

Et kritisk blik på opgaverne i PISA med særlig vægt på matematik

Inge Henningsen¹

Afdeling for Anvendt Matematik og Statistik, Københavns Universitet

Artiklen præsenterer en kritik af PISA 2003 gennem en analyse af nogle af de offentliggjorte opgaver. Artiklen argumenterer for at PISA's intentioner om at teste elevernes kompetence i virkelige situationer ikke er opfyldt. Mange opgaver er i stedet typiske test- eller skoleopgaver hvor eleverne f.eks. skal kende reglerne for "opgavekontrakten", i stedet for at inddrage ægte hverdagsræsonnementer. Artiklen viser eksempler på manglende forbindelse mellem opgaven og den virkelighed den skal modellere, eksempler på manglende oplysninger i opgaven, knudret sprog, usammenlignelige oversættelser, uklarhed i kravene til elevernes svar, samt fejl i opgavernes krav til elevernes ræsonnement. Der argumenteres for at hvis ikke alle elever har samme forhåndsviden om hvordan man rent teknisk skal forholde sig for at score flest muligt point, er det et problem for validiteten af PISA, f.eks. når man sammenligner mellem lande.

Hvad er PISA?

PISA er en forkortelse for Programme for International Student Assessment der er et OECD-program hvor der i 2003 var 41 deltagende lande (30 OECD-lande og 11 partnerlande). Hensigten med PISA er at teste 15-16-åriges kundskaber inden for læsning, naturvidenskab, matematik og problemløsning. PISA blev afholdt første gang i 2000 med hovedvægt på læsning. I 2003 lå hovedvægten på matematik, og i 2006 på naturvidenskab. I PISA 2003 deltog 276.165 elever, heraf 4.218 elever fra 207 danske skoler. Alle var født i 1987. I Danmark udføres PISA af et konsortium bestående af Amternes og Kommunernes Forskningsinstitut, Danmarks Pædagogiske Universitet og Socialforskningsinstituttet. Den danske del af PISA 2000 er afrapporteret i (Andersen et al, 2001), og en redegørelse for PISA 2003 findes i (Mejding, 2004). De to rapporter vil i det følgende blive refereret til som (PISA, 2000) og (PISA, 2003).

¹ Jeg vil gerne indføje en tak til de to kritikere der har foretaget review, og redaktionen for den gode og konstruktive tilbagemelding jeg har fået på artiklen undervejs.

Allerede før resultaterne fra PISA 2003 blev offentliggjort havde projektet skabt debat, da der ud fra resultaterne i PISA 2000 var blevet rejst tvivl både om PISA-undersøgelseernes pålidelighed, deres konklusioner og deres politiske rolle. Efter offentliggørelsen af resultaterne fra PISA 2003 har bl.a. professor Erik Jørgen Hansen (2005) i Dansk Pædagogisk Tidsskrift sat spørgsmålstegn ved PISA's videnskabelighed, og professor Peter Allerup (2005) har samme sted vist hvorfor man ikke umiddelbart kan sammenligne resultaterne fra PISA 2000 og PISA 2003, og han har samtidig vist at visse af de brugte skalaer er skæve f.eks. i forhold til køn.

Der har derimod ikke været megen offentlig interesse for opgavekvaliteten i PISA, hvad der kan hænge sammen med at opgaverne ikke i deres helhed er offentligt tilgængelige. Skal man dømme ud fra de offentliggjorte opgaver, er der imidlertid god grund til at sætte spørgsmålstegn ved selve opgaveteksterne og inddrage dette i vurderingen af de resultater og anbefalinger som PISA 2003 har resulteret i. I den følgende gennemgang af udvalgte af PISA-opgaver giver jeg eksempler på direkte fejl i opgaverne og på oversættelser der ændrer opgaverne og derved gør sammenligning mellem lande problematisk. Jeg diskuterer problemer med den "real-world"-kontekst for matematik og naturvidenskabelige opgaver som er PISA's mål, og undersøger om opgaverne i matematik giver anledning til den matematiseringsproces som PISA intenderer.

Modellering og "real world"

Jeg er meget enig i de intentioner for undervisning som PISA tager udgangspunkt i: At undervisningen skal give eleverne et grundlæggende handleberedskab og fremme deres evne til at bruge f.eks. matematik og naturvidenskab i deres fremtidige liv. I konsekvens heraf er opgaverne i PISA ikke tænkt at skulle være traditionelle skoleopgaver i modsætning til f.eks. Third International Mathematics and Science Study 1995-96 (TIMSS), der havde til formål at teste de unges skolekundskaber i matematik og naturvidenskab (se f.eks. Allerup et al, 1998). De skal dreje sig om "det virkelige liv". Det hedder f.eks. om matematikopgaverne i PISA 2003:

Derfor formuleres opgaverne i PISA i, hvad der betegnes som "real-world situationer" således, at eleverne får mulighed for at demonstrere matematisk viden og kunnen i sammenhænge, der for manges vedkommende afviger fra opgaver i skolen, og dermed vil være en udfordring, eleven skal handle på. Elevernes problembehandling forsøges initieret af beskrivelser af en problemstilling i en sammenhæng, der formodes at være relevant for et ungt menneske. (PISA, 2003, s. 39)

Tilsvarende formuleringer findes om opgaverne i naturvidenskab og i læsning.

I matematik lægges stor vægt på at eleverne ikke bare skal “sætte ind i formler”, men at der skal foregå en aktiv matematiseringsproces, hvor man oversætter frem og tilbage mellem matematik og hverdagsproblem:

Problemstillinger i opgaverne i PISA er formuleret således, at de er tilgængelige for matematisk behandling i en “real-world” sammenhæng. Hensigten med sammenhængen er at aktivere eleven til at undersøge, hvilke matematiske begreber og processer, der kan indgå i en matematisk behandling af det stillede problem. Det grundlæggende begreb i denne proces er “matematisering”, som i hovedtræk dækker over tre delprocesser. Først oversættelsen af det givne problem i “real world” til matematikkens verden, for eksempel gennem opstilling af en model. Andet trin er arbejdet med den matematiske model ved hjælp af matematikken inden for matematikkens verden. Tredje trin er oversættelse af resultatet af den matematiske bearbejdning af modellen til et resultat om fænomenet i “real-world”, afsluttende med en refleksion over resultatets pålidelighed. (PISA 2003, s. 42)

I det følgende vil jeg tage udgangspunkt i PISA’s egne målformuleringer og undersøge hvordan opgaverne forholder sig til intentionerne.

Opgavekontrakt og autentiske opgaver

PISA opgaverne skal som nævnt ovenfor tage udgangspunkt i “real-world” situationer. Så selv om opgaverne er konstruerede, er de opbygget som om det drejede sig om autentiske tekster f.eks. ved at bruge formuleringer som “En geolog udtaler”, “Care Medical gruppen giver følgende oplysninger”, etc. Et særligt træk er at der indføres distraktorer, irrelevant information, i teksten (se f.eks. PISA 2003, s. 55). Men i sådanne konstruerede tekster hvor man ikke kender afsenderen, kan man ikke bruge de læsekompetencer som i det virkelige liv ofte gør det muligt at styre helskindet gennem en tekst, selv om den ikke er særlig klart formuleret. Hvor man f.eks. kan manøvrere uden om de irrelevante oplysninger fordi man kan se den sammenhæng i hvilken de er relevante. I PISA-opgaverne er de irrelevante oplysninger sat ind alene for at være irrelevante, men der er ingen pejlemærker i teksten til at skelne mellem hvad man kan bruge, og hvad man ikke kan bruge i forhold til den givne opgave.

De principielle problemer som ligger indbygget i de konstruerede opgaver, træder tydeligt frem hvis man ser på den model for læsefærdighedens delaspekter som opstilles i rapporten om PISA 2003. Her ses læsning som et samspil mellem at bruge information fra teksten og anvende forhåndsviden, at reflektere over indhold og form, at have forståelse for delelementers sammenhæng, at vurdere teksten som helhed etc. (PISA, 2003, s. 134). Dette kan ikke bruges i tekster der giver sig ud for at være noget andet end de egentlig er. I den praktiske undervisningssituation tackles disse

problemer gennem opbygningen af en “opgavekontrakt”, hvor eleverne samler en – måske implicit – viden om hvad der menes med forskellige opgaveformuleringer, og om hvordan man skræller den “realistiske” indpakning af de konstruerede problemer. De elever der deltager i PISA, opererer i forhold til en lang række lokale og nationale opgavekontrakter der stemmer mere eller mindre overens med de konventioner der ligger i PISA opgaverne. I analysen af PISA’s resultater savner jeg en problematisering af dette forhold.

I det følgende vil jeg imidlertid tage PISA på ordet. Jeg vil ikke læse opgaverne som skoleopgaver, men som opgaver hvor virkelige problemstillinger formaliseres og analyseres af 15-årige der både har logisk sans, forhåndsviden og lyst og evne til at bruge deres kompetencer.

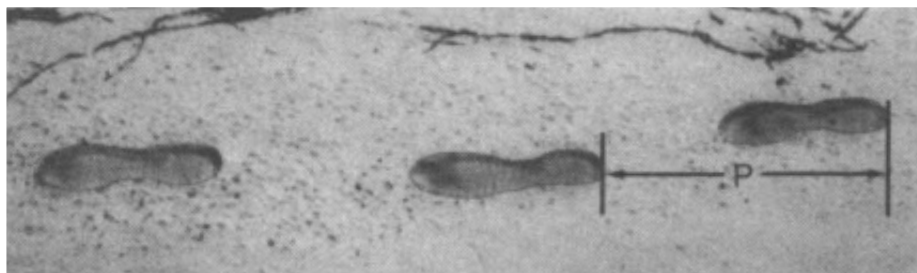
Analyse af PISA-opgaver

Eksempel 1: Vi skal gå hånd i hånd?

I denne opgave opstilles en model for mænds gang, og eleverne skal bruge den til at beregne visse størrelser knyttet til gang. I min kommentar problematiseres den anvendte model og den valgte illustration. Samtidig påpeges det at opgaven i den valgte form er en ren “skoleopgave” som slet ikke forholder sig til den matematiseringsproces som skulle være et grundlæggende element i PISA’s opgaver, og hvor eleven kan opstille en matematisk model og kan gå frem og tilbage mellem “virkeligheden” og modellen.

I opgaven modelleres mænds gang ved en ligefrem proportionalitet mellem antal skridt pr minut og skridtlængde med en proportionalitetsfaktor på 140. I omtalen af opgaven skriver ACER, et af de fire uddannelseskonsortier der er hovedansvarlige for PISA, imidlertid:

Gang



Billedet viser fodsporene fra en mand, der er ude at gå. Skridtlængden P er afstanden mellem det bageste af to fodaftryk, der følger lige efter hinanden.

Formlen $\frac{n}{P} = 140$ angiver et forhold for mænd mellem to størrelser n og P ,

hvor

n = antallet af skridt pr. minut og

P = skridtlængden i meter.

Spørgsmål 1: GANG

Hvis formelen gælder, når Henrik går, og Henrik tager 70 skridt i minuttet, hvad er Henriks skridtlængde så? Vis, hvordan du nåede frem til dit resultat.

Spørgsmål 2: GANG

Benny ved, at hans skridtlængde er 0,80 meter. Formlen gælder for Bennys gang.

Beregn Bennys ganghastighed i meter pr. minut og i kilometer pr. time. Vis, hvordan du nåede frem til dit resultat.

Figur 1. Gang.

Students would be familiar with seeing their footprints in sand or soil but probably would not have given much thought to the relationship between the “number of steps taken per minute” and “pace length”. (Thomson, 2004, s. 64)

Men har PISA's opgavestillere selv tænkt over forholdet mellem antal skridt pr minut og skridtlængde? Den formel der i opgaven angives for mænds gang, medfører at antallet af skridt pr minut er ligefremt proportionalt med skridtlængden. Hvis en mand tager *længere* skridt, så stiger *antallet* af skridt pr. minut. Tager han skridt på $\frac{1}{2}$ meter, tager han 70 skridt i minuttet. Forøges skridtlængden til 1 meter, tager han 140 skridt i minuttet. Heraf følger f.eks. at to mænd der har forskellig skridtlængde, ikke kan følges ad når de går. Er dette en realistisk model for mænds gang?

At det slet ikke var ligefrem proportionalitet mellem skridtlængde og skridttakt, men en omvendt proportionalitet der var tænkt på i opgaven, fremgår af fortsættelsen af ovenstående citat:

Students needed to recognize that as the pace length increases, so the number of steps per minute will decrease, and in order to gain credit for this item needed to carry out the actual calculation. (Thomson, 2004, s. 64)

Her er simpelthen en modstrid mellem opgavekonsortiets beskrivelse af det fænomen der modelleres (“as the pace length increases, so the number of steps per minute will decrease”), og den formel der bruges i opgaven.

Går man fra modellen videre til spørgsmålene, så skulle det være et særkende ved PISA at eleverne ikke som i den gængse skolematematik bare løser opgaver ved at sætte ind i en formel, men aktivt skal oversætte frem og tilbage mellem “virkelighed” og matematik. Det er dog ikke tilfældet i opgaven om *Gang*. Her får man tværtimod eksplicit at vide at formlen gælder for Henriks og Bennys gang, og skal så bare sætte ind i formlen og vise at man kan løse en ligning. Det betyder ikke noget for besvarelsen om man har tænkt over formlen. Snarere tværtimod. Der er nærliggende fare for at eleven vil begynde at tvivle på sine egne overvejelser hvis vedkommende afslutter “med en refleksion over resultatets pålidelighed.” Det var i hvert fald min oplevelse da jeg regnede ud at Henrik i opgaven gik med en hastighed på 2,1 km/t.

Hvordan forholder opgaven sig til den matematiseringsproces som skulle være et grundlæggende element i PISA's opgaver, hvor eleven opstiller en matematisk model, arbejder i den opstillede model, går frem og tilbage mellem “virkeligheden” og modellen, og forholder sig kritisk til konklusionerne. I opgaven om “gang” er modellen imidlertid opstillet på forhånd, og eleverne inviteres ikke til at reflektere over den. Tværtimod.

Illustrationen er et kapitel for sig selv. Her defineres skridtlængde som “afstanden mellem det bagerste af to fodaftryk, der følger lige efter hinanden”. Men blot et overfladisk blik på fotografiet viser at de to afbildede skridt ikke er lige lange. Så hvad er “skridtlængden” hos den mand der har gået der? Havde man målt hans andet skridt, ville man have fået et helt andet resultat (og med skridt der er 35% længere, ville han gå næsten dobbelt så hurtigt). Giver det mening at lave en formel for sammenhæng mellem skridtlængde, defineret som anført, og skridthastighed hvis personer går så uregelmæssigt som angivet på illustrationen? Et vigtigt led i en matematiseringsproces er at vurdere om den valgte model giver mening i forhold til det forelagte problem. Der må findes tusinder af fotografier af regelmæssige fodspor, hvorfor så vælge en illustration der vil forvirre de elever der reflekterer over fotografiet, og belønne dem der bare går i gang med at sætte tal ind i formlerne uden overhovedet at tænke over hverken tekst eller illustration? Hvordan svarer det til den ovenfor citerede beskrivelse af PISA’s matematiseringsproces?

Eksempel 2: Vækst eller opvækst?

I forbindelse med en opgave om pigers/kvindes vækst vil jeg diskutere validitetsproblemer affødt af usammenlignelighed mellem oversættelser og uerkendte begrænsninger i inferens fra tværsnitsgrafer. Spørgsmålet om autentiske materialer og det tilladelige i at ændre på data i opgavesammenhæng rejses.

Begynder man med et kritisk blik på overskriften, så handler opgaven faktisk om “Vækst” og ikke om “Opvækst”. Den næste linie, “De unge bliver højere”, har ikke noget med grafen at gøre, for den foreliggende tværsnitsgraf, der består af forskellige fødselsårsgange, siger intet om den tidsmæssige udvikling af højden for unge i en given alder. Hvad er begrundelsen for at sætte scenen for opgaven med fejlagtige overskrifter?

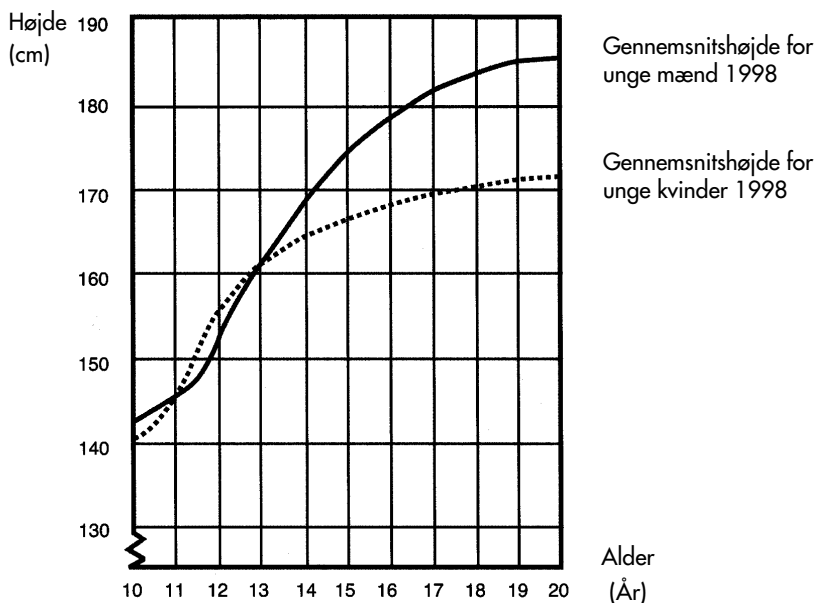
Det første spørgsmål i opgaven tager udgangspunkt i påstanden: “Siden 1980 er gennemsnitshøjden for 20-årige kvinder forøget med 2,3 cm til 170,6 cm.” Udsagnet kan forekomme mærkeligt løsrevet fra konteksten, da grafen viser at de 20-årige hollandske kvinder i 1998 har en gennemsnitshøjde på ca. 172 cm. Selvfølgelig kan eleven trække 2,3 fra 170,6 uden at vide hvad opgaven drejer sig om. Men er det så ikke ren skolematematik? Og hvis det drejer sig om de hollandske kvinder, hvorfor passer figur og tekst så ikke sammen?

Spørgsmål 2 illustrerer oversættelsesproblemer. Formuleringen “den hastighed, hvormed pigernes gennemsnitshøjde vokser, aftager” er meget knudret og svært gennemskuelig. På norsk er den tilsvarende formulering “at veksthastigheden for jenter i gjennomsnitt avtar”. Dette er en langt mere mundret formulering. Til gengæld er norske forskere bekymrede over brugen af fagtermen “veksthastighed”. Det kan være

Opvækst

De unge bliver højere

Denne graf viser gennemsnitshøjden i 1998 for henholdsvis unge mænd og unge kvinder i Holland.



Spørgsmål 1

- Siden 1980 er gennemsnitshøjden for 20-årige kvinder forøget med 2,3 cm til 170,6 cm. Hvad var gennemsnitshøjden for 20-årige kvinder i 1980?

Spørgsmål 2

- Forklar, hvordan man ud fra grafen kan se, at den hastighed, hvormed pigernes gennemsnitshøjde vokser, aftager efter 12 års alderen.

Spørgsmål 3

- I hvilket tidsrum i deres liv er kvinder ifølge denne graf højere end mænd på samme alder?

Figur 2. Opvækst.

vanskeligt teoretisk at afgøre hvilken formulering der er sværest, og det kan være at det er forskellige elever der vil have vanskeligheder med hver af de to formuleringer. Men det er helt klart ikke "præcis den samme opgavetekst" i de to lande, hvad der efter PISA's egen vurdering er "en fare for validiteten" (PISA 2003, s. 27). Dette gælder i endnu højere grad spørgsmål 3 hvor ordet "gennemsnitlig", der er forbindelsen mellem spørgsmålet og grafen, helt er faldet ud af den danske tekst, mens det f.eks. indgår både på norsk og på engelsk.

Set fra et modellerings synspunkt er det meget problematisk at opgavestillerne i spørgsmål 3 går ud fra at man uden videre kan slutte fra en tværsnitsgraf der er stykket sammen af højden i året 1998 hos piger/kvinder født i forskellige år, til udsagn der vedrører forløb over flere år hos generationer af kvinder. I opgaven taler man om "tidsrum" i kvinders liv, men den afbildede tværsnitsgraf siger intet om dette, fordi hver kvinde kun indgår i et år. Man må i hvert fald eksplicitere de forudsætninger der skal gøres, for at man på grundlag af grafen kan svare på det stillede spørgsmål, og det har eleverne hverken tid eller lejlighed til. At vide hvilke spørgsmål man kan stille til en given model, fremhæves ofte som vigtigt i matematik. Det er derfor uheldigt at PISA vælger opgaver hvor netop dette aspekt skal ignoreres for at opgaven kan løses. Som ovenfor nævnt bliver opgaven for de danske elever yderligere kompliceret af at oversætterne har udeladt henvisningen til at det drejer sig om gennemsnitshøjde.

Endnu en sproglig note: Det er altid vanskeligt at finde ud af hvornår man skal sige piger, og hvornår man skal sige kvinder, men det virker ganske ulogisk, at man i Danmark har valgt at kalde de 12-20 årige for piger (spørgsmål 2) og de 11-13 årige for kvinder (spørgsmål 3). Selv om det ikke behøver at være et bevidst valg for at forvirre eleverne, så er det igen et problem når f.eks. norsk PISA taler om "jenter" hele vejen igennem og dermed ikke inkluderer dette distraherende element.

Endelig er der spørgsmålet om "realisme". Grafen antyder at de 20-årige hollandske piger/kvinder er 4 cm højere end de 16-årige. Dette forekommer umiddelbart usandsynligt i forhold til den medicinske viden om kvinders vækst. En litteratursøgning giver da heller ingen hollandske højdeundersøgelser fra 1998, men en artikel (Fredriks, 2000) giver data fra 1997 og viser et lidt andet billede (se tabel 1). Her vokser pigerne f.eks. kun 2 cm fra de er 16 til de er 21. I denne undersøgelse er pigerne i gennemsnit højere end drengene fra de er ca. 10 år til de er ca. 13 år, og forskellene er noget mindre end på PISA's figur. Man genkender imidlertid tallet 170,6, dog som her er gennemsnitshøjde for 21-årige kvinder. Er det denne undersøgelse der er udgangspunkt for PISA's opgave, hvor man så bare har ændret på tallene?

Alder	Dreng	Piger
9	138.3	137.5
10	143.2	143.3
11	148.2	149.2
12	154.0	155.3
13	160.9	160.8
14	168.2	164.7
15	174.4	167.1
16	178.7	168.6
17	181.3	169.3
18	182.6	169.8
19	183.2	170.2
20	183.6	170.5
21	184.0	170.6

Tabel 1. Gennemsnitshøjde for hollandske drenge og piger 1997.

Kilde: (Fredriks, 2000)

Hvis det er rigtigt at data stammer fra (Fredriks, 2000), rejser det spørgsmålet om opgavestilleren bevarer forbindelsen til “the real-world” når data ændres for at få pæne resultater. Personligt mener jeg at det ikke hænger sammen med et erklæret mål om at arbejde i en virkelighedskontekst. En af de færdigheder som matematikundervisningen burde bibringe eleverne, hvis den skal bidrage til at “det enkelte menneske kan fungere som en konstruktiv, engageret og reflekterende borger” (PISA 2003, s. 38-39), må være at kunne se om grafer og tabeller ser rigtige ud, og kunne vurdere om forfattere af undersøgelser og artikler har ændret på data. Det burde være et vigtigt element i matematikundervisningen at eleverne lærer at forholde sig kritisk og reflektivt til kvantitative fremstillinger. For at løse opgaven om vækst skal man nærmest gøre det modsatte.

Naturvidenskabelig kompetence

PISA har defineret grundlaget for vurdering af det naturfaglige område ved hjælp af begrebet “scientific literacy” der i oversættelse er defineret således:

Færdighed i at kunne anvende naturvidenskabelig baseret viden; at kunne genkende naturvidenskabelige spørgsmål og kunne drage slutninger på grundlag af naturvidenskabelige kendsgerninger i bestræbelsen på at forstå og være med til at træffe afgørelser om den naturgivne omverden og de påvirkninger af den, som menneskers aktiviteter medfører. (PISA 2003, s. 155)

PISA opstiller et temmelig ambitiøst program for elevernes naturvidenskabelige kompetence der bl.a. inkluderer følgende om tolkning af naturvidenskabelige kendsgerninger og konklusioner

Dette betyder at kunne finde mening i naturvidenskabelige resultater som argumenter for påstande og konklusioner. Det kan omfatte vurdering af naturvidenskabelig information og at kunne formulere og formidle konklusioner baseret på naturvidenskabelig argumenter. Det kan også handle om at vælge mellem alternative konklusioner, og om at give argumenter for eller imod en given konklusion ved hjælp af de givne informationer, eller om at identificere de antagelser, der er gjort for at nå en konklusion, og at overveje og formidle mulige samfundsmæssige implikationer af en naturvidenskabelig konklusion. (PISA 2003, s. 158)

Imidlertid er det min opfattelse at mange af PISA-opgaverne er præget af en manglende indsigt i hvad der konstituerer en naturvidenskabelig undersøgelse. Jeg vil eksemplificere dette ved en gennemgang af en opgave fra PISA 2000. For yderligere dokumentation kan henvises til (Braams, 2004) der foretager en kritisk gennemgang af science-opgaverne fra The 2003 PISA Assessment Framework (OECD, 2003). Se også (Henningsen, 2005) der indeholder uddybende kommentarer.

Eksempel 3: Om barselsfeber og jordskælv

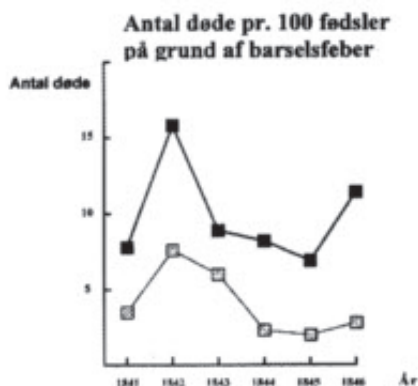
Denne opgave, der handler om fødselslægen Semmelweis, blev brugt i PISA 2000 og er offentliggjort i forbindelse med afrapporteringen af denne (PISA 2000, s. 238). Opgaven citeres og diskuteres i en lang række publikationer om og fra PISA og kan derfor anses for at være en central opgave. Et i opslag i Google (marts 2005) på "Semmelweis + PISA" gav 830 hit).

Det spørgsmål der blev stillet på grundlag af ovenstående tekst, var:

Forestil dig, at du var Semmelweis. Giv en grund, som er baseret på de oplysninger, Semmelweis indsamlede, til, at det er usandsynligt, at barselsfeber kan skyldes jordskælv.

SEMMELWEIS' DAGBOG: TEKST 1

"Juli 1846. I næste uge skal jeg tiltræde stillingen som 'Hr. Doktor' ved Første Fødeafdeling på Wiens kommunehospital. Jeg blev forfærdet over at høre om den procentdel af patienterne, som døde på denne afdeling. I denne måned døde der ikke mindre end 36 ud af 208 mødre, alle af barselsfeber. At føde børn er lige så farligt som en førstegrads lungebetændelse."



Disse linjer fra Ignaz Semmelweis' (1818-1865) dagbog illustrerer barselsfeberens katastrofale virkninger. Barselsfeber er en smitsom sygdom, som mange kvinder, der lige havde født, døde af. Semmelweis indsamlede oplysninger fra hospitalets 1. og 2. afdeling. (Se diagrammet).

Diagram

Lægerne, deriblandt Semmelweis, vidste så godt som intet om årsagen til barselsfeberen. Semmelweis' dagbog fortsætter:

"December 1846. Hvorfor dør så mange kvinder af denne feber efter en fødsel uden nogen som helst komplikationer? I århundreder var videnskaben af den opfattelse, at årsagen til mødrenes død kunne være en usynlig epidemi. Denne kunne skyldes luftforandringer eller påvirkninger stammende ude fra verdensrummet eller kunne hænge sammen med geologiske fænomener, som f.eks. jordskælv."

Nu om dage er der ikke mange mennesker, som ville mene, at påvirkninger stammende ude fra verdensrummet eller jordskælv kan være mulige feberårsager. Men dengang Semmelweis levede, troede mange mennesker, selv videnskabsmænd, at det hang sådan sammen. Imidlertid vidste Semmelweis, at det var lidet sandsynligt, at feberen kunne skyldes påvirkninger stammende ude fra verdensrummet eller jordskælv. Han brugte forskellene i dødelighed mellem de to fødeafdelinger (se diagrammet) til at overbevise sine kolleger.

Figur 3. Semmelweis Dagbog. Tekst 1.

I opgaven vises en graf over antal tilfælde af barselsfeber pr. 100 kvinder på hver af to afdelinger på kommunehospitalet i Wien for årene 1841 til 1846. De to kurver forløber for alle praktiske formål parallelt, men med betydelige årlige udsving (f.eks. er døds-hyppigheden i 1842 omtrent dobbelt så stor som i 1845). En naturlig model vil derfor være at antage at døds-hyppigheden af barselsfeber afhænger af en afdelingseffekt plus en årstalseffekt, således at man kan modellere

antal døde pr 100 kvinder = effekt af afdeling + årstalseffekt + tilfældig variation

Her dækker “årstalseffekten” over forhold der varierer fra det ene år til det andet. For at opnå en dækkende beskrivelse af data er der her brugt en model hvor der er en flerhed af kilder til smitteoverførsel (Semmelweis var snarere ude på at finde smittekilderne end at finde årsagen til barselsfeber).

Semmelweis var ikke helt uvidende om hvordan smitten overførtes. Han havde tværtimod en teori om at de mange tilfælde af barselsfeber skyldtes at kvinderne på første afdeling blev smittet af de lægestuderende der kom lige fra obduktionsstuerne uden at vaske hænder. Når han også indsamlede data fra afdeling 2, var det fordi man der ikke uddannede læger, men jordemødre. Han kunne så sammenligne smittebilledet på de to afdelinger. Dette nævnes imidlertid ikke i opgaven. Man kan i øvrigt anføre at Semmelweis ikke bare prøvede at overbevise sine kolleger ved statistiske beviser for jordskælvs manglende effekt. Han lod dem der assisterede ved fødslerne vaske hænderne i klorkalk og bragte i løbet af to år dødeligheden ned fra 9,02 % til 1,27 % (dog uden at overbevise kollegerne om nytten af håndvask).

Man kan se at data viser den af Semmelweis’ forventede forskel på afdelingerne, men de viser samtidig også en “årstalseffekt”, en variation fra år til år, som data ikke giver mulighed for at analysere nærmere. Det er således klart fra data at der er mere end en smittekilde til barselsfeber, idet incidensen varierer både med tid og afdeling. Men selv om afdelingseffekten var Semmelweis ærinde, så er det ikke opgavestillerens. De tager derimod udgangspunkt i det dagbogscitat hvor der spekuleres over om barselsfeber skyldes en “usynlig epidemi”, eller måske jordskælv, og beder eleverne om at forklare hvorfor data viser at det er meget usandsynligt at barselsfeber kan forklares ved forekomsten af jordskælv. Fuldt pointtal gives for svar der “henviser til forskellen mellem antallet af døde (pr. 100 fødsler) på de to afdelinger” (PISA, 2000, s. 241).

Det angiveligt “rigtige” svar bygger på et argument om at når der er forskel mellem afdelingerne, viser det at der ikke kan være andre (medvirkende) årsager. Argumentet sammenfattes således:

Diagrammet viser tilsvarende variation over tid, men dødsraten er hele tiden højere på den 1. afdeling end på den 2. afdeling. Hvis jordskælv var årsagen, skulle dødsraten på de to afdelinger være den samme. Diagrammet antyder, at noget på de to afdelinger må kunne forklare forskellen. (PISA 2000, s. 240)

Dette er ikke gyldig naturvidenskabelig argumentation. Et gyldigt (men virkelighedsfjernt) argument kunne være: Hvis vi antager at jordskælv er den eneste årsag til barselsfeber, og at jordskælv virker ens på de to afdelinger, så skal dødsraten på de to afdelinger være den samme. Men heller ikke dette argument er tilstrækkeligt. Når dødsraten ikke er den samme på de to afdelinger, betyder det at mindst en af de to

forudsætninger er falsificeret. Men deraf kan man ikke slutte at det er usandsynligt at barselsfeber kan skyldes jordskælv. Det svarer til at påstå at når piger både i Finland og Danmark læser bedre end drenge, så viser det at læsefærdigheder ikke har noget med uddannelsessystemet at gøre.

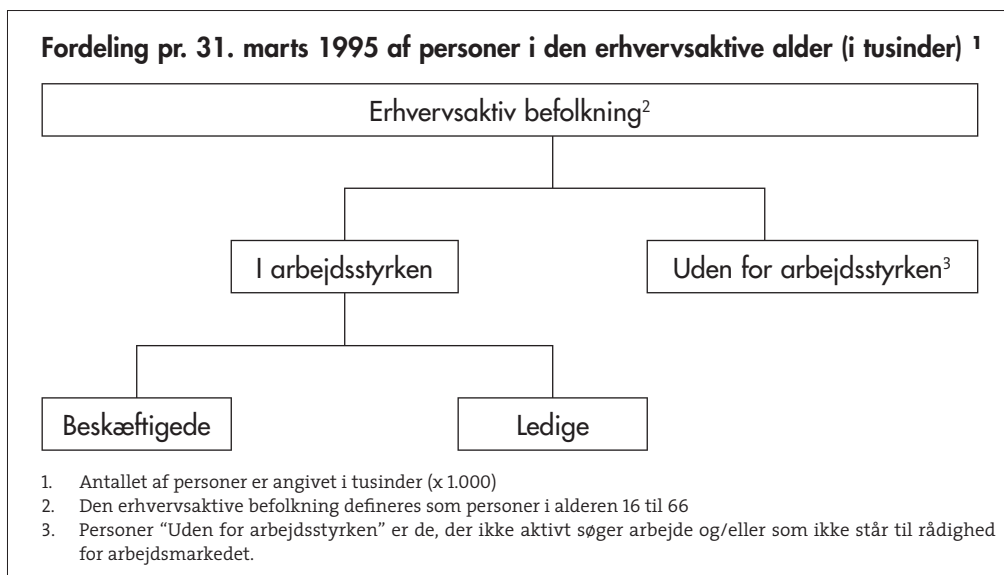
Som to andre eksempler på manglende virkelighedstilknytning kan nævnes en opgave hvor en bil kører 45 kilometer i timen (OECD, 2003). Den bliver overhalet af en bil der kører 60 kilometer i timen. Spørgsmålet lyder hvor hurtigt det ser ud som om bilen kører, når man selv sidder i bilen der kører 45 kilometer i timen. I denne opgave er det pointgivende svar "15 km i timen". Et andet eksempel er en opgave om jordskælv (www.dpu.dk), hvor PISA lader en geolog udtale: "I løbet af de næste 20 år er sandsynligheden to ud af tre for, at der vil komme jordskælv i Zed by." Hvilken geolog udtaler sig om sandsynligheden for jordskælv uden at præcisere styrken på jordskælvne? Flere eksempler på hvordan PISA's opgaver og rettevejledninger indeholder fejlagtige eller mangelfulde konklusioner baseret på naturvidenskabelige argumenter, kan findes hos (Brahms, 2004). Alt i alt efterlader opgaverne et indtryk af at opgavestillerne ikke har nogen erfaring med anvendte problemer.

Eksempel 4: Lost in translation

I opgaven om opvækst blev der givet et par eksempler på hvordan spørgsmål blev ændrede i oversættelsesprocessen. PISA er i princippet opmærksom på at dette kan være problematisk. Derfor går oversættelserne gennem mange led, og der argumenteres for at denne "omstændige oversættelsesprocedure er anvendt, fordi det kan udgøre en fare for validiteten, hvis eleverne ikke svarer på præcis den samme opgavetekst i de enkelte lande" (PISA 2003, s. 27).

De nævnte eksempler viser at opgaveteksterne i forskellige lande ikke altid er "præcis den samme opgavetekst". Et andet eksempel på hvordan oversættelsen betyder noget for sværhedsgraden, er hentet fra bilag 4 (om læsning) i PISA 2000. Opgaven består af et grendiagram hvor en befolkning er inddelt i forskellige beskæftigelsesgrupper. Nedenfor ses en rekonstruktion af den øverste del af diagrammet, samt de tilhørende noter (PISA, 2000, s. 216). Antallene i grendiagrammet er udeladt, da de ikke spiller nogen rolle i de spørgsmål jeg behandler.

Sammenholder man overskriften, betegnelsen for den øverste gruppe i grendiagrammet og note 2, ser man at dansk PISA i grendiagrammet bruger betegnelsen *erhvervsaktive befolkning* for det man i almindelighed (og i overskriften) kalder *befolkningen i den erhvervsaktive alder*. Hermed kommer *den erhvervsaktive befolkning* til at omfatte en stor gruppe der ikke er erhvervsaktive, fordi man inkluderer alle mellem 16 til 66 år, uanset deres tilknytning til arbejdsmarkedet. Dette ordvalg er meget lidt intuitivt, og at der ikke er tale om en pædagogisk pointe fra PISA's side, kan man se i den engel-



Figur 4. Fordeling af personer i den erhvervsaktive alder.

ske original der konsekvent bruger betegnelsen "working age". At ordvalget gør den danske version af opgaven sværere, illustreres f.eks. af følgende spørgsmål:

Hvilke to hovedgrupper er den erhvervsaktive befolkning inddelt i?

- A. Dem, der er i arbejde og de arbejdsledige
- B. Dem, der er i den erhvervsaktive alder og de andre aldersgrupper
- C. Heltids- og deltidsansatte
- D. Dem, der er i arbejdsstyrken, og dem, der er uden for arbejdsstyrken.

Her er det meget lidt intuitivt at D. er det pointgivende svar. Det havde været enklere at vælge D. som svar hvis spørgsmålet havde lydt som i Norge

Hvilke to hovedgrupper er befolkningen i den erhvervsaktive alder inddelt i?

Bemærk i øvrigt at spørgsmålet opererer med begrebet *befolkning i den erhvervsaktive alder*, som slet ikke bliver defineret i den danske tekst. Dette er særligt ironisk i lyset af følgende kommentar i selve rapporten

Opgaven kræver at eleven kan identificere den korrekte kategori for alle de beskrevne tilfælde. Noget af informationen er gemt i fodnoter, og er derfor vanskeligt tilgængeligt. (PISA, 2000, s. 217)

Der kan nævnes andre eksempler som antyder noget om den manglende omhu med oversættelsen. I en problemløsningsopgave optræder ordet “smertestillere” der viser sig at være en lidt for direkte oversættelse af det engelske “painkillers” (PISA 2003, s. 103). I en opgave om CD’er præsenteres vi for begrebet “lyrisk forfatter” der viser sig at være “lyric writer”, altså det der på dansk hedder “tekstforfatter” (PISA 2003, s. 108).

Eksempel 5: Gætteleg

I flere af de opgaver der af PISA præsenteres som lukkede afkrydsningsopgaver, er der ikke oplysninger nok til at afgøre hvad der er det rigtige svar. Som eksempel kan man tage et af de andre spørgsmål i opgaven om “den erhvervsaktive befolkning”. Her skal eleven afgøre om forskellige personer er “I arbejdsstyrken; beskæftigede”, “I arbejdsstyrken; ledige”, “Uden for arbejdsstyrken” eller “Tilhører ikke nogen af de nævnte grupper”. En af dem er en “fuldtidsstuderende på 21 år”. Hvis man ikke ved mere end det, kan man imidlertid ikke klassificere vedkommende, da der i hvert fald i Danmark er ganske mange fuldtidsstuderende der er i arbejdsstyrken. Oplysningerne hører simpelthen hjemme i et andet klassifikationsskema. Man kan heller ikke med sikkerhed klassificere “En mand, 28 år, som for nylig har solgt sin butik og som søger arbejde”. Han kan jo i princippet også være i arbejdsstyrken. Den prøvetrænede elev kan måske godt gætte hvilket svar PISA vil have i de to tilfælde. (“Uden for arbejdsstyrken”). Men hvis eleven tager spørgsmålet alvorligt og lader være med at svare, fordi man ikke *kan* svare på grundlag af de givne oplysninger, så tildeles der ingen point. Man kunne spørge hvorfor eleven ikke bare gætter, men vedkommende kunne jo faktisk være i god tro. Man kunne forestille sig at det var en del af prøven at der var spørgsmål der ikke kunne besvares, og at PISA forventede at man indså dette og viste det ved ikke at besvare spørgsmålet. Det kunne være både legitimt og pædagogisk. Empirien omkring opgaverne synes dog at vise at opgavestillernes intention har været at alle opgaver skulle kunne besvares. Men hvilke dele af elevpopulationen har fået disse oplysninger? Hvis man ved at man under alle omstændigheder skal sætte et kryds, kan man formentlig godt finde ud af hvad PISA havde tænkt i det pågældende spørgsmål. Men et pointgivende svar kræver at man kender “opgavekontrakten”.

En lidt anden version af “gæt hvad jeg tænker på” illustreres af opgaven “Bier” fra Pilottesten i 1999 (www.dpu.dk). Her lød et af spørgsmålene: “Angiv tre af hovedkilderne til nektar”, og følgende citat var tænkt som udgangspunktet for besvarelsen: “Nogle af de vigtigste kilder til nektar er frugttræer, kløver og blomstrende træer.” Af scorevejledningen fremgår at fuldt pointtal opnåedes for svarene

- Frugttræer, kløver og blomstrende træer
- Frugttræer, kløver og blomster
- Kløver, blomster og træer

Strengt taget kan ingen af svarene siges at være *tre* kilder, da kløver jo er en blomst, lige som frugttræer er blomstrende træer, der igen er træer. Tager man imidlertid udgangspunkt i scorevejledningen, så ser det ud som om *frugttræer, kløver, blomstrende træer, blomster* og *træer* i nogle kombinationer kan optræde som pointgivende kilder til nektar. Til gengæld synes det at være ganske arbitrært hvilke kombinationer der giver point. Det er f.eks. vanskeligt at forstå hvorfor “Kløver, blomster, træer” skulle være et bedre svar end “Kløver, blomster, blomstrende træer”. Se tabel 2.

Svar	PISA score
Frugttræer, kløver, blomstrende træer	+
Frugttræer, kløver, blomster	+
Kløver, blomster, træer	+
Frugttræer, kløver, træer	0
Frugttræer, blomstrende træer, blomster	0
Frugttræer, blomster, træer	0
Kløver, blomstrende træer, blomster	0
Kløver, blomstrende træer, træer	0
Kløver, blomster, træer	0
Blomstrende træer, blomster, træer	0

Tabel 2. Kombinationer af frugttræer, kløver, blomstrende træer, blomster og træer og deres pointstatus i PISA-spørgsmål om kilder til nektar.

Alle står lige – men nogle står mere lige end andre

I enhver form for undervisning skabes en opgavekontrakt hvor lærer og elever opbygger en mere eller mindre fælles forståelse af hvordan opgaver skal læses og besvares, og ligesom andre test undersøger PISA en enhed af faglige færdigheder og testfærdighed hvor det er ikke er muligt at skelne bidrag fra henholdsvis det ene og det andet i det endelige resultat. For at internationale sammenligninger som PISA skal være til nogen nytte, burde det derfor være en væsentlig pointe at alle elever så godt som muligt var forberedt på prøveformatet. Her advokerer jeg ikke for at danske elever skal bruge tid på test bare for at klare sig godt i PISA, men for at (også) de danske elever skal kende den række enkle opgavetekniske fif som kan betyde ganske meget for resultatet. Hvis ikke alle elever har samme forhåndsviden om hvordan man rent teknisk skal forholde sig for at score flest muligt point, er det et problem for validiteten af PISA, f.eks. når man sammenligner mellem lande. De følgende tommelfingerregler, der er formuleret på grundlag af PISA's opgaveformuleringer og scorevejledninger, antyder arten af den information som alle elever burde have fået på ensartet måde.

- **Besvar alle opgaver.** Det giver i forventning flere point at gætte end ikke at svare.
- **Giv mange begrundelser.** I opgaver med åbne svar har PISA på forhånd udpeget et antal rigtige begrundelser. Man får point, hvis disse er indeholdt i svaret, men der må også stå andet.
- **Brug tekstens oplysninger. Også når de er forkerte.**
- **Husk at teksterne er konstruerede.** De skal derfor ikke læses på samme måde som man læser autentiske tekster.
- **Husk at der kan være indsat irrelevant information.**
- **Tænk på at teksten er oversat fra engelsk.**

PISA og virkeligheden

Artiklen tager kun fat i et hjørne af den internationale PISA-undersøgelse, nemlig opgaverne, og beskæftiger sig f.eks. ikke med den statistiske analyse af PISA-prøverne. Den beskæftiger sig heller ikke med de mere overordnede uddannelsespolitiske konsekvenser af PISA-undersøgelsen, men leverer nogle præmisser om selve måleinstrumentet som må tages i betragtning når man vurderer hvad man kan slutte på grundlag af PISA 2003. I artiklen har jeg givet eksempler på en række problematiske opgaver som ikke lever op til PISA's høje ambitioner, men også på opgaver som indeholder banale fejl i selve teksten og/eller i oversættelsen. Jeg mener derfor ikke at det er korrekt når der i PISA 2003 siges:

Opgaverne er udvalgt og udviklet af ekspertpaneler i samarbejde med forskere i de enkelte lande. De valgte opgaver er derved det *tætteste man kan komme på hvad eksperter inden for disse områder anser for at kunne måle elevernes kompetence* (min udhævning). (PISA 2003, s. 27)

Udviklingen af PISA er et stort og ambitiøst projekt som næppe har fundet sin endelige form med PISA 2003. Det er derfor overraskende at den danske PISA-gruppe på denne måde prøver at lukke dialogen omkring PISA's opgaver næsten før den er kommet i gang.

Et mere grundlæggende problem er imidlertid virkelighedsdimensionen i PISA. I indledningen rejstes f.eks. spørgsmålet om hvorvidt PISA-opgaverne lever op til intentionerne om at problemstillingerne skal være formuleret så de er tilgængelige for matematisk behandling i en "real-world" sammenhæng. Dette må for mange opgavers vedkommende besvares negativt. Jeg har tidligere peget på at der

på grund af lærernes manglende praksiserfaringer er en tendens til, at anvendelsesaspektet i matematikundervisningen bliver meget formelt. Anvendelserne bliver "en

slags "lege"-problemer, hvor den refleksive praksis, der kendetegner virkelighedens matematiske anvendelser, forsvinder helt ud af billedet. (Henningsen, 2001)

Dette præger også opgaverne i PISA trods intentionerne om det modsatte. I (Henningsen, 2001) argumenteres for at man i højere grad inddrager praktikere i udarbejdelse af opgaver i matematikundervisningen. Selv om det måske kunne hjælpe, vil det være naivt at tro at det løser alle problemer med opgavernes virkelighedsforbindelse. Ud fra en socio-kulturel synsvinkel har matematikdidaktikere undersøgt en lang række matematikopgaver som inddrager "real world" kontekst (se f.eks. Zevenbergen et al., 2002), og vist en række faldgruber. Derfor kan det sidste kritikpunkt rettes generelt mod alle matematikopgaver af denne type, og PISA kan således blive en anledning til at huske på at brug af kontekster udefra i matematikundervisningen langt fra er en ukompliceret sag.

Et sidste spørgsmål som ikke er blevet berørt direkte handler om gyldigheden af PISA som "universelt" projekt. Det hedder i indledningen til matematikafsnittet:

PISA fokuserer på matematik-kompetencer, som kan siges at være relevante for ethvert voksent menneske i et højteknologisk demokratisk samfund. (PISA 2003, s. 38)

I denne sammenhæng kan det siges at være et problem at PISA bliver brugt af en lang række udviklingslande uden for OECD-kredsen, og at flere og flere ser ud til at komme til. Men selv inden for kredsen af OECD-lande, er det så præcis den samme matematikviden der er relevant for alle? Kan man med en og samme test undersøge om matematikviden hos en pige i et portugisisk landdistrikt, en dreng i København eller et barn i slummen i Mexico City svarer til det der kræves for at "det enkelte menneske kan fungerer som en konstruktiv, engageret og reflekterende borger"? Her siger PISA, så vidt jeg kan se, selv nej, fordi der i det samme afsnit om disse "funktionelle matematikfærdigheder og -forståelser som alle mennesker principielt har brug for at have", siges at de "ændrer sig med tid og sted: samfundsudvikling og teknologisk udvikling". Mathematical literacy, den kompetence som PISA har til hensigt at måle, er kulturelt, historisk og samfundsmæssigt relativ. Det er således ikke bare i opgavesammenhæng at virkelighedsdimensionen er problematisk, det er også i den indbyggede reference til virkeligheden.

Læs også kommentarer til denne artikel side 83 i dette nr.

Referencer

- Allerup P., Bredo O. og Weng P. (1998). *Matematik og naturvidenskab i ungdomsuddannelser*. København: DPI 1998.04.
- Allerup, P. (2005). PISA præstationer – målinger med skæve målestokke? *Dansk Pædagogisk Tidsskrift* 53(1), 68-81.
- Andersen, A. M. et al (2001). *Forventninger og færdigheder – danske unge i en international sammenligning*. AKF, DPU & SFI-Survey.
- Braams, B.J. (2004). <http://www.math.nyu.edu/mfdd/braams/links/pisa0207.html>
- Fredriks, A.M. et al. (2000). Continuing Positive Secular Growth Change in the Netherlands 1955-1997. *Pediatric Research* 47(3), 316-24.
- Hansen, E.J. (2005). Pisa – et svag funderet projekt. *Dansk Pædagogisk Tidsskrift* 53(1), 64-67.
- Henningsen, I. (2001). Dobbeltblik. I: Niss, M. (red) *Matematikken og Verden* (s. 147- 69). København: Fremad.
- Henningsen, I. (2005). *Noter om PISA*. <http://www.math.ku.dk/~inge/pisa>.
- Meiding, J. (red) (2004). *PISA 2003 – Danske unge i en international sammenligning*. Danmarks Pædagogiske Universitets Forlag.
- OECD (2003). *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. OECD.
- Thomson, S. et al (2004). *Facing the Future*. Australian Council for Educational Research.
- Zevenbergen, R.; Sullivan, P. og Mousley, J. (2002). Contexts in mathematics education: Help? Hindrance? For whom? I: Valero, P. og Skovsmose, O. (Red.)(2002). *Mathematics Education and Society. Proceedings of the Third International Mathematics Education and Society Conference. Part 2* (pp.575-582). København: Center for forskning i matematiklæring, DPU.