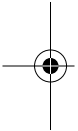


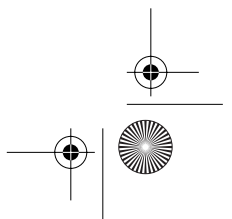
**Marit Kjærnsli, Svein Lie,
Rolf Vegar Olsen og Astrid Roe**

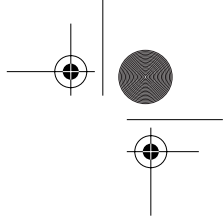
TID FOR TUNGE LØFT

Norske elevers kompetanse i naturfag,
lesing og matematikk i PISA 2006



Universitetsforlaget





© Universitetsforlaget 2007

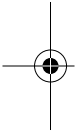
ISBN 978-82-15-01146-2

Materialet i denne publikasjonen er omfattet av åndsverklovens bestemmelser. Uten særskilt avtale med rettighetshaverne er enhver eksemplarfremstilling og tilgjengeliggjøring bare tillatt i den utstrekning det er hjemlet i lov eller tillatt gjennom avtale med Kopinor, interesseorgan for rettighetshavere til åndsverk. Utnyttelse i strid med lov eller avtale kan medføre erstatningsansvar og inndragning og kan straffes med bøter eller fengsel.

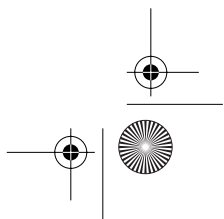
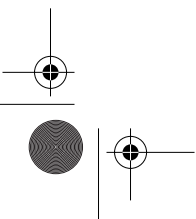
Henvendelser om denne utgivelsen kan rettes til:

Universitetsforlaget AS
Postboks 508 Sentrum
0105 Oslo

www.universitetsforlaget.no



Omslag: AIT Trykk Otta AS
Sats: Laboremus Prepress AS
Trykk og innbinding: AIT Trykk Otta AS
Boken er satt med: Times 11/13 pkt.
Papir: 90 g G-print



Vedlegg 1

Kort om gjennomføringen av PISA 2006

Utvalg

Populasjonen i undersøkelsen omfattet 15-åringer, definert som alle elever født i 1990 som gikk på skolen. I motsetning til i mange andre undersøkelser valgte man alder og ikke årstrinn som grunnlag for utvelgelsen. Det vil si at man undersøkte et helt årskull uavhengig av antall år på skolen. I Norge er forskjellen mellom aldersbasert og trinnbasert populasjon minimal fordi over 96 prosent av populasjonen går på 10. trinn. Omtrent 2 prosent befinner seg på 9. trinn, og litt over 1 prosent går på videregående skole.

De norske skolene ble tilfeldig trukket fra alle landets skoler basert på offisiell skolestatistikk for skoleåret 2004/2005, og selve trekkingen ble foretatt av et internasjonalt senter, som da også fungerte som en kvalitetskontroll. Totalt ble 227 skoler trukket ut, fordelt på fire strata. For hver skole som ble trukket ut, ble det også trukket to reserveskoler med omtrent de samme karakteristika. Tabell 1 viser en oversikt over antall skoler og antall aktuelle elever innen hvert stratum.

Tabell 1: Oversikt over aktuelle skoler og elever samt uttrukne skoler

Stratum	Antall aktuelle skoler	Antall aktuelle elever	Antall uttrukne skoler
1. Store ungd.skoler (> 29 elever)	667	54133	176
2. Små ungd.skoler (15–29 elever)	191	4102	19
3. Veldig små ungd.skoler (< 15 elever)	284	1956	14
4. Videregående skoler	372	770	18
Sum	1514	60961	227



I stratum 3 og 4 er det veldig få elever, men relativt mange skoler. For å få et mer effektivt utvalg ble det trukket spesielt få skoler fra disse to strataene. Av de 18 videregående skolene som ble trukket ut, var det for eksempel bare fire skoler som hadde elever født i 1990. I vedlegg 2 er det gjort rede for de konsekvensene det hadde at det i noen strata ble trukket forholdsvis få skoler. På de skolene som ble trukket ut til å delta, ble 30 tilfeldige elever trukket enkeltvis. Det vil si at det ikke var hele grupper eller klasser som deltok. På skoler med færre elever enn 30 deltok alle.

Den norske undersøkelsen ble gjennomført i april 2006, og med i alt nesten 4700 elever fra 203 skoler. Det var svært høy norsk deltakelse både på skolenivå (ca. 90 %) og på elevnivå (ca. 85 %), noe som er viktig for å tilfredsstille de internasjonale kravene til deltakelse. I noen land blir skolene og elevene pålagt å delta, mens i land der deltakelse er frivillig, kan det oppstå problemer, slik tilfellet var i Nederland i 2000 og i Storbritannia i 2003.

Fritak av elever

Selv om det stilles strenge krav til utvalg og høy deltakelse, reises det ofte spørsmål om hvorvidt resultatene i ulike land kan sammenliknes. For eksempel trekkes det fram at elever som i andre land ville gått på spesialskoler, er integrert i vanlige klasser i norsk skole. Dette er noe man internasjonalt har tatt hensyn til. Alle elever født i 1990, og som gikk på skole, skulle i utgangspunktet være med i utvalget i alle land, uansett hvilken type skole elevene gikk på. Det var imidlertid mulig å holde noen skoler eller enkeltelever utenfor, men det var strenge internasjonale kriterier for hvilke elever eller skoler dette kunne gjelde. Stort sett dreide det seg om elever som var så fysisk eller psykisk funksjonshemmet at de ikke var i stand til å besvare oppgavene, eller fremmedspråklige elever som hadde vært for kort tid i landet. I tekstboks 1 er det gitt en mer detaljert oversikt over kriteriene for deltakelse. Et hovedprinsipp var at undersøkelsen skulle være så inkluderende som mulig. For hvert land er det beregnet hvor mange prosent av elevene som ble ekskludert, og disse tallene varierer lite mellom land.

Design

I praksis innebar undersøkelsen at elevene fikk en to timers prøve med oppgaver i de tre fagområdene lesing, matematikk og naturfag. Deretter besvarte de et spørreskjema, blant annet for at man skal kunne se hvilke karakteristiske trekk ved skoler og elever som kan ha betydning for prestasjonene. Av samme grunn var det også et spørreskjema til skolelederne. Det ble derimot ikke utviklet spørreskjema til lærerne, først og fremst fordi



Tekstboks 1: Kriterier for fritak av elever

Hvilke elever deltok ikke i undersøkelsen?

- **Elever med fysisk funksjonshemming.** Dette gjelder bare elever med en type fysisk handikap som kan hindre dem i å gjennomføre prøven.
- **Elever med psykisk og/eller emosjonell funksjonshemming.** Dette er elever som er vurdert av PP-tjenesten eller andre fagpersoner som psykisk eller emosjonelt funksjonshemmet. Disse skal bare utelukkes hvis de er uskikket til å forstå og følge instruksjonene i undersøkelsen. Elever må ikke utelukkes bare fordi de presterer dårlig på skolen eller har generelle disiplinproblemer.
- **Elever med begrensede norskkunnskaper.** Dette er elever som ikke er i stand til å lese norsk, og som derfor vil få problemer med å forstå språket i oppgavene, for eksempel elever som har hatt mindre enn ett års undervisning i norsk.

elevene ble trukket ut enkeltvis og ikke klassevis, og på de fleste skoler ville det være svært få elever som kunne koples til hver enkelt lærer.

Oppgavene

Oppgavene i PISA tok utgangspunkt i det internasjonale rammeverket som var utviklet for undersøkelsen (OECD 1999b, 2000, 2003, 2006). Oppgavene var stort sett organisert i oppgaveenheter, der hver enhet besto av en tekst med en rekke oppgaver knyttet til teksten. Tekstene skulle være virkelighetsnære og var derfor så langt det er mulig hentet fra aviser, tidsskrifter, brosjyrer eller liknende.

For å dekke et så stort faglig område som mulig var det en rotasjon av oppgavene i 13 forskjellige oppgavehefter, men hver elev besvarte bare ett hefte. Naturfag hadde hovedfokus i denne undersøkelsen og dominerte derfor de fleste heftene. Likevel inneholdt alle heftene forskjellige kombinasjoner av de ulike fagområdene.

Oppgavene hadde flere formater, men kan deles inn i følgende to hovedgrupper:

Flervalgsoppgaver: I disse oppgavene skulle elevene markere svaret med avkryssing. Det var to forskjellige typer flervalgsoppgaver. I det enkleste og vanligste formatet skulle elevene velge riktig svar blant flere alternati-



ver. Det andre formatet krevde at elevene krysset av for ja/nei eller riktig/galt på en serie av påstander eller spørsmål.

Åpne oppgaver: I de åpne oppgavene måtte elevene selv skrive svar. Graden av «åpenhet» varierte mye fra oppgave til oppgave. I noen oppgaver skulle eleven bare skrive et tall eller noen få ord, mens i andre oppgaver ble det stilt krav om fullstendige resonneringer, lengre begrunnelser, utregninger, forklaringer eller andre vurderinger.

I vedlegg 3 er det gitt flere eksempler på oppgaveenheter. Alle de frigitte oppgavene er også lagt ut på prosjektets hjemmeside www.pisa.no.

Utvikling av oppgaver er en lang og tidkrevende prosess. I denne fasen ble alle oppgavene i lesing og matematikk hentet fra tidligere undersøkelser, mens det var behov for å lage mange nye oppgaver i naturfag. Alle deltakerlandene ble oppfordret til å sende inn forslag til oppgaver. En internasjonal gruppe, bestående av noen nasjonale sentre, deriblant Norge, var ansvarlig for å utvikle nye oppgaver og bearbeide nye forslag. Under denne prosessen ble oppgavene også prøvd ut i små elevgrupper. Oppgavene ble deretter vurdert av representanter fra hvert land. De ble vurdert etter hvor relevante de ble antatt å være for en 15-åring, hvorvidt det var trekk ved oppgaven som gjorde at spesielle elevgrupper hadde fordeler ved å få denne oppgaven, og hvorvidt oppgaven kunne oversettes likt med samme vanskegrad til alle språk. Etter denne vurderingen ble det gjennomført en generalprøve der to til tre ganger så mange oppgaver som skulle være med i den endelige undersøkelsen, ble prøvd ut blant elever i alle land.

På grunnlag av resultater fra generalprøven ble så det endelige oppgaveutvalget foretatt. Kriteriene for utvalget av oppgaver var først og fremst at oppgavene til sammen skulle danne en god operasjonalisering av det man ønsket å måle i PISA (se OECD 2006). Dernest var det viktig at de skulle ha en passende vanskegrad, at det ikke skulle være for store kulturelle skjevheter, og at hver oppgave i rimelig grad skulle skille (diskriminere) mellom sterke og svake elever.

Parallelt med utvikling av oppgaver ble det også utviklet koder for vurdering av de åpne oppgavene. En viktig hensikt med generalprøven var også å prøve ut disse foreløpige kodene, og ikke minst å gi et grunnlag for videreutvikling av disse kodene basert på elevsvarene. For å få innsikt i hvordan elevene hadde tenkt, hvilke løsningsstrategier de hadde brukt og eventuelle alternative forestillinger de hadde, ble det på bakgrunn av elevsvar fra generalprøven utarbeidet et tosifret kodesystem for noen av oppgavene i naturfag og matematikk. I tillegg til å gi informasjon om riktige svar vil det for slike oppgaver være mulig å gå tilbake og analysere i detalj hvordan elevene har løst oppgavene.

Spørreskjemaene

Elevspørreskjema

Elevene ble spurt om kjønn, alder, nasjonalitet, hvilket språk de snakket hjemme, og hvilke aktiviteter de var med på. Videre fikk de spørsmål om hjemmebakgrunn, mors og fars utdanning og yrke, og de fikk spørsmål som skulle fange opp familiens kulturelle og økonomiske kapital. De fikk også en rekke spørsmål om naturfagundervisningen, om holdninger til naturfag, om bruk av IKT og om skolen og forhold ved den. Disse spørsmålene var obligatoriske i alle land. I tillegg fikk de norske elevene de samme spørsmålene om lesevaner som ble gitt i 2000 samt om deres bruk av læringsstrategier i naturfag. Elevspørreskjemaet er gjengitt i sin helhet på den norske hjemmesiden www.pisa.no.

Skolespørreskjema

Skolelederne på hver deltakerskole svarte på et spørreskjema om skolens beliggenhet, skole- og klassestørrelse, ressurser, organisering av skolen, ansvarsområder, antall lærere i hel- og deltidsstillinger, undervisningsforhold samt lærerklima. Internasjonalt er det lagt stor vekt på å studere sammenhenger mellom skoledata og elevprestasjoner, for på den måten å prøve å finne ut hva som utgjør «den gode skole».

Det er stor forskjell på skolekulturen i de ulike land, og det var mange spørsmål i det internasjonale skolespørreskjemaet som ikke passet like bra til norske forhold. Skolespørreskjemaet er lagt ut på hjemmesiden www.pisa.no.

Kvalitetssikring

For å få gode og sammenliknbare data ble det internasjonalt lagt stor vekt på å kvalitetssikre alle prosedyrer; som utvalg av skoler og elever, oversetting av oppgaver og spørsmål, kontakt med skoler, gjennomføring av undersøkelsen på skolene, vurdering av åpne oppgaver samt innskriving av data (OECD 2005b). For å illustrere hvordan dette ble gjort i praksis, er noen av områdene beskrevet litt mer i detalj i det følgende.

Krav til deltakelse

Det var strenge internasjonale krav til deltakerprosenten både på skolenivå og på elevnivå, og land som ikke innfridde disse kravene, ble som nevnt tatt ut av de internasjonale sammenlikningene. Utvalget av skoler var basert på skolestatistikk i hvert enkelt land. Selve trekkingen av de enkelte skolene som skulle delta, ble foretatt internasjonalt. Deltakerskolene sendte deretter



inn lister over alle elever ved skolen som var født i 1990. Ut fra disse listene ble elevene trukket enkeltvis og tilfeldig.

Øversetting

I internasjonale undersøkelser har man erfart at språklige forskjeller og kulturelle tilpasninger har vært et problem når det gjelder å framskaffe pålitelig informasjon (se for eksempel Grisay 2003). Samme tekst kan oversettes på ulike måter og likevel være «riktig» oversatt, men ulike oversettelser kan gjøre vanskegraden forskjellig. For å gjøre disse problemene så minimale som mulig og for å imøtekomme mye av den kritikken som har vært reist mot komparative studier når det gjelder oversetting, har det vært svært strenge krav til prosedyrene for oversetting. Alle oppgavetekster og spørsmål var i utgangspunktet utviklet på både engelsk og fransk, og alle land måtte ha uavhengige oversettelser fra begge språk. De nasjonale sentrene i hvert land var ansvarlig for å oversette og ferdigstille testmaterialet. Oversettelsene ble deretter kontrollert av språkekspertene i de respektive språk. Her ble både tekst og layout i forhold til den originale versjonen kontrollert.

Gjennomføring

For å være sikker på at undersøkelsen ble gjennomført likt i alle land, ble det holdt samlinger for alle deltakerskolene før gjennomføringen. Det ble også laget detaljerte veiledninger internasjonalt, og disse måtte oversettes og brukes i alle land. Her var det informasjon om hva som måtte gjøres før og etter undersøkelsen, samt en detaljert beskrivelse for selve gjennomføringen. Det var blant annet krav til nøyaktig tidtaking og om at instruksjonen til elevene skulle leses ordrett.

Vurdering og koding av elevsvar

PISA-undersøkelsen besto som nevnt av både flervalgsoppgaver og åpne oppgaver. Elevsvarene på de åpne oppgavene måtte vurderes og gis en kode, og for at det skulle vurderes så likt som mulig både innen hvert land og mellom alle land, var det internasjonalt utviklet en detaljert kodeveiledning for hver oppgave. Kodeveiledningene ble utarbeidet på bakgrunn av elevsvar fra generalprøven. For å få en felles forståelse av kodene måtte alle land delta på en internasjonal samling. De som vurderte de norske besvarelsene, satt samlet under hele prosessen og fikk detaljert veiledning før hver ny oppgave. For å kontrollere samsvaret mellom dem som vurderte oppgavene, ble en bestemt andel av heftene kodet av fire uavhengige personer. Samsvaret mellom koderne viste seg å være svært bra både for Norges vedkommende og for de fleste andre land. Data på dette vil bli publisert senere i en teknisk rapport.