi er dermed forsiktige med å tolke de forskjellene vi finner.

### **5.3.1. Effekter for alle elever (målgruppeelever + øvrig elever)**

Vi har også studert effekter for alle elever under ett (målgruppe + øvrige elever). Resultatene for alle elever gir imidlertid lite innsikt utover hva vi får av resultatene for målgruppen og øvrige elever separat, og vi viser derfor ikke disse i rapporten. Kort oppsummert ligner resultatene for alle elever på resultatene for øvrige elever i dette avsnittet. Dette er som forventet, ettersom et mindretall av elevene er i målgruppen. Unntaket er når vi ser på grupper som er dominert av målgruppeelever, for eksempel elever med svake resultater på NP8. I disse tilfellene ligner resultatene på resultatene til målgruppeelevene.

## **5.4. Oppsummering og kost/nytte – grunnskole**

I dette avsnittet oppsummerer vi først effektestimater for grunnskolen. Vi forsøker å sammenligne disse med enkelte øvrige effektestimater. Vi oppsummerer deretter kostnadene ved gjennomføring av tiltaket, samt diskuterer hvordan kostnader ved framtidig gjennomføring i Oslo-skolen eller andre steder vil kunne være annerledes. Basert på disse tallene gjør vi anslag på forholdet mellom kostnader og nytte.

### **Oppsummering effektestimater**

Vi finner ikke noe klart grunnlag for å konkludere med at tiltaket har ført til økte gjennomsnittlige faglige ferdigheter, verken i målgruppen eller blant øvrige elever. I målgruppen finner vi at gjennomsnittlig resultat på NP9 er omtrent 0,2 skalapoeng høyere i tiltaksskolene i vår hovedanalyse, der vi tar hensyn til forskjeller i NP8. Dette er en liten forskjell, klart mindre enn hva som ble vurdert som realistisk ifm utformingen av tiltaket basert på tidligere forskning og erfaring (1 skalapoeng). Forskjellen er heller ikke statistisk signifikant, dvs., den er så liten at den kan skyldes tilfeldigheter, og gir ikke noe grunnlag for en klar konklusjon om at det i det hele tatt finnes noen effekt.

Vi finner imidlertid større forskjeller når vi tar hensyn til elevenes forutsetninger på andre måter. Når vi bruker NP5 som mål på elevenes forutsetninger finner vi en forskjell på omtrent 0,4 skalapoeng, som er signifikant på 10 prosent nivå. Vi finner også at en litt høyere andel gjennomfører NP9 i tiltaksskolene, noe som normalt vil bidra til å redusere snittresultatet blant de som har resultat fra prøven. I sum mener vi det er sannsynlig, om enn ikke sikkert, at snittresultatet på NP9 til elevene i målgruppen er bedret med 0,2-0,4 skalapoeng.

Vi finner at andelen elever på laveste mestringsnivå i målgruppen er redusert med omtrent 2 prosentpoeng. Dette resultatet er statistisk signifikant på 10 prosent nivå, men ikke 5 prosent. Vi har dermed noe grunnlag for å konkludere, selv om det er svakere enn det ville vært om effekten var signifikant på 5 prosent nivå.

For øvrige elever finner vi effektestimater som er svært nær null for gjennomsnittlige resultater på NP9. Vi finner at andelen på lave mestringsnivåer øker litt, men forskjellen er mindre enn forskjellen i andel elever som gjennomfører prøven. I sum mener vi ikke det er tegn til effekter på øvrige elever, verken positive eller negative.



I Kirkebøen m.fl. (2018) fant vi ingen tegn til effekt første året, verken av fullt tiltak (kursing og smågrupper) eller av bare smågrupper. Ettersom disse resultatene baserte seg på bare ett årskull og henholdsvis 8 og 16 skoler var disse resultatene mindre presise og mer usikre enn resultatene i denne rapporten. Det gjelder særlig resultatene for fullt tiltak, der vi hadde store utfordringer med balansering av tiltaks- og kontrollgrupper.

### **Økonomisk nytte**

Gitt et anslag på effekter kan vi vurdere nytten av tiltaket. Vi finner ikke klare effekter av tiltaket, men finner likevel tegn til en effekt på omtrent 0,2-0,4 skalapoeng på NP9 i favør tiltaksskolene blant målgruppeelevene i de to siste årene av tiltaket. Selv om denne forskjellen er for liten til at vi kan konkludere med at den ikke skyldes tilfeldigheter er dette det beste anslaget på en effekt på NP9.

Vi har ikke noen direkte verdivurdering av evt. effekter på resultater på NP9. Basert på sammenhengen mellom resultat på NP9 og fullføring av VGO finner vi at en effekt på 0,2 skalapoeng på NP9 svarer til en forventet økning i fullføring av VGO på 0,25 prosentpoeng. For omtrent 1000 målgruppeelever totalt i 2017-18 og 2018-19 svarer dette til 2-3 elever. En effekt på 0,4 skalapoeng svarer tilsvarende til omtrent 5 ekstra elever som fullfører og består VGO.

Falch m.fl. (2009) vurderer gevinster for samfunnet av at elever fullfører videregående opplæring, og kommer fram til en anslått gevinst på 900 000 kr per ekstra elev. En stor del av denne gevinst er økt arbeidsinntekt, og dersom vi justerer denne for lønnsvekst vil et oppdatert anslag være omtrent 1,25 millioner kr. Gevinsten av at 2 (5) elever fullfører er dermed omtrent 2,5 (6,25) millioner kroner.

### **Kostnader**

Gjennomføringen av tiltaket har samlet hatt en kostnad på omtrent 20,5 millioner kroner (ikke inkludert forskningen). Omtrent 9,5 millioner har gått til finansiering av smågruppeundervisning, alt dette har gått til tiltak på ungdomstrinnet. De øvrige 11 millionene har gått til administrasjon, utvikling av kursopplegg og materiell og gjennomføring av kurs (inkl. finansiering av vikarer). Dette fordeler seg på ungdomstrinn og videregående. Et anslag på kostnaden i grunnskolen er dermed 14 millioner kroner.

Tiltaksskolene i grunnskolen har omtrent 2500 elever/år, hvorav omtrent 500 er i målgruppen. I løpet av tre år har det dermed vært omtrent 1500 målgruppeelever i tiltaksskolene. Hele tiltaket i grunnskolen har dermed kostet i underkant av 5 millioner per år, eller 10 000 kr per elev i målgruppen i tiltaksskolene. Smågruppeundervisning har utgjort i overkant av 6000 kr av kostnaden per elev. Ikke alle elevene i målgruppen har fått undervisning i smågrupper. Kostnaden per smågruppe har vært omtrent 60 000 kr. Første år, med åtte elever per smågruppe var kostnaden omtrent per elev som fikk smågruppeundervisning 7500 kr ved fulle grupper. Med grupper på seks elever økte kostnaden per elev i gruppen til 10 000 kr.

Dersom et lignende tiltak skal videreføres eller gjennomføres av en annen skoleeier er det rimelig å vente at kostnaden til smågruppeundervisning vil være tilsvarende. Øvrige kostnader vil kunne være annerledes. Større gjenbruk av kurs- og undervisningsmateriell kan redusere kostnadene knyttet til dette. Prosjektet har også fått ekstra administrative ressurser for å samordne med og bidra med data til forskningen. Administrasjonskostnader vil også avhenge av organisering og omfang på prosjektet.

### **Kost/nytte og sammenligning med øvrige tiltak**

Basert på tallene over kommer vi til et anslag på samlet økonomisk nytte på 2,5-6,25 millioner kroner for tiltaket i grunnskolen. Samlet kostnad har vært 14 millioner kr. Tiltaket har, gitt antagelsene gjort her, altså sannsynligvis ikke vært samfunnsøkonomisk lønnsomt. Den økonomiske gevinsten for samfunnet skal svare til 14 millioner kr, kostnaden av tiltaket, må antall som fullfører videregående opplæring øke med omtrent 11 personer. Med 500 tiltakselever per år i tre år krever det en øking i fullføring av VGO på omtrent 0,75 prosentpoeng. Dette svarer igjen til en effekt på NP9 på 0,6 skalapoeng, som er omtrent øvre grense av konfidensintervallet for hovedmålet vårt på effekter, og godt innenfor konfidensintervallene for estimatene der vi måler elevenes forutsetninger med andre mål enn NP8. Det vil si at selv om våre effektestimater er klart mindre enn forventet, og vi ikke kan konkludere klart med at tiltaket har hatt effekt på gjennomsnittlige resultater, kan vi heller ikke utelukke at tiltaket de siste to årene har vært omtrent i samfunnsøkonomisk balanse.

Vi kan sammenligne resultatene fra denne effektanalysen med andre nylige norske satsinger på ungdomstrinnet. Dette omfatter Overgangsprosjektet i Ny GIV og ekstra lærere på ungdomstrinnet. Evalueringene av begge satsingene konkluderer med at det ikke er tegn til effekter, og utelukker store effekter (Huitfeldt m.fl., 2018; Kirkebøen m.fl., 2017). Det er heller ingen tegn til at redusert klassestørrelse på ungdomstrinnet påvirker elevprestasjoner positivt (Leuven m.fl., 2008; Leuven og Løkken, 2020). I tråd med dette fant vi ikke effekter i skolene som kun fikk ressurser til smågrupper, men ikke kursing første år (Kirkebøen m.fl., 2018). Dette kan tyde på at det er vanskelig å oppnå store effekter på ungdomstrinnet, i tråd med forskningen som viser at det er viktig med tidlig innsats.

Samtidig har vi vist til Cook m.fl. (2015) og Cortes m.fl. (2014) som finner effekter av satsinger i ungdomsskole/videregående. Fryer (2017) konkluderer i en omfattende metaanalyse av eksperimenter med tiltak for bedre skolerresultater at forskjellige tiltak bedrer elevenes resultater i matematikk uavhengig av alder (i motsetning til for lesing, der tidlige tiltak er mer effektive). Fryer konkluderer også at smågrupper («high-dosage tutoring») er en av få grupper tiltak som gjennomgående bedrer elevenes resultater.

For at et tiltak skal klassifiseres som smågruppeundervisning av Fryer må det være maksimalt seks elever per gruppe og et omfang tilsvarende 50 timer i løpet av 32 uker. Vårt tiltak i grunnskolen har omtrent denne gruppestørrelsen og omfanget. Cook m.fl. (2015) bruker svært små grupper i matematikkundervisningen, to elever per veileder. Vi har ikke tilstrekkelig mange elever i grupper mindre enn seks elever til at vi kan vurdere om dette gir større effekt, men det er mulig.

En viktig forskjell mellom vårt tiltak og Cortes m.fl. (2014) er at tiltaksundervisning Cortes m.fl. studerer kommer i tillegg til, ikke i stedet for, ordinær matematikkundervisning (den erstatter i stedet andre fag). Dette betyr at samlet omfang på matematikkundervisning øker, og kan også unngå problemene noen skoler og elever har opplevd med tilbakeføring til ordinær undervisning etter tiltaket. En slik omfordeling av timer mellom fag ble imidlertid ikke vurdert som aktuelt i vårt tiltak.

## 6. Effektestimater – videregående

I dette kapittelet presenterer vi de kvantitative analysene for videregående. Vi starter med å presentere datagrunnlaget. Deretter undersøker vi hvorvidt tiltak- og kontrollgruppen er sammenliknbare, før vi til slutt presenterer resultatene samlet for alle årene, 2016/17 - 2018/19. I motsetning til for grunnskolen er det ikke vesentlige forskjeller i organiseringen av tiltaket underveis, så vi vil derfor se alle årene under ett for å få mest mulig presise resultater. Vi avslutter kapitlet med å oppsummere resultatene, samt å vurdere kostnader av tiltaket og anslag på framtidige gevinster av evt. effekter.

### 6.1. Datagrunnlag

Som for grunnskolen bruker vi forskjellige datakilder, både data fra nasjonale registre og fra Oslo kommunes administrative data. De viktigste datakildene er resultater fra tester før og etter tiltaket. I tillegg har vi data for elevenes tidligere resultater fra grunnskolen, deres familiebakgrunn og gjennomstrømning i videregående opplæring.

Resultater for familiebakgrunn kommer fra NUDB, som tilsvarende data brukt i analysene av grunnskolen. NUDB inneholder også karakterer ved avsluttet grunnskole, samt data for gjennomstrømning i videregående opplæring. På lengre sikt gjør dette oss i stand til å studere hvorvidt elevene fullfører og består videregående opplæring. Foreløpig har vi bare mulighet til å studere hvorvidt elevene fra de to første tiltaksårene (2016/17 og 2017/18) har fullført og bestått første året på videregående (Vg1).

For grunnskolen bruker vi testresultater fra NP8 og NP9 for å måle ferdigheter før og etter tiltaket. I videregående måler vi ferdigheter før tiltaket med læringsstøttende prøve i regning (tidligere kalt kartleggingsprøve), som gjennomføres tidlig i Vg1. Denne er utviklet av UDIR, og har likheter med nasjonale prøver. Blant annet plasseres elever på mestringsnivåer, som på nasjonale prøver. Disse mestringsnivåene bruker vi til å finne målgruppeelevene. I motsetning til nasjonale prøver i grunnskolen er dette imidlertid prøver skolene står fritt til å gjennomføre eller ikke, og det er større variasjon i gjennomføringen. For å måle ferdighetene til elevene etter at tiltaket er gjennomført benytter vi Oslo-prøven i matematikk, som gjennomføres sent på våren i Vg1. Denne er utviklet av Oslo kommune, og gjennomføres bare i Oslo. Skolene har betydelig frihet også angående gjennomføring av Oslo-prøven.

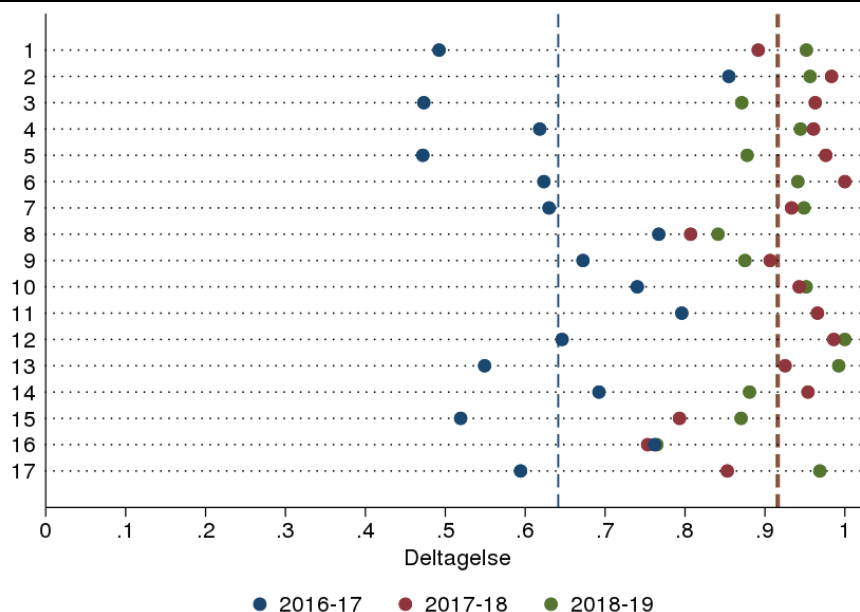
Figur 6.1 og 6.2 viser andel av aktuelle elever som gjennomfører henholdsvis læringsstøttende prøve og Oslo-prøve ved hver skole (anonymisert) hvert av tiltaksårene. Punktene gir gjennomføringsandeler for hver skole og år, mens de stiplede vertikale linjene viser gjennomsnitt for hvert år. Gjennomføringsandelen for Oslo-prøven er i snitt mye høyere enn for læringsstøttende prøve. Og variasjonen på tvers av skoler i gjennomføringsandelen er mye mindre for Oslo-prøven enn for læringsstøttende prøve. Høsten 2016 varierte andelen som gjennomførte læringsstøttende prøve fra nesten 90 prosent til under 50 prosent, mens snittet lå på 64 prosent.

I forkant av prøvene i 2017 og 2018 jobbet vi sammen med Utdanningsetaten i Oslo kommune for å motivere skolene til å gjennomføre prøver, ved å understreke prøvenes betydning for forskningen. Det var uklart hvordan prøvene i utgangspunktet ble presentert til elevene. Oslo kommune bidro derfor med manus skolene kunne bruke ved introduksjon av prøvene. For å legge til rette for en mest mulig lik gjennomføring i tiltak- og kontrollskoler ble motivasjon for og tilrettelegging for læringsstøttende prøver og Oslo-prøver i størst mulig grad kommunisert tilsvarende

til tiltak- og kontrollskoler og uavhengig av tiltaket. Fra figurene ser vi at 92-93 prosent av elevene gjennomførte læringsstøttende prøve i 2017 og 2018, og at det var mindre forskjeller mellom skolene.

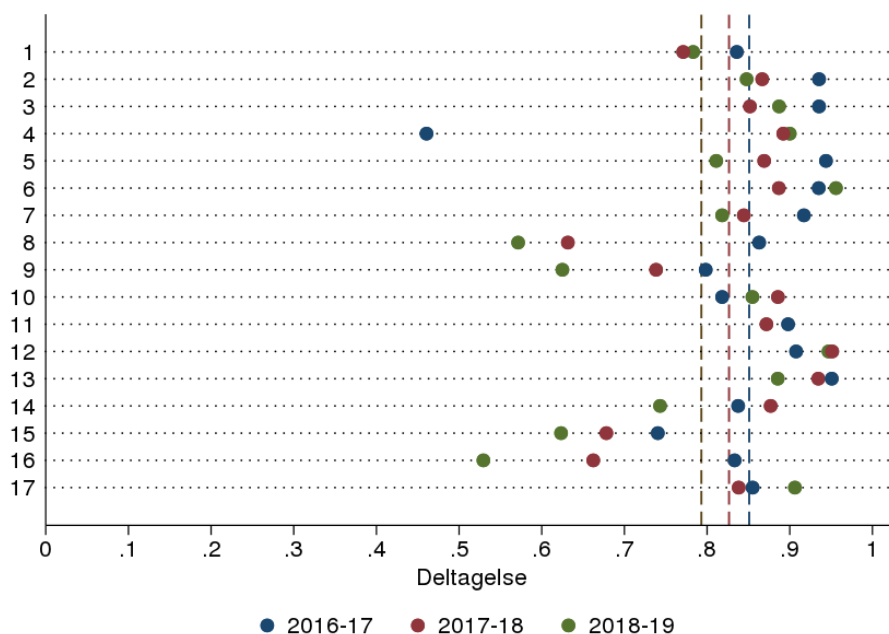
80-85 prosent av elevene gjennomførte Oslo-prøven, men for enkelte skoler og år er andelen vesentlig lavere.<sup>23</sup> I de videre analysene vil vi undersøke om det er forskjeller mellom tiltaks- og kontrollskolene i gjennomføring av prøver.

**Figur 6.1 Gjennomføring av læringsstøttende prøve på skolenivå (andel som har gjennomført prøve)**



Kilde: Statistisk sentralbyrå, basert på data fra Oslo kommune.

**Figur 6.2 Gjennomføring av Oslo-prøve på skolenivå (andel som har gjennomført prøve)**



Kilde: Statistisk sentralbyrå, basert på data fra Oslo kommune.

<sup>23</sup> Merk at vi fordeler elever til skolen de var registrert ved starten av skoleåret, ettersom det er på dette tidspunktet vi klassifiserer etter tiltak og kontroll. Det betyr at elever som har sluttet underveis i skoleåret teller med, og trekker andelen med prøve ned. Elever som har endret skole underveis i skoleåret er tilordnet sin opprinnelige skole.

## Utfallsvariable – videregående skole

På sikt er det ønskelig å studere hvorvidt tiltaket har påvirket fullføringen av videregående opplæring. Disse resultatene vil først bli tilgjengelig om noen år. Vårt viktigste resultatmål i denne rapporten for elever på videregående skole er derfor resultatet på Oslo-prøven som tas sent om våren. I tillegg ser vi på fullføring av første år på videregående. I Tabell 5.1 gir vi en deskriptiv oversikt over disse variablene samt andre resultatmål som bestemmes i forkant av tiltaket. Som forventet presterer målgruppeelevene svakere på alle utfall sammenliknet med alle elever under ett. Læringsstøttende prøve og standpunkt 10. trinn er normalisert til å ha gjennomsnitt lik 0 og standardavvik lik 1.

**Tabell 6.1** Deskriptiv oversikt over variabler som måler elevprestasjoner på videregående – årene 2016/17 - 2018/19

	Alle		Målgruppen	
	Snitt	Ant. obs	Snitt	Ant. obs
Utfallsvariable				
-Oslo-prøve	-0.13	4002	-0.78	1280
-Fullføring VG1	0.81	3381	0.71	1065
Pre-determinerte resultat				
-Læringsstøttende prøve (standardisert)	-0.03	4782	-0.94	1703
-Snitt standpunkt 10.trinn (standardisert)	-0.19	4554	-0.64	1519
-Standpunkt matematikk 10. trinn	3.06	4596	2.44	1533

Kilde: Statistisk sentralbyrå.

### 6.1.1. Andre bakgrunnskjennetegn ved elevene

I Tabell 6.2 gir vi en oversikt over elevens bakgrunn. Som i Tabell 4.1 deler vi opp i tre grupper: analysesamplet, målgruppen og øvrige elever. Også med tanke på elevsammensetningen i de ulike gruppene er det store likhetstrekk med tallene i Tabell 5.1. Det er flere gutter enn jenter blant målgruppeelevene (større forskjell enn på grunnskolen). Foreldres utdanning er også lavere blant målgruppeelevene og flere har utenlandskfødte foreldre.

**Tabell 6.2** Deskriptiv oversikt over variable som måler familiebakgrunn på videregående – årene 2016/17 - 2018/19

	Analysesamplet		Målgruppen	
	Snitt	Ant. obs	Snitt	Ant. obs
Kjønn				
-Gutt	0.546	4782	0.636	1617
-Jente	0,454	4782	0,364	1617
Foreldres utdanning				
-Lav	0,551	4924	0,632	1703
-Høy	0,449	4924	0,368	1703
Innvandrerbakgrunn				
-Norskfødte foreldre	0,674	4924	0,417	1703
-Utenlandsfødte foreldre	0,326	4924	0,583	1703

Kilde: Statistisk sentralbyrå.

### 6.2. Balansering av tiltak- og kontrollskoler

Ettersom skoler er tilfeldig fordelt til tiltak og kontroll forventer vi ikke systematiske forskjeller uavhengig av tiltaket. Analysene av grunnskolen viste hvordan det likevel kan oppstå tilfeldige forskjeller. Vi har også sett at det er vesentlige forskjeller i gjennomføring av prøver mellom skoler, som kan gi forskjeller mellom tiltaks- og kontrollskolene. Vi starter derfor, som for grunnskolen, med å undersøke balansering, i hvilken grad skolene var sammenliknbare før tiltaket, før vi vurderer effekter ved å sammenligne resultater etter tiltaket.

Disse resultatene er rapportert i Tabell 6.3. Vi viser en tilsvarende balanseringstest for videregående skole som for grunnskolen. I kolonne (1) og (2) ser vi at det er ingen systematisk forskjell mellom tiltak- og kontrollgruppen i gjennomføringen av

læringsstøttende prøve (LS-prøve) eller andelen elever i målgruppen.<sup>24</sup> Videre ser vi i kolonne (3) og (4) at elevene i tiltaksskolene har litt svakere resultater på læringsstøttende prøve og har i litt større grad gjennomført Oslo-prøven. Alle disse forskjellene er så små at de kan skyldes tilfeldigheter. Læringsstøttende prøve gjennomføres på høsten, og i forkant av den tilrettelagte matematikkundervisningen.

**Tabell 6.3 Balansering av tiltak- og kontrollskoler for videregående, årene 2016/17 - 2018/19<sup>1</sup>**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Har resultat LS	Inngår i målgruppen	Resultat LS-prøve	Har resultat Oslo-prøve	Har resultat grunn-skole	Snitt st.pkt 10.trinn	St.pkt matte 10.trinn
Tiltak	-0.00 (0.03)	-0.00 (0.02)	-0.03 (0.04)	0.04 (0.03)	-0.02** (0.01)	-0.06 (0.06)	-0.10** (0.05)
Ant observasjoner	4782	4782	4782	4782	4782	4537	4563
Snitt avhengig variabel	0.93	0.41	-0.03	0.85	0.95	-0.21	3.04

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (2) med forskjellige uavhengige variable (y) i kolonner. Alle estimater kontrollerer for årskull og strata (gruppe i randomiseringen). Klyngerobuste beregnede standardfeil i parentes under hvert effektestimater. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

Vi finner en tydelig forskjell i resultater fra grunnskolen (kolonne (5) – (7)): Færre elever i tiltaksskolene har registrert grunnskolerresultater, og de som har, har i snitt svakere resultater i matematikk. Disse forskjellene er signifikante på 5 prosent nivå. I sum er det ikke store forskjeller mellom tiltaks- og kontrollskolene før tiltaket, men det er tegn til svakere forutsetninger blant tiltakselevne, som vi vil ta hensyn til i effektanalysene.

### 6.3. Effektestimat – videregående

Tabell 6.4 rapporterer punktestimatet av tiltaket på videregående. I tillegg til å se på resultatene fra Oslo-prøven (kolonne (1) – (3)) ser vi også på hvorvidt elevene har fullført første året på videregående (kolonne (4) og (5)). Også i denne tabellen representerer hver celle med et punktestimat en separat regresjon. Som for grunnskolen tar vi også hensyn til elevenes forutsetninger ved å kontrollere for elevenes familiebakgrunn, resultater fra læringsstøttende prøve (gjennomført i forkant av tiltaket) og resultater fra grunnskolen. Resultatene i første rad tar kun hensyn til forskjeller i elevenes familiebakgrunn. I andre rad betinger vi i tillegg på forskjeller i resultatet fra læringsstøttende prøve. I tredje rad kontrollerer vi ytterligere for forskjeller i grunnskolerresultater.

**Tabell 6.4 Forskjeller i Oslo-prøven og fullføring av VG1 mellom målgruppeelever i tiltak- og kontrollskoler, årene 2016/17 - 2018/19<sup>1</sup>**

	(1)	(2)	(3)	(4)	(5)
	Oslo-prøve Målgruppe	Oslo-prøve Utvidet målgruppe	Resultat Alle elever	Fullført VG1 Målgruppe	Fullført VG1
Gitt fam.bakgr	-0.10 (0.06)	-0.11 (0.07)	-0.08 (0.05)	0.04 (0.03)	0.01 (0.03)
+ Gitt LS	-0.09 (0.06)	-0.09 (0.06)	-0.05 (0.05)	0.04 (0.03)	0.02 (0.02)
+ Gitt gsk.resultat	-0.05 (0.05)	-0.04 (0.05)	0.00 (0.04)	0.05* (0.03)	0.03* (0.02)
Ant observasjoner	1280	1843	4002	1065	3381
Snitt avhengig variabel	-0.78	-0.76	-0.13	0.71	0.81

<sup>1</sup> Hver celle gir forskjell fra estimering av ligning (2) med forskjellige uavhengige variable (y) i kolonner, og forskjellige sett av kontrollvariable (x) i rader. Alle estimater kontrollerer for årskull og strata (gruppe i randomiseringen).

Klyngerobuste beregnede standardfeil i parentes under hvert effektestimater. Statistisk signifikans: \*\* 5 prosent, \* 10 prosent.

Kilde: Statistisk sentralbyrå.

<sup>24</sup> Målgruppen omfatter elever på mestringsnivå 1-2, totalt omtrent 40 prosent av elevene, i tillegg har vi en utvidet målgruppe som også omfatter elever uten prøveresultat. Balanseringen endrer seg ikke nevneverdig ved å utvide målgruppen.

I kolonne (1) ser vi at målgruppeelever i tiltaksskoler scorer lavere på Oslo-prøven enn målgruppe-elever i kontrollskoler. Ettersom den tilrettelagte undervisningen for målgruppeelever foregår i ordinære klasserom på videregående, er det naturlig å estimere effekten for hele elevgruppen samlet sett. I kolonne (3) ser vi at det gir et punkttestimat som er nokså likt det for målgruppen. I ingen av tilfellene er det en forskjell mellom tiltak- og kontrollgruppen utover det vi kan forvente vil oppstå som følge av tilfeldig variasjon. Vi ser også at forskjellene reduseres når vi tar hensyn til forskjeller i elevenes forutsetninger. Dette gjelder uavhengig av hva vi kontrollerer for. Som allerede diskutert ovenfor, er det ikke obligatorisk å ta læringsstøttende prøve. Når vi inkluderer elever som står uten resultat på læringsstøttende prøve, ved å gi de en lav verdi (i stedet for manglende) er resultatene omtrent uendret (se kolonne (2)).

I de to siste kolonnene viser vi et foreløpig mål på gjennomføring, fullført og bestått Vg1. Her ser vi at det er en forskjell i favør tiltaksskolene, på 1-3 prosentpoeng for alle elever, og på 4-5 prosentpoeng for elever i målgruppen. Forskjellen når vi tar hensyn til resultater fra grunnskolen er omtrent på grensen av hva vi kan vente av tilfeldige forskjeller (signifikant på 10 prosent nivå, men ikke på 5 prosent).

Som for grunnskolen har vi også sett på forskjellige effekter for grupper av elever. I grunnskolen så vi også på effekter for grupper av skoler. I videregående har vi færre skoler, og vil ikke forsøke å analysere slike forskjeller. Som for grunnskolen er det ikke grunnlag for sterke konklusjoner om effekter for noen elevgrupper framfor andre.

#### **6.4. Oppsummering og kost/nytte – videregående**

I videregående har vi ikke tilsvarende problemer med balansering av skolene som for ungdomsskolene: resultater på læringsstøttende prøve er tilsvarende i tiltaks- og kontrollskolene. For alle elevgrupper og sett av kontrollvariable finner vi effekter nær null og som ikke er statistisk signifikante. Det er likevel en tendens til at effekttestimatene reduseres når vi inkluderer mer omfattende kontroller for forskjeller mellom tiltaks- og kontrollskolene. I sum tyder dette ikke på effekter på matematikkferdigheter målt ved Oslo-prøven.

Når vi studerer fullført og bestått Vg1 finner vi imidlertid en forskjell i favør tiltaksskolene. For målgruppen er denne omtrent 4-5 prosentpoeng. Dette er en forholdsvis betydelig forskjell. Når vi tar hensyn til forskjell i resultater fra grunnskolen er forskjellen signifikant på 10 prosent nivå, ellers er den ikke signifikant. Det er overraskende at vi finner en forskjell i gjennomføring, når vi ikke finner forskjeller i resultater fra Oslo-prøve. Ettersom vi også har et mindre datagrunnlag i analyser av gjennomføring (to årskull, ikke tre), er vi foreløpig litt avventende med å konkludere at dette er en effekt av tiltaket.

#### **Kostnader**

Når vi vurderte effekter i grunnskolen oppsummerte vi også kostnader av tiltaket. Vi anslo at gjennomføringen av tiltaket totalt kostet omtrent 20,5 millioner kr, hvorav omtrent 14 i grunnskolen og dermed 6,5 i videregående skole, dvs. omtrent 2,2 millioner per år. Kostnaden i videregående omfatter ikke smågrupper som i grunnskolen, men kun administrasjon, materiell og kursing (inkludert noe finansiering av vikarer til skolene). Kostnadene ved en evt. videreføring/kopiering vil kunne være annerledes, for eksempel dersom det er mulig å gjenbruke materiell og kursopplegg framfor å utvikle nytt.

I videregående er skillet mellom målgruppen og øvrige elever ikke like tydelig. Dersom vi fordeler kostnadene på de omtrent 1000 målgruppeelevene i tiltaksskolene blir det en kostnad på omtrent 6500 kr per elev. Dersom vi fordeler kostnadene på alle de omtrent 2500 elevene i tiltaksskolene blir det en kostnad på omtrent 2600 kr per elev.

### **Kost/nytte**

For videregående finner vi ingen effekt på Oslo-prøven, som er målet på ferdigheter i matematikk. Vi finner imidlertid at fullføring av Vg1 øker med omtrent 5 prosentpoeng for målgruppen og 3 prosentpoeng for alle elever. Som for grunnskolen er denne forskjellen for liten til at vi kan si sikkert at den skyldes tiltaket og det er usikkerhet knyttet til manglende samsvar mellom resultater på Oslo-prøve og fullføring av Vg1, men det er likevel vårt beste anslag på effekt. En forskjell på 3-5 prosentpoeng i fullføring av Vg1 er assosiert med en forventet forskjell på omtrent 1 prosentpoeng i fullføring av VGO. I tråd med resultatene fra Falch m.fl. (2009), justert for lønnsvekst (jf. diskusjon ifm. analyse av resultatene fra grunnskolen), svarer en økning i fullføring av VGO på 1 prosentpoeng til en forventet gevinst på omtrent 12 500 kr per elev som deltar i tiltaket (1 prosent av gevinsten på 1,25 millioner kr per ekstra elev som fullfører VGO).

Ettersom vi ikke finner noen effekt på Oslo-prøven og estimatet for fullføring av Vg1 ikke er signifikant i flere spesifikasjoner har vi ikke noe sterkt grunnlag for å faktisk vente at fullføring øker med 1 prosentpoeng. Det vil sannsynligvis heller ikke være mulig å noenlunde sikkert påvise en såpass liten effekt på fullføring. Det er likevel interessant å merke seg at en slik effekt, som vil være i tråd med forskjellen vi finner i fullføring av Vg1, vil være tilstrekkelig til å gjøre tiltaket i videregående samfunnsøkonomisk lønnsomt.



## 7. Kvalitativ analyse

### «En plutselig opplevelse av mestring». Aktørenes erfaringer med implementering og praktisering av tiltaket

Temaet for dette kapitlet er aktørenes erfaringer med å implementere tiltaket, og deres opplevelser av å delta. Med utgangspunkt i funnene fra de kvantitative beregningene, er det også relevant å prøve å forklare hvorfor tiltaket ikke har hatt større målbare effekter. Kapitlet skal svare på tre spørsmål, hvorav det første legger det empiriske grunnlaget for de to andre:

- (i) Hvilke erfaringer har rektorer, lærere og elever med å implementere og praktisere tiltaket?
- (ii) Hvorfor har ikke tiltaket hatt større direkte effekter (jf. kapittel 5 og 6)?
- (iii) Hvordan skal og kan skolene forvalte kompetansen de har fått når satsingsperioden er over?

Enkelte av resultatene er presentert i kapitlet *Analyse av implementering, endring og tilpasning: Fra implementering til stabilisering* (Kirkebøen mfl. 2018: 15-30) og artikkelen *Lojalitet i spill: Eierskap i implementering av et matematikdidaktisk tiltak i Oslo-skolen* (Reegård og Rogstad 2019). Begge disse arbeidene er basert på data fra de to første årene. Dette kapitlet bygger videre på funnene i disse publikasjonene, men inkluderer også data fra det tredje og siste året tiltaket ble gjennomført.

### 7.1. Analytisk rammeverk: Implementering og endring

En viktig del av evalueringen har vært å analysere aktørenes egen forståelse og oppfatning av tiltaket. I de kvalitative analysene rettes derfor søkelyset på aktørenes *subjektive erfaringer*. Formålet med disse analysene er å peke på mulige årsaker til funnene som er gjort med kvantitative data, samt løfte fram betydningen av kontekstuelle forhold.

Det er rimelig å anta at aktørenes oppfatninger om tiltaket også vil virke inn på deres lojalitet overfor gjennomføringen av tiltaket (Reegård og Rogstad 2019). Det er imidlertid ikke like opplagt at aktørenes subjektive vurderinger samsvarer med det som faktisk kan observeres (Kirkebøen mfl. 2017 – se fotnote 1).

Når det gjelder tiltaket vi studerer i denne rapporten, kan man tenke at rektorer og lærere, som har arbeidet for å implementere tiltaket, også antar at det har stor betydning. Dette kan for eksempel skyldes en type ønsketenkning. De fleste vil at egen innsats skal ha betydning.

De kvalitative dybdeintervjuene har alltid, for en gitt elevkohort, funnet sted i forkant av den kvantitative effektevalueringen. Dette skyldes hovedsakelig at utfallsvariablene våre ikke har vært tilgjengelig før utpå vårparten. Imidlertid framgikk det av foreløpige kvantitative analyser som ble gjennomført midtveis i prosjektet, at enkelte skoler lykkes bedre enn andre med tiltaket. I samråd med Utdanningssetaten ble det da besluttet at vi skulle vie ekstra oppmerksomhet til skolene som tilsynelatende hadde best uttelling. Var det mulig å peke på noe de gjorde særskilt? Dette valget ble gjort ut fra en antakelse man hadde da om at tiltaket ville ha større effekter enn det viste seg å ha.

Det sentrale temaet i dette kapitlet er *implementering*. Et emne hvor det finnes mye litteratur, noe vi presenterte i forrige rapport (Kirkebøen mfl. 2018). Her vil vi derfor bare rekapitulere noen sentrale elementer. Et felles trekk for studier av tiltak

og skoleutvikling, er vektlegging av kompleksitet og hvordan det kan oppstå koordineringsproblemer (Durlak og DuPre 2008; Blossing mfl. 2012). Aktuelle faktorer kan være forhold som lærernes lojalitet, kvalitet på kurs og støttemateriell, deltakernes respons, og hvorvidt det er tilstrekkelige ressurser i form av penger og tid. Det sistnevnte, at endring er tidkrevende, er et tema som vies mye plass i litteraturen (se for eksempel Datnow 2002; Tyack og Cuban 1995). En forklaring på tidsbruk er antall aktører som må involveres på en skole. I tillegg må ofte læremateriell skiftes ut, noe som kan være krevende prosesser når endringene skal underlegges tilstrekkelig kvalitetssikring. Dette gjelder også for materiell til bruk i tiltak av den typen vi analyserer her.

Et velkjent analytisk perspektiv for å forstå implementeringsprosesser, er basert på Lipskys (1980) begrep «street level bureaucrats» eller «bakkebyråkrater» på norsk. Begrepet reflekterer at menneskene som fyller rollene nært brukerne ofte må fortolke og tilpasse sentralt fattede beslutninger på måter som de opplever som adekvate for å gjennomføre oppgavene. Ofte kan dette være nyttige og gode tilpasninger, men lokale justeringer kan også ha som resultat at tiltak praktiseres ulikt. Inspirert av Lipsky er det relevant å følge styringslinjen, og se hvordan aktører i forskjellige posisjoner gjør ulike typer av tilpasninger. Når det for eksempel gjelder tiltaket vi her skal studere, er det Kunnskapsdepartementet som fastsetter målet, mens Utdanningsetaten (skoleeier), forskerne og de enkelte skolene påvirker organisering, innhold og arbeidsmåter. I det daglige arbeidet er det likevel lærerne som praktiserer, noe som aktualiserer deres skjønnsvurderinger (Grimen 2006).

Den vertikale styringslinjen kan følges ovenfra eller nedenfra. Det førstnevnte refereres gjerne til som et «ovenfra og ned-perspektiv». I studier med dette perspektivet er det viktig å identifisere hvordan sentralt vedtatte tiltak endres gjennom fortolkning og praktisering – fra beslutningstakere til praktikere (Hoyle og Wallace 2007). I skolen vil aktuelle spørsmål være hvordan et tiltak formidles og tilpasses fra rektorer, via avdelingsledere og ut til den enkelte lærer: Hvordan tolker, operasjonaliserer og handler de ulike aktørene i lys av målsettinger og det didaktiske innholdet som er foreskrevet gjennom tiltaket? Dette perspektivet aktualiserer den vertikale styringslinjen, der politiske ideer og visjoner omsettes til samfunnsendringer gjennom styring og kontroll. Det andre perspektivet, «nedenfra og opp», tar utgangspunkt i aktørene. I denne sammenheng vil det si hvordan elever, lærere og rektorer opplever tiltaket. I Reegård og Rogstad (2019), ble et slikt nedenfra og opp-perspektiv brukt for å analysere aktørenes lojalitet i gjennomføringen av tiltaket. Denne studien var inspirert av Arvidson og Axelsson (2014), som argumenterer for at lojalitet kan være horisontal og vertikal, samt frivillig og ufrivillig. Mens vertikal lojalitet er knyttet til oppgave, setter horisontal lojalitet relasjonelle forhold i sentrum. I begge tilfeller kan lojalitet være noe man er påført utenfra eller noe som bygger på opplevelse av et mer indrestyrt ønske om utvikling og endring.

## 7.2. Data og metode

Dette kapitlet er basert på kvantitative og kvalitative data. De kvantitative dataene er samlet inn med bruk av websurvey, det vil si at vi sender et spørreskjema per e-post. I vår undersøkelse brukte vi websurvey til rektorer, avdelingsledere og lærere. Formålet med de kvantitative dataene var å få en bred oversikt over hvordan tiltaket var implementert og om det var avgjørende forskjeller mellom skolene i gjennomføringen av tiltaket. I tillegg ønsket vi å avdekke eventuell variasjon når det gjaldt oppfatninger og vurderinger av tiltakets form og innhold.

Samtidig ønsket vi å åpne for at det er erfaringer og vurderinger vi ikke hadde tenkt ut forut for datainnhenting, som vi dermed ikke inkluderte i form av

svaralternativer. Noe av dette ble ivaretatt med åpne svarkategorier, men vi ønsket også at informantene skulle få mulighet til å gi mer utfyllende svar. Av den grunn gjennomførte vi også kvalitative intervjuer med et mindre antall informanter. Vi så også at det var viktig å intervjuer elevene kvalitativt. Av dem var det viktig å få innblikk i hvordan de vurderte tiltaket, og hva de la vekt på i sin begrunnelse for de oppfatningene de hadde. Vi var særlig interessert i å få svar på hvorfor tiltaket virket eller ikke, sett fra deres ståsted, noe man heller ikke kan spørre om direkte. Blant tiltakselevne var det lite som indikerte at intervjuer om deltakelsen var sensitiv. Vi valgte derfor å benytte oss av gruppeintervjuer, som vi mente var gunstig for å la elevene kunne reflektere over enkelte av spørsmålene i fellesskap. Dessuten var det strategisk lurt med gruppeintervjuer fordi vi på den måten fikk truffet et større antall elever enn vi hadde gjort om vi intervjuet dem én og én.

En viktig presisering når det gjelder kvalitative data, er at de ikke er representative. Hensikten er således ikke å trekke slutninger om populasjonen. Snarere ønsker vi å komme i dybden av hva informantene mener, hvorfor de har bestemte vurderinger og gjennom dette hvilke mekanismer som synes å være viktige ut fra de informantene vi har snakket med.

Utover de over nevnte betraktningene var det tre andre hensyn som var avgjørende i datainnsamlingen. For det første å identifisere eventuell *endring* i løpet av tiltaket. Dette ble gjort ved å spørre tiltakslærerne hvert av de tre årene tiltaket varte. Vi har i tillegg intervjuet rektorene på to tidspunkter, samt møtt tiltakselevne alle tre tiltaksårene. For det andre ville vi ha en eksplorerende tilnærming, noe som åpner for at informantene kan fremme *subjektive* vurderinger av forhold som vi ikke hadde tenkt ut på forhånd. Et tredje forhold var hensynet til å innarbeide et styringsperspektiv, noe som fordret at vi intervjuet aktører på ulike *nivåer* i styringskjeden.

Et moment som følger med designet av intervensjonen er som allerede nevnt at tiltaket var betydelig klarere definert for ungdomsskolen enn for videregående skole. To viktige forskjeller var at man i ungdomsskolen for det første hadde en klar arbeidsplan (økt-planer) som inneholdt detaljerte beskrivelser av hva læreren skulle gå gjennom i hver time, og for det andre var tiltakselevne delt opp i store og små grupper, som følgelig var identifiserbare. I videregående var tiltakselevne en del av den ordinære klassen. Den viktigste forskjellen var at læreren hadde gått på kurs, og følgelig kunne de didaktiske grepene som metoden foreskrev. Denne forskjellen resulterte i at vi valgte å konsentrere studien av erfaringer til ungdomsskolen i tiltakets tredje år.

### **Websurvey**

Den første websurveyen ble sendt til alle rektorene i ungdomsskolen og videregående skole våren 2016. Det vil si til de skolene som var valgt ut som henholdsvis tiltaks- og kontrollskoler. Rektorene svarte altså på spørreskjemaet før tiltaket var igangsatt. Dette var en form for nullpunktsdata, som var viktig for å kunne identifisere forventninger de hadde, og fordi vi med utgangspunkt i data før tiltaket startet hadde en baseline som kunne brukes for å kartlegge hvorvidt gjennomføringen av tiltaket faktisk resulterte i noen typer endringer. Rektorene fikk websurveyen på to måletidspunkter, mens lærerne fikk det hvert av de tre tiltaksårene.

I tabell 7.1 har vi satt inn en oversikt over når websurveyene ble sendt til rektorer og lærere, samt svarprosenten for hver av undersøkelsene. Det skal her også trekkes fram at vi purret på svar i flere runder. I arbeidet med å få opp andelen som besvarte undersøkelsen bidra også Utdanningsetaten, gjennom å sende e-poster til rektorer og lærere med oppfordring til å svare. Samtidig som det framgikk at det

var viktig at svarprosenten ble så høy som mulig, framgikk det også at deltakelsen var frivillig og at det enkelte svar ville forbli anonymt for både Utdanningsetaten og offentligheten.

**Tabell 7.1** År for gjennomføring av ulike websurveyer etter type skole og stilling. Rapporter i hver celle er svarprosenten

	2016		2017		2019	
	Ungdomsskole	VGS	Ungdomsskole	VGS	Ungdomsskole	VGS
Rektor	61.5	100	-	-	100	100
Lærere	79	79	50	78	57	62

Av tabellen framgår det at svarprosenten har vært høyere blant rektorene enn for lærerne. Dette er trolig ikke så underlig gitt at de har mer direkte kontakt med Utdanningsetaten, og ikke minst at rektorene i større grad er pålagt å ta ansvar for implementering av tiltak for å sikre en god skoleutvikling. Når det gjelder lærerne var det tidvis krevende å få inn tilstrekkelig andel svar, men etter purring og henstilling om å besvare undersøkelsen er andelen tilfredsstillende.

### Casestudieskoler

Ved hver av caseskolene intervjuet vi rektor, avdelingsleder(e), tiltakslærere og grupper med elever. Hvert intervju varte fra 45 minutter til 1,5 time. Gruppeintervjuer med elever tok lengst tid. Der sikret vi at hver av elevene fikk tid til å prate om egen opplevelse, samtidig som vi også satte av tid til felles refleksjon omkring tiltaket, matematikkundervisningen generelt – og ikke minst det å være med i en slik gruppe. Alle intervjuene ble tatt opp på bånd og vi skrev en logg umiddelbart etter at intervjuene var gjennomført.

Totalt har vi vært på åtte skoler. I 2016 og 2017 gjennomførte vi intervjuer ved fire ungdomsskoler og to videregående skoler. I disse to årene besøkte vi de samme skolene. Med utgangspunkt i resultatene som ble presentert i Kirkebøen mfl. 2018, ble det besluttet at vi skulle fokusere på ungdomsskolen og i tillegg sette søkelys på de skolene som synes å ha fått mest igjen for tiltaket. Resultatet var at vi beholdt to av de fire ungdomsskolene, men supplerte utvalget med fire nye. Samlet gjennomførte vi dermed intervjuer på seks skoler også i den siste runden. Totalt har vi følgelig gjennomført intervjuer på 10 tiltaksskoler.

Skolene som ble valgt ble trukket ut på bakgrunn av en vurdering av en normalfordeling av andelen elever i målgruppa, samt geografisk beliggenhet innenfor Oslo. Geografisk beliggenhet ble her valgt ut fra etablerte analyser av skillet mellom skoler øst og vest i Oslo. Mer konkret vil det si at valget av beliggenhet speiler to andre og antatt viktige hensyn, nemlig sosioøkonomisk bakgrunn og andel med etnisk minoritetsbakgrunn på skolen og i nærområdet.

I tillegg var det viktig å matche caseskole med en kontrollskole hvor den geografiske beliggenheten og andelen i målgruppen ligner. Alle intervjuene ble tatt opp på bånd. I tillegg skrev vi feltlogg etter skolebesøkene. Vi gjennomførte også observasjon og deltok på samlinger og kurs i regi av Utdanningsetaten. Disse dataene ble ikke brukt direkte i analysene, men var nyttige i intervjuene fordi vi kunne bruke innsiktene i hva deltakerne hadde lært når vi senere gjennomførte intervjuer.

### 7.3. Resultater: Normalisering – fra Singapore til Oslo-skolen

I prosjektet har vi samlet inn data på flere tidspunkter. Dette er gjort for å få innblikk i eventuelle endring i aktørenes oppfatninger av implementering og praktisering av tiltaket. Basert på surveydata til skolene, og gjennom kvalitative

intervjuer, har det vært mulig å følge endringer i de involverte aktørenes vurderinger fra første til andre tiltaksår, oppsummert som «fra implementering til stabilisering» (Kirkebøen mfl. 2018: 22). Konklusjonen på de kommende sidene er at tiltaket fra det andre til det tredje tiltaksåret gikk fra en *stabilisering til en normalisering*. Med det mener vi at sentrale deler av de didaktiske grepene i tiltaket, gjerne forklart med referanse til matematikken slik den undervises i Singapore, ble innarbeidet som en del av det ordinære didaktiske opplegget for flere av tiltaksskolene. Normalisering innebærer altså at denne didaktikken synes å ha blitt integrert som del av den generelle undervisningen. Kort sagt, fra å være noe ytre som ble presset på skolene ovenfra, er dette nå noe de selv gjør. Singapore-skolen har blitt en selvfølgelig del av Oslo-skolen. I siste runde med intervjuer var det derfor mer et spørsmål om hvordan de skulle videreføre de grepene de hadde innarbeidet i matematikkundervisningen. Dette var langt fra det inntrykket vi fikk i første runde med intervjuer på tiltaksskolene (jf. Kirkebøen mfl. 2018).

I analysene har vi valgt å samle inn data fra rektorer, lærere og elever – altså involverte aktører som befinner seg på ulike nivåer i den vertikale styringslinjen. Vi starter med vurderingene fra ledelsen ved skolene, altså rektorene, før vi ser på lærerne og til slutt elevene. For rektorer og lærere har vi både kvantitative og kvalitative data, mens vi kun har kvalitative data når det gjelder elevene.

#### **7.4. Rektorene – implementeringsansvarlige på skolen**

Rektor er bindeleddet mellom skoleeier og skole. I kraft av å være skoleleder er også rektorene formelt ansvarlige for implementeringen av et sentralt fattet tiltak. Lokalt ved skolene er det gjerne avdelingsleder som får ansvaret for den daglige oppfølgingen. Det strategiske utviklingsarbeidet ved skolene er likevel lagt til rektorene. De er pådrivere og ansvarlige for å hente inn tilstrekkelig informasjon, samt å formidle målsettinger til stab, foreldre og elever ved skolen. Når det gjelder tiltaket vi studerer her, hadde ikke rektorene innvirkning på tiltakets utforming. Den var definert av forskerne og vedtatt implementert av Utdanningsetaten. Sagt på en annen måte, rektorene hadde ansvar for implementering av et tiltak de selv ikke kunne mene om de ville være med på eller hvordan skulle gjennomføres.

Det var imidlertid stor forskjell mellom tiltak rettet inn mot videregående skole og ungdomsskole. Oppdeling av grupper, gjennomføring av timer og temaer som skulle behandles var klart definert for ungdomsskolen, mens tiltaket var mer et didaktisk verktøy som ble implementert som en del av den ordinære undervisningen blant tiltaksskolene på videregående nivå. Av den grunn hadde tiltakslærerne i videregående større frihet.

I Kirkebøen mfl. (2018:20) har vi presentert rektorenes første reaksjon på å innføre tiltaket, slik det framkom i dataene som ble samlet inn med bruk av websurvey til alle tiltaksskolene. Resultatene fra høsten 2016 indikerte at det var en forskjell mellom rektorer i ungdomsskolen og i videregående skole. Mens ni av ti rektorer i ungdomsskolen sa seg helt eller delvis enige i at de hadde fått god informasjon, var andelen seks av ti i videregående skole. Da vi intervjuet rektorene etter tredje tiltaksperiode, så de kunne sammenlikne informasjonen de hadde fått i hvert av tiltaksårene, ble informasjonen fra første tiltaksår omtalt i ganske andre ordelag.

*Det første året var jo mye litt kaotisk. Jeg tror vi kjørte i gang for tidlig. For meg står det som en slags pilot. Senere ble det mye bedre. Dessuten var materiellet mer utviklet og vi fikk det i tide. (rektor ungdomsskole, tiltak, 2019)*

I Kirkebøen mfl. (2018) satte vi søkelys på rektorenes vurderinger med hensyn til om det var behov for tiltak for de svakest presterende i matematikk, og hvorvidt de mente at det aktuelle tiltaket var et adekvat svar på utfordringene enkelte elever har i faget. Rektorenes vurdering av dette er interessant i lys av en antakelse om at de vil være mer lojale i implementeringen av et tiltak de har tro på enn et tiltak de først og fremst opplever som de er påført fra skoleeier.

I rektorenes svar på den første websurveyen, gikk det fram at de generelt var positive til tiltak i matematikk. Av den grunn er det heller ikke overraskende at de også ga uttrykk for at de var positive til, og lojale i gjennomføringen av, dette tiltaket. I den første websurveyen la rektorene vekt på fire forhold (Kirkebøen mfl. 2018:21):

- Øke motivasjonen blant matematikklærerne
- Økt faglig kompetanse blant matematikklærerne
- Økt kompetanse blant de svakest presterende elevene
- Økt motivasjon hos de svakest presterende elevene

Disse forventningene var forhold rektorene la vekt på også i den siste websurveyen (2019). Forventningene var følgelig preget av stabilitet til tross for at det var betydelige organisatoriske endringer i gjennomføringen av tiltaket fra 2016 til 2019. Disse endringene var først og fremst knyttet til at informasjon og materiell kom raskere ut og var bedre gjennomarbeidet. Allerede andre skoleår kunne denne endringen spores gjennom svar vi fikk på spørsmål vi stilte rektorene. Beskrivelsen av tiltaket og samarbeidet med Utdanningsetaten var noe annet enn den vi fikk første året, noe som også kom fram i det forrige sitatet. Dette var også funn som lå til grunn for konklusjonen i Kirkebøen mfl. 2018, når vi der pekte på at det allerede andre året hadde inntrådt en stabilisering i praktiseringen av tiltaket.

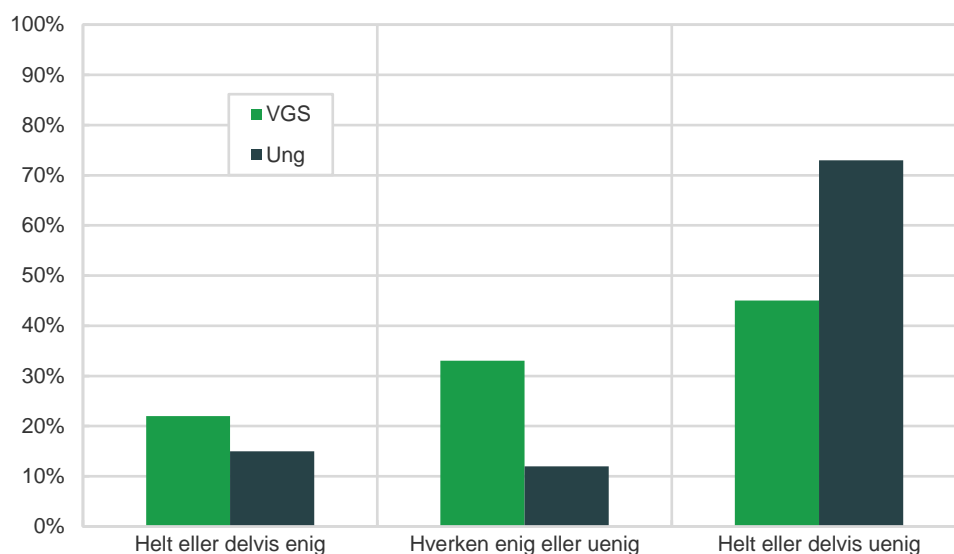
Hovedforklaringen på at de var mer positive fra og med år to var informasjon, og at de dermed visste mer om hva tiltaket innebar. Mer kunnskap gjorde det lettere å planlegge året, noe som i neste omgang gjorde det enklere å motivere lærerne. Trolig var også motiveringen enklere når det var lærere i kollegiet som hadde erfaring etter å ha gjennomført tiltaket en gang. Uavhengig av årsak, når flere av informantene ga uttrykk at de didaktiske grepene som var tatt første år, så ut til å bli en del av det normale i år to og tre.

*Hos oss har vi tatt til oss mye av det vi lærte på tiltaket. Jeg vet at lærerne her har kopiert oppgavene og økt-planene. De er så gode og gjennomarbeidet at de sier at de vil bruke dem videre.*

*For min del er jeg mest opptatt av hvordan vi organisatorisk kan gå videre og sikre at nye lærere også får kurs eller får lære de grepene tiltakslærerne hos oss har fått. (rektor ungdomsskole, tiltak 2019)*

Det er verdt å dvele litt ved rektorenes utfordring knyttet til å motivere lærerne. Denne utfordringen handlet dels om å sikre at lærerne gjennomførte tiltaket i henhold til oppsatt plan, dels om å få dem med på en måte hvor de fikk eierskap til denne undervisningen. I figur 8.1 ser vi fordelingen over rektorenes opplevelse av hvor vanskelig det har vært å motivere lærerne til å delta på kurs. Dette spørsmålet stilte vi ikke i første runde, men tok det inn ettersom flere av rektorene beskrev dette da vi i det første intervjuet spurte om de hadde noe ekstra å legge til. Av figuren framgår det at særlig rektorene på ungdomsskolen sier seg uenige i denne påstanden. Samtidig er det forskjell mellom ungdomsskolen og videregående skole. På videregående skole er under halvparten av rektorene uenige, mot noe over 70 prosent av rektorene i ungdomsskolen.

**Figur 7.1** Påstand: «Det har vært vanskelig å motivere lærerne til å delta på kurs». Rektor ved tiltaksskoler. Videregående opplæring og ungdomsskole. (N vgs=9, N ung=26) 2019



Kilde: Statistisk sentralbyrå.

La oss her igjen minne om at tiltaket var ganske klart definert i ungdomsskolen, mens grensene mellom tiltaksgruppen og de øvrige elevene var mer flytende i videregående skole.

Da vi gjennomførte siste intervjurunde i ungdomsskolene 2018/2019 fulgte vi opp spørsmål knyttet til organisering. Dette er innsikter som er interessante for å vurdere dette tiltaket isolert sett, men også for å kunne si noe generelt om gode grep for implementering av senere tiltak.

Et første tema dreier seg om gangen i en tiltaksperiode. Som allerede påpekt flere ganger var første år preget av at ikke alt var på plass. I ettertid framstår dette året mest som en pilot. Men allerede andre tiltaksår var situasjonen en annen, kjennetegnet ved det vi i Kirkebøen mfl. (2018) refererte til som stabilitet. Noe vi ikke hadde forutsett før siste intervjurunde, var at rektorene allerede var på vei videre. Nå var de i gang med å planlegge nye tiltak. Det som skulle skje etter at PBG ble avsluttet. De gjennomførte opplegget som de skulle, men flere fortalte at de ved skolen nå hadde bestemt seg for å gå for andre satsinger. På noen skoler var dette innen matematikk, på andre skoler mente de at det var tid for å prioritere andre fag.

*Vi skal i gang med et annet tiltak i matematikk fra neste høst. I farten husker jeg ikke helt hva det heter. Men du vet. Sånn går nå dagene. (rektor, ungdomsskole, tiltak 2019)*

Nå er det ikke gitt at disse tiltakene avviker så mye fra hverandre. Det synes mer som om det har vært hva vi kan referere til som en «Singapore-vending». I de kvalitative intervjuene med rektorer og lærere ble det for eksempel vektlagt at Engebråten-modellen i praksis også minnet mye om en del av hovedgrepene i tiltaket vi hadde presentert.<sup>25</sup> Kort sagt: at det er ulike navn, men at de gjennomgående grepene har mye til felles.

*Vi har jo vært veldig inspirert av Engebråten-modellen, og ser at disse grepene inneholder veldig mye av det samme. Jeg tror kanskje at for en del*

<sup>25</sup> <https://engebraten.osloskolen.no/fagtilbud/satsingsomrader/matematikk/>

*av de ganske nyere lærere, så er dette en kjent måte å undervise på (avdelingsleder, ungdomsskole, tiltak 2019)*

Gjennom de kvalitative intervjuene med rektorene etter siste tiltaksperiode kom det også fram andre forskjeller mellom skolene. I disse samtalen framsto rektorene ganske tydelige med hensyn til hva de mente at deres skole hadde behov for. Det kan også tenkes at lojaliteten i gjennomføringen av tiltaket ble lavere på slutten av tiltaksperioden som følge av at mange av tiltaksskolene da var i gang med å planlegge nye satsinger og tiltak.

Innslaget av skjønn synes å være større på slutten av tiltaksperioden enn det vi hadde fått informasjon om når det gjaldt de to første årene. En større åpenhet trenger selvsagt ikke bare å skyldes endringer i praksis, men kan også være at rektorene opplevde at de kunne snakke friere om det de gjorde nå som tiltaket i praksis var over. Et eksempel fra ungdomsskolene var i hvilken grad de tilpasset hvilke elever som fikk være med på tiltaket.

*Vi skjønte etter hvert at vi måtte bruke noe mer skjønn når vi satt sammen gruppene (rektor, ungdomsskole, tiltak 2019)*

Enkelte vil mene at stort innslag av subjektive vurderinger og skjønn er problematisk fordi det innebærer manglende forutsigbarhet og videre kan medføre forskjellsbehandling. Vi har imidlertid ikke grunnlag for å si at denne typen skjevheter var et problem når det gjaldt implementeringen av dette tiltaket. Bruk av skjønn var viktig fordi det tilsynelatende bidro til å øke eierskapet til lærerne, som dermed fikk noe rom til selv å gjennomføre tiltaket på den måten de synes var best.

Det var større debatt knyttet til hvilke kriterier som skulle legges til grunn for å selektere ut tiltakselevne. En oppfatning som vi fikk høre ved flere skoler var at nasjonale prøver ikke fungerte godt nok. Et moment i den sammenheng var at årets oppgave bestod av mange oppgaver med mye tekst og begreper. Tidligere forskning har vist at det er systematiske forskjeller knyttet til elevenes forutsetninger for å forstå språk og begreper som brukes på prøver i matematikk (Andresen mfl. 2017). Forskjeller i språkforståelse henger blant annet sammen med elevens sosioøkonomiske bakgrunn, men kan i en del tilfeller også tilbakeføres til om elevene og/eller foreldrene har innvandret til Norge. Siste runde med kvalitative intervjuer indikerte at skolene hadde varierende praksis med hensyn til å velge ut tiltakselever. Mens enkelte skoler forholdt seg strengt til retningslinjene de hadde fått (om å bruke resultatene fra nasjonale prøver), var det andre skoler som tok seg noe frihet til å justere gruppene ut fra hva de mente var hensiktsmessig. Det er også interessant at lærerne og rektorene vi intervjuet var forholdsvis samstemte på dette punktet. På direkte spørsmål slo de gjerne fast at det både var elever på tiltaket som ikke burde vært der, mens andre som ikke var med på tiltaket absolutt burde vært med. En typisk uttalelse var at en kandidat nok «bare hadde en dårlig dag», som en lærer sa. Og på spørsmål om andel som var feilplasserte, fikk vi høre følgende:

*Jeg tror kanskje at en av fire er feil. Enten at det er elever som burde være med eller elever som egentlig ikke trenger å være en del av tiltaksgruppa. (rektor, ungdomsskole, tiltak 2019)*

Gitt at denne rektoren forteller om en trend ved alle skolene, kan det muligens være en forklaring på hvorfor tiltaket ikke fikk større effekt enn det gjorde. Dette vil vi komme tilbake til i slutten av kapitlet.



Først må vi si litt om lærerne, som skal gjennomføre tiltaket i det enkelte klasserom. Mens rektorene kan og skal legge til rette gjennom organisering og drift av skolene, er det lærerne som praktiserer tiltaket. De er følgelig i en nøkkelposisjon med hensyn til om tiltaket blir en suksess eller ikke. Mens rektorene måtte ta stilling til om de skulle satse på dybde eller bredde i den kollegiale utviklingen, var det lærerne som måtte iverksette planene. Med bredde og dybde menes det strategiske valget om skolen skulle satse på å få noen lærere som hadde vært tiltakslærere flere ganger (dybde) eller om alle skulle få sjansen, slik at kollegiet utviklet seg mer samlet (bredde). I praksis viste det seg over tid at mange gikk fra idealer om dybde til en realitet som i beste fall sikret bredde. I begge tilfeller var det en utfordring at tiltakslærerne sluttet, byttet jobb eller var opptatt med annen undervisning. Det var derfor ikke bare et spørsmål om hva rektorene ville, men hvem de kunne sette til å gjennomføre tiltaksundervisningen. Dette kom fram i de kvalitative studiene, hvor både rektorer og lærere var opptatt av denne tematikken. I praksis endte man ofte med verken bredde eller dybde, men snarere at det var få lærere som var mulig å benytte på grunn av kravene om formell kompetanse i form av vekttall i matematikk. I enkelte tilfeller var man heldige som når en lærer som hadde vært tiltakslærer ved én skole, byttet arbeidssted og kunne da være tiltakslærer ved en annen skole.

Til det organisatoriske kan det også legges til at vi så markante forskjeller på de store gruppene. Ved én skole ble for eksempel verken økt-planer eller bøker benyttet. Dette ga lærerne mye arbeid, men ikke minst så bidro det til at selve opplegget ble litt omtrentlig i tiltaksperioden. Dette handlet om rektorenes forståelse av tiltaket, noe som først og fremst hadde konsekvenser for lærerne. I det følgende skal vi derfor se mer på hva lærerne forteller.

#### 7.4.1. Lærerne: Fra instruks til oppslutning

Så langt har vi sett på rektorene, som har ansvaret for å implementere tiltaket på skolene. Praktiseringen av tiltaket skjer imidlertid ikke på rektors kontor, men i det enkelte klasserom. Der er tiltakslæreren sjef. Dermed er det ikke sagt at de kan legge opp undervisningen fritt. Et første spørsmål er derfor hvorvidt de opplever tiltaket som et press ovenfra eller om de tar eierskap til tiltaket og gjør det til sitt eget.

I Kirkebøen mfl. (2018:23) siterer vi en lærer som høsten 2016 uttalte:

*Det gjelder å finne fram til begrunnelser i tiltaket som gjør at du kan fortsette akkurat som før. (lærer, ungdomsskole, tiltak, 2016)*

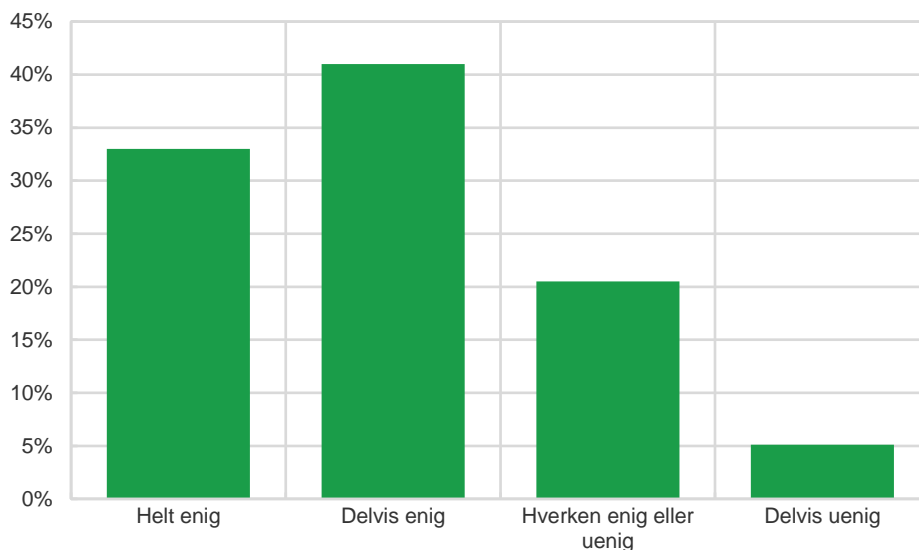
Da vi vendte tilbake til skolen to år senere, var det ingen av lærerne vi intervjuet som sa noe liknende – tvert om. De fleste ga uttrykk for at de var fornøyde med tiltaket når det gjaldt det didaktiske, det innholdsmessige og kursene. De kvantitative dataene ga et liknende bilde. Mer enn sju av ti var helt eller delvis enige i at tiltaket ville føre til at elevene presterte bedre i matematikk, mot tre av ti året før (Kirkebøen mfl. 2018:24, figur 5.5).<sup>26</sup>

I websurveyen hadde vi også et spørsmål om hvordan lærerne vurderte metodene i kurset, og hvorvidt de antok at de didaktiske metodene ville være nyttige i framtidig undervisning. Resultatene for ungdomsskolen er framstilt i figur 8.2. Fordelingen vitner om at et stort flertall (nær åtte av ti) av de tiltakslærerne som

<sup>26</sup> Det var en positiv utvikling også blant tiltakslærerne i videregående skole, emn endringen var noe mindre. Mens om lag tre av ti (på linje med andelen i ungdomsskolen) sa seg helt eller delvis enige i at tiltaket ville føre til at elevene presterte bedre i 2016, var andelen fem av ti i 2017 (Kirkebøen mfl 2018:25, fig 5.6).

svarte på undersøkelsen i 2019 er positive til metodene, og gir uttrykk for at de vil bruke dem framover.

**Figur 7.2** Påstand: «Jeg tror alle er positive til metodene vi har lært gjennom tiltaket, og det kommer til å være nyttig framover» (N=39, 2019)

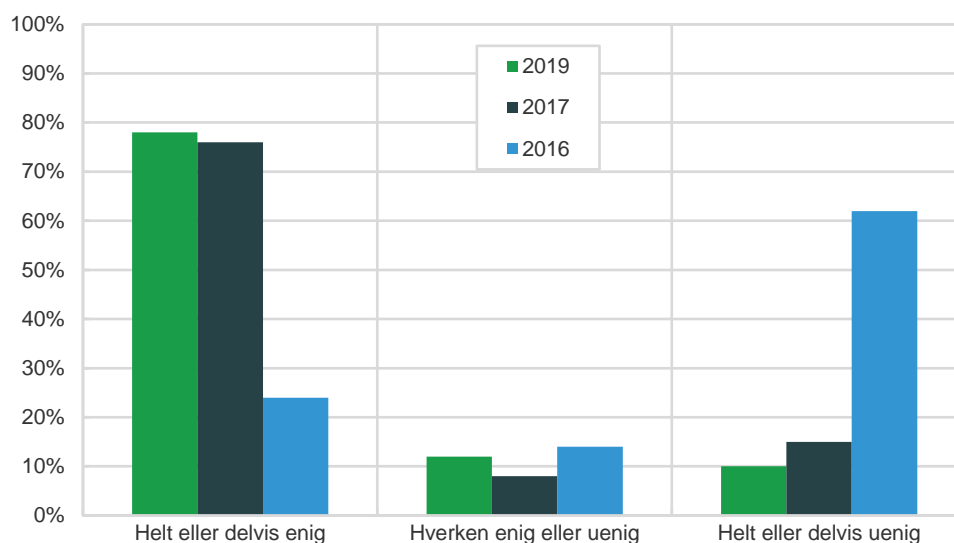


Kilde: Statistisk sentralbyrå.

Hva forteller disse tallene? For det første vitner fordelingen om oppslutning. Lærerne er gjennomgående svært positive til metoden, og ikke minst så ønsker de å bruke den framover i egen undervisning. For det andre forteller den positive holdningen at endring er mulig, noe som ikke var opplagt. I det analytiske rammeverket innledningsvis i kapitlet, refererte vi til forskning som indikerer at skolen er hva man ofte referer til som en treg institusjon, noe som gjerne resulterer i at skoleutvikling tar tid. Satt på spissen kan vi hevde at det er mer oppsiktsvekkende hvor raskt tiltaket fikk oppslutning enn at det, som et nytt tiltak, møtte noe motstand første tiltaksår. Dette funnet er det empiriske grunnlaget for det vi her har referert til som normaliseringen av tiltaket.

Når det er sagt skal det trekkes fram at tiltakslærerne, i likhet med rektorene, ga uttrykk for frustrasjon knyttet til selve gjennomføringen av tiltaket det første året. Men de viktigste frustrasjonene var betydelig redusert allerede det andre tiltaksåret (Kirkebøen mfl 2018:23). En mulig forklaring på denne endringen er at mange av skolene brukte de samme tiltakslærerne flere ganger. De hadde allerede hatt en bratt læringskurve, og visste trolig mye om hva de kunne og burde gjøre for å nå fram til elevene med de metodene som lå i tiltaket. En annen forklaring avtegnert seg i de kvalitative intervjuene og under observasjonen på kursene. Da ble det pekt på at opplegget fra Utdanningsetaten var bedret betydelig: Informasjonen var klarere og kom i tide. Kursene var også gjennomgående gode og konkretere, noe som ga lærerne større klarhet i hvordan de skulle undervise i tiltaket. Opplegget som helhet framsto følgelig mer gjennomarbeidet. At en slik utvikling fant sted underbygges av dataene som ligger til grunn for figur 8.3, hvor lærerne i ungdomsskolen ble bedt om å si seg enig eller uenig i en påstand: «Jeg har fått tilstrekkelig informasjon om hvordan tiltaket skal gjøres og dets omfang». Mens litt over to av ti sa seg helt eller delvis enige i at de hadde fått tilstrekkelig informasjon i 2016, var andelen åtte av ti i 2017 og 2019. Med andre ord: andre og tredje tiltaksår var nær like gode, ut fra lærernes vurdering, mens det første året skilte seg ut negativt.

**Figur 7.3** Påstand: «Jeg har fått tilstrekkelig informasjon om hvordan tiltaket skal gjennomføres og dets omfang». Ungdomsskole (N høst 2016=20, høst 2017=55, vår 2019=40)



Kilde: Statistisk sentralbyrå.

I samtalene med lærerne kom det fram at mange ikke bare opplevde at dette var et tiltak som de var pålagt ovenfra. I løpet av tiltaksperioden erfarte de selv at metoden både inspirerende og ga resultater. Da vi intervjuet tiltakslærere siste gang var bekymringen først og fremst knyttet til hva som vil skje når tiltaksperioden snart var over.

*De har sagt at jeg skulle sende inn økt-planene, men jeg har tatt en kopi for å si det slik. Det er jo veldig bra oppgaver her, så de må da alle bare være glade for at vi bruker videre (lærer, ungdomsskole, tiltak)*

*Men nå er det litt opp til dere. Dere kan jo bare bruke det? (Fafo)*

*Utfordringen tror jeg er å koordinere det hele. Alle lærerne må nesten være med for at vi skal få det til å gå opp. Hvordan skal vi lage mindre grupper. Jeg tror vi får det til, men kanskje ikke i fem ukers bolker. Men dette handler ikke om didaktikk, men organisasjon. Og kanskje mer organisasjon enn ressurser, noe som forutsetter ledelse. (lærer, ungdomsskole, tiltak)*

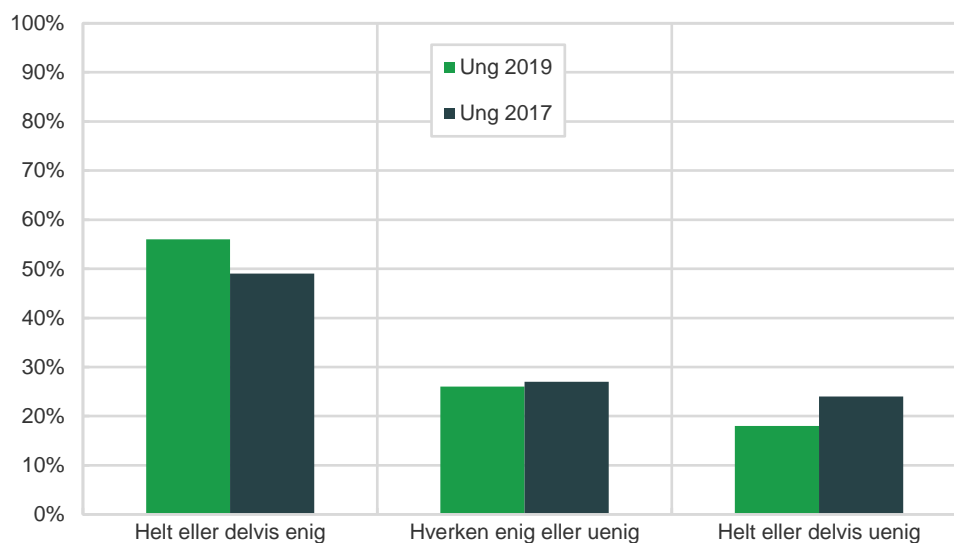
Innledningsvis skrev vi om ulike former for lojalitet. Det vi kunne se er en lojalitet som gikk i flere i retninger, knyttet til oppgaver og til kolleger. Like avgjørende er forholdet mellom pålagt og frivillig lojalitet. Våre data indikerer at det har vært en glidning fra lojalitet som var pålagt til en lojalitet som var mer frivillig. Og fra at lojaliteten var vertikal, knyttet til å løse arbeidsoppgaver man ble pålagt, ble lojaliteten også horisontal, i betydningen av at den ble kollegial (Reegård og Rogstad 2019).

Til tross for økt lojalitet, er det ikke urimelig å anta at gjennomføringen av tiltaket i enkelte tilfeller kom i konflikt med idealer om frihet til å tilpasse matematikktimene slik lærerne ønsket. I intervjuene var det flere av informantene som pekte på at selve gjennomføringen av tiltaket ble opplevd som rigid. De fikk retningslinjer for hva de skulle gjøre, hvordan det skulle undervises, og ikke minst hvilken rekkefølge temaene skulle behandles. Dette var en instruks, hvor de i tillegg opplevde at de ble kontrollert av forskere. Til tross for at det ikke var hensikten med noen form for kontroll av lærerne, var det flere som opplevde observasjonen som kontroll, noe de synes var ubehagelig. At prosjektet ble fulgt av

forskere skjønnte imidlertid de fleste nytten av, noe som trolig bidro til at de tok tiltaket og gjennomføringen på alvor. Det var derfor heller ikke uvanlig at det kom noen kommentarer knyttet til at de skjønnte at de måtte følge forskningen, men at tiltaket jo var veldig rigid. I denne delen av intervjuene ble også forskjellige lærertyper anskueliggjort: De som ble omtalt som regelryttere på den ene siden, og lærere som tilpasset undervisningen til behovene i klassen på den andre.

I ungdomsskolen var det nær halvparten av lærerne som sa seg helt eller delvis enige i at de ønsket å tilpasse timene. Denne andelen var svakt økende fra 2016 til 2019 (jf. figur 8.4).

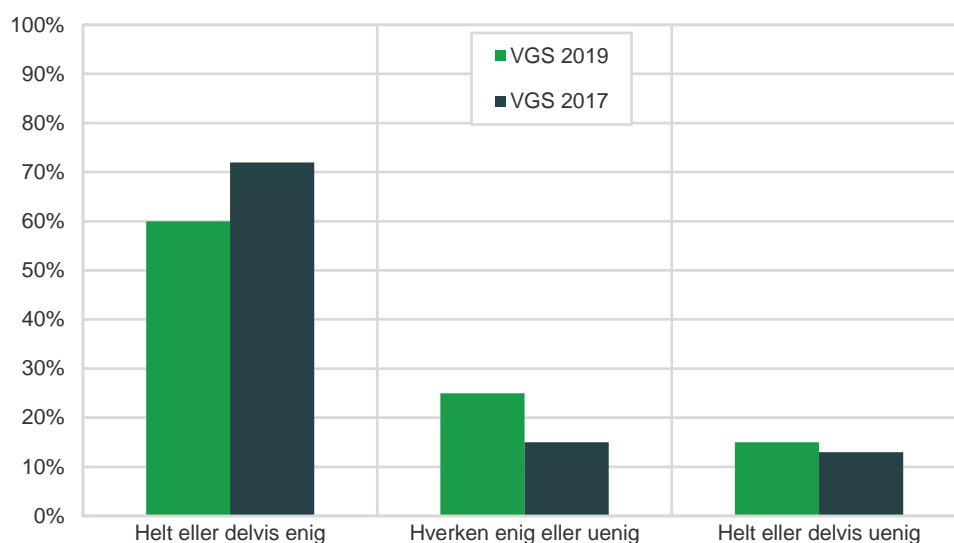
**Figur 7.4** Påstand: «Jeg ønsker frihet til å tilpasse matematikktimene selv». Ungdomsskolen. (N høst 2017=55, N vår 2019=39)



Kilde: Statistisk sentralbyrå.

I løpet av prosjektperioden avtegnet det seg et mønster som indikerte at dette ikke bare handlet om personlighetstyper, men kanskje mer at det systematisk er forskjell mellom lærertypene i ungdoms- og videregående skole. I figur 8.5 har vi satt opp oppfatningene om lærernes frihet for videregående skoler. Vi ser at andelen er noe høyere, men også at den gikk tilbake i løpet av tiltaksperioden. I siste intervjurunde var det derfor ikke så stor forskjell mellom lærere i ungdomsskole og i videregående skole, når det gjaldt andelen som sa seg helt eller delvis enig i en påstand om at de ønsket seg frihet til å styre egen undervisning.

**Figur 7.5** Påstand: «Jeg ønsker frihet til å tilpasse matematikktimene selv.» Videregående skole (N høst 2017=38, N vår 2019=20)



Kilde: Statistisk sentralbyrå.

Nå kan det tenkes at tiltaket ble opplevd som pålagt og ufritt i begynnelsen når de så tiltaket som noe de ble presset på ovenfra, mens det senere ikke ble opplevd på samme måte fordi de da trodde på tiltaket. Med tro på tiltaket fulgte også motivasjon, og derigjennom en opplevelse av at denne undervisningen var frivillig.

For å gå inn i dette problemkomplekset i de kvalitative intervjuene, spurte vi tiltakslærerne et åpent spørsmål om hvorfor de valgte å bli lærer, og hvorvidt de hadde noe lærerkall. Svarene vi fikk aktualiserer spørsmål om betydningen av eventuell seleksjon, noe som i denne sammenheng ville innebære at det er forskjellige lærere som søker seg til ungdomsskolen og videregående skole. At dette kan være tilfelle kan illustreres med to sitater fra andre runde med intervjuer i 2017:

*For meg er det viktig at alle skal lære noe. De færreste av elevene kommer til å jobbe med matematikk, men mange har bruk for matematikk i deres framtidige liv. Min oppgave er å bidra til å gjøre det litt mer forståelig for alle. (lærer tiltak, ungdomsskole 2017)*

*At en elev hos meg ble tatt med i Abel-konkurransen, er jo noe jeg synes er skikkelig gøy. Det er en bra dag på jobben (lærer tiltak, videregående skole, 2017)*

Mens mange av tiltakslærerne i ungdomsskolen la vekt på sosiale forhold og verdien av å løfte fram de svakest presterende elevene, var det flere av lærerne i videregående som framsnakket de best presterende elevene.

Vi kan selvsagt ikke trekke noen generelle slutninger basert på vårt avgrensede datamateriale. Samtidig kan funnene herfra lede til spørsmål for videre forskning. Dersom det er en slik forskjell mellom lærerne kan det muligens forstås i lys av at ungdomsskolen er obligatorisk, mens elevene selv velger videregående skole. Og i første klasse på videregående velger de mellom P- og T-matte. Mens den første er praktisk orientert er T-matte teoretisk. Alle elevene på tiltaket tok P-matte. Allerede i valget mellom P og T synes det som om en del av lærerne i matematikk mente at dette var en elevgruppe som det ikke var så faglig interessant å jobbe med.

Samtidig var det en annen dimensjon som lærere la mer vekt på, og det handlet om relasjonelle forhold, dels kollegene i mellom, dels mellom lærer og elever. Kollegialt aktualiserer tiltak spørsmål om hvordan et matematikk-team ved en skole kan utvikles best mulig. Dette spørsmålet har vi allerede tatt opp under de strategiske vurderingene en rektor må gjøre: Er det best å satse på noen enkeltlærere som blir tiltakslærere over flere år, eller snarere mer fordelaktig å bruke flere lærere, slik at flere får gå på kurs og lære om tiltaket direkte? Denne todelingen var likevel mer teori enn praksis. Det skjedde noe utskifting av lærerne i løpet av de tre årene tiltaket varte, så de fleste rektorene hadde derfor begrenset mulighet til å velge å ha noen særlig dedikerte tiltakslærere.

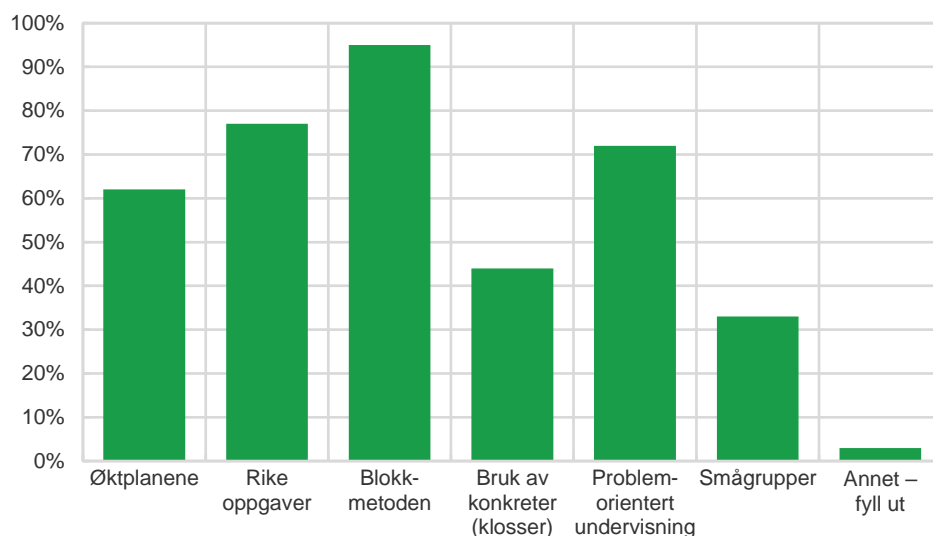
Et annet element er betydningen av lærerens kjennskap til elevene som er med på tiltaket. Flere av våre informanter pekte på at det var stor fordel å kjenne elevene fra den ordinære undervisningen. Dette handlet i noen grad om å vite hva de slet med i matematikken. Viktigere synes det å være å ha kjennskap til hvem de var som personer. Mest avgjørende var likevel det å sikre gode arbeidsforhold og å holde ro i klassen.

*Jeg har to små grupper. En hvor jeg kjenner elevene og en hvor jeg knapt vet hvem de er. Det er stor forskjell. De jeg kjenner stoler på meg og jeg vet hvordan jeg skal nærme meg dem. De jeg ikke kjenner. Hva skal jeg si. Det er liksom ekstra viktig å ha en fortrolighet når vi skal ta opp en annen type undervisning. Aller viktigst er det likevel at det er mye lettere å skape en ro i en klasse du kjenner enn i en klasse hvor du ikke kjenner elevene. (lærer tiltak, ungdomsskole 2017)*

Betydningen av å skape ro, og gjennom dette betydningen av hvor mange det skulle være på gruppene, var et tilbakevendende tema. Samtidig var det ikke noe ønske om at det skulle være veldig få elever. I intervjuene ble vi fortalt at de med «veldig få» mente ned mot fem elever. Det ble sagt at man gjerne trengte å nå over en kritisk masse på seks til åtte elever for å legge til rette for gode diskusjoner, ha læringspartnere og ikke minst var dette viktig for å gi elever litt frirom mens de var inne i timene. Det sistnevnte ble vi fortalt var viktig ettersom tiltakselevne gjennomgående var lite motivert for matematikk i utgangspunktet. Når det ble satt søkelys på dem og matematikk særskilt var det umulig å koble litt ut dersom det var veldig få elever til stede.

Avslutningsvis kan man spørre om hvordan tiltaket vil bli brukt i framtidig undervisning. Hva vil lærerne ta med fra tiltaket i sitt videre virke som lærere i matematikk? I figur 8.6 har vi satt opp en oversikt over hvilke deler lærerne i ungdomsskolen svarer at de vil ta med videre. Det er verdt å kommentere at ingen av de spurte lærerne krysset av på svaralternativet «ikke noe. PBG er ikke aktuelt for oss». Av figuren framgår det at en stor andel av tiltakslærerne vil ta med seg sentrale elementer. En særlig stor andel, over ni av ti, har for eksempel krysset av på at de vil fortsette med å bruke blokkmetoden. Det kan her legges til at dette var også noe av det elevene husket best, når vi ba dem fortelle oss om hva de lærte på tiltaket.

**Figur 7.6** Påstand: «Nå er tiltaksperioden over. Hva fra PBG vil dere ta med i videre matematikkundervisning ved skolen?» N=95



Kilde: Statistisk sentralbyrå.

Utover blokkmetoden, er det også en betydelig andel (mer enn 50 prosent) som har krysset av på økt-planene, rike oppgaver og problemorientert undervisning. Bruk av konkrete, som byggeklosser og smågrupper, er det derimot noe mindre andel som svarer bekreftende på at de vil ta med seg videre. Smågrupper er trolig ikke så overraskende gitt at dette er en beslutning som ikke kan fattes av den enkelte lærer. Det krever ekstra ressurser i form av flere lærere, eventuell en organisering av undervisningen som gjør dette mulig. Bruk av konkrete var noe problematisert. Blant annet var det tiltakselever i videregående skole som mente at bruk av konkrete var for barnslig. Det ble blant annet gitt uttrykk for at de følte at de ble sett på som dumme – at det var babyoppgaver (jf. Kirkebøen mfl 2018:28).

Et spørsmål er hvorvidt dette var en opplevelse som bare gjaldt bruk av konkrete. For å få mer innblikk i hvordan elevene opplevde tiltaket, skal vi i den neste del se nærmere på deres vurderinger.

#### 7.4.2. Elevenes opplevelser

Vi vil her konsentrere oss om elevene i ungdomsskolen ettersom flere av tiltakselevne i videregående skole ikke selv var klar over at de var med på tiltaket. Det skal også legges til at elevene utelukkende ble intervjuet kvalitativt, og med bruk av gruppeintervjuer. Det sistnevnte var fordelaktig med hensyn til at elevene da kunne ha en refleksjon om hva de hadde fått ut av tiltaket og hvordan det eventuelt kunne ha fungert enda bedre. Ettersom vi snakket sammen i grupper fikk alle muligheten til å respondere på andres opplevelser og erfaringer. Dette var heller ikke intervjuer hvor det ble tatt opp veldig sensitive temaer. Alle de som var til stede hadde til felles at de slet med matematikken, og at de hadde blitt selektert til å være med på et særskilt tiltak. Ettersom alle hadde en felles erfaring som tiltakselever var det heller ikke forbundet med stigmatisering å bli intervjuet.

Erfaringene kan deles i to, med trivsel og opplevelse av mestring på den ene siden og faren for stigma og organisatoriske utfordringer, knyttet til det å være utenfor ordinær undervisning, på den andre. I første delrapport tematiserte vi at elevene gjennomgående var positive til å være med på tiltaket (Kirkebøen mfl. 2018:27). Kort oppsummert kan dette beskrives som den gjeldende situasjonen alle de tre årene vi utførte intervjuer av tiltakselever. Det er kanskje ikke så overraskende, gitt at tiltaket er en måte å bli sett og møtt av skolene. Spørsmålet er om det også var

andre årsaker til at tiltakselevne var positive. Et tema var små grupper, men enda viktigere var at dette var grupper hvor alle var på omtrent samme nivå.

*Det er så bra å være på en liten gruppe. Da har lærerne bedre tid til hver av oss og kan forklare mye grundigere enn ellers (elev 1, ungdomsskole, tiltak 2019)*

*Og så føler jeg meg ikke dum når jeg spør om noe (elev 2, ungdomsskole, tiltak 2019)*

*Hva er viktigst, syns dere. At gruppene er små eller at alle er på omtrent samme nivå? (Fafo)*

*At alle kan like lite som meg, ha, ha (elev 1 ungdomsskole, tiltak 2019)*

*(Flere nikker og sier seg enig)*

Etter denne samtalen, spurte vi om forholdet mellom gruppestørrelse og mestringsnivå. Og svarene denne gruppen ga synes å være gjennomgående for dem vi intervjuet. Det å slippe å føle seg dum bidro til at elevene turte å åpne seg opp og tenke matte igjen. De kunne diskutere det de ikke fikk til i stedet for å bruke mye energi på å skjule sine fraværende ferdigheter. Dette ble omtalt som en stor lettelse.

*Det er ikke noe gøy når alle ler av deg. Da prøver jeg heller å la være å si noe. Sånn er det ikke på gruppa. (elev 4, ungdomsskole, tiltak 2019)*

I intervjuene kom det også fram at mange av tiltakselevne opplevde gruppa de var på som veldig hyggelig sosialt. Det var derfor flere som ønsket seg at gruppene skulle være permanente. I det minste håpet de å kunne fortsette å ha tiltaksperioder.

Et tema som vi tok opp med lærere og elever, var hvorvidt de trodde at et avgrenset tiltak kunne ha stor betydning. Blant lærerne var det flere som la vekt på betydningen av å knekke en kode, som å gå opp et trinn i forståelsen. I denne fortellingen ble det lagt vekt på at de mente at matematikk var ulikt andre fag, som i større grad var modningsfag. I matematikk, ble vi forklart, kunne en elev gå ut av en time og faglig sett være på et helt annet nivå enn før timen. At det i beste fall kunne være som å «knekke koden med å svømme. Du går aldri tilbake». Også elevene var opptatt av mestring, og at tiltaket kunne gi dem opplevelsen av å knekke koder.

*For første gang opplevde jeg å få til matte. Jeg var jo egentlig blant de beste på kurset. Skulle egentlig ikke vært der. Men så er det litt digg da. Plutselig å få til alt. Kunne forklare de andre hva de skal gjøre (elev ungdomsskole, tiltak 2019)*

Da vi spurte lærerne om denne typen av elever ble vi fortalt at tiltaket ofte var særlig positivt for dem. Delvis fordi de fikk en type «relativ mestringsopplevelse» under tiltaksperioden. Gjennom å sammenlikne seg med de andre tiltakselevne i stedet for klassen opplevde de plutselig å være gode, relativt sett. Det er lett å tenke at denne mestringsopplevelsen kun var en korttidseffekt, som forsvinner straks de er tilbake i den ordinære undervisningen. For en del av tiltakselevne skjedde sikkert også det. Men både elever og lærere fortalte også om andre erfaringer. At det var elever som hadde opplevd mestring på tiltaket, og som så på seg selv på en fornøyd måte etterpå.



*Det blir litt som ringer i vannet. De gode erfaringene fra tiltaket tar de med seg. Jeg synes det var litt rart, men jeg tror alle som har opplevd mestring, som har sett at det kan gå, at du kan få det til, tar det med seg over i de ordinære timene (lærer ungdomsskole, tiltak 2019))*

*For meg ga kurset en plutselig opplevelse av mestring. Og jeg vil ikke tilbake (elev ungdomsskole, tiltak 2019)*

Mange av deltakerne opplevde altså at kurset hadde en positiv effekt for dem. Vi ble også fortalt at det særlig var rimelig å vente at 2-erne ville få utbytte av kurset, mens det tilsynelatende er langt vanskeligere å løfte svake 1-ere, ble vi fortalt. De kvantitative analysene indikerer imidlertid at dette ikke er tilfelle. Vi har ikke data til fullt ut å forklare dette. En mulighet kan likevel være at nivået på undervisningen evnet å trekke med de svakest presterende, mens det utfordret de best presterende elevene i for liten grad.

Uavhengig av hvor godt man presterer, var et tema for oss om det er slik at det å være med på tiltaket ble opplevd som stigmatiserende i miljøet på skolen. Noe som underbygget en slik antakelse var uttalelser det første året, hvor vi ble fortalt at tiltakselever i noen sammenhenger var blitt omtalt som forskerbarna. Senere hørte vi imidlertid ikke noe slikt igjen. Selvsagt trenger ikke det å bety at stigmatiseringen var fraværende. Men tiltaket var trolig mer kjent. På spørsmål om det var belastende å ikke være en del av den ordinære undervisningen en periode, framsto forskerne tidvis mer bekymret enn elevene.

*Det er elever som går inn og ut på mindre grupper hele tiden. Det går egentlig ganske bra. (Elev ungdomsskole, tiltak 2019)*

*Jeg tenker egentlig ikke så mye på det (elev ungdomsskole, tiltak 2019)*

Elevene var mer kritiske til organiseringen av tiltaket. En viktig innvending handlet om at de mistet undervisning mens de var med på tiltaket. Mens tiltakselevne gikk sakte fram i pensum, gikk de ordinære klassene raskere fram enn tidligere. Sagt på en annen måte: Det at de svakest presterende elevene ble tatt ut av klassen, innebar samtidig at det ble dannet en ny gruppe med de gjenværende elevene. Denne nye gruppen bestod av færre elever enn vanlig. Og trolig viktigere: Blant dem var det kun middels og høyt presterende elever, som i mange tilfeller også var mer motiverte for matematikk. Det er ikke urimelig at dette hadde noen konsekvenser for læringsmiljøet. Mest påfallende var det likevel når tiltakselever møtte til tentamen til jul, som var umiddelbart etter den første tiltaksperioden, men hvor de ble prøvd i emner de ikke hadde vært gjennom på skolen. Tiltaket bidro følgelig til at elevgruppen stilte ulikt til start på tentamensdagen. Det kan argumenteres for at de i mange tilfeller neppe hadde klart oppgavene innenfor det aktuelle emnet, men det forhindrer ikke at deres opplevelse var at situasjonen var ugrei. Vi fikk også vite at enkelte foreldre hadde engasjert seg i denne tematikken, noe som samlet utfordret tiltakets legitimitet.

Da vi spilte inn disse synspunktene tidlig i prosjektet, opplevde vi at dette ble tatt på stort alvor i Utdanningsetaten. I siste runde var det også flere av skolene som hadde tatt høyde for denne situasjonen. Blant annet var det skoler som ga en egen tentamen til elevene som hadde vært med på tiltaket. Samlet viser denne erfaringen både at det er nyttig med følgeforskning som gir feedback for mulige justeringer underveis, og ikke minst at man trenger informasjon som ikke framkommer ved utelukkende å fokusere på de rene utfallsmålene.

Et siste moment vi vil nevne kort er knyttet til det didaktiske opplegget og hva det gjør med læringsmiljøet. Flere av elevene i tiltaksgruppa var lite motivert for skole generelt og matematikk spesielt. Støy og uro var derfor et moment som ble trukket fram som en utfordring av både elever og lærere.

*Når selve opplegget er at vi skal prate sammen blir det jo en del bråk i timene. Det er litt irriterende. Ganske mange starter med å snakke om matte, men det går raskt over til å handle om andre ting. (elev ungdomsskole, tiltak 2019)*

Det samme sies av en lærer, men da med litt andre ord:

*PBG-metoden fremmer samtale, noe som fører til en støy i en del tilfeller. Jeg syns det tidvis har vært vanskelig å følge metoden, men samtidig sikre et godt læringsmiljø (lærer ungdomsskole, tiltak 2019)*

Det er ikke noe enkelt svar på utfordringen knyttet til læringsmiljø. Det kan tenkes at det er en trade-off mellom metoden som skal benyttes og å sikre læringsmiljøet. I praksis fortalte elever om lærere som tok noen grep for å sikre at de fleste av tiltakselevne fikk så godt utbytte av tiltaket som mulig.

## 7.5. Hvordan nyttiggjøre erfaringene fra tiltaket i framtiden?

I denne siste delen vil vi ta opp to spørsmål: For det første om det er mulig å identifisere noen kjennetegn ved de skolene som i særlig grad har lyktes. For andre hva som skjer framover, etter at tiltaket snart er over. Før siste runde med kvalitative intervjuer ble det besluttet i prosjektets styringsgruppe at Fafo skulle se særskilt på skoler hvor tiltaket synes å være vellykket, basert på foreløpige analyser av første og andre tiltaksår. Vi valgte da å se særskilt på seks ungdomsskoler i det tredje tiltaksåret (se under metode i dette kapitlet).

Hva var det med disse skolene, og er det forhold som lar seg overføre til andre skoler? Etter å ha gått gjennom materialet er det vanskelig å trekke noen entydige slutninger om forhold som skiller disse skolene fra andre. Det som derimot er mulig, er å peke på elementer de har lagt særlig vekt på, og av den grunn kan antas å være sentrale.

- Godt utarbeidet kursmaterieell som kommer i tide
- En leder som prioriterer tiltaket, og gir lærerne tid til å gå på kurs og drive fagutvikling på skolen
- Forankring gjennom hele organisasjonen
- Lærere som slutter opp om tiltaket – ikke pålagt, men som har en indre motivasjon
- Lærere som fokuserer på å løfte de svakest presterende elevene

Det er ikke utenkelig at også skolene som ikke har lyktes like godt vil si at de også prioriterte en del av disse elementene. Trolig er spørsmålet heller ikke hvilke deler som blir prioritert, men med hvilken styrke innsatsen gjøres. Som vi har pekt på tidligere i kapitlet, var det stor frustrasjon første tiltaksår fordi undervisningsmateriellet ikke kom til skolene før etter at tiltaket hadde startet. Dette skapte umiddelbare utfordringer, men like viktig var muligens lærernes opplevelse av at tiltaket ikke var tilstrekkelig gjennomarbeidet og planlagt før det ble igangsatt. Dette var en erfaring som utfordret lojaliteten i begynnelsen av prosjektet (Reegård og Rogstad 2019).

I tiltaksår to og tre var informasjonen og kursene gode, dessuten kom materiellet i god tid før tiltaket startet. I siste tiltaksår hadde tiltaket bred oppslutning, og både lærere og elever ønsket at det skulle vare videre. På direkte spørsmål i web-

surveyen svarte en betydelig andel av tiltakslærerne at de ville bruke oppgaver og grep fra tiltaket i videre undervisning. Men det er ikke gitt at dette vil bli så enkelt. Flere av elementene i tiltaket fordrer planlegging, organisering og prioritering. Et eksempel er bruk av grupper, noe som synes som en viktig forutsetning for å gjennomføre tiltaket (jf. foregående punkt hvor vi sammenliknet tiltaket på videregående skole og ungdomsskolen).

Vi har allerede påpekt at fra første til andre år var det utvikling fra implementering til stabilisering. År tre har vi her beskrevet som en type normalisering. Med det mener vi at en stor andel av lærerne nå ser de didaktiske prinsippene som en selvfølgelig del av undervisningen. Spørsmålet nå er hvorvidt det framover blir en avvikling eller fortsatt normalisering. Gitt manglende direkte effekter, kan det argumenteres for at de didaktiske grepene i tiltaket bør avvikles. Samtidig er det ikke tvil om at de involverte aktørene har stor tro på prosjektet og opplever at det er meningsfullt. Vi spurte derfor lærerne om hva de mente var avgjørende for en videreføring. To forhold som gikk igjen var kursing og organisering. Behovet for faglig påfyll gjennom kurs ble trukket fram som avgjørende – både for å vedlikeholde kompetansen og fordi det er betydelig turnover blant lærerne. Med bruk av webbaserte løsninger trenger heller ikke slike kurs å være veldig kostbare.

Sistnevnte poeng handler også om prioriteringer på den enkelte skole. Flere lærere vi intervjuet fortalte at skolene allerede var i gang med planlegging av nye tiltak. I enkelte tilfeller var dette rettet mot andre fag, som skolen nå mente det var på tide å prioritere. På andre skoler ble det arbeidet på flere prosjekter samtidig.

Spørsmålet er hvorvidt det er hensiktsmessig å sette i gang et nytt tiltak i matematikk eller om det hadde vært bedre å justere og tilpasse et tiltak mange har tro på og kjenner godt til. I utgangspunktet er det nærliggende å konkludere med at man ikke bør fortsette med et tiltak med begrenset effekt. Samtidig kan det likevel være at det er viktige grep fra tiltaket som bør ivaretas. Vi har tidligere referert til Engebråten-modellen, som har dokumenterte effekter på læringsutbytte. De didaktiske prinsippene var også inspirert av erfaringer internasjonalt, gjerne referert til som Singapore-skolen. Det kan derfor hevdes at det er for tidlig å avslutte å bruke de didaktiske grepene som ligger i tiltaket. Man kan i stedet argumentere motsatt, og peke på at tiltaksperioden kunne bli utvidet og at grepene ikke bør reserveres for de lavest presterende elevene. Dette er imidlertid ikke noe vi kan trekke slutninger om basert på erfaringer med dette tiltaket.

## 7.6. Avslutning

Dette kapitlet har handlet om aktørenes erfaringer med å implementere tiltaket. Kapitlet har vært organisert ut fra tre problemstillinger, hvorav det første legger det empiriske grunnlaget for å besvare de to andre.

- (i) Hvilke erfaringer har rektorer, lærere og elever med å implementere og praktisere tiltaket?
- (ii) Hvorfor har ikke tiltaket hatt større direkte effekter (jf. kapittel 5 og 6)?
- (iii) Hvordan skal og kan skolene forvalte kompetansen de har fått når satsingsperioden er over?

Det empiriske materialet som er lagt til grunn i kapitlet har bestått av en websurvey til rektorer og lærere, samt kvalitative intervjuer med rektorer, lærere og elever. Vi har hentet inn informasjon fra alle tre tiltaksårene.

De innsamlede dataene indikerer at det har vært en stor endring i aktørenes vurdering av tiltaket – fra implementering til stabilisering og normalisering. Dette er tre begreper som kan brukes for å betegne hvert av de tre tiltaksårene. Mens

første året ble oppfattet som noe rotete, med en oppstart som skjedde før skoleeier og skolene var klar for å gå i gang, var situasjonen en helt annen allerede fra andre tiltaksår. Da var materiellet klart i tide og kursenes kvalitet var betydelig bedre. Dette resulterte i bred oppslutning fra rektorer og lærere. Også elevene ga uttrykk for at det å være med på tiltak var gunstig for deres utvikling og motivasjon for faget. Både i de kvantitative og de kvalitative dataene framgår det at lærerne antar at elevene vil prestere bedre i matematikk. Dessuten forteller lærerne at de vil ta med elementer i de didaktiske grepene i framtidig undervisning.

## 8. Sammenstilling av resultater

Evalueringen av tiltaket baserer seg på forskjellige typer analyser: Effektevaluering, kvalitativ prosessevaluering, samt analyser av undervisningen. De to første er diskutert i de foregående kapitlene. Analyser av undervisningen, basert på klasseromsobservasjoner og lærernes egen beskrivelse av undervisningen i websurveyer, ble presentert i underveistrapporten (Kirkebøen m.fl., 2018). I dette kapitlet ser vi de forskjellige analysene i sammenheng. En viktig motivasjon for å kombinere forskjellige evalueringsmetoder har vært at disse kan utfylle hverandre. Effektevalueringen er egnet til å beskrive effekter på forskjellige resultatmål. Deltakernes egen opplevelse gir et ikke-fullstendig bilde av faktisk effekt. Effektevalueringen gir imidlertid mindre innsikt i hva som har fungert godt eller dårlig, og hvordan det er evt. er mulig å forbedre gjennomføringen. Dette kan vi lære mer om gjennom tettere observasjon og dialog med deltakerne.<sup>27</sup>

Dersom vi fant tydelige forskjeller i effekt mellom grupper av skoler eller elever ville det vært et godt utgangspunkt for å vurdere hva som skulle til for at tiltaket virket. Når vi ikke finner slike forskjeller (jf. omtale av heterogene effekter i tidligere kapitler) kan vi ikke trekke sterke konklusjoner. Basert på de forskjellige analysene kan vi likevel diskutere hva som kan forklare våre funn, eller hva som kunne bidratt til større effekter, men dette blir mer å betrakte som hypoteser enn som klare konklusjoner.

Tidligere forsøk med ekstra ressurser/smågrupper har ikke gitt effekter på elevenes resultater (jf. oppsummeringen i kapittel 5). Det er derfor liten grunn til å vente effekter på elevenes resultater dersom det ikke har vært en effekt på undervisningen. Som diskutert i kapittel 2 har flere tidligere studier funnet at det er krevende å endre måten lærerne underviser på.

Hvorvidt lærerne tar i bruk metoder og prinsipper fra tiltaket vil blant annet avhenge av hvorvidt de oppleves som velbegrunnede og nyttige. Fra implementeringsstudien, diskutert i kapittel 7, vet vi at lærerne opprinnelig var positiv til motivasjonen for tiltaket, men at de ikke opplevde å få tilstrekkelig informasjon. I andre og tredje tiltaksår var informasjonen tilstrekkelig, lærerne er generelt positive til innholdet i tiltaket, og uttrykker at de ønsker å benytte flere tiltakselementer i videre undervisning.

I Kirkebøen m.fl. (2018) diskuterer vi undervisningen i tiltaksskolene i mer detalj, basert på observasjon av undervisningsøkter og lærernes svar på spørreundersøkelsen. Observasjonene viser at det er betydelige variasjon i tiltaksundervisningen, både mellom økter og i hvilken grad de forskjellige elementene i tiltaket er til stede.. For eksempel er det utstrakt bruk av *samtale som støtter forståelse*, mens lærerne i mindre grad *skaper sammenheng mellom økter*. Basert på observasjonene var tiltakselementene i større grad til stede i det andre tiltaksåret enn i det første tiltaksåret. Samlet sett var det et betydelig innslag av elementer fra tiltaket i undervisningen, men undervisningen gjenspeilet heller ikke fullt ut prinsippene i tiltaket. Vi har ikke tilstrekkelig observasjoner fra tredje tiltaksår til å vurdere utviklingen fra år to til år tre. Lærernes svar i spørreundersøkelsene gir et lignende bilde. Lærerne oppgir å bruke de forskjellige metodene fra tiltaket, men i noe varierende grad, og ikke i alle timene.

Effektevalueringen baserer seg på å sammenligne resultater i tiltaks- og kontrollskolene. Vi vil ikke finne noen effekt dersom vi ser samme utvikling i tiltaks- og kontrollskolene. Vi har ikke hatt kapasitet til å gjennomføre omfattende

<sup>27</sup> En ytterligere motivasjon har vært å få et løpende innblikk i gjennomføringen av tiltaket. Resultater fra klasseromsobservasjoner og prosessevalueringen har gitt informasjon underveis som har blitt brukt til å gjøre løpende justering i organisering og kursing.

observasjoner av undervisningen i kontrollskolene. Og de begrensede observasjonene som er gjennomført tyder på at det er stor variasjon i undervisningen, med noe innslag av tiltakspreget undervisning. Ergo kan vi ikke konkludere med at de didaktiske prinsippene og metodene er eksklusive for tiltaket. I Kirkebøen m.fl. (2018) diskuterer vi forskjeller i undervisning mellom tiltaks- og kontrollskoler, basert på spørreundersøkelsene til lærerne. Det er krevende å sammenligne på en god måte, ettersom tiltaks- og kontroll lærere ikke nødvendigvis har samme vokabular for å beskrive undervisningen. Når vi stiller konkrete spørsmål oppgir imidlertid tiltakslærerne å ha mer undersøkende undervisning, der de bruker lengre tid på flere løsningsstrategier for noen få oppgaver, og at flere elever er aktive i undervisningen.

Selv om prinsippene fremmet gjennom tiltaket verken er fullt gjennomført i tiltaksskolene eller eksklusive for tiltaksskolene, tyder forskjellene i lærernes svar på at tiltaket har ført til en forskjell i hvordan undervisningen gjennomføres. Vi er imidlertid ikke i stand til å gi en eksakt eller uttømmende beskrivelse av hvor stor forskjellen er. I den grad vi har kartlagt undervisningen i kontrollskolene har vi også gjort det med utgangspunkt i didaktikken i tiltaket. Vi vet fra prosess-evalueringen at det til enhver tid er flere pågående satsninger. Det er mulig at kontrollskolene har deltatt i eller gjennomført andre, omtrent tilsvarende effektive satsninger.

Implementeringsstudien peker på endringer over tid i opplevelsen og gjennomføringen av tiltaket. I effektevalueringen kan vi ikke konkludere med at det er statistisk signifikante forskjeller mellom år, men vi ser tegn til høyere effekt i andre tiltaksår enn i tredje. Implementeringsstudien finner at tiltaket i denne perioden «normaliseres»: Det blir mindre fokus på tiltaket som sådan, og det gjøres i større grad lokale tilpasninger. Det er mulig at disse endringene har ført til lavere faktisk eller målt effekt. Faktisk effekt kan ha blitt redusert i mindre grad ble preget av tiltakselementer som var effektive. Målt effekt kan ha blitt redusert dersom det skjer endringer blant kontrollskolene, som gjør at disses undervisning blir likere tiltaksundervisningen.

Implementeringsstudien peker også på at flere skoler byttet tiltakslærere underveis samt generell turnover blant lærere fra det ene skoleåret til det andre. Dette kan påvirke faktisk effekt, fordi det blir vanskeligere å gjennomføre tiltaksundervisningen. Lærere som bytter jobb kan også påvirke målt effekt, dersom lærere går fra tiltaks- til kontrollskoler, og dermed påvirker undervisningen i kontrollskolene. Læreres jobbskifter kan undersøkes videre med administrative data, men har ikke vært mulig innen rammene av denne rapporten.

## Referanser

- Arvidson, Markus & Axelsson, Jonas (2014). Lojalitetens sociala former om lojalitet och arbetsliv. *Arbetsmarknad & Arbetsliv*, 20(1), s. 55–64.
- Aughinbaugh, A. (2012). The effects of high school math curriculum on college attendance: Evidence from the NLSY97. *Economics of Education Review*, 31(6):861–870.
- Cook, P. J., Dodge, K., Farkas, G., Fryer, J. R. G., Guryan, J., Ludwig, J., Mayer, S. (2015). Not Too Late: Improving Academic Outcomes for Disadvantaged Youth. University of Chicago, Northwestern University Working Paper Series WP-15-01.
- Cortes, K., Goodman, J., & Nomi, T. (2014). Intensive math instruction and educational attainment: Long-run impacts of double-dose algebra. HKS Faculty Research Working Paper
- Durlak, Joseph A. & DuPre, Emily P. (2008). Implementation matters: review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3), s. 327–350.
- Eielsen, G., Kirkebøen, L., Leuven, E., Rønning, M., & Raaum, O. (2013). Effektevaluering av intensivoppøringen i overgangsprosjektet, Ny GIV. Rapport 54/2013. Statistisk sentralbyrå
- Faragher, R., Brady, J., Clarke, B., & Gervasoni, A. (2008). Children with Down syndrome learning mathematics: can they do it? Yes, they can! *Australian Primary Mathematics Classroom*, 13(4), 10–15.
- Fixsen, Dean L., Naoom, Sandra, Blase, Karen A., Friedman, Robert M. & Wallace, Frances (2005). *Implementation research: A synthesis of the literature*. Florida: University of South Florida.
- Forgazs, H. (2010). Streaming for mathematics in years 7-10 in Victoria: An issue of equity? *Mathematics Education Research Journal*, 22(1), 57-90.
- Boaler, J. (2011). Changing students' lives through the de-tracking of urban mathematics classrooms. *Journal of urban mathematics education* 4(1) 7-15.
- Fryer R.G. (2017) The Production of Human Capital in Developed Countries: Evidence from 196 Randomized Field Experiments. In: *Handbook of Field Experiments*. Vol. 2. Amsterdam: North-Holland. pp. 95-322.
- Fullan, M. (2001, third edition). *The NEW Meaning of Educational Change*. London: Routledge Falmer.
- Goodman, J. (2019). The Labor of Division: Returns to Compulsory High School Math Coursework. *Journal of Labor Economics*, 37(4).
- Gvozdic, K., Sander, E. (2018). When intuitive conceptions overshadow pedagogical content knowledge: Teachers' conceptions of students' arithmetic word problem solving strategies. *Educational Studies in Mathematics* 98, 157–175.
- Hargreaves, A., Fullan, M. (2012). *Professional Capital: Transforming Teaching in Every School*. New York: Teachers college press.
- Heyd-Metzuyanim, E., Smith, M., Bill, V., Resnick, L.B. (2019). From ritual to explorative participation in discourse-rich instructional practices: a case study of teacher learning through professional development *Educational Studies in Mathematics* 101, 273–289.

- Huitfeldt, L.J. Kirkebøen, S. Strømsvåg, G. Eielsen og M. Rønning (2018) Fullføring av videregående opplæring og effekter av tiltak mot frafall: Sluttrapport fra effektevalueringen av Overgangsprosjektet i Ny GIV, Rapport 2018/08. Statistisk sentralbyrå.
- Joensen, J. S. and Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources*, 44(1):171–198.
- Kaur, B. et al. (2015). The Mathematics Education in Singapore, p. 311-316. In: Sung Je Cho (Ed.), *The Proceedings of the 12th International Congress on Mathematical Education Intellectual and Attitudinal Challenges 8 July – 15 July 2012, COEX, Seoul, Korea*. Springer.
- Kazemi, E., Wæge, K. (2015). Learning to teach within practice-based methods courses. *Mathematics Teacher Education and Development* 17(2), 125–145.
- Kirkebøen, L.J., G. Eielsen, M. Rønning, S. Strømsvåg, S. Andresen, K. Reegård, J. Rogstad, J.E. Berge og L. Lindenskov (2018). Matematikdidaktisk etterutdanning av lærere og målrettet strukturert matematikkundervisning ved overgang til 8. trinn og VG1. Foreløpig beskrivelse av utforming og gjennomføring av tiltak. Notater 2018/15. Statistisk sentralbyrå.
- Kirkebøen, L.J., A. Kotsadam, O. Raaum, S. Andresen og J. Rogstad (2017) Effekter av satsing på økt lærertetthet, Rapport 2017/39. Statistisk sentralbyrå.
- Leuven, E. og S. Løkken (2020) Long term impacts of class size in compulsory school. *Journal of Human Resources* 55 (1). 309-348.
- Leuven, E, H. Oosterbeek og M. Rønning (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *Scandinavian Journal of Economics* 110 (4). 663-693
- Lipsky, Michael (1980). *Street-level Bureaucracy: Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation.
- Okazaki, M., Okamoto, K., Morozumi, T. (2019). Characterizing the quality of mathematics lessons in japan from the narrative structure of the classroom: “mathematics lessons incorporating students’ ‘questions’ as a main axis” as a leading case Japan. *Hiroshima journal of mathematics education* 12, 49-70.
- Reegård, K. & J Rogstad (2019). Lojalitet i spill: Eierskap i implementering av et matematikdidaktisk tiltak i Oslo-skolen. *Norsk pedagogisk tidsskrift*, nr 4 2019.
- Rose, H. and Betts, J. R. (2004). The effect of high school courses on earnings. *The Review of Economics and Statistics*, 86(2):497–513.
- Rønning, W., Hodgson, J., Tomlinson, P. (2013). Å SE OG BLI SETT. Klasseromsobservasjoner av intensivopplæringen i Ny Giv. NF-rapport Nr. 6.
- Westergård, Elsa & Roland, Pål. (2015). Ulike perspektiver på implementering. I Roland, Pål & Westergård, Elsa (red.). *Implementering*.



## Figurregister

Figur 5.1	Fordeling av skalapoeng (NP8) og mestringsnivå – høst 2017.....	35
Figur 6.1	Gjennomføring av læringsstøttende prøve på skolenivå (andel som har gjennomført prøve) .....	50
Figur 6.2	Gjennomføring av Oslo-prøve på skolenivå (andel som har gjennomført prøve) .....	50
Figur 7.1	Påstand: «Det har vært vanskelig å motivere lærerne til å delta på kurs». Rektor ved tiltaksskoler. Videregående opplæring og ungdomsskole. (N vgs=9, N ung=26) 2019 .....	61
Figur 7.2	Påstand: Jeg tror alle er positive til metodene vi har lært gjennom tiltaket, og det kommer til å være nyttig framover (N=55, 2019).....	64
Figur 7.3	Påstand: Påstand: "Jeg har fått tilstrekkelig informasjon om hvordan tiltaket skal gjennomføres og dets omfang". Ungdomsskole (N høst 2016=20, høst 2017=55, vår 2019=40).....	65
Figur 7.4	Påstand: "Jeg ønsker frihet til å tilpasse matematikktimene selv". Ungdomsskolen. (N høst 2017=55, N vår 2019=39).....	66
Figur 7.5	Påstand: «Jeg ønsker frihet til å tilpasse matematikktimene selv.» Videregående skole (N høst 2017=38, N vår 2019=20).....	67
Figur 7.6	Påstand: Nå er tiltaksperioden over. Hva fra PBG vil dere ta med i videre matematikkundervisning ved skolen? N=95.....	69

## Tabellregister

Tabell 5.1	Elevprestasjoner på ungdomsskolen – analysesamplet <sup>1</sup> skoleårene 2017/18 og 2018/19 .....	36
Tabell 5.2	Deskriptiv oversikt over variable som måler familiebakgrunn på ungdomsskolen – årene 2017/18 og 2018/19 .....	36
Tabell 5.3	Balansering av tiltak- og kontrollskoler for grunnskolen, årene 2017/18 og 2018/19 <sup>1</sup> .....	37
Tabell 5.4	Balansering av målgruppen i tiltak- og kontrollskoler for grunnskolen, årene 2017/18 og 2018/19 <sup>1</sup> .....	38
Tabell 5.5	Effektestimater målgruppen i grunnskolen: Forskjeller i NP9 mellom målgruppeelever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19 <sup>1</sup> .....	40
Tabell 5.6	Heterogene effekter på NP9 av tiltaket for målgruppeelever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19 <sup>1</sup> .....	43
Tabell 5.7	Balansering av øvrige elever i tiltak- og kontrollskoler for grunnskolen, årene 2017/18 og 2018/19 <sup>1</sup> .....	44
Tabell 5.8	Effektestimater øvrige elever i grunnskolen: Forskjeller i NP9 mellom øvrig elever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19 <sup>1</sup> .....	44
Tabell 5.9	Heterogene effekter på NP9 av tiltaket for øvrig elever i tiltak- og kontrollskoler, årene 2017/18 og 2018/19 <sup>1</sup> .....	45
Tabell 6.1	Deskriptiv oversikt over variabler som måler elevprestasjoner på videregående – årene 2016/17 - 2018/19 .....	51
Tabell 6.2	Deskriptiv oversikt over variable som måler familiebakgrunn på videregående – årene 2016/17 - 2018/19 .....	51
Tabell 6.3	Balansering av tiltak- og kontrollskoler for videregående, årene 2016/17 - 2018/19 .....	52
Tabell 6.4	Forskjeller i Oslo-prøven og fullføring av VG1 mellom målgruppeelever i tiltak- og kontrollskoler, årene 2016/17 - 2018/19 .....	52
Tabell 7.1	År for gjennomføring av ulike websurvey etter type skole og stilling. Rapporter i hver celle er svarprosenten .....	58



© Statistisk sentralbyrå, 2020

Ved bruk av materiale fra denne publikasjonen skal Statistisk sentralbyrå oppgis som kilde.

ISBN 978-82-587-1068-1 (trykt)

ISBN 978-82-587-1069-8 (elektronisk)

ISSN 0806-2056