

PISA-succén - och sedan?

Ett diskussionsunderlag kring de internationella kunskaps- och kompetensmätningarna

Datum: 2008-03-06

Version 1.4

Jan Hylén

Innehållsförteckning

1	VARFÖR DENNA STUDIE?.....	3
2	METOD OCH MATERIAL.....	4
3	LÄGET I ANDRA EUROPEISKA LÄNDER	4
4	VAD SÄGER OLIKA AKTÖRER?	5
4.1	HISTORISK BAKGRUND.....	5
4.2	INNEHålLET I MÅTNINGARNA	6
4.3	METODFRÅGOR.....	8
4.4	BEROENDEN PÅ KORT OCH LÅNG SIKT	9
4.5	EKONOMISKA KONSEKVENSER	10
4.6	RISKANALYS	11
5	BÖR NÅGOT GÖRAS?	12
6	BILAGOR	16
1	ÖVERSIKT ÖVER INTERNATIONELLA KUNSKAPSMÅTNINGAR	16
2	LISTA ÖVER INTERVJUPERSONER (I INTERVJUORDNING):	17

1 Varför denna studie?

Metamatrix erhöll hösten 2007 uppdraget från Utbildningsdepartementet att med utgångspunkt i den starkt ökade mängden av internationella kunskaps- och kompetensmätningar analysera de långsiktiga konsekvenserna av Sveriges deltagande i dessa mätningar och lägga perspektiv på framtiden i dessa frågor. Studien har genomförts i samarbete med Utbildningsdepartementets analysfunktion.

OECD:s PISA-studier har under 2000-talet blivit en internationell expertbaserad standard med betydelse för vad som definieras som väsentlig kunskap och kompetens för framtiden. Allt fler länder deltar och resultaten påverkar i hög grad den politiska debatten i länderna. Utöver PISA har Sverige även deltagit och deltar i IEA-studier som PIRLS och TIMSS som under senare år blivit intermittenta. Dessutom planeras och genomförs ytterligare en rad relativt omfattande, komplexa och kostsamma kunskaps- och kompetensmätningar, betingade både av krav på legitima resultatmätt och av den succé som fr.a. PISA-studierna får anses ha blivit.

Framtiden för dessa internationellt jämförande studier torde vara starkt beroende av

- att kvaliteten fortsatt är – och uppfattas vara – hög i testinstrument och genomförande (strikt urval, begränsat bortfall etc.);
- att det finns en nationell beredskap, motivation och resurser i vid mening att hantera en tät agenda av internationella surveys;
- att det som *inte* mäts uppfattas i den nationella politiken vara av sekundär betydelse, för att inte relevansen och betalningsviljan ska minska; och
- att de länder som till slut beslutar sig för att delta i varje studie utgör prioriterade jämförelseobjekt.

Om någon av dessa faktorer fallerar så kommer det att bli svårare att genomföra studierna på det sätt som hittills skett. Rapporten kretsar kring förutsättningarna för dessa faktorer i ett framtida perspektiv av 5-10 år.

Den nationella utvärderingen av skolsystemen och dess resultat har i ökad utsträckning blivit beroende av de internationella jämförelserna. Detta hänger samman med rationell användning av resurser och att de internationella jämförelserna uppfattas vara av hög kvalitet vad gäller test och genomförande. Denna studie hanterar inte frågan om Sverige bör delta i de internationella jämförelserna eller inte. Studien utgår ifrån att det svenska deltagandet i de internationella

mätningarna kommer att kvarstå och vara högt under överskådlig tid. Därför finns det skäl att dels titta på förutsättningarna för att dessa mätningar ska fortsätta fungera och dels diskutera vilka brister eller svagheter mätningarna har och om dessa skulle kunna motverkas. Studien avser som nämnts att utgöra ett diskussionsunderlag och har ett begränsat perspektiv, nämligen att lyfta fram och diskutera de faktorer som kan vara av intresse för att göra bedömningar när det gäller framtidens internationella kunskapsmätningar.

2 Metod och material

Metamatrix har mot den beskrivna bakgrunden gjort en genomgång för att fördjupa kunskapen om konsekvenserna av denna utveckling och möjliggöra medvetna val av berörda myndigheter och departement.

Under hösten 2007 skickades en förfrågan ut från departementet till företrädare för ministerier i andra europeiska länder, för att utröna om något liknande analysunderlag tagits fram. En första förfrågan riktades till ledamöter i EU-kommissionens expertorgan Standing Group on Indicators and Benchmarks. Eftersom svarsfrekvensen var låg sändes samma fråga ut inom Eurydike-nätverket.

Andra underlag till analysen utgörs bl.a. av Skolverksrapporter. Eftersom tillgången till material totalt sett ändå är tämligen begränsad har huvudunderlaget till studien utgjorts av intervjuer. Dessa har genomförts med fr.a. svenska experter och aktörer inom området men också med ett par personer som arbetat länge inom OECD. De frågeställningar som diskuterades var ganska öppna och allmänna och hade i de flesta fall skickats ut till intervjupersonerna före samtalet. I och med att intervjuformen var mycket öppen fanns det goda möjligheter för informanterna att ta upp egna frågor. Sex (fler nu?) av intervjuerna genomfördes per telefon.

3 Läget i andra europeiska länder

Förfrågan till utbildningsministerierna i övriga EU-länder enligt ovan innehöll två frågor:

- Har det gjorts någon analys i bred mening av landets samlade deltagande i internationella kunskapsmätningar?
- Har landet antagit någon form av strategi för att möta utvecklingen med ett växande antal sådana?

Svar på rundskrivelsen inkom från ett 15-tal länder. Svaren är kortfattade. Inte i något av länderna har man utvecklat eller antagit en strategi för sitt deltagande i internationella kunskapsmätningar.

Deltagandet sägs i allmänhet avgöras från fall till fall. Samtidigt instämmer flera länder i behovet av en mer genomtänkt strategi för sitt deltagande. Storbritannien svarar att de valt att delta i samtliga undersökningar av elevers kunskaper, medan deltagandet i andra studier avgörs *ad hoc*. Finland har tillsatt en arbetsgrupp inom ministeriet för att följa arbetet med indikatorer och internationella jämförelser framför allt inom OECD och EU. Spanien överläter motsvarande arbete till ett relativt

fristående forskningsinstitut. Från Skottland framförs att man för närvarande genomför en översikt över sitt deltagande i olika mätningar, vilket möjligen kan komma att resultera i en strategi. Frankrike lyfter fram att stora metodproblem i undersökningarna gjort att resultaten inte alltid upplevts vara jämförbara. De metodologiska aspekterna tillmäts därför stor vikt i besluten om man ska delta eller inte.

4 Vad säger olika aktörer?

Det här avsnittet bygger främst på nio intervjuer som genomfördes i slutet av hösten 2007. De flesta är aktörer /forskare inom Sverige men två av dem har sina erfarenheter huvudsakligen inom OECD-kontexten. Informanterna har förtecknats i bilaga.

4.1 Historisk bakgrund

Enligt Jarl Bengtsson, tidigare chef för OECD:s Centre for Educational Research and Innovation (CERI), var ett av de historiska skälen till att man började genomföra mätningar av detta slag att förflytta fokus från input-sidan i utbildningssystemen, dvs. hur mycket resurser som lades in, till output eller resultat. Inom OECD-länderna ökade intresset att få en indikation på utbildningssystemets kvalitet i förhållande till andra länder och i förhållande till de kostnader man hade. De mätningar som hade startats på initiativ av forskare kom att få en ökad politisk tyngd. Även idag känns det naturligt att uppmärksamma resultat mer än insatta resurser. Pedagogikprofessor Roger Säljö menade att talet om resultat i denna mening inte får leda till en förytligad och förenklad bild av vad kunskap är och hur den kan mätas. Han framhöll också starkt att skolan har i uppgift att utveckla egenskaper och förmågor hos eleverna som kan vara mycket svåra eller kanske omöjliga att mäta och jämföra internationellt. Vi håller på att få en olycklig fokusförskjutning i vad skolan uppfattas vara till för.

De intervjupersoner som själva deltar i mätningens arbetet, som Astrid Pettersson professor i pedagogik, Caroline Liberg, professor i utbildningsvetenskap och KG Karlsson, forskare i fysik och nationell ansvarig för PISA, höll alla med om vikten av att lyfta fram de begränsningar som finns i mätningarna. Men media tar sällan in den typen av nyanseringar och tilläggsinformation i sin rapportering. De vill ha entydiga resultat, raka svar och gärna fokusering på problem och negativa tendenser .

En viktig fråga gäller vad resultaten av de internationella testerna egentligen visar och vilka resultat som är viktigast att följa över tid. I media uppmärksammas vanligen ranglistorna mest och små förskjutningar tillmäts stor betydelse. KG Karlsson menade att det är *inhemska* trender och förskjutningar över tid som man bör uppmärksamma mest, även om stora negativa förändringar i förhållande till andra länder också kan ge viktiga varningssignaler om att åtgärder behöver sättas in.

Varför har det växt fram ett så starkt politiskt intresse för internationella kunskapsmätningar? Daniel Pettersson, som under våren 2008 disputerar på en avhandling om effekterna av internationella kunskapsmätningar på den svenska utbildningspolitiska debatten, menade att det finns en acceptans för dessa mätningar eftersom de skapar en arena för att diskutera skolpolitik. De ger vad som uppfattas som en valid nulägesbeskrivning som kan användas för att motivera den egna skolpolitiska visionen (oavsett politisk färg). De kan användas både av regering och opposition, antingen till att kreditera eller diskreditera den förda politiken.

Pettersson framförde också att undersökningarna tycks påverka hur vi resonerar om kvalitetsmätning inom skolektorn. Skolverkets SIRIS och SALSA, men också deras nationella utvärdering, NU -03, menar han, är uppbyggda av indikatorer på likartat vis som OECD:s indikatorer. Likaså fanns det en matematikdiskussion på 1980-talet och vissa insatser runt lärares mattekunskaper, som påverkades av dåliga svenska resultat i de internationella mätningarna.

4.2 Innehållet i mätningarna

En av de centrala aspekterna i kunskapsmätningar av detta slag är vad det är för kunskaper eller kompetenser som man avser mäta och om dessa uppfattas som viktiga och relevanta. PISA mäter 15-åringars läsförståelse, matematikkunskaper samt kunnande i naturvetenskap. TIMSS mäter matematikkunskaper och naturkunskap i olika åldersgrupper. PIRLS mäter läsförståelse. EU-kommissionen planerar ett test i moderna europeiska språk och inom OECD planerar man för en omfattande mätning av vuxnas kompetens. Denna lista kan kommenteras på flera sätt. För det första kan man utgå från de kunskaps- och kompetensmätningar som görs och studera hur väl de passar in i det svenska skolsystemet – hur väl de kunskaper som testerna representerar svarar mot de mål som slagits fast i svenska kurs- och läroplaner. Skolverkets analyser tyder på en "tämligen god" överensstämmelse.¹ För det andra kan man titta på vad som inte mäts i testerna. Det kan då konstateras att testerna endast omfattar några få ämnesområden. Stora fält, t.ex. inom de samhällsvetenskapliga och praktiskt-estetiska områdena, omfattas inte. Likaså är det främst 15-åringar som testas, i något fall 10-åringar. Andra åldersgrupper saknas i stort. Dessa begränsningar utgör ingen kritik mot de existerande mätningarna i sig men blir till en kritisk faktor om kostnaderna för deltagandet i de internationella projekten tränger undan nationella kunskapsmätningar som täcker fler ämnen och åldrar. (vem säger alltså går det att lägga in något från någon av de intervjuade?)

Hur är det med den framtida relevansen för de kompetenser som undersöks i mätningarna? En kommentar som framfördes av flera informanter, t.ex. Johnny Nilsson och Jan-Erik Gustavsson, var att läsförståelse och grundläggande matematiska kunskaper är centrala förmågor för ett fortsatt kunskapsinhämtande och i så motto viktiga kunskaper att mäta. Andra, t.ex. Roger Säljö, menade att

¹ Skolverket (2006) Rapport 286, sid 40

testerna i huvudsak mäter repetitiva eller reaktiva kunskaper som är kopplade till att vara "duktig på att sitta still i 1-2 timmar och ensam svara på förelagda frågor med av papper, penna och minneskunskaper". Det kan, menade t.ex. Jarl Bengtsson,, ifrågasättas hur relevanta sådana kunskaper är för dagens arbetsliv. Det brukar påpekas att det är svårt att påvisa någon koppling mellan hur länder klarar sig i dessa mätningar och hur deras ekonomiska tillväxt ser ut.. Finland som ofta har toppresultat i mätningarna har trots detta relativt hög arbetslöshet. Å andra sidan kan man t.ex. fråga sig hur den finska ungdomsarbetslösheten skulle se ut om man hade mediokra skolresultat... Det är sannolikt generellt svårt att finna orsakssamband på denna höga nivå.

Flera av informanterna menar att det egentligen vore intressantare att mäta kompetenser som problemlösning – gärna i grupp – analysförmåga, interaktion med andra, förmåga att skapa ny kunskap etc. Skälet till att denna typ av tester sällan förekommer i internationella sammanhang är svårigheten att utforma och genomföra dem på ett likartat och kulturneutralt sätt. Det är också svårt att utforma testerna så att de mäter breda kompetenser utan att bli allmänna intelligenstester. Enligt Jarl Bengtsson skulle Singapores premiärminister i en intervju i Economist ha sagt att man kommer att dra sig ur PISA med hänvisning till att de kunskaper som mäts inte är tillräckligt intressanta för landets framtida kunskapsbaserade ekonomi. Tom Schuller, avgående chef för OECDs Centre for Educational Research and Innovation (CERI), konfirmerar ryktet. Hans egen bedömning är att Singapore kan ha två skäl för detta: För det första är Singapore extremt beroende av att framstå som en toppnation inom kunskapsekonomin. Man skulle förlora sin goda image om man föll ner ett par pinnhål i rankingen – därför har man mer att förlora än att vinna på att delta. För det andra har Singapore en ytterst välutbyggd utbildningsforskning inom landet och behöver därför inte PISA. Men Schuller framhåller också att det bara är topprankade nationer (som Singapore och möjligen Finland) som kan lämna PISA utan att få en förödande inhemsk debatt.

Oscar Öqvist, tidigare ansvarig för Skolverkets nationella utvärderingar, menade att erfarenheterna från de nationella utvärderingarna och de programvisa utvärderingarna av gymnasieskolan visar att testerna påverkar lärarnas syn på sitt ämne och på vad som är viktig kunskap. Samma sak gäller den nationella provbanken. Detta bör innebära att om nationella utvärderingar ersätts med internationella, så kommer på sikt de svenska lärarnas uppfattning om sitt ämne och vad som är relevanta kunskaper att påverkas.

Samtliga informanter menar att internationella kunskapsmätningar inte kan mäta allt av vikt för ett land. Det finns, och bör finnas, nationella särdrag och aspekter som inte kommer med i de gemensamma testerna. En möjlighet finns att ett enskilt land eller en grupp länder själva utformar och betalar för en särskild modul med egna frågor. Möjligheten och intresset från andra länder, t.ex. i Norden, att samarbeta runt sådana tilläggsmoduler kan vara värt att undersöka närmare, menade t.ex. NN. Det kan också diskuteras hur stor vikt som ska läggas vid de nationella särdragen i förhållande till internationella jämförelser. Fördelen med ett nationellt fokus är möjligheten att fokusera en läroplan

som återspeglar nationella värderingar, t.ex. ifråga om demokratisk och social fostran. Fördelen med att delta i internationella mätningar är möjligheten att jämföra det egna systemets prestationer med andra länders och därmed få en indikation på hur välfungerande det är.

En svaghet i mätningarna är att de i sig inte ger någon information om vad som orsakar undersökningsresultaten eller förslag till lösningar i det fall resultaten är mindre goda. För att få ut maximal nytta av det svenska deltagandet i dessa mätningar menade KG Karlsson att man skulle behöva genomföra nationella uppföljningsstudier. Det skulle ge bättre kunskaper om varför det finns kunskapsbrister hos svenska elever inom vissa områden och vad som skulle kunna göras för att åtgärda dem. Men uppföljningsstudier kräver ytterligare resurser vilket tränger undan andra uppgifter och medför att de internationella mätningarna ännu mera kommer att påverka den svenska utbildningspolitiska agendan. Samtidigt måste man komma ihåg att en viktig poäng med deltagande i internationella studier är att många länder delar på utvecklingskostnaden. Det gör att medverkan kan vara kostnadseffektiv jämfört med om man skulle utveckla liknande tester i det egna landet, samt att det kan vara svårt för ett litet land att över huvud taget ha kompetensen att utveckla nya avancerade mätinstrument. Vidare hävdade Karlsson att de internationella studierna kan användas till lokal skolutveckling, om kommunala skolpolitiker och lärare använder sin lokalkännedom om och relaterar sina problem till de nationella och internationella resultaten.

4.3 Metodfrågor

De internationella kunskapsmätningarna är utvecklade av internationellt välrenommerade forskare och experter inom området. Frågebatteriernas utformning och den efterföljande statistiska analysen präglas av hög kvalitet. Man strävar efter att utforma frågor och testsituationer så att de inte påverkas av att de genomförs i många länder med olika språk och kultur. Detta medför, som redan framhållits, att man i vissa av testerna utgår från en minsta gemensamma nämnare mellan de deltagande ländernas kursplaner eller olika "kulturer". PISA utgör ett försök att ange centrala kompetenser för det moderna samhällslivet, oavsett land. Båda dessa metoder kan leda till en utslätning och kan – i princip - minska undersökningens relevans för det enskilda landets del.

Roger Säljö är av bland annat det skälet kritisk till denna typ av studier och menar att strävan efter att göra testsituationen likartad i alla länder är fruktlös. Ländernas divergerande traditioner ifråga om vilken typ av tester eleverna är vana att utsättas för, får betydelse för resultaten. Om deltagande lärare och elever uppfattar sig som nationella representanter i en internationell tävling där de försöker prestera bästa möjliga resultat eller ser deltagandet i testet som en oönskad börda pålagd ovanifrån, spelar stor roll. Astrid Petterson instämmer i denna kritik men menar ändå att det finns ett värde i att delta i mätningarna, givet att man är medveten om deras begränsningar. KG Karlsson höll med om att vissa resultat inte går att jämföra internationellt. Det gäller t.ex. attityderna till naturvetenskap som han menar visar sig vara alltför kulturellt bundna. Däremot menade han att man kan studera

variationer inom landet och över tid. Å andra sidan framhöll Jan-Eric Gustavsson att när man jämför resultat från olika typer av tester, t.ex. traditionella matematiktester respektive mer komplexa och problemlösande uppgifter, så blir resultaten länderna emellan relativt likartade vilket möjligen indikerar att resultaten visar på reella skillnader mellan länderna.

Caroline Liberg, som är svensk delegat i PIRLS, framhåller av egen erfarenhet betydelsen av att delta aktivt i arbetsgrupper och liknande och att medverka i arbetet redan från början. Kommer man in sent i processen är möjligheterna att påverka både innehåll och metoder betydligt mindre. Vidare påpekar hon vikten av att delegaterna är metodmässigt kunniga – inte bara kunniga om ämnesområdet i allmänhet. Hon framhåller också betydelsen av att skapa allianser med andra länder för att kunna få igenom sina krav. Erfarna och kunniga delegater som skapar internationella allianser och redan från starten deltar aktivt i arbetsgrupper, är med andra ord viktiga ingredienser i ett framgångsrecept.

4.4 Beroenden på kort och lång sikt

Vad innebär i olika avseenden det internationella beroende som skapas genom en kraftigt ökad förekomst av policyrelevanta internationella standards i form av kunskaps- och kompetensmätningar? Är det möjligt, och vad skulle det i så fall innebära, att ställa sig vid sidan av dessa mätningar?

Kunskapsmätningarna skapar två typer av beroenden inom landet. Det första är ett politiskt beroende: det finns ett starkt politiskt tryck att delta, särskilt när många andra näraliggande länder deltar. Det är svårt att ställa sig vid sidan, även om det kan finnas kostnadsmässiga eller andra skäl för att ibland avstå. Skolverket bedömer det t.ex. som omöjligt att inte delta i PISA (Rapport 286/2006). Politiskt tryck från oppositionen gentemot regeringen kom också att ändra Sveriges beslut angående deltagande i PIRLS 2006. Flera informanter, bl.a. tidigare statssekreteraren vid Utbildningsdepartementet Johnny Nilsson, framhåller att så länge som många jämförbara länder deltar och så länge undersökningarna åtnjuter en stark legitimitet, så är det politiskt omöjligt att helt ställa sig vid sidan. Däremot finns alltid möjlighet att diskutera vilka projekt man ska delta i.

Den andra typen av beroende uppstår som en följd av själva deltagandet. Flera av informanterna framhåller vikten av god "mottagarkompetens", dvs. att det finns tillräckligt kvalificerad personal i Sverige för att kunna analysera och använda resultaten för den svenska skolans utveckling. Ett led i bevarandet av sådan kompetens framhålls t.ex. vara att Sverige genomför egna undersökningar av liknande slag. I diskussionerna skilde man inte mellan mottagarkompetens och konstruktörskompetens. Det antogs vara samma sak, vilket inte är självklart men mycket väl kan vara fallet. Vissa menade att Skolverket självt behöver delta i detta arbete för att ha egen personal med rätt kompetens. Andra ansåg att jobbet hellre bör utföras av fristående forskare. Men eftersom deltagandet i internationella undersökningar är kostsamt så tvingas Skolverket paradoxalt nog att för närvarande avstå från att genomföra motsvarande nationella undersökningar, vilket i sin tur kan innebära kompetensförluster.

Det starka politiska intresset för kunskapsmätningar har hittills inte lett till motsvarande intresse bland svenska forskare. Nyrekryteringen av unga forskare inom fältet beskrevs av t.ex. KG Karlsson och Jan-Erik Gustavsson som relativt svag och man kan vänta stora pensionsavgångar inom de närmaste 5-10 åren vilket kan komma att leda till en kompetensbrist inom området.

En aspekt av mottagarkompetensen är journalisternas kunnande och deras förmåga att självständigt läsa och analysera resultaten från undersökningarna. Här efterlyste en informant större insatser för att öka presskårens kompetens.

4.5 Ekonomiska konsekvenser

Som redan framhållits så har det politiska trycket att delta de internationella kunskapsmätningarna ökat under senare år. Det finns också en tydlig trend att allt fler av dessa mätningar blir periodiskt återkommande. PISA genomförs nu vart tredje år, PIRLS vart femte och TIMSS vart fjärde år. Till detta kommer att EU planerar att påbörja egna kunskapsmätningar. Trycket att delta, kombinerat med att mätningarna blir allt fler, innebär ett ökat ekonomiskt åtagande. Enligt beräkningar utifrån underlag från Skolverket kommer kostnaderna för svensk del att uppgå till i storleksordningen 100 miljoner kronor för deltagande i olika internationella kunskapsmätningar under perioden 2006 – 2012, exklusive två studier som EU planerar samt ytterligare en planerad OECD-studie.²

Skolverket fick 2004 i uppdrag av regeringen att utreda förutsättningarna för att införa ett mer systematiskt, rullande stickprovsbaserat system för kunskapsutvärdering för samtliga ämnen i grundskolan. Verket uppdrog i sin tur åt Jan-Eric Gustavsson att lämna förslag till hur ett sådant system skulle kunna byggas upp. Förslaget bedömdes ta alltför mycket resurser i anspråk och istället föreslog Skolverket att de internationella kunskapsmätningarna skulle integreras i den nationella strategin för att "helt eller delvis ersätta nationella mätningar inom valda kunskaps- eller ämnesområden".³ Samtidigt föreslog verket att de nationella utvärderingar som genomförts 1992, 1995 och 2003 inte ska upprepas. För de ämnesområden som inte täcks av internationella mätningar föreslogs...

² Detta kan jämföras med anslaget till utbildningsvetenskaplig forskning som Vetenskapsrådet delar ut, som var 128 miljoner kronor under 2006.

³ Skolverket (2006) Rapport 286.

4.6 Riskanalys

Utöver några risker som redan har lyfts fram i diskussionen så finns ytterligare ett par osäkerheter som förtjänar att belysas.

Det kan diskuteras om mätningarnas popularitet innebär att de ökar i antal och frekvens till en sådan nivå att de till slut blir för många och för dyra för länderna att medverka i och därmed faller under sin egen tyngd. Samtliga informanter håller med om att det finns en sådan risk men den bedöms genomgående som begränsad. Jarl Bengtsson framhöll att det pågår en diskussion om vilket internationellt organ som på sikt ska driva PISA. Antalet deltagare har växt kraftigt och för närvarande är majoriteten av länderna inte OECD-medlemmar. Därför har det höjts röster för att flytta över PISA till t.ex. Världsbanken och låta OECD koncentrera sig på andra uppgifter som PIAAC. Om detta skulle ske, vilket i och för sig inga officiella uppgifter tyder på, så skulle sannolikt prestige och det politiska intresset för PISA minska kraftigt.

Tom Schuller konstaterar att diskussionen om att hitta ett annat huvudmannaskap för PISA inte förs aktivt inom OECD, vilket han beklagar. Schuller hyser stor beundran för vad PISA-programmet åstadkommit, men framhåller att dess framgång medför en rad risker för PISA-programmet såväl som för OECD och dess medlemsländer. PISA får hela tiden ökande uppmärksamhet i media. Kostnaderna för programmet är stora och ökande, men det går att ifrågasätta hur mycket nya relevanta data man har att komma med vid varje cykel. Det finns en risk att PISA-ledningen pressas att lova mer än vad som är realistiskt att hålla, vad gäller användbara resultat. Rankinglistorna dominerar debatten, framför allt i media, vilket är olyckligt och riskerar att snedvrider den utbildningspolitiska debatten (se nedan). PISA kan inte undkomma detta, menar Schuller, trots att de ansvariga för PISA säger sig försöka. Diskussionen om rankinglistor riskerar att störa eller påverka utbildningsdebatten i deltagarländerna i fel riktning. Små förändringar i positioner leder till stora diskussioner, medan betydligt viktigare frågor kommer i bakgrunden. För OECD som organisation finns en fara att PISA kommer att överskugga allt annat arbete inom OECDs utbildningsavdelning.

En risk som är förknippad med det starka politiska trycket på att allt fler ska delta i ett växande antal mätningar är att nationella prioriteringar ges mindre vikt, att mätningarna och jämförelserna åstadkommer en internationell nivellering där olika länders skolsystem allt mindre speglar det egna landets särdrag. Oscar Öqvist menade att det finns en risk för "regression mot mitten", dvs. att mätningarna gör ländernas utbildningssystem allt mer lika – själva jämförelsen ökar likheterna. Det blir också en stark koncentration på sifferresultaten från mätningarna. Samtidigt som upprepade mätningar över tid minskar skillnaderna i resultat och gör jämförelserna mindre intressanta, så finns en risk att man överdriver de små skillnader som finns.

Jan-Erik Gustavsson framhöll att han ur ett forskningsperspektiv emellanåt önskar en avmattning i det politiska intresset. Det skulle öka möjligheterna att föra en saklig diskussion om mätningarnas kvaliteter och brister vilket sannolikt skulle öka möjligheterna att förbättra dem. Å andra sidan, vilket hävdades i intervjun med KG Karlsson, om studierna visar ett ökat genomslag för t.ex. elevernas socioekonomiska bakgrund, så är det ett resultat av stort politiskt intresse eftersom det under lång tid varit en skolpolitisk målsättning att minska sådana skillnader.

Det har också förts upp att mängden tester och mätningar orsakar trängsel i årskurs 8-9 och risk för en provtrötthet hos både elever och lärare. Bortfallet är litet i de internationella kunskapsmätningarna och det är avgörande att detta förhållande inte rubbas.

5 Bör något göras?

Som framhölls inledningsvis så utgår vi ifrån att de internationella mätningarna finns kvar under överskådlig tid och att Sverige kommer att delta aktivt i de flesta av dem. Det finns då skäl att diskutera hur framtiden kan se ut, och vilka brister eller svaga sidor mätningarna har. Det bör också vara värt att se hur dessa svaga sidor skulle kunna motverkas eller åtminstone hur en beredskap att möta dessa skulle kunna byggas upp. I följande avsnitt görs ett försök att lyfta några övergripande frågor för vidare diskussion, framför allt frågor kring om något bör göras med anledning av det som hittills sagts.

Konsekvenser

Den succé som PISA och de andra internationella kunskapsmätningarna hittills rönt har :

- bidragit till ett ökat internationellt beroende i utvärderingen av utbildningen ;
- inneburit ett stort genomslag för resultaten av dessa mätningar i den utbildningspolitiska debatten och i bilden av skolan;
- lett till en tyngdpunktsförskjutning från nationella till internationella tester av kunskaper och att en större del av resurserna för nationell utvärdering läggs på internationella mätningar”
- inneburit ett fokus på vissa kunskaper och på vissa åldersgruppers kunskaper;
- bidragit till en intresseförskjutning från input till output i utbildningssystemen.

PISA – barn av sin tid?

PISA och idén om de internationella kunskapsmätningarna hänger inte i luften. De är idémässigt kopplade till en ”epok” med vissa dominerande synsätt i fr.a. de ledande ekonomierna i världen (jmf:

"Idéer som reser", Røvik, 1998⁴). Det är idéer om "resultat" och "resultatstyrning", om att allt viktigt går att mäta, om *accountability*, om indikatorer och nyckeltal som ska säga något väsentligt, om vikten av jämförelser mellan enheter, *best practice* etc. PISA är en del av en sådan "mental epok", en idéströmning som starkt påverkar agendan i många länder samtidigt. Inte så att förhållandena är lika i alla länder men så att grundidéerna, retoriken, begreppen för att definiera var vi befinner oss på kartan och vilka problemen är, är likartade. Mer decentralisering, fler externa tester, mer entreprenörskap, mer av privat finansiering, mer fokus på kunskap, mer inspektion etc. är exempel på vad som kan ses som en "tankebur" utanför vilken det är svårt att lägga fram förslag och synsätt som bygger på en alternativ logik. Dessa idéer genomströmmar på olika sätt alla länder och utgör en form av standarder för just problemidentifiering. Därför blir också lösningarna av samma typ, dvs. standardmässiga retoriskt sett och anslutna till problemidentifieringen. Detta är en form av organisationsteoretiska konstateranden, men de betyder också att det som är "dominerande idéer" i en epok inte är det i en annan. Om ett antal år så kommer idealen troligen att se litet annorlunda ut. Det finns något av "mode" över de dominerande tanketraditionerna, och detta mode växlar över tid. Vår tanke skulle alltså kunna söka sträcka sig förbi dagens självklarheter just med konstaterandet att det inte kommer att förbli så som det ser ut (så som det verkar) idag. Det kan vara nyttigt – om ock svårt – att tänka sig alternativa framtider som en första utgångspunkt.

Hur länge finns PISA kvar?

En kanske spekulativ men inte oviktig fråga är om intresset för fortsatta internationella jämförelser av detta slag kommer att fortsätta vara starkt. Man kan söka svaret på denna fråga både på utbudssidan – dvs. utbudet av undersökningar av detta slag – och på efterfrågesidan. För att börja med utbudssidan så är det som Jarl Bengtsson kallar "den internationella mätindustrin" en stark aktör. Det är forskare och institut med internationella kunskapsmätningar som specialitet. Dessa forskare och institut kommer sannolikt att fortsätta att erbjuda denna typ av studier och agera för att få länder att betala för att få delta. Den risk som redan diskuterats, att dessa projekt skulle duka under av sin egen popularitet och tyngd, är ju i första hand en organisatorisk fråga – vissa aktörer kan komma att få större eller mindre inflytande. Det innebär inte nödvändigtvis att denna typ av studier försvinner. Om t.ex. PISA blir ett för stort och ohanterligt program kan man tänka sig att olika grupper länder, t.ex. nordamerikanska och europeiska, fortsätter att samarbeta samtidigt som andra grupperingar också uppstår.⁵ En uppsplittring i mindre kluster av länder skulle öka utvecklingskostnaderna för varje land, men sannolikt också öka relevansen i mätningarna, vilket möjligen väger upp kostnadsökningen. Tom Schuller bedömer dock att sannolikheten för en uppdelning i kluster är liten.

⁴ Røvik, K-A, 1998, *Moderne organisasjoner. Trender i organisasjonstenkingen ved tusenårsskiftet*. Fagbokforlaget.

⁵ Forskarna bakom rapporten *Northern Lights on PISA* pekar t.ex. på att samtidigt som Finland och Sydkorea har de bästa resultaten sammantaget så är det svårt att hitta någon förklarande faktor som gäller båda länderna. Vitt skilda vägar tycks kunna leda till liknande (PISA)resultat. Det blir därmed svårt att få ett lärande i jämförelserna, i form av överföring av lösningar.

Ser man till efterfrågesidan så utgörs den – eller företräds den – av regeringar och nationella myndigheter. Olika skäl kan göra att något land – eller någon delstat/provins – skulle vilja avstå från deltagande vilket skulle kunna få en dominoeffekt. Egentligen kan man i nuläget bara spekulera i detta. Men kan studierna komma att "gå ur modet" mer på bredden? I den mån de internationella jämförelserna är en del av en större förvaltningspolitisk trend som förordar tydligare ansvar, uppföljning och transparens (som brukar benämnas "accountability"), så har andra liknande trender kommit och gått. Men förvaltningspolitik rör stora och ofta trögrörliga system och trendcyklerna är jämförelsevis långa. Vidare har Sverige och sannolikt även andra länder, börjat anpassa sina nationella system till att kunna finansiera, delta i och ta emot resultaten av PISA och andra liknande mätningar. Därmed skapas också en organisatorisk tröghet som också talar för att dessa mätningar kommer att efterfrågas??. Sammantaget talar argument för att vi under överskådlig tid kommer att ha kvar internationella kunskapsmätningar som en viktig del av det svenska utvärderingssystemet. Däremot är det inte givet med vilken periodicitet och med vilken täckning av nationell curriculum.

Resursmässigt torde PISA vara särskilt krävande genom att det är den studie som genomförs mest tätt, vart tredje år. Det finns skäl att fråga om denna periodicitet är nödvändig framåt. Förändringarna från en studie till nästa är små och fyller kanske i första hand ett uppmärksamhetsbehov. Studierna tar kraft från annat och inte minst verkar det svårt att både hinna med själv testcykeln och tillräckligt med analyser av allt det material som PISA alstrar, internationellt och nationellt.

Avslutande frågor

Underlaget i denna rapport ger anledning till ett antal avslutande frågor. En del är av den karaktären att de bör bli föremål för utredning /beredning eller redan är det. Andra är mer öppna och tjänar mer som underlag för diskussion.

- Hur följer man upp kunskapsutvecklingen i de *ämnesområden och delområden* inom ämnen som *inte omfattas* av de internationella kunskapsmätningarna? Skolverket nämner i sin rapport (286/2006) att det bör utredas närmare vilka olika typer av kompletteringar som behövs, men detta arbete är enligt uppgift ännu inte påbörjat.
- Hur ser möjligheterna ut att bl.a. genom allianser med andra länder *skapa tilläggsmoduler till internationella mätningar*? Är detta ett realistiskt alternativ till nationella studier, med avseende på kostnader, genomförbarhet m.m.? Om det är realistisk, hur läggs en strategi för att genomföra detta fast?
- Hur är det med *behovet av uppföljningsstudier* och kostnaderna för dessa? Vilken typ av studier är mest lämpliga och hur många behövs för att optimalt ta tillvara de investeringar som gjorts i deltagandet i PISA och PIRLS?
- Hur kan man hantera en hypotetisk situation där utredningar visar att Sverige av ekonomiska skäl genomför färre uppföljningsstudier än önskvärt och att man antingen borde allokera mer

resurser till uppföljningsstudier eller delta i färre internationella mätningar? Är något av alternativen realistiska?

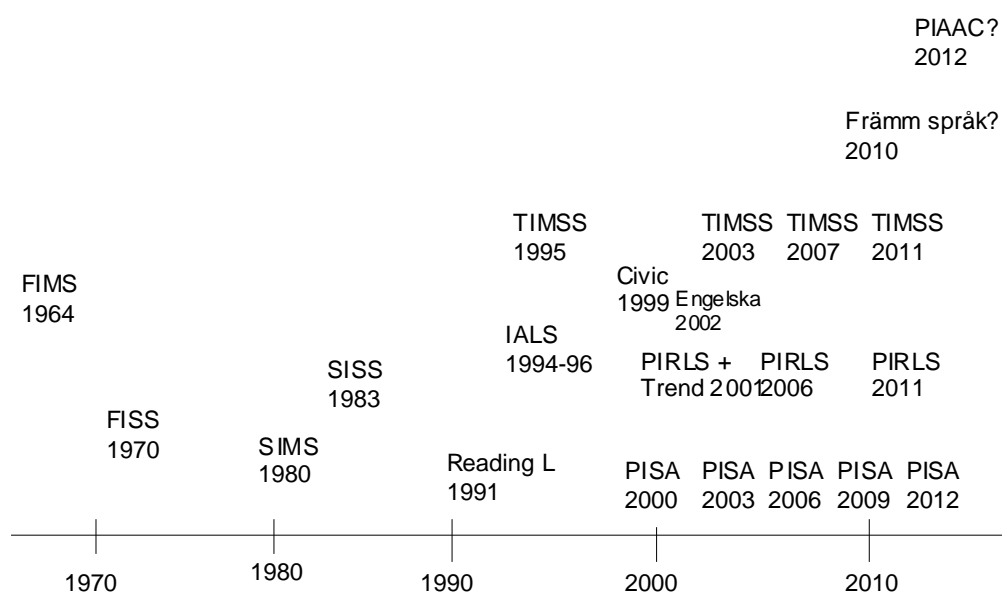
- Är det möjligt och önskvärt att utöka den svenska forskningen på basis av resultat och deltagande i de internationella kunskapsmätningarna? Behöver forskningsmiljöer för detta byggas upp el stärkas? Hur skapa intresse för sådana uppdrag bland forskare, och vem kan vara beställare av detta?
- Är det möjligt och önskvärt att utveckla tydligare riktlinjer för hur de svenska representanterna/delegaterna ska agera inom ramen för respektive internationellt projekt?
- Är det möjligt och önskvärt att samla upp resultaten av dessa överväganden och undersökningar i en övergripande strategi för det svenska deltagandet i internationella kunskapsmätningar? En sådan strategi kunde t.ex. visa hur en övergripande fördelning av resurser mellan nationella och internationella utvärderingar borde se ut. Den skulle också kunna ge riktlinjer för beredningsprocessen mellan Utbildningsdepartementet och Skolverket.

Utan att avge ett säkert svar på den sista frågan så ser det ut att finnas en del skäl som talar för att det finns behov av att utveckla en sammanhållen nationell strategi för det långsiktiga deltagandet i de internationella kunskaps- och kompetensmätningarna. Det synes inte som om något annat EU-land tagit fram en sådan, men många representanter framhåller att man med intresse skulle ta del av en sådan strategi om t.ex. Sverige gick i bräschen för ett sådant arbete. Många av de svenska aktörer vi intervjuat ger också uttryck för behovet av en långsiktig och övergripande strategi. Förhoppningsvis kan detta diskussionsunderlag tjäna som en inledning på et sådant arbete.

6 Bilagor

1 Översikt över internationella kunskapsmätningar

Översikt över internationella kunskaps- och kompetensmätningar 1964-2013



Beteckningar:

FIMS: First International Mathematics and Science Study

FISS: First International Science Study

SIMS: Second International Mathematics Study

SISS: Second International Science Study

Reading L: IEA:s läskunnighetsundersökning 1990/91

IALS: International Adult Literacy Survey

TIMSS 95: Third International Mathematics and Science Study

Civic 99: IEA-studie om elevers kunskaper och uppfattningar i demokrati- och samhällsfrågor

PISA: Programme for International Student Assessment

PIRLS: Progress in International Reading Literacy Study

Trend: Parallell till PIRLS; en jmf i nio länder med resultaten från 1991

Engelska: The assessments of pupils' skills in eight European countries

TIMSS (efter 1995): Trends in International Mathematics and Science Study

PIAAC: OECD-studie om vuxnas kompetens, under planering. Inga beslut om deltagande är fattade

2 Lista över intervjupersoner (i intervjuordning):

- Jarl Bengtsson, tidigare chef för OECDs Centre for Educational Research and Innovation (CERI)
- Oscar Öqvist, tidigare ansvarig för Skolverkets nationella utvärderingsprogram
- Roger Säljö, professor i pedagogik vid Göteborgs universitet
- Jan-Erik Gustavsson, professor i pedagogik vid Göteborgs universitet
- Johnny Nilsson, tidigare statssekreterare vid Utbildningsdepartementet, numera socialdemokraternas riksdagskansli
- Daniel Petterson, doktorand i pedagogik som inom kort lägger fram en avhandling om effekterna av internationella kunskapsmätningar på den svenska utbildningspolitiska debatten.
- Andreas Mörck, Lärarnas Riksförbund
- Karl-Gunnar Karlsson, forskare vid Mittuniversitetet och nationell projektledare för PISA
- Tom Schuller, chef för OECDs Centre for Educational Research and Innovation (CERI) till mitten av februari 2008.
- Caroline Liberg, professor i utbildningsvetenskap, Umeå universitet, svensk delegat i PIRLS
- Astrid Pettersson, professor i matematik vid Stockholms universitet