

DAMAD



05/09/14

PISA-relatering af de kriterie- baserede nationale test

Delrapport 1 – formidling af resultater

For information on obtaining additional copies, permission to reprint or translate this work, and all other correspondence, please contact:

DAMVAD

info@damvad.com
damvad.com

Copyright 2014

Indhold

1	Indledning og sammenfatning	4
1.1	Hovedkonklusioner	5
1.2	Læsevejledning	6
2	Sammenhænge i de nationale test og til PISA-undersøgelserne	7
2.1	Indplacering af elever ud fra den normbaserede og den kriteriebaserede skala	7
2.2	Sammenhængen mellem kriteriebaserede nationale test og niveauer i PISA-undersøgelserne	9
2.3	Sammenhængen mellem PISA-resultater og de forskellige profilområder	11
2.4	Opsamling	12
3	Betydningen af baggrundsvariable	13
3.1	Baggrundsvariables gennemsnitlige betydning for nationale test og PISA	13
3.2	Baggrundsvariables rolle i forudsigelse	14
3.3	Opsamling	15
4	Forudsigelse af PISA-resultater på baggrund af de nationale test	16
4.1	Den statistisk baserede model	16
4.2	Sammenhæng mellem PISA-resultater og de nationale test	17
4.3	Eksempler på modellens forudsigelser	17
4.4	Forudsigelser på baggrund af en samlet kriteriebaseret kategorisering	20
4.5	Opsamling	21

1 Indledning og sammenfatning

Undervisningsministeriet har for at styrke opfølgningen af de faglige mål i aftalen om den seneste folkeskolereform, "Et fagligt løft af folkeskolen", igangsat et arbejde med at kriteriebasere de nationale test. Da de nationale test første gang blev indført i 2010, var tilbagemeldingerne af testresultaterne normbaserede og derfor uden faste kriterier for, hvornår en elevs resultater i et givent fag på et givent klassetrin var tilfredsstillende. I stedet blev resultaterne for den enkelte elev sammenlignet med andre elevers resultater, opgjort som landsgennemsnittet i 2010.

I "Et fagligt løft af folkeskolen" er der formuleret en række konkrete faglige mål for folkeskolen. Målene er bl.a., at mindst 80 pct. af eleverne skal være gode til at læse og regne i de nationale test, og at andelen af de allerdygtigste elever i dansk og matematik skal stige år for år. Derudover skal andelen af elever med dårlige resultater i de nationale test for læsning og matematik reduceres år for år. For at der kan følges op på disse nationale mål, kriteriebaseres de nationale test, så der fremover også er en fast skala, der indikerer om en elevs resultater er gode eller dårlige på baggrund af fagligt fastsatte kriterier.

Arbejdsgrupper af personer med faglig indsigt i de enkelte testfag (opgavekommissionerne) har i første halvdel af 2014 formuleret en række foreløbige kriterier for at vurdere den enkelte elevs resultater i de nationale test. På baggrund af kriterierne inddeles elevernes testresultater i en række kategorier, der spænder fra *utilstrækkelig* til *fremragende*. På tidspunktet for afrapporteringen er de foreløbige kriterier ved at blive pilotafprøvet.

Den nærværende undersøgelse skal ses i sammenhæng med, at der ud over de nationale test samtidig gennemføres PISA-undersøgelser i Danmark, der også indeholder kategoriseringer af elevernes faglige niveau. I forbindelse med dette projekt er der

derfor blevet lagt vægt på at undersøge sammenhængen mellem de foreløbige kriteriebaserede kategorier og PISAs kategorier.

Det overordnede formål med projektet er at opstille en sandsynlighedsbaseret model for at relatere resultaterne i de nationale test til resultaterne i PISA-undersøgelsen fra 2012. Undersøgelsens formål er nærmere beskrevet i boks 1.1 nedenfor.

Undersøgelsen tager dermed sit udgangspunkt i de foreløbige kriterier for vurdering af elevers præstationer i de nationale test og sammenholder dem med resultaterne i PISA-undersøgelsen fra 2012.

Boks 1.1 Formålet med opgaven

Formålet med opgaven er at klarlægge, hvorvidt resultater i de nationale test kan relateres til PISA-resultater. Målsætningen med afdækningen er således at udarbejde:

- En sandsynlighedsbaseret model for relateringen af resultater i de nationale test på test- og profilområdeniveau til PISA. Testniveauer, som skal inddrages, er:
 - matematik (herunder profilområderne: (1) tal og algebra, (2) geometri og (3) matematik i anvendelse)
 - læsning (herunder profilområderne: (1) sprogforståelse, (2) afkodning og (3) tekstforståelse)
- Relatering af de kriteriebaserede kategorier i de nationale test med kategorier i PISA

Derudover ønskes det afdækket:

- hvor stort et sammenfald, der er blandt elever, der ikke har deltaget i de nationale test og PISA.
- eventuelt hvorvidt de socioøkonomiske effekter er ens i de nationale test og PISA.

Kilde: Kvalitets- og Tilsynsstyrelsen 2014

Grundet undersøgelsens tekniske natur afrapporteres den i to delrapporter: Delrapport 1 – formidling af resultater og Delrapport 2 – teknisk rapport og dokumentation. Hvor Delrapport 2 fokuserer på opstillingen af den statistiske model og metodiske og datamæssige udfordringer i den forbindelse, afrapporterer vi i denne rapport (Delrapport 1) resultaterne af tre spørgsmål:

1. Er der en sammenhæng mellem resultaterne i de nationale test og resultaterne i PISA-undersøgelsen, og hvordan ser denne sammenhæng i givet fald ud?
2. Hvilken rolle spiller baggrundsvariable som elevens køn, herkomst og socioøkonomiske status samt sen eller normal skolestart for evnen til at forudsige testresultaterne?
3. Kan vi – og i så fald hvordan – fremover forudsige PISA-undersøgelserne på baggrund af de nationale test?

Undersøgelsen er gennemført for Kvalitets- og Tilsynsstyrelsen i et samarbejde mellem DAMVAD og professor emeritus ved Københavns Universitet, cand.stat. Svend Kreiner.

1.1 Hovedkonklusioner

Nedenfor opsummeres analysens hovedkonklusioner:

1. Der er en tydelig sammenhæng mellem resultaterne fra de nationale test og resultaterne fra PISA-undersøgelserne. Sammenhængen kan observeres på tværs af profilområder i både læsning og matematik, men er ikke nødvendigvis jævnt fordelt.

Analyserne viser, at der er en statistisk robust sammenhæng mellem, hvordan eleverne klarer sig i de nationale test, og hvordan eleverne klarer sig i PISA-undersøgelserne. Dette betyder, at der er en høj grad af overensstemmelse imellem, hvilke kategorier eleverne indplaceres i i henholdsvis de nationale test og PISA. Elever, der indplaceres i de lavere kategorier i de nationale test som *utilstrækkelig* og *mangelfuld*, vil som regel også være indplaceret i de laveste kategorier i PISA-undersøgelserne, hhv. under 1b, 1a, 1 og 2. Sammenhængen mellem resultaterne fra de nationale test og PISA-undersøgelserne er især stærk for elever, der klarer sig mindre godt.

I de nationale test måles elevernes niveau i læsning og matematik ud fra tre såkaldte profilområder inden for begge testfag. Analysen viser, at elevernes præstationer i PISA-undersøgelsen for læsning stemmer bedst overens med præstationerne inden for "tekstforståelse" i de nationale test. Når det kommer til matematik, er sammenhængen mellem PISA-undersøgelsen og profilområderne "tal og algebra", "geometri" og "matematik i anvendelse" mere jævnt fordelt med en svag overvægt for "matematik i anvendelse".

Den tydelige sammenhæng mellem resultaterne fra de nationale test og PISA betyder samtidig, at de to test uafhængigt af hinanden når til relativt enslydende vurderinger af elevens faglige niveauer. Det er en bekræftelse af, at de nationale test siger noget relevant om elevernes faglige niveau i de områder, hvori de testes.

2. Køn, herkomst og socioøkonomisk status samt sen eller normal skolestart har betydning for resultaterne i de nationale test og i PISA-undersøgelserne. Endvidere bliver forudsigelserne af resultaterne i PISA-undersøgelserne mere præcise, hvis man tager højde for oplysninger om elevens baggrund.

Oplysninger om elevens baggrund som køn, herkomst og socioøkonomisk status samt sen eller normal skolestart har en betydning for resultaterne i de nationale test og i PISA-undersøgelserne. Generelt påvirker baggrundsvariablene elevernes gennemsnitlige resultater på samme måde i de to test. F.eks. opnår piger i gennemsnit bedre testresultater i læsning, mens drenge i gennemsnit opnår bedre testresultater i matematik i begge testsystemer. Det gælder dog, at nationale test foretages på et bestemt klassetrin, hvorimod PISA-undersøgelserne er rettet mod 15-16-årige elever uanset klassetrin. Dette betyder, at en sen eller normal skolestart sær-

ligt får betydning for resultatet af PISA-undersøgelserne, idet elever med en sen skolestart har haft et års mindre skolegang.

Beregninger foretaget i Delrapport 2 viser samtidig, at de baggrundsvariable, der er nævnt ovenfor, også har en selvstændig betydning, når vi forsøger at forudsige PISA-resultatet på baggrund af elevens præstationer i de nationale test. Dette gælder særligt køn, herkomst samt sen eller normal skolestart. Med andre ord hænger PISA-resultatet i læsning og matematik ikke kun sammen med elevernes resultater i de nationale test, men påvirkes også selvstændigt af de nævnte baggrundsvariable.

3. Der kan fremover foretages en statistisk baseret forudsigelse af, hvad en given elev vil score i en PISA-undersøgelse på baggrund af vedkommendes resultater fra den nationale test. Forudsigelser bygger på den statistiske model, der er opstillet i projektet og tager også højde for oplysninger om elevens baggrund¹.

Det er muligt på baggrund af en given elevs præstation i de nationale test at komme med en statistisk baseret forudsigelse af, hvordan vedkommende vil klare sig i en PISA-undersøgelse. På baggrund af resultaterne fra et enkelt profilområde inddelt i kategorier kan de gennemsnitlige PISA-resultater (med begrænset præcision) forudsiges. Man kan ligeledes (med langt højere præcision) forudsige resultaterne i læsning og matematik ved at foretage en række beregninger på baggrund af matematiske modeller, der viser sammenhængen mellem de to test, ved at benytte baggrundsvariablene, køn, herkomst samt sen eller normal skolestart.

Hvis vi f.eks. tager udgangspunkt i en dreng af dansk herkomst med normal skolestart, kan der på baggrund af disse oplysninger og hans resultater fra den nationale test foretages en statistisk baseret forudsigelse. Forestiller vi os, at hans testresultater i læsning i 8. klasse kan kategoriseres som *god* i "sprogforståelse", *god* i "afkodning" og *jævn* i "tekstforståelse", vil forudsigelsen være, at han i en senere PISA-undersøgelse vil score omkring 465. Dette er lidt under middel i forhold til PISAs kategoriseringer.

1.2 Læsevejledning

Den resterende del af rapporten er disponeret efter de tre hovedspørgsmål i analysen.

I *kapitel 2* viser vi, at der er en sammenhæng mellem de nationale test og PISA-undersøgelserne, og gennemgår, hvordan denne ser ud.

I *kapitel 3* viser vi, hvilken rolle baggrundsvariable som køn, herkomst og socioøkonomisk status mv. og sen eller almindelig skolestart spiller for evnen til at forudsige testresultaterne.

I *kapitel 4* viser vi, hvordan vi kan forudsige PISA-undersøgelserne på baggrund af de nationale test.

¹ Det bør for en god ordens skyld tilføjes, at da formålet med PISA-undersøgelserne alene er at måle resultaterne for større grupper af elever, skal forudsigelserne kun anvendes på aggregeret niveau, som f.eks. landeni-

veau, og ikke for enkelte elever. Derudover gælder det, at den bagvedliggende statistiske model ikke kan anvendes til sammenligninger med andre lande, blandt andet på grund af mangel på sammenlignelige data.

2 Sammenhænge i de nationale test og til PISA-undersøgelserne

I det følgende kapitel sammenlignes først indplaceringen af eleverne ud fra de nationale test, når man anvender de nye foreløbige kriterier med indplaceringen af eleverne i forhold til de gamle normbaserede mål. Dernæst vises, hvordan de kriteriebaserede indplaceringer af elever relaterer sig til PISAs kategoriseringer. Dette gøres ved at se på, hvordan elever konkret har klaret sig i hver af de to test. Endelig vises sammenhængen mellem de forskellige profilområder inden for læsning og matematik (med og uden de kriteriebaserede resultater) og de tilsvarende PISA-undersøgelser.

2.1 Indplacering af elever ud fra den normbaserede og den kriteriebaserede skala

Som nævnt indledningsvis, er der tidligere i år (2014) blevet formuleret en række foreløbige kriterier for præstationerne i de nationale test. Som det fremgår af tabel 2.1 og tabel 2.2, der viser resultaterne for henholdsvis "tekstforståelse" og "matematik i anvendelse", er der en fin sammenhæng mellem den nye kriteriebaserede og den gamle normbaserede skala. Samtidig viser tabellerne, at de kriteriebaserede resultater er velegnede til at sondre

mellem de fagligt stærkeste og de fagligt svageste elever. Dette betyder, at de nye kriteriebaserede kategorier styrker mulighederne for at differentiere mellem forskellige grupper af særligt dygtige og særligt svage elever.

For "tekstforståelse" i 8. klasse (tabel 2.1) fordeles elever, som scorer blandt de laveste 10 pct. på den normbaserede skala, f.eks. i de tre laveste kriteriebaserede kategorier fra *utilstrækkelig* til *jævn*, med hovedparten i kategorien *mangelfuld*. For midtergruppen, der er mellem 35 og 65 pct. på den normbaserede skala, gælder det, at samtlige elever indplaceres i den kriteriebaserede kategori *god*. Blandt de 10 pct. bedste elever på den normbaserede skala indplaceres knap 80 pct. sig i kategorien *rigtig god*, mens de resterende godt 20 pct. er i kategorien *fremragende*.

For matematik i 6. klasse er der ligeledes en god sammenhæng mellem placeringen i kriteriebaserede kategorier og elevernes resultater på den normbaserede skala. Det vises for profilområdet "matematik i anvendelse" i tabel 2.2.

TABEL 2.1
Profilområde "tekstforståelse" i 8. klasse

Kriteriebaseret skala	Utilstrækkelig	Mangelfuld	Jævn	God	Rigtig god	Fremragende	Total
Interval, normbaseret skala							
Op til 10 pct.	14,3	61,9	23,8	-	-	-	100
10 - 35 pct.	-	-	73,0	27,0	-	-	100
35 - 65 pct.	-	-	-	100,0	-	-	100
65 - 90 pct.	-	-	-	49,1	50,9	-	100
90 pct. og derover	-	-	-	-	78,3	21,7	100
Samlet fordeling	1,6	6,7	21,7	46,3	21,4	2,3	100

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: Antal elever=7.112

Note: Kun elever, der både har deltaget i de nationale test og i PISA-undersøgelsen, indgår i opgørelsen

Elever indplaceret i de forskellige normbaserede intervaller placeres i to kriteriebaserede kategorier. Blandt de svageste 10 pct. af eleverne placeres omkring 2/3 af eleverne i *mangelfuld*, mens godt 1/3 placeres i *utilstrækkelig*. For gruppen fra 10 til 35 pct. er der ligeledes ca. 2/3 i *mangelfuld*, mens godt 1/3 placeres i *jævn*. Sammenlignet med "tekstforståelse" i 8. klasse (tabel 2.1) er der således en tendens til, at eleverne i højere grad koncentrerer sig i kategorierne *mangelfuld* og *god*, mens *jævn* fanger ca. 20 pct. af eleverne.

Afsnittet har dermed vist, at den normbaserede og den kriteriebaserede skala hænger fint sammen, og at de nye foreløbige kriteriebaserede kategorier er bedre til at nuancere forskelle blandt de svageste og stærkeste elever. De næste afsnit ser nærmere på sammenhængen mellem elevernes præstationer målt på den kriteriebaserede skala og PISA.

Boks 2.1 De nationale test

De nationale test (DNT) er it-baserede adaptive test, der siden 2010 årligt er blevet afholdt i folkeskolen. Der er tale om i alt 10 obligatoriske test fordelt på fagene dansk (læsning), matematik, engelsk, geografi, biologi og fysik/kemi. Der testes i læsning på 2., 4., 6. og 8. klassetrin og i matematik på 3. og 6. klassetrin.

Testenes funktion er først og fremmest at understøtte lærernes løbende evaluering af elevernes faglige niveau og derfor give en indikator på elevernes faglige niveau inden for hvert af de delområder, som hver test afdækker. Disse delområder kaldes profilområder. Hver af de nationale test består af tre profilområder. Eksempelvis er profilområderne i læsning "sprogforståelse", "afkodning" og "tekstforståelse". I matematik er de "tal og algebra", "geometri" og "matematik i anvendelse".

Med den seneste folkeskolereform er det blevet besluttet, at resultaterne af de nationale test fremover skal gøres kriteriebaserede. Det betyder, at de nationale test ikke længere skal opgøres i forhold til landsgennemsnittet, men i forhold til fastlagte kriterier for, hvornår en elev f.eks. er god til geometri. Fremover vil elevernes præstationer inden for hvert profilområde derfor kunne kategoriseres som enten *utilstrækkelig*, *mangelfuld*, *jævn*, *god*, *rigtig god* eller *fremragende*.

TABEL 2.2
Profilområde "matematik i anvendelse" i 6. klasse

Kriteriebaseret skala							
	Utilstrækkelig	Mangelfuld	Jævn	God	Rigtig god	Fremragende	Total
Interval, normbaseret skala							
op til 10 pct.	34,5	65,5	-	-	-	-	100
10 - 35 pct.	-	63,9	36,1	-	-	-	100
35 - 65 pct.	-	-	28,8	71,2	-	-	100
65 - 90 pct.	-	-	-	97,7	2,3	-	100
90 pct. og derover	-	-	-	-	86,2	13,8	100
Samlet fordeling	5,6	32,5	19,5	35,4	6,1	0,9	100

Kilde: Delrapport 2 – teknisk rapport og dokumentation.
Note: Antal elever=1.420

2.2 Sammenhængen mellem kriteriebaserede nationale test og niveauer i PISA-undersøgelserne

Ud over de nationale test testes danske elever også i regi af OECD i de såkaldte PISA-undersøgelser. I dette afsnit belyses sammenhængen mellem de kriteriebaserede nationale test og PISA-undersøgelsen fra 2012.

I PISA-undersøgelserne opdeles eleverne i niveauer fra 1 til 6 baseret på baggrund af det opnåede PISA-resultat. Der anvendes forskellige skæringspunkter af niveauerne for matematik og læsning, ligesom der er forskellige niveauer i bunden af skalaen (f.eks. er niveau 1 opdelt i flere underniveauer for læsning). PISA-undersøgelsen er nærmere beskrevet i boks 2.2.

Boks 2.2 PISA-undersøgelsen

PISA er en verdensomspændende undersøgelse af kompetencer i matematik, læsning og naturfag. Testene udføres i OECD-regi med en række ikke-OECD-lande som samarbejdspartnere. Første test var i år 2000, og de foretages hvert tredje år på 15-16-årige, der fortsat går i skole. Seneste test er udført i år 2012. Resultaterne opgøres i PISA-rapporterne på nationale niveauer og ikke på elev-, skole- eller regionalt niveau.

Et resultat af en PISA-undersøgelse rapporteres typisk som en såkaldt PISA-score, der spænder i intervallet fra 0 til 1000. Gennemsnittet i OECD er derfor omkring (men ikke nødvendigvis lig) 500 for alle de tre testområder. PISA-scoren er beregnet ud fra værdier fra den Rasch-model, der ligger bag ved testene (se boks 4.1). Som en følge af denne statistiske model vil elevernes resultater koncentrere sig om de midterste værdier. Der er eksempelvis langt flere elever, der scorer mellem 400 og 500, end elever, der scorer mellem 600 og 700.

Denne PISA-score kan oversættes til nogle på forhånd fastsatte kategorier i intervallet fra 1 til 6, evt. med underopdelinger af kategori 1 i 1b og 1a i læsning. Kriterierne for, hvilken PISA-score det kræver at opnå en bestemt kategori, varierer lidt fra område til område. Eksempelvis kræver det i PISA 2012 en score på 607 i matematik og 626 i læsning at opnå et elevresultat i kategori 5, og 669 i matematik og 698 i læsning for at opnå kategori 6. Kategori 5 og 6 benævnes af PISA som "top-performers". PISA betegner kategori 2 som et baseline-niveau for, hvad der skal til for at deltage "effektivt og produktivt i livet". For at nå dette niveau skal eleven have scoren 420 i matematik og 407 i læsning.

TABEL 2.3

Profilområdet "tekstforståelse" i 8. klasse i 2011–2012 sammenholdt med PISAs læseresultater fra 2012

Kriteriebaseret skala, DNT	PISA-kategori								Total
	<1b	1b	1a	2	3	4	5	6	
Utilstrækkelig	10,5	26,3	26,3	28,9	5,3	2,6	-	-	100
Mangelfuld	4,5	17,9	35,0	34,6	6,9	1,2	-	-	100
Jævn	2,4	6,2	25,4	39,9	23,2	2,6	0,2	-	100
God	0,3	1,4	8,5	32,5	37,8	14,6	4,7	0,1	100
Rigtig god	0,1	0,2	1,5	12,1	36,1	38,2	10,3	1,4	100
Fremragende	-	-	1,0	3,1	15,6	47,9	31,3	1,0	100
Samlet fordeling	1,0	3,3	11,9	28,9	31,8	17,3	5,3	0,4	100

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: Antal elever=4.156

Note: Elever under niveau 2 har i PISA-undersøgelserens terminologi "utilstrækkelige læsefærdigheder"

I tabel 2.3 ses, at der er en stærk sammenhæng mellem placeringen i de kriteriebaserede kategorier i de nationale test for profilområdet "tekstforståelse" og placeringen i kategorierne for resultaterne af PISA-undersøgelsens læsetest. Da tendensen er den samme for alle tre profilområder, præsenteres her kun en tabel for "tekstforståelse", som har den stærkeste korrelation med PISA. Tabeller for de resterende profilområder findes i Delrapport 2.

Tabel 2.3. viser, at elever, som får den kriteriebaserede score *utilstrækkelig*, fordeler sig i PISA-niveauerne fra under 1b til 4, dog ligger størstedelen i 1b til 2. Det samme mønster gør sig gældende for elever som får placeringen *mangelfuld*, hvor der dog er en større koncentration af eleverne, som placeres i PISA-niveauerne 1a og 2, og færre i de laveste kategorier (1b og derunder).

Elever, der placeres i kategorien *rigtig god* ud fra de kriteriebaserede nationale test indplaceres over hele spektret i PISA-niveauerne, men med en klar hovedvægt i kategorierne 3 og 4 i den efterfølgende PISA-test i læsning. Overordnet set viser tabel 2.3 dog stadig den fine sammenhæng, at de elever,

som bliver placeret i de lave kategorier i de nationale test, også indplaceres i de lavere kategorier i PISA.

Der kan være flere årsager til, at nogle elever scorer meget lavere – eller højere – i den efterfølgende PISA-undersøgelse i læsning, end indplaceringen på den kriteriebaserede skala ved den nationale test tilsagde. For det første kan det skyldes reelle bevægelser i elevernes faglige niveau, fordi der er en vis tidsforskel mellem gennemførelse af de to test. For det andet kan det skyldes forskelle i, hvad og hvordan der måles i de to test. Endelig er det velkendt, at elever nogle gang har en 'dårlig dag' på testningstidspunktet, og at testresultatet derfor ikke afspejler elevens egentlige faglige niveau.

Det er i den sammenhæng betryggende, at der er så relativt få elever, som scorer 'helt ved siden af'. Det indikerer, at de to første forklaringsmuligheder formentligt ikke skal tillægges for meget vægt. Dette er vigtigt for muligheden for at forudsige et fremtidigt PISA-resultat på baggrund af resultaterne fra de nationale test, som vi vil se nærmere på i kapitel 4.

TABEL 2.4

"Matematik i anvendelse" fra sene skolestartere 6. klasse i 2010 og PISAs matematikresultater for 2012

PISA-kategori	1	2	3	4	5	6	Total
Kriteriebaseret skala, DNT							
Utilstrækkelig	85,4	7,3	2,4	4,9	-	-	100
Mangelfuld	56,4	31,8	9,4	2,4	-	-	100
Jævn	35,8	32,8	25,9	5,5	-	-	100
God	20,2	34,1	27,7	16,0	2,0	-	100
Rigtig god	7,9	14,5	28,9	28,9	19,7	-	100
Fremragende	10,0	10,0	-	30,0	40,0	10,0	100
Samlet fordeling	35,9	30,5	20,5	10,4	2,5	0,1	100

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: Antal elever=1.063. Niveau under 1 findes ikke matematik-testen.

Der er den samme tendens mellem placeringen på den kriteriebaserede kategorisering af den nationale test i matematik i 6. klasse og PISAs matematiktest i 2012, som der var for læseområdet. I tabel 2.4 ses sammenhængen for profilområdet "matematik i anvendelse".

2.3 Sammenhængen mellem PISA-resultater og de forskellige profilområder

De nationale test i henholdsvis matematik og læsning består af hver tre profilområder, der måles for at vurdere elevernes præstation. Af pladshensyn viste det ovenstående afsnit alene resultaterne for det profilområde med den stærkeste sammenhæng til PISA-resultatet. I dette afsnit illustreres sammenhængen mellem hvert enkelt af de tre profilområder og det tilsvarende PISA-resultat i henholdsvis læsning og matematik.

Sammenhængen illustreres ved hjælp af såkaldte korrelationskoefficienter, der på en skala fra 0 (ingen sammenhæng) til 1 (perfekt lineær sammenhæng) viser, hvor stor sammenhæng der er mellem profilområderne og det tilsvarende PISA-resultat. Jo højere korrelationskoefficienten er, desto større er sammenhængen.

Det fremgår af tabel 2.5, at der er en stærkere sammenhæng mellem profilområdet "tekstforståelse" og PISAs læsetest, end der er for de to andre profilområder, "sprogforståelse" og "afkodning". Endvidere ses, at "tekstforståelse" også er mere korreleret med de to andre profilområder, end disse to er med hinanden. Dette resultat understøtter, at der er en konsistent sammenhæng mellem målingen af profilområder i de nationale test og resultaterne i PISA-undersøgelsen.

TABEL 2.5

Korrelation mellem værdier fra Rasch-modellen på profilområder i de nationale test og PISA-undersøgelsen for læsning

	Sprogforståelse	Afkodning	Tekstforståelse	PISA læsning
Sprogforståelse	1	0,55	0,61	0,46
Afkodning		1	0,62	0,49
Tekstforståelse			1	0,62
PISA læsning				1

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: Alle korrelationer er statistisk signifikante på et 1 pct. niveau. 7.118 elever har resultater for DNT og 5.156 for PISA. Af disse har 4.162 resultater for begge test.

For matematik har profilområderne en lidt højere indbyrdes korrelation. Korrelationen til PISA er af nogenlunde samme styrke som for læsning, men korrelationerne er mere ensartede. Det fremgår af tabel 2.6, at "matematik i anvendelse" har den største korrelation til såvel de andre profilområder som til PISA-undersøgelsen. Dette billede er altså tilsvarende, hvad vi så for læsning.

TABEL 2.6

Korrelation mellem værdier fra Rasch-modellen på profilområder i de nationale test og PISA-undersøgelsen for matematik.

	Tal og algebra	Geometri	Matematik i anvendelse	PISA matematik
Tal og algebra	1	0,58	0,67	0,48
Geometri		1	0,62	0,51
Matematik i anvendelse			1	0,55
PISA matematik				1

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: Alle korrelationer er statistisk signifikante på et 1 pct. niveau. 1.420 elever har resultater for DNT og 7.441 for PISA. Af disse har 1.063 resultater for begge test. Den store forskel skyldes, at det alene er sene skolestartere, der har DNT-data i matematik. Se i Delrapport 2 for uddybning.

2.4 Opsamling

Dette kapitel har vist, at der er en tydelig sammenhæng mellem resultaterne fra de nationale test og PISA-undersøgelserne. Sammenhængen kan observeres på tværs af profilområder i både læsning og matematik, men er dog generelt set lidt stærkere i matematik. I det følgende kapitel ses nærmere på, hvilken betydning baggrundsvariable som køn, herkomst, socioøkonomisk status og sen eller normal skolestart har for resultaterne i de nationale test og PISA-undersøgelserne.

3 Betydningen af baggrundsvariable

Dette kapitel ser nærmere på, hvilken betydning oplysninger om elevens baggrund har for resultaterne i de nationale test og i PISA-undersøgelsen. Hermed får vi en indikation af, hvorvidt baggrundsvariable som køn, herkomst, sen eller normal skolestart samt socioøkonomisk status kan bruges til at forudsige PISA-resultatet på baggrund af resultaterne i de nationale test.

3.1 Baggrundsvariables gennemsnitlige betydning for nationale test og PISA

Det er velkendt, at både køn, herkomst, socioøkonomisk status samt sen eller normal skolestart har en betydning for elevernes præstationer i folkeskolen. Derfor er det ingen overraskelse, at dette også ses i resultaterne fra de nationale test og PISA-undersøgelserne. Tabel 3.1 viser de gennemsnitlige PISA-resultater fordelt på ovennævnte baggrundsvariable set i forhold til det samlede gennemsnit for de danske elever i testen. Som det fremgår af tabellen, er der en tendens til, at drenge klarer sig lidt bedre end piger til matematik, mens piger i gennemsnit klarer sig bedre i læsning.

Det ses endvidere, at elever af dansk herkomst i gennemsnit klarer sig en del bedre end elever med en anden herkomst både i matematik og læsning. Sene skolestartere, som er testet et klassetrin lavere end normale skolestartere i PISA-undersøgelserne, klarer sig ligeledes dårligere.

Som mål for socioøkonomisk status er anvendt PISA's ESCS, der er et samlet indeks for socioøkonomisk status, som tager udgangspunkt i forældres uddannelse, beskæftigelse og et mål for besiddelser i hjemmet (økonomiske som kulturelle faktorer).² I tabel 3.1 vises de gennemsnitlige PISA-resultater

for de børn, som ligger blandt de øverste 20 pct., de midterste 20 pct. og de nederste 20 pct. på dette ESCS-indeks. Ikke overraskende ses det, at jo højere socioøkonomisk status en elev har, desto bedre klarer vedkommende sig i PISA-undersøgelserne. Effekten af socioøkonomisk status er ganske betydelig.

TABEL 3.1
Effekt af baggrundsvariable på gennemsnitlige resultater i PISA-undersøgelserne

		PISA matematik	PISA læsning
Gennemsnitlig PISA-score i datasættet		489	487
Køn	Dreng	494	476
	Pige	484	498
Herkomst	Dansk herkomst	506	503
	Anden herkomst	442	442
Skolestart	Sen	451	443
	Normal	498	497
Socioøkonomisk status (PISA-ESCS)	Høj (øverste 20 pct.)	532	527
	Middel (midterste 20 pct.)	494	492
	Lav (nederste 20 pct.)	449	447

Kilde: Delrapport 2 – teknisk rapport og dokumentation
Note: Kun elever, der både har deltaget i de nationale test og i PISA-undersøgelsen, indgår i opgørelsen
Note: Data om socioøkonomisk status stammer fra PISA (ESCS).

Et tilsvarende mønster kan genfindes i resultaterne for de nationale test. Af tabel 3.2 fremgår, at retningen for de forskellige baggrundsvariable er den samme som i PISA-undersøgelserne: Drenge har generelt bedre resultater på de tre matematiske profilområder, piger på to af de tre profilområder for

² Se PISA Results 2012: "What Students Know and Can Do" (OECD, 2014).

læsning, ligesom elever af dansk herkomst generelt har bedre resultater end de af anden herkomst osv.

3.2 BaggrundsvARIABLES rolle i forudsigelse

Som det fremgik af forrige afsnit, har baggrundsvARIABLENE stor betydning for såvel resultaterne i PISA-undersøgelserne som resultaterne i de nationale test. Et oplagt spørgsmål er derfor, om variableNE

påvirker resultaterne lige kraftigt? Dette har en stor betydning for, hvordan en statistisk baseret forudsigelsesmodel, der relaterer resultater fra de nationale test til PISA-resultaterne, skal tage sig ud.

Er effekten helt ensartet i de to test, vil baggrundsvARIABLENE ikke bidrage med yderlige forklaringskraft, når PISA-resultaterne skal forudsiges ved

TABEL 3.2

Betydningen af baggrundsvARIABLE i de nationale test illustreret ved det gennemsnitlige resultat på hvert profilområde for forskellige grupper af elever

Profilområde	----- Matematik -----				----- Læsning -----			
	Antal elever	Tal og algebra	Geometri	Matematik i anvendelse	Antal elever	Tekstforståelse	Sprogforståelse	Afkodning
Gennemsnit for:								
De nationale test (i det anvendte data)	1.420	15,0	15,6	14,1	7.118	26,6	15,7	24,5
Køn*								
Dreng	670	16,0	16,3	15,0	3.038	26,7	16,2	24,8
Pige	399	14,7	15,4	13,7	3.029	27,7	15,8	25,1
Herkomst								
Dansk herkomst	687	16,1	16,8	15,2	4.459	28,6	16,8	25,7
Anden herkomst	382	14,4	14,3	13,3	1.608	23,5	13,8	23,0
Skolestart								
Sen		-	-	-	1.508	24,2	14,5	23,1
Normal		-	-	-	5.610	27,3	16,0	24,9
Socioøkonomisk status (målt ved PISA ESCS)								
Høj (top 20 pct.)	106	17,3	18,5	16,3	1.093	31,0	18,0	26,5
Middel (20 pct.)	188	16,0	16,1	14,9	1.210	27,3	16,2	25,1
Lav (bund 20 pct.)	311	14,2	14,8	13,2	1.224	23,9	13,8	23,4

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: *Da der mangler oplysninger om køn, herkomst, skolestart og socioøkonomisk status for en række elever, indgår disse ikke i beregningen af gennemsnitsresultatet fordelt på disse baggrundsvARIABLE. De indgår dog i beregningen af det samlede gennemsnitlige resultat. For opdelingen på køn betyder det, at det gennemsnitlige resultat for såvel drenge som piger er højere end landsgennemsnittet. Samme problem gør sig gældende for herkomst, skolestart og socioøkonomisk status, men problemet er ikke på samme måde tydeligt for disse variable, da det samlede gennemsnit falder mellem de to gennemsnit for baggrundsvARIABLENE.

Note: De gennemsnitlige kriteriebaserede scorer for matematik bygger alle på resultater for elever med sen skolestart. Det følger heraf, at cellerne for matematik med normal og sen skolestart er tomme (og gennemsnittet for de med sen skolestart er det fra de nationale test i øverste række).

hjælp af de nationale test. Er effekten derimod forskellig, vil baggrundsvariablene bidrage til forudsigelsen af PISA-resultaterne, ud over hvad resultaterne i de nationale test kan forklare. Som det fremgår af det følgende kapitel, er baggrundsvariablene vigtige, for at den statistiske model så præcist som muligt kan forudsige et PISA-resultat på baggrund af resultaterne i de nationale test.

Dette kan grundlæggende skyldes to forhold: 1) at baggrundsvariablene af forskellige årsager påvirker resultaterne i de nationale test mere eller mindre end resultaterne i PISA-undersøgelserne, og/eller 2) at baggrundsvariablene påvirker den faglige udvikling, der er sket for eleven, fra de nationale test er gennemført, til at PISA-undersøgelserne foretages – altså med andre ord forskelle i, hvor stort udbytte der opnås af skolegangen, imellem de to test foretages. Med det nuværende datagrundlag er det ikke muligt at undersøge nærmere, hvilken forklaring der dominerer, og det er heller ikke af særspecial interesse i denne rapport. I forhold til en model, der relaterer de nationale test til PISA-undersøgelserne, kan vi blot konstatere, at baggrundsvariablene bidrager til modellens præcision og dermed til evnen til at forudsige et PISA-resultat på baggrund af resultater fra de nationale test.

3.3 Opsamling

I dette kapitel har vi vist, at køn, herkomst og socioøkonomisk status samt sen eller normal skolestart har betydning for de gennemsnitlige resultater i de nationale test og i PISA-undersøgelserne. Beregninger på den statistiske model, der er foretaget i Delrapport 2 viser, at når der er kontrolleret for køn, herkomst og sen/normal skolestart, så bidrager den socioøkonomiske variabel ESCS kun med en begrænset forklaringskraft til modellen. Da det samtidig er en variabel, der kommer fra PISA-undersø-

gelserne, og derfor ikke er tilgængelig fra de nationale test, er denne variabel ikke medtaget i den foretrukne statistiske model til forudsigelse af PISA-resultater på baggrund af de nationale test, som præsenteres i det følgende kapitel.

4 Forudsigelse af PISA-resultater på baggrund af de nationale test

Som det fremgik af kapitel 2, er det muligt at sammenligne resultaterne fra de nationale test med resultaterne fra PISA-undersøgelserne. Der er således generelt set overensstemmelse mellem elevernes testresultater i de to test. Spørgsmålet, der besvares i dette kapitel er derfor, hvordan vi fremover kan forudsige PISA-resultaterne ud fra resultater i de nationale test og oplysninger om elevens baggrund?

4.1 Den statistisk baserede model

Det tekniske grundlag for at kunne forudsige PISA-resultaterne bygger på beregninger foretaget i Delrapport 2. Både de nationale test og PISA-undersøgelserne er baseret på en "Rasch-skala", hvorfra de oprindelige testresultater konverteres til en ny skala, hvor elevernes præstationer kan opdeles i kategorier. Rasch-modellen er nærmere beskrevet i nedenstående boks.

Boks 4.1 Rasch-modellen

Rasch-modellen er en statistisk model udviklet af den danske matematiker Georg Rasch omkring 1960. Modellen bruges til analyse af 'duelighedstest', f.eks. intelligencetest eller spørgeskemaer, der måler folks holdninger til et givent emne.

Rasch-modellen er en sandsynlighedsmodel. Et særligt kendetegn ved den er sammenhængen mellem elevernes dygtighed og opgavernes sværhedsgrad, som beregnes på samme skala. Sværhedsgraden af en opgave defineres som lig med dygtigheden af den elev, der har præcis 50 procent sandsynlighed for at svare korrekt. Både PISA og de nationale test benytter Rasch-modellen til at beskrive effekten af elevernes dygtighed på svarene på spørgsmålene i de pædagogiske test.

I modellen måles elevernes færdigheder på såkaldte logit-skalaer. Man kan derfor tale om en "PISA-logit" for matematik og læsning og en "logit-værdi" for hver profilområde i de nationale test. En af fordelene ved logit-skalaer er, at de har intervallskalaegenskaber. Det er derfor værdier på disse skalaer, som er blevet anvendt i forbindelse med forudsigelserne af, hvad en given elev scorer i en PISA-undersøgelse.

I Delrapport 2 kan man læse mere om Rasch-modellen, og hvordan den finder anvendelse i forbindelse med de nationale test og PISA-undersøgelserne.

Indholdsmæssigt er der ikke større forskel på en "Rasch-skala" og de skalaer, som de nationale test og PISA-undersøgelserne anvender, end der er på Celsius og Fahrenheit. Begge skalaer fortæller, hvad temperaturen er. På samme måde viser både et resultat målt på "Rasch-skalaen" og PISAs egen skala, hvor god en elev er til f.eks. læsning.

Imidlertid viser beregninger foretaget i Delrapport 2, at resultaterne i de nationale test og i PISA-undersøgelserne hænger lineært sammen, hvis man måler dem på "Rasch-skalaen". Der kan derfor med fordel opstilles en lineær regressionsmodel, der beskriver den matematiske sammenhæng med udgangspunkt heri.

Boks 4.2 Regressionsmodellerne for læsning og matematik

Med udgangspunkt i Rasch-modellen kan der opstilles en regressionsmodel for de forventede resultater i PISA-undersøgelsen for henholdsvis matematik og læsning. Modellerne viser sammenhængen mellem logit-værdierne for de tre profilområder i enten læsning eller matematik og den forventede PISA-logit givet ved oplysningerne om elevens baggrund.

Modellen for læsning er givet ved:

$$PISA_L = -0,745 + 0,119 * S + 0,099 * A + 0,438 * T - 0,180 * køn - 0,282 * herkomst + 0,350 * skolestart$$

hvor $PISA_L$ er resultatet af læsning i PISA-målingen, S er resultatet af "sprogforståelse" i den nationale test, A er resultatet af "afkodning" i den nationale test, og T er resultatet af "tekstforståelse" i den nationale test ($R^2=0,442$).

Modellen for matematik er givet ved:

$$PISA_M = -0,658 + 0,200 * T + 0,310 * G + 0,380 * A + 0,130 * køn - 0,455 * herkomst + 0,406 * skolestart$$

hvor $PISA_M$ er resultatet af læsning i PISA-undersøgelsen, T er resultatet af "tal og algebra" i den nationale test, G er resultatet for "geometri" i den nationale test, og A er resultatet for "matematik i anvendelse" i den nationale test ($R^2=0,412$).

For begge modeller gælder det, at "køn" er 1 for dreng og 0 for pige; "herkomst" er 1, hvis der er tale om en elev med anden etnisk herkomst end dansk, og 0, hvis eleven er etnisk dansk af herkomst; "skolestart" er 1, hvis eleven har en sen skolestart, og 0, hvis eleven har haft en normal skolestart.

Indtil videre er der kun blevet set på sammenhængen ud fra et profilområde ad gangen eller på den gennemsnitlige forskel på f.eks. drenges og pigers præstationer i de to test. Med en statistisk baseret model stilles et mere præcist spørgsmål, der benytter alle oplysningerne på et og samme tidspunkt til at forudsige, hvad resultatet af en PISA-undersøgelse vil blive.

4.2 Sammenhæng mellem PISA-resultater og de nationale test

Antag, at der foreligger testresultater fra de nationale test i læsning i 8. klasse, og der ønskes et bud på, hvor eleven vil placere sig, hvis eleven senere skulle deltage i en PISA-undersøgelse.

Udgangspunktet kunne f.eks. være, at der var tale om en etnisk dansk dreng med normal skolestart.

Forestiller vi os, at den pågældende elevs testresultater kan kategoriseres som *god* i "sprogforståelse", *god* i "afkodning" og *jævn* i "tekstforståelse", kan vi forvente, at han i en senere PISA-undersøgelse vil score omkring 465, givet at hans scorer i den nationale test ligger midt i kategorierne. Dette svarer til niveau 2 i PISAs kategoriseringer i 2012, som går fra 407 til 480.

For at benytte regressionsmodellen for læsning i det konkrete tilfælde er det nødvendigt at omregne værdierne fra den skala, der anvendes i de nationale test, til "Rasch-skalaen". Vi bliver med andre ord nødt til at gå fra Celsius til Fahrenheit. Dette er muligt, da vi i Delrapport 2 har udviklet en række metoder, der bruges til at omregne testværdier fra én skala til en anden.

Når regressionsmodellerne skal bruges i virkeligheden, vil processen i praksis ikke starte med, at scoren fra den nationale test omregnes til en værdi på "Rasch-skalaen". Testresultater fra de nationale test

og PISA-undersøgelserne beregnes altid først på "Rasch-skalaen" og konverteres derefter til skalaer, der er lettere at tolke på. Boks 4.3 beskriver, hvordan værdier målt på "Rasch-skalaen" omregnes til PISA-scorer.

Boks 4.3 Omregning af logit-værdier til PISA-scorer

For at kunne gennemskue, hvordan et testresultat målt på "Rasch-skalaen" skal tolkes, kan det blive nødvendigt at konvertere en række værdier fra én skala til en anden. I Delrapport 2 er derfor udarbejdet funktioner, som kan konvertere logit-værdier fra PISA-undersøgelserne til PISA-scorer, der er lettere at tolke på.

For sammenhængen mellem testresultater målt på "Rasch-skalaen" og de scorer, som PISA anvender i egne målinger i henholdsvis læsning og matematik, benyttes følgende funktioner:

$$PISA_L = 515,188 + 68,185 * \text{logit} - 2,155 * \text{logit}^2$$

$$PISA_M = 471,374 + 79,663 * \text{logit} - 5,719 * \text{logit}^2$$

hvor "logit" angiver logit-værdien målt på "Rasch-skalaen".

I Delrapport 2 kan man læse mere om funktionerne, og hvordan det rent teknisk er muligt at opstille en matematisk formel herfor.

Endvidere bør det bemærkes, at PISA-undersøgelserne ikke har til hensigt at måle den enkelte elevs præstationer, hvorfor forudsigelserne i praksis også vil fokusere på den samlede population af elever omfattet af begge test.

4.3 Eksempler på modellens forudsigelser

For at give et mere nuanceret billede af, hvordan modellens forudsigelser fungerer, kan det være nyttigt at tage udgangspunkt i to illustrative eksempler, selvom forudsigelserne i praksis vil have fokus på grupper af elever og ikke den enkelte elev.

Første eksempel

Aisha er 13 år og går i 6. klasse. Aisha er af anden herkomst end dansk og begyndte i skole et år senere end normalt. I den nationale test i matematik

klarede Aisha sig ud fra de kriteriebaserede kategorier som følger: *jævnt* i “tal og algebra”, *mangelfuldt* i “geometri” og *mangelfuldt* i “matematik i anvendelse”.

På baggrund ovenstående beskrivelse kan vi forvente, at Aisha ca. to år senere vil score omkring 290 i PISA-undersøgelsen i matematik. Dette svarer til niveauet “under 1” i PISAs kategorisystem. Da PISA-målingen i matematik er rettet mod 15-16 årige elever, vil Aisha gå i 8. klasse, på det tidspunkt testen gennemføres.

FIGUR 4.1
Eksempel på forventede ændringer i PISA-resultatet i matematik

	Forventet PISA-score
Person fra eksemplet (Aisha)	290
<i>Ændring fra eksemplet</i>	<i>Forventet ændring (+/-)</i>
DNT-resultater	
Tal og algebra <i>Jævn -> god</i>	+14
Geometri <i>Mangelfuld -> god</i>	+40
Matematik i anvendelse <i>Mangelfuld -> rigtig god</i>	+99
Alle 3 profilmråder <i>-> Rigtig god</i>	+183
Baggrundsvariable	
Køn <i>Pige -> dreng</i>	+13
Skolestart <i>Sen -> normal</i>	+41
Etnicitet <i>Anden herkomst -> dansk</i>	+45

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

I tabel 4.1 er vist, hvordan det forventede resultat i PISA-undersøgelsen i matematik vil ændre sig, hvis vi ændrer på Aishas præstationer i den nationale test. Af tabellen fremgår også, hvilke ændringer vi kan forvente, hvis vi ændrer på oplysninger om Aishas baggrund - altså hvis hun f.eks. havde været en dreng eller af dansk herkomst.

Tabellen tager udgangspunkt i den score, som Aisha ifølge modellen for matematik statistisk set ville opnå. Herfra viser tabellen, hvor meget denne score påvirkes af, at der ændres på oplysninger fra eksemplet (“Forventet ændring”).

Som vi kan se af tabellen, er de baggrundsvariable, der betyder mest for resultatet af Aishas forventede PISA-score i matematik, den sene skolestart og spørgsmålet om herkomst.

Andet eksempel

Jesper er 14 år, dansk af herkomst og går i 8. klasse. Jesper begyndte i skole som normalt. I den nationale test i læsning klarede Jesper sig ud fra de kriteriebaserede kategorier som følger: *Rigtig godt* i “sprogforståelse”, *godt* i “afkodning” og *godt* i “tekstforståelse”.

På baggrund af ovenstående beskrivelse kan vi forvente, at Jesper ca. et år senere vil score omkring 506 i PISA-undersøgelsen i læsning. Dette svarer til niveau 3 i PISAs kategorisystem. Da PISA-undersøgelsen i læsning er rettet mod 15-16 årige elever, vil Jesper gå i 9. klasse, på det tidspunkt testen gennemføres.

I tabel 4.2 er vist, hvordan det forventede resultat i PISA-undersøgelsen i læsning vil ændre sig, hvis vi ændrer Jespers præstationer i den nationale test og oplysninger om hans baggrund.

Ligesom for Aisha er det også tidspunktet for skolestarten og spørgsmålet om herkomst, der er de baggrundsvariable, som betyder mest for den forventede PISA-score i læsning for Jesper. De nævnte variable påvirker dog i højere grad Aishas matematik-score end Jespers score i læsning.

Boks 4.4 Eksempel på forudsigelser på baggrund af intervaller

TABEL 4.2

Eksempel på forventede ændringer i PISA-resultatet i læsning

	Forventet PISA-score
Person fra eksemplet (Jesper)	506
<i>Ændring fra eksemplet</i>	<i>Forventet ændring (+/-)</i>
DNT-resultater	
Sprogforståelse <i>Rigtig god -> fremragende</i>	+10
Afkodning <i>God -> mangelfuld</i>	-17
Tekstforståelse <i>God -> fremragende</i>	+72
Alle 3 profilområder <i>-> Jævn</i>	-70
Baggrundsvariable	
Køn <i>Dreng -> pige</i>	+13
Skolestart <i>Normal -> sen</i>	-27
Etnicitet <i>Dansk -> anden herkomst</i>	-22

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Som beskrevet i denne rapport, er den mest præcise måde, hvorpå man kan forudsige et PISA-resultat, at opstille statistisk baserede modeller, der benytter alle relevante oplysningerne på et og samme tidspunkt.

Modellerne kan også give indblik i, hvordan eleverne inden for et bestemt interval af logit-værdierne for et bestemt profilområde forventes at klare sig i en senere PISA-undersøgelse. Sådanne beregninger giver muligheder for mere enkle fremstillinger af relationen mellem profilområder og PISA-resultater, men forudsigelserne er selvsagt upræcise og usikre, idet de både inkluderer oplysninger fra de andre profilområder og oplysninger om elevens baggrund.

Nedenstående tabel viser således et eksempel på, hvilke resultater i PISA-undersøgelsen i læsning der kan forventes inden for bestemte intervaller i de nationale test for "sprogforståelse". Tabellen viser intervallerne som logit-værdier målt på "Rasch-skalaen" og de dertilhørende forventede PISA-scorer. I parenteser ud for PISA-scoren er endvidere angivet det PISA-niveau, som scoren placerer sig indenfor.

Forventede PISA-scorer på baggrund af opnået testresultat i "sprogforståelse" i de nationale test i 8. klasse

Sprogforståelse (logits)	Forventet PISA-score
-4 til -3	253 (niveau <1b)
-3 til -2	345 (niveau 1a)
-2 til -1	394 (niveau 1a)
1 til 0	442 (niveau 2)
0 til 1	490 (niveau 3)
1 til 2	526 (niveau 3)
2 til 3	537 (niveau 3)
3 til 4	534 (niveau 3)

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

4.4 Forudsigelser på baggrund af en samlet kriteriebaseret kategorisering

De nationale test giver et mere nuanceret billede af elevernes færdigheder end PISA-undersøgelserne. Resultatet af de nationale test består således af en 'elev-profil' med oplysninger om, hvordan eleven har klaret sig inden for tre forskellige profilområder. I modsætning hertil giver PISA-undersøgelsen kun en samlet score for elevens færdigheder i henholdsvis læsning og matematik.

Forudsigelser af PISA-resultaterne forudsætter, at der tages højde for alle tre profilområder på et og samme tidspunkt. Med andre ord udnyttes, at de nationale test giver et mere nuanceret billede til at forbedre forudsigelserne. Imidlertid vanskeliggør denne metode formidlingen, da 'elev-profilerne' omfatter mange forskellige kombinationer af resultater profilområderne imellem.

For at råde bod herpå, viser tabel 4.3 og 4.4 de forventede PISA-resultater i forhold til en samlet kriteriebaseret kategorisering, som er defineret i nedenstående boks.

Boks 4.5 En samlet kriteriebaseret kategorisering

Testresultatet opfattes som:

“Fremragende”, hvis præstationen har været *fremragende* på mindst 2 ud af tre profilområder og mindst *god* i det tredje.

“Rigtig god”, hvis præstationen har været *rigtig god* eller bedre på mindst 2 ud af tre profilområder og mindst *jævn* i det tredje.

“God”, hvis præstationen har været *god* eller bedre på mindst 2 ud af tre profilområder og mindst *mangelfuld* i det tredje.

“Jævn”, hvis præstationen har været *jævn* eller bedre i mindst 2 ud af tre profilområder.

“Mangelfuld”, hvis præstationen har været *mangelfuld* eller bedre i mindst 2 ud af tre profilområder.

“Ikke tilstrækkelig”, hvis præstationen har været *utilstrækkelig* på to eller tre profilområder.

Kilde: Kategoriseringen er foreslået af Kvalitets- og Tilsynsstyrelsen

Da der er gradforskelle inden for hver af de samlede kriteriebaserede kategorier (og da der også er andre forhold, der har betydning for resultatet af en PISA-undersøgelse), kan der forventes flere forskellige PISA-niveauer til hver kategori. Der er i de fleste tilfælde tale om et eller højst to dominerende PISA-niveauer til hver kriteriebaserede kategori.

Tabellerne viser det PISA-niveau, som det forventes, at flest elever (se procentsats) vil placere sig indenfor, på baggrund af en samlet kriteriebaseret kategorisering for henholdsvis læsning (tabel 4.3) og matematik (tabel 4.4). Resultaterne vises særskilt for elever med sen og normal skolestart, idet PISA-resultatet som illustreret i forrige afsnit særligt afhænger af, om eleven testes af PISA i 8. eller 9. klasse.

TABEL 4.3

Forudsigelse af PISA-resultatet i læsning ud fra en samlet kriteriebaseret kategorisering

DNT-kategori	Sen skolestartere	Normale skolestartere
Fremragende	4 (44 %)	4 (76 %)
Rigtig god	3 (73 %)	3 (71 %)
God	2 (78 %)	3 (61 %)
Jævn	1a (64 %)	2 (85 %)
Mangelfuld	1a (56 %)	1a (70 %)
Ikke tilstrækkelig	<1b (78 %)	1b (67 %)

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: Parentesen angiver, hvor mange pct. af eleverne der forventes at opnå det pågældende PISA-resultat på baggrund af den samlede kriteriebaserede kategorisering.

Af tabel 4.3 ses, at der er god overensstemmelse mellem den samlede kriteriebaserede kategorisering for læsning og de forventede PISA-resultater for både sene og normale skolestartere. Det ses dog også af både tabel 4.3 og 4.4, at det samme PISA-niveau kan forventes at være det hyppigst forekommende resultat inden for flere kategorier.

TABEL 4.4

Forudsigelse af PISA-resultatet i matematik ud fra en samlet kriteriebaseret kategorisering

DNT-kategori	Sene skolestartere	Normale skolestartere
Fremragende	4 (67 %)	5 (90 %)
Rigtig god	3 (67 %)	4 (75 %)
God	3 (49 %)	3 (77 %)
Jævn	2 (73 %)	2 (62 %)
Mangelfuld	1 (79 %)	2 (59 %)
Ikke tilstrækkelig	1 (100 %)	1 (100 %)

Kilde: Delrapport 2 – teknisk rapport og dokumentation.

Note: Parentesen angiver, hvor mange pct. af eleverne der forventes at opnå det pågældende PISA-resultat på baggrund af den samlede kriteriebaserede kategorisering. Forudsigelserne af PISA-undersøgelsen i matematik for normale skolestartere baserer sig på modelberegninger, idet testresultaterne fra DNT ikke er tilgængelige for disse elever.

Af tabel 4.4 ses, at der også for matematik kan observeres en god overensstemmelse mellem de forventede PISA-resultater og den samlede kriteriebaserede kategorisering.

Både i matematik og læsning gør det sig også gældende, at de forventede PISA-resultater for de normale skolestartere generelt set er bedre. F.eks. forventes 90 pct. af de normale skolestartere, hvis præstation ud fra den samlede kriteriebaserede kategorisering kan betegnes som *fremragende*, at opnå PISA-niveauet 5 i matematik. For de sene skolestartere forventes det mest hyppige resultat (67 pct.) ud fra en tilsvarende præstation i den nationale test for matematik at være PISA-niveau 4.

4.5 Opsamling

Dette kapitel har vist, hvordan det fremover er muligt at foretage en statistisk baseret forudsigelse af, hvordan en given elev vil score i en PISA-undersøgelse på baggrund af vedkommendes resultater fra den nationale test og oplysninger om elevens baggrund. Endvidere er der givet eksempler på, hvordan forudsigelserne fungerer.

Såfremt man er interesseret i at læse mere om undersøgelsens tekniske baggrund, henvises læseren til Delrapport 2.

Boks 4.6. Rapportens datagrundlag

Datagrundlaget for analyserne i denne rapport består af resultater fra de nationale test i matematik og læsning i 6. klasse i 2010 og læsning i 8. klasse fra 2011 og 2012 samt testresultater i læsning og matematik fra PISA 2012.

Datamaterialet omfatter resultaterne fra de nationale test, såfremt disse forelå fra de elever, der deltog i PISA i 2012. De fleste af de elever, der deltog i PISA i 2012 gik i 6. klasse i 2009 og 2010. Idet der kun foreligger de nationale testresultater fra 6. klasse i 2010 og fra 8. klasse i 2012, da de blev testet af PISA, har det været nødvendigt at estimere effekten af undervisningen i 9. klasse, som de fleste PISA-elever havde modtaget, for at kunne opstille en statistisk model til forudsigelse.

Ud over testresultater fra de nationale test ligger oplysningerne om elevens køn, herkomst, alder og socioøkonomiske status også til grund for analyserne.



DAMVAD
COPENHAGEN . OSLO . STOCKHOLM

Klarabergsviadukten 63,
SE-101 23 Stockholm

Badstuestræde 20
DK-1209 Copenhagen K

Grensen 13
N-0159 Oslo