

# Att undersöka kunskapstrender med hjälp av PISA

Likvärdighet, förståelse och kunskapssyn

*Anders Jakobsson*

The purpose of this article is to highlight different ways of using data from large-scale-studies to interpret trends in the Swedish results in science during the last decade. A majority of the PISA- and TIMSS-studies during this period indicates a clear descending trend in the Swedish national mean values and in some studies the Swedish results display the most significant downward results of all countries. This article discusses how this development may be explained and to what extent the national mean values may be understood as trustworthy. In alternative studies some researchers have used the data to conduct secondary analyzes and show that the trend only are valid in some performance groups and that the reduced results in the group of low- and mid-performers explains the main part of the descending national trend. In addition, these studies indicate that increased school segregation and the development of decreased equity between schools may explain the trend in science during the last decade.

Keywords: Large-scale studies, science education, school segregation, equity.

## Bakgrund

Huvudsyftet med denna artikel är att bidra till att fördjupa förståelsen av hur man kan använda data från storskaliga studier för att undersöka trender när det gäller svenska elevers kunskaper i och om naturvetenskap under 2000-talet. De internationella resultatbeskrivningarna från PISA och TIMSS-studierna indikerar en relativt kraftig och generell nedgång av svenska elevers kunskaper inom ämnesområdet under de senaste decenniet (se t.ex. OECD 2004,

---

*Anders Jakobsson* är professor i naturvetenskapernas didaktik vid fakulteten för Lärande och Samhälle vid Malmö Högskola. Malmö högskola, Lärande och samhälle, 20506 Malmö. E-post: anders.jakobsson@mah.se

OECD 2007, OECD 2010, Skolverket 2010). I några av studierna och inom vissa ämnesområden visar resultaten till och med att de svenska eleverna uppvisar den mest omfattande nedgången av samtliga deltagande länder (TIMSS 2007, Skolverket 2008). I relation till denna nedåtgående trend blir därför viktigt att diskutera hur denna utveckling kan förstås och eventuellt förklaras men också undersöka i vilken utsträckning resultaten kan räknas som tillförlitliga med tanke på studiernas validitet och reliabilitet. Det är i detta sammanhang angeläget att komma ihåg att de storskaliga studierna framförallt syftar till att jämföra skolsystem och att beskriva nationella medelvärden men att dessa vanligtvis inte kan säga något om de kausala sambanden bakom resultaten. Tanken med denna artikel är därför att diskutera de svenska elevernas resultat utifrån från några forskningsstudier som använder alternativa eller komplementära analysmetoder för att re-analysa PISA-materialet under perioden 2000-2009 (t.ex. Davidsson, Karlsson & Oskarsson 2013, Jakobsson et al 2013, Serder & Jakobsson 2013). Dessa analyser visar att grunddata från de storskaliga studierna kan användas för att relatera kunskapsstrender till faktorer som likvärdighet, kunskapssyn och elevers förståelse av ämnesområdet. Syftet med att använda dessa studier är framförallt att analysera hur man kan förstå de nedåtgående trender som finns beskrivna under 2000-talet men också diskutera hur denna typ av data kan användas på alternativa eller komplementära sätt för att undersöka andra fenomen än vad som normalt avses. Ett annat, men relaterat syfte är att undersöka och beskriva hur forskarsamhället uppfattar och förhåller sig till de storskaliga studierna och då speciellt till PISA-studierna. I internationella sammanhang är det tydligt att det existerar en omfattande granskning när det gäller studiernas validitet, reliabilitet och syfte och att den framförda kritiken formuleras utifrån olika utbildningsvetenskapliga perspektiv och paradig. Detta ger i sin tur uttryck för vilka anspråk man utgår ifrån när det gäller att representera och beskriva individers kunskaper och kompetenser i naturvetenskap.

## The Programme for International Student Assessment

The Programme for International Student Assessment (PISA) skapades 1997 gemensamt av OECD-ländernas regeringar med syftet att övervaka utbildningssystem i termer av elevers prestationer inom ett internationell överenskommet ramverk. Det explicita syftet med kunskapsmätningarna har framförallt varit att utvärdera svagheter

och styrkor med olika skolsystem och "allow national policy makers to compare their education system with those of other countries" (OECD 1999, s 7). Kortfattat kan man beskriva testerna utifrån att de är individuella "paper-and-pencil" tester för 15-åriga elever bestående av en blandning av flervalsfrågor och uppgifter som kräver egenkonstruerade lösningsförslag. Omkring 4500–10 000 slumpvis utvalda elever från varje land deltar och resultaten beskrivs som ett nationellt medelvärde som sedan jämförs med alla deltagande OECD-länder. Antalet nationer som deltar i undersökningarna har under senare år ökat från 43 länder år 2000 till 75st år 2009 och olika mätningar under denna period har haft olika ämnesområden såsom matematik, naturvetenskap och läsning som huvudfokus.

För att förstå hur de nationella medelvärdena beräknas är det viktigt att notera att studiernas statistiska metodologi är uppbyggd utifrån en s.k. roterad design inspirerad av Raschmodellen (1960). Detta innebär konkret att varje enskild elev deltar i maximalt två delområden (t.ex. läsning och matematik) och svarar på omkring 30 % av samtliga frågeställningar med olika svårighetsgrad (se teknisk rapport, OECD 2006; 2012). Utifrån lösningsfrekvensen på dessa och korrelationen mellan delområdena beräknas sedan ett konstruerat individuellt resultat som om eleven i fråga faktiskt hade besvarat samtliga frågeställningar inom alla tre delområden (s.k. plausibla värden). Genom denna procedur skapas ett sannolikt resultat för varje elev som sedan används för att beräkna det nationella medelvärdet. Eftersom OECDs medelvärde för samtliga deltagande länder vid varje mätning skall vara exakt 500 poäng korrigeras de enskilda nationernas värde i relation till detta utifrån en statistisk formel.

I PISA-studien 2006 var naturvetenskap huvudämnesområde och tre naturvetenskapliga kompetenser relaterade till OECD:s definition av naturvetenskaplig allmänbildning var innehållet som skulle mätas. Dessa beskrivs som "students' ability to identify scientific issues, scientifically explain phenomena, and use scientific evidence" (OECD 2006). Enligt samma ramverk är målet framförallt inriktat på att utvärdera elevernas förmåga att använda sina kunskaper på ett sätt som motsvarar behoven och kraven från det omgivande samhället och utvärdera elevernas förståelse av naturvetenskapliga begrepp, processer och kontexter.

## Storskaliga studiers påverkan på skolpolitiken i ett globalt perspektiv

Det finns flera faktorer som pekar på att de storskaliga studiernas påverkan på skolsystemen över hela världen gradvis har ökat under det senaste decenniet (se t.ex. Biesta 2010). Flera forskare menar att resultatbeskrivningarna i form av nationella medelvärden och nationsrankingar används alltmer av skolmyndigheter och media för att övervaka utbildning och beskrivs som indikatorer för institutionell effektivitet, argument för kvalitetssäkring och som instrument med vilka lärare och skoladministratörer kan hållas ansvariga (t.ex. Jakobsson, Mäkitalo & Säljö 2009, Brunner et al, 2007). I en intervjustudie (Grek 2009) av europeiska skolpolitiker framkommer att PISA-studierna har blivit en indirekt men alltmer inflytelserik påverkansfaktor för en ny skolpolitisk strategi när det gäller att styra utbildningssystem. Det fungerar genom att politiker på olika nivåer utnyttjar resultatbeskrivningar från OECD för att legitimera sina egna förslag om förändringar utifrån argumentet att testerna utgör robusta och entydiga bevis. I Sotiria Greks studie beskrivs exempelvis vilken betydelse PISAs nationsrankingar från 2000 och 2003 har haft för dagens tyska utbildningssystem. Resultaten indikerade en relativt kraftig nedgång i samtliga delområden vilket innebar att landet hamnade på 20:e plats av de 32 deltagande länderna år 2000. De svaga resultaten skapade en PISA-chock (Grek 2009) som dominerade tysk media och skolpolitik under flera år och så småningom resulterade i ett antal nya nationella skolprojekt, fler tester och prov samt förändringar av skolsystemen i olika delstater. Joachim Wuttke (2006) menar att alla dessa förändringar genomfördes trots att det framfördes en relativt samstämmig kritik från forskarsamhället när det gällde PISAs ramverk och slutsatsernas validitet.

Användningen av de svenska resultaten utgör ett liknande exempel. Som tidigare nämnts har Sverige upplevt en nedåtgående trend i de nationella medelvärdena i naturvetenskap enligt PISA och TIMSS-mätningarna under perioden 2000-2009. Detta har i sin tur inneburit att resultatbeskrivningarna har används som politiska argument (t.ex. SOU 2008:109) för att genomföra förändringar av lärarutbildningen, revideringar av skollagen och ämneskursplanerna trots att det är högst tveksamt om man kan dra denna typ av långtgående slutsatser utifrån testresultaten (se t.ex. Goldstein 2004, Allerup 2007). Ytterligare ett exempel utgörs av Norge där de svaga resultaten från PISA-mätningarna i början på 2000-talet resulterade i "war-like headings in most national newspapers" (Sjøberg 2007) och en situation där media drar egna slutsatser om skolsystemets effektivitet och kvaliteten i lärarnas arbete.

Enligt Sotiria Grek (2009) utgör dessa exempel ett generellt mönster där det utbildningsvetenskapliga samhället inledningsvis kritiserar studierna utifrån deras bristande validitet och reliabilitet men att resultaten ändå successivt accepteras och slutligen resulterar i förändringar av skolsystemen. Michael Uljens (2007) menar att denna utveckling utgör exempel på en global marknadsföring av en neo-liberal skolpolicy som styrs och kontrolleras av OECD och en världsomspännande ”Educational-Assessment-Movement”. Enligt författaren är tanken att stödja en tävlingsinriktad mentalitet mellan länderna samt verka för en utveckling av tvärnationella måldokument och homogena utbildningssystem som kan tänkas gynna den ekonomiska marknaden. Han argumenterar vidare att denna utveckling stöds genom att göra kunskap och lärande till ”marknadsvaror” och utbildningssystemen till handelsplats för de studerande.

## Storskaliga studier och kritiken från forskningssamhället

I det internationella forskningssamhället pågår sedan flera år en diskussion avseende det vetenskapliga värdet av de storskaliga studierna och flera forskare uttrycker kritiska synpunkter om validitet och reliabilitet i testerna och då speciellt i relation till nationsrankingarna. Exempelvis menar Svein Sjøberg (2007) och Elisabeth Bautier & Patrick Rayou 2007 att PISA och TIMSS-studierna inte utgör valida representationer av elevernas kunskaper och att det är svårt att dra några slutsatser på nationell nivå. De pekar också på den existerande skillnaden mellan olika länders läroplaner och OECDs och IEAs ramverk samt risken att de storskaliga studierna uppfattas som ”dolda läroplaner” för undervisning i naturvetenskap på en internationell nivå. Andra forskare framhåller att testerna innehåller en kulturell bias och implicita antaganden (t.ex. Olsen 2005) eller diskuterar problem som uppstår när proven skall översättas till nästan 60 språk (Puchhammer 2007). I en studie visar Martin Brunner et al (2007) att elevernas prestationer kan förbättras signifikant genom att låta dem genomföra för-tester och handleda dem inför provtillfället (Coaching for the PISA-test). Dessutom pekar andra (Bottani & Vrignaud 2005) på den inneboende konflikten mellan politiska och vetenskapliga intressen där forskarna strävar mot att urskilja komplexa mönster om människors förståelse och attityder till naturvetenskap samtidigt som det politiska samhället kräver enkla svar och snabba lösningar. Vidare argumenterar de att de storskaliga studierna framförallt fokuserar endimensionella rankingar

mellan länder vilka syftar till att tillfredsställa politiska krav och inte utbildningsvetenskaplig forskning.

Harvey Goldstein (2004) har också genomfört en analys av PISA-testernas reliabilitet och statistiska metodologi och tankarna bakom den internationella rankingen. I sammanfattningen konstaterar han att "the reality of comparing countries is a complex multi-dimensional issue that is well beyond PISA's ineffectual and one-dimensional attempt" (s 328). Han menar vidare att den typ av "cross-sectional data" (tvärsnittsdata) som används för att genomföra nationsrankingerna gör det omöjligt att dra tillförlitliga slutsatser om de deltagande ländernas skolsystem och föreslår istället en utveckling mot mer longitudinella perspektiv. Även Michael Fertig (2004) menar att man endast kan dra slutsatser från nationsrankingerna om man utgår från starka antaganden om att det existerar stora likheter i ländernas utbildningssystem. Han argumenterar också att eftersom skolsystemen vanligtvis skiljer sig i flera avseenden blir det svårt att identifiera vad som utgör den avgörande faktorn bakom resultaten. Peter Allrup (2007) visar dessutom att de statistiska skalorna som används i PISA inte är homogena med tanke på kön och etnicitet vilket leder till en misstolkning av Rasch-modellen som konstituerar utgångspunkten för metodologin bakom nationsrankingarna.

## Förslag om alternativa eller komplementära modeller och analyser

Det politiska inflytandet, de kontroversiella egenskaperna i testerna och resultatbeskrivningarna i form av nationsrankingarna har uppmuntrat forskare över hela världen att genomföra alternativa eller komplementära analyser/studier baserade på PISA-data eller på uppgifter från PISA-tester. Flera av dessa använder tillgänglig data för att genomföra re-analyser. I detta sammanhang är det viktigt att framhålla att all data från PISA-studierna är tillgänglig för forskare och allmänhet vilket innebär att det är möjligt att genomföra mer utförliga analyser (se <http://www.oecd.org/pisa/pisaproducts/>). Ett exempel på en sådan är Peter Lietz (2009) studie som visar att genuskillnaderna som framkommer i lästesterna har en avgörande betydelse för elevernas prestationer i naturvetenskap. Det vill säga att elevernas läsförmåga i hög utsträckning tycks styra vilka resultat eleverna får på provet i naturvetenskap (s.k. reading load). I en annan studie jämför Jari Lavonen et al (2006) de nordiska ländernas läroplaner i de naturvetenskapliga ämnena i relation till PISAs internationella ramverk. Studien pekar på att det existerar relativt stora skillnader mellan ländernas läroplaner

och vad PISA-studierna avser att mäta ur ett innehållsperspektiv. Marit Kjaernsli och Svein Lie (2011) använder också PISA-data för att analysera om det existerar internationella skillnader när det gäller elevernas attityder och förhållningssätt till naturvetenskap och andra (Olsen & Lie 2011) för att undersöka elevernas uppfattningar om en eventuell framtida karriär som naturvetare. Ett annat förslag är att använda PISA-data för att hitta nya mönster som kan ge lärare explicita råd och förslag till förbättringar av undervisningen. Exempelvis argumenterar Gerard Bonnet (2002) att det skulle varit mer informativt att fokusera vilka frekventa fel eleverna gör, vilka specifika ämnesområden som orsakar generella problem och om det existerar skillnader mellan länderna i dessa avseenden.

Även Peter Fensham (2009) pekar på att det är viktigt att genomföra sekundära analyser genom att undersöka materialet utifrån kontextuella uppsättningar av testuppgifter. Analysen avslöjar att genom att beskriva de nationella resultaten i form av ”means of percentage of correctness” kommer han fram till helt andra resultat än vad de nationella rankingarna från OECD rapporterna visar. Studien indikerar att det existerar relativt små skillnader mellan topprankade länder och signifikant lägre rankade länder vilket motsäger resultatbeskrivningarna från OECD. Peter Mortimore (2009) utgår från ett mer övergripande perspektiv och föreslår radikala förändringar av PISAs metodologi. Han menar att testerna och analyserna borde utvecklas mot en mer nyanserad tolkning av ländernas styrkor och svagheter i relation till medborgarnas livslånga lärande och i vilken utsträckning skolsystemen förmår att främja elevernas prestationer, stödja pedagogisk utveckling och en likvärdig utbildning. Mortimers förslag förutsätter en förändring eller utvidgning av metodologin i testerna mot mer longitudinella inslag, analyser av trender, undersökningar av lärarnas uppfattningar och insamling av fördjupande kontext- och bakgrundsdata.

## Vad kan de storskaliga studierna säga om elevernas kunskaper i och om naturvetenskap?

Som tidigare beskrivits existerar det i vetenskapliga sammanhang diametralt motsatta uppfattningar om det vetenskapliga värdet av de storskaliga studierna och vilka slutsatser man kan dra om elevernas kunskaper utifrån testerna. Dessa uppfattningar är naturligtvis formulerade utifrån mer eller mindre explicit uttryckta epistemologiska antaganden om människors kunskaper och i vilka situationer dessa kan representeras och göras synliga. I PISA-sammanhang utgår OECD

från antagandet att det möjligt att ”mäta” elevernas kunskaper i och om naturvetenskap utifrån ett nationellt medelvärde som sedan kan jämföras med andra länder och att dessa produkter kan säga något om utbildningssystemens effektivitet (se t.ex. tekniska rapporter OECD 2006; 2012). Hela tanken bakom de nationella rankingarna utgår i flera led från denna typ av implicit eller explicit uttryckta antaganden om relationen mellan uppgifterna, testsituationen, beräkningen av resultaten, slutsatser om elevernas kunskaper och ländernas utbildningssystem.

Utifrån helt andra antaganden om hur människors kunskaper kan förstås eller representeras uttrycker Thomas Popkewitz (2011) en generell kritik mot nationsrankingarna i PISA. Han menar att beräkningarna utgår från premissen att resultaten och siffrorna (the numbers) uttrycker ”sanningen” om elevernas kunskaper och skolsystemen och att det är möjligt att mäta kunskaper om skolämnen utifrån frågeställningar om vardagliga situationer. Med andra ord inriktar sig Popkewitz kritik framförallt mot proceduren bakom nationsrankingarna som, enligt honom, innebär att man reducerar beskrivningen av elevernas kunskaper till ett endimensionellt tal. Han uttrycker detta genom att påstå:

Numbers standardize and relocate the local and personal in abstract systems of knowledge that at the same time operate in the spaces of personal knowledge. (s 36)

Även Sjøberg (2007) utgår från andra antaganden än OECD när det gäller frågan hur människors kunskaper kan representeras eller förstås utifrån provsituationer. Enligt honom utgör det ett allvarligt validitetsproblem att PISA-organisationen hävdar att testerna mäter ”real-life-situations” med skolämnesinnehåll samtidigt som eleverna svarar på frågeställningarna med hjälp av ”paper-and-pencil” tester. Han hävdar att livet som medborgare i moderna samhällen kräver ett stort antal kompetenser och färdigheter av olika slag vilka inte enkelt låter sig mätas utifrån PISA-uppgifter.

Kritiken i detta fall handlar alltså framförallt om de kunskapsanspråk OECD utgår ifrån när man konstruerar uppgifter och när man drar slutsatser om relationen mellan testresultaten och livet som medborgare i samhället. Dessutom kan man förstå Sjøbergs kritik utifrån en diskussion om hur väl ett individuellt PISA test kan representera elevernas kunskaper i helt andra situationer. I detta avseende är det möjligt att se vissa likheter med en kritik utifrån sociokulturella perspektiv. Tyvärr finns det i dessa sammanhang inte utrymme att mer i detalj utreda hur en sådan kritik skulle kunna utformas och hur



alternativ skulle kunna formuleras. Några möjliga utgångspunkter i detta perspektiv är att människors kunskaper framförallt blir explicita och synliggjorda när de aktivt använder dem under problemlösning ”in action” och i autentiska situationer som betyder något för deltagarna och deras liv. I dessa situationer har de dessutom tillgång till relevanta artefakter och hjälpmedel av olika slag (datorer, böcker, grafräknare etc.), varandras argument och resonemang samt möjligheter att förhandla innebörder och betydelser med andra (se t.ex. Jakobsson, Mäkitalo & Säljö 2009, Schoultz, Säljö & Wyndhamn, 2001). Huvuddragen i en sociokulturell kritik av de storskaliga studierna inriktar sig alltså framförallt mot premissen att människor ”har” kunskaper lagrade som kan aväckas eller avslöjas vid enskilda provtillfällen, att dessa är avskilda från de situationer då kunskaperna normalt används och att eleverna inte har tillgång till de hjälpmedel de brukar använda. Enligt detta perspektiv riskerar därför resultaten bli artificiella produkter som är otillräckliga eller ofullständiga när det gäller att beskriva elevernas kunskaper eller vad elever klarar av i olika situationer.

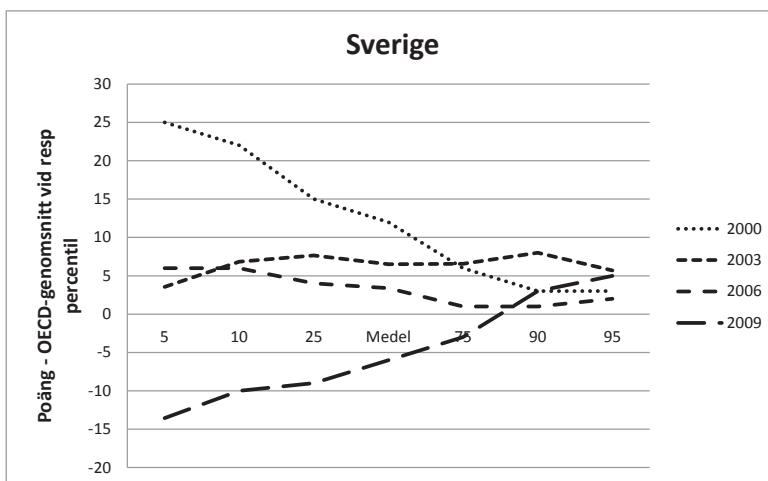
## Att använda storskalig data för att analysera trender i likvärdighet i svensk skola

En möjlighet när det gäller att undersöka likvärdighetstrender i de svenska resultaten består i att analysera elevernas kunskaper i och om naturvetenskap under perioden 2000-2009 utifrån hur de faktiskt svarar på enskilda PISA-uppgifter. En viktig slutsats i relation till den tidigare beskrivna kritiken blir då att inte använda resultat som bearbetats utifrån psykometriska metoder med syftet att åstadkomma nationsrankingar utan att istället använda PISA-data på andra sätt. Dessutom blir målsättningen med ett sådant förfarande inte att uttala sig om elevernas kunskaper på individnivå utan att studera generella mönster som framträder i resultaten. Exempelvis kan man dela upp de svenska elevernas resultat i naturvetenskap under 2000-talet i olika prestationsgrupper (percentiler) för att analysera om och i så fall hur prestationerna i olika grupper förändras under perioden. En viktig fråga blir då om den nedåtgående trenden i naturvetenskap gäller alla elever eller om det endast är vissa grupper som står för nedgången. När man genomför en sådan analys framträder ett intressant mönster som inte blir synligt om man endast använder de nationella medelvärdena eller kvalitativ data.

I figur 1 beskrivs svenska elevers resultat i naturvetenskap från fyra PISA-studier under 2000-talet utifrån olika prestationsnivåer. X-axeln

beskriver olika gruppers resultat i form av percentiler vilket innebär att ett resultat exakt på X-axeln motsvarar OECDs medelvärde för motsvarande percentil. Om resultatet exempelvis för den 25-percentilen (25% lägst presterande) vid en mätning ligger ovanför X-axeln har dessa 25% av eleverna ett högre medelvärde än motsvarande grupper i OECD-länderna. Y-axeln beskriver spridningen i antal poäng jämfört med det internationella medelvärdet. Ett negativt Y-värde visar ett resultat som är lägre än OECDs medelvärde för den gruppen elever.

Kurvan som visar resultaten från PISA-undersökningen år 2000 visar att de svenska eleverna i samtliga prestationsgrupper ligger över OECDs medelvärde för motsvarande grupper och att de lägst presterande eleverna (< 25:e percentilen) i Sverige presterar signifikant högre än motsvarande grupper i andra OECD-länder. Med andra ord är det möjligt att använda denna graf för att visa att den svenska skolan år 2000 lyckades relativt bra när det gäller att erbjuda svenska elever en hög grad av likvärdighet relaterat till resultaten i PISA. Utifrån figuren är det också möjligt att se att gruppen högpresterande elever (>75 percentilen) presterar försumbart högre än motsvarande högpresterande inom OECD. Om man dessutom jämför med de andra kurvorna är det möjligt att utläsa vad de olika prestationsgrupperna får för resultat vid olika mättillfällen under 2000-talet. En sådan jämförelse indikerar att den lägst presterande gruppen och de medelpresterande eleverna successivt har försämrat sina resultat i naturvetenskap under 2000-talet och att de vid den senaste mätningen presterar signifikant under OECDs medelvärde för motsvarande grupper medan den högpresterande gruppen presterar på ungefär samma nivå under hela 2000-talet.



Figur 1: Svenska elevers resultat på PISA – naturvetenskap under 2000-talet uttryckt i prestationsgrupper i form av percentiler (från: Davidsson, Karlsson & Oskarsson 2013)

Utifrån kurvorna i grafen har det alltså blivit möjligt att upptäcka ett distinkt och tydligt mönster när det gäller de svenska elevernas resultat i naturvetenskap på PISA-testerna under 2000-talet. Analysen indikerar att det framförallt är gruppen lågpresterande och i viss mån de medelpresterande elevernas resultat som försämrats relativt kraftigt under perioden 2000-2009. När det gäller den fjärdedel av eleverna som är högst presterande existerar endast marginella förändringar under samma tidsperiod.

I detta sammanhang kan man framhålla att den mediala bilden om svenska elevers allt sämre kunskaper i naturvetenskap endast delvis tycks ha giltighet och att den generella nedgången framförallt kan förklaras utifrån att det är de lågpresterande elevernas resultat som sjunker. Med andra ord blir det utifrån analysen ovan svårt att diskutera svenska elevers prestationer utifrån perspektivet att de utgör en homogen grupp med liknande kunskapsutveckling. Men om man endast använder de nationella medelvärdena och nationsrankingarna för att tolka och förstå elevernas kunskaper riskerar man att erhålla en reducerad bild av en mer komplicerad verklighet. Dessutom riskerar man att missförstå vad resultaten verkligen indikerar vilket i sin tur kan medföra att skolmyndigheter och kommuner fattar beslut på bristfälliga premisser (se exempelvis SOU 2008:109).

En viktig fråga i sammanhanget är emellertid hur resultaten kan förklaras. Varför har just de lågpresterande elevernas resultat försämrats under perioden och varför tycks gapet mellan låg- och högpresterande elever i Sverige öka under 2000-talet? Tyvärr är det inte möjligt att dra några enkla slutsatser om orsakssamband utifrån grafen men andra typer av analyser och studier kan underlätta förståelsen ytterligare. Exempelvis gör Eva Davidsson, Karl-Göran Karlsson och Magnus Oskarsson (2013) i detta tidskriftsnummer en jämförande analys av utvecklingen i olika prestationsgrupper i åtta europeiska länder genom att återanvända internationell PISA-data. Resultaten av deras analys visar att utvecklingen skiljer sig relativt kraftigt åt i olika länder under 2000-talet. Exempelvis minskar skillnaderna mellan låg- och högpresterande elever i Polen vilket tycks resultera i att de nationella resultaten totalt sett förbättras och i Tyskland ökar resultaten inom samtliga prestationsgrupper. Författarna pekar på att det existerar ett samband mellan en hög grad av likvärdighet i landets skolsystem och höga resultat på PISA-testerna och att de svenska resultaten indikerar en kraftig försämring av likvärdigheten under 2000-talet (se vidare Davidsson, Karlsson och Oskarsson, 2013). Även OECDs (2010) internationella rapport stödjer en sådan slutsats och pekar på att en hög grad av likvärdighet är en stark framgångsfaktor för att nå höga resultat.

En annan möjlighet att öka förståelsen av fenomenen är att undersöka resultaten på skolnivå genom att analysera hur *mellanskolevariansen* har förändrats under 2000-talet, det vill säga, att undersöka elevernas prestationer i naturvetenskap utifrån vad elever på olika skolor presterar på skolnivå och analysera hur skillnaden mellan skolor har förändrats under perioden. Figur 2 beskriver förändringen av mellanskolevariansen i de nordiska länderna i naturvetenskap under perioden 2000-2009. Ett positivt tal visar på en ökning av mellanskolevariansen vilket innebär att prestationsskillnaden mellan landets skolor ökar.

Länder	2000	2003	2006	2009	procentuell förändring av mellanskolevariansen
Finland	5,5	3,5	4,7	5,8	+5%
Norge	10,2	8,1	9,9	9,1	-11%
Danmark	19,4	12,7	14,8	15,3	-21%
Island	6,4	3,5	9,3	16,4	+156%
Sverige	8,0	9,9	11,5	16,1	+101%

Figur 2: Mellanskolevariansen utifrån resultaten i PISA-naturvetenskap i de nordiska länderna år 2000-2009 samt den procentuella förändringen under samma period.

Vid en närmare granskning av figur 2 visar det sig att mellanskolevariansen när det gäller resultaten i naturvetenskap i Sverige i princip har fördubblats under 2000-talet. Det vill säga att skillnaden mellan prestationerna uttryckta på skolnivå indikerar en kraftig förändring under perioden. Utifrån figuren är det dessutom möjligt att visa att Sverige tillsammans med Island år 2009 har den största skillnaden i resultaten mellan skolor i de nordiska länderna. När det gäller motsvarande jämförelse av svenska elevers resultat i läsning visar analysen en tredubbling av mellanskolevariansen under samma period (Skolverket 2010). En viktig slutsats blir alltså att om man analyserar likvärdighet i svensk skola utifrån perspektivet hur elever på olika skolor presterar hamnar Sverige (tillsammans med Island) i en nordisk bottennivå och på en genomsnittlig nivå i OECD sammanhang. Resultaten indikerar dessutom att en trolig förklaring till de försämrade nationella resultaten i naturvetenskap tycks kunna relateras till en förstärkt skolsegregation under perioden vilket också får ett generellt stöd i ett flertal andra studier (Andersson & Malmberg 2010, Bunar 2010; Davidsson, Karlsson & Oskarsson 2013; Jakobsson et al, 2013; Öst 2010, Kallstenius, 2010, Trumberg 2011).

## Att använda storskalig data för att analysera epistemologiska trender i resultaten

I en studie av Anders Jakobsson et al (2013) visar författarna att det är möjligt att använda data från PISA-studierna för att genomföra trendanalyser när det gäller elevers kunskapssyn och epistemologiska förståelse av den naturvetenskapliga undervisningen. Epistemologi skall i detta sammanhang förstås som elevernas förståelse av kunskapsinnehållet, hur de förstår syfte och mening med ämnesområdet och vilket innehåll de fokuserar. I analysen handlar det om att undersöka om det existerar trender eller andra förändringar i resultaten med utgångspunkt i hur elevernas prestationer på olika typer av uppgifter varierar över tid. För att kunna genomföra dessa jämförelser mellan olika år redovisas resultaten på samtliga återkommande PISA-uppgifter (s.k. link items) från tre nordiska länder under hela 2000-talet. Dessa uppgifter förekommer i exakt samma utformning vid minst tre testtillfällen. Studien har använt Douglas Roberts' (1982, 2007) Curriculum Emphases som analysinstrument för att kategorisera frågeställningar med avseende på kunskaper och kompetenser. En fördel är att emfaser ger en tydlig indikation och beskrivning vilket innehåll eller mål de olika uppgifterna i PISA-testet avser mäta. Roberts' emfaser konstruerades ursprungligen utifrån läroplaners eller skolbäckers olika intentioner och beskriver implicita eller explicita budskap i naturvetenskapliga texter eller undervisning. Det vill säga vad texterna eller undervisningen egentligen handlar om, dess avsikt och mening. Genom att kategorisera frågeställningarna i PISA utifrån emfaser och sedan analysera elevernas svar på dessa uppgifter skapas förutsättningar att förstå hur elevsvaren förändrats över en viss period. Detta kan i sin tur skapa möjlighet att föra en diskussion om det existerar kunskapstrender men också om elevers förståelse av ämnesområdet har förändrats under perioden. I denna artikel beskrivs elevernas resultat endast utifrån två emfaser nämligen *correct explanations* och *structure of science*. I studien (Jakobsson et al 2013) redovisas elevernas resultat på alla uppgifter som kategoriserats utifrån samtliga sju emfaser.

PISA-uppgifter som har kategoriserats som *Correct explanations* avser framförallt att mäta elevers kännedom om naturvetenskapliga kunskapsprodukter och fakta. I denna typ av uppgifter handlar det ofta om att eleverna skall kunna återge vissa termer och begrepp eller ge korrekta förklaringar med några enstaka ord utan några egentliga krav på att de har förstått innebörden av dem. Uppgifter inom denna kategori behandlar vanligtvis kunskapsprodukter som relaterade till ett skolämneperspektiv och är relativt ofta av flervalstyp. I Tabell 1

beskrivs den faktiska lösningsfrekvensen i tre påföljande PISA-mätningar från fyra exempel som kategoriserats tillhörande denna grupp. Dessutom återges motsvarande resultat från Danmark, Finland och ett medelvärde från samtliga OECD-länder. Om förändringen är signifikant och konsistent genom samtliga mätningar markeras differensen med en asterisk (\*).

Tabell 1: Elevernas prestationer uttryckt som procent korrekt svar på uppgifter i kategorin *correct explanations*

Item	Year	OECD	Den	Swe	Fin	Item	Year	OECD	Den	Swe	Fin
213Q1	2000	40.30	43.20	35.90	53.80	326Q3	2003	56.72	57.27	61.07	73.89
	2003	41.98	47.18	42.59	59.25		2006	58.27	57.79	60.38	78.18
	2006	47.86	50.54	48.51	67.88		2009	60.62	57.83	63.31	77.72
diff		+7.56*	+7.34*	+12.61*	+14.08*	diff		+3.90*	+0.56	-	-
213Q2	2000	75.90	68.60	81.80	90.40	326Q4	2003	22.33	15.57	23.50	34.52
	2003	76.22	72.36	80.37	91.66		2006	23.31	22.01	26.72	37.68
	2006	79.37	75.91	83.99	94.93		2009	25.32	23.89	28.58	37.63
diff		+3.47	+7.31*	-	+4.53*	diff		+2.99*	+8.32*	+5.08*	-

I tabellen (1) beskrivs andelen elever i olika länder som svarar rätt på uppgifter som kategoriserats som *correct explanations*. Resultaten av analysen indikerar generellt en tydligt uppåtgående trend eller ett oförändrat resultat i samtliga presenterade länder. När det gäller de svenska eleverna visar tabellen dessutom att i två av uppgifterna är den uppåtgående trenden signifikant och i de andra i stort oförändrad från första till sista mättillfället. Noterbart i detta sammanhang är att de nationella medelvärdena i naturvetenskap för de svenska eleverna under samma tidsperiod visar en tydlig nedåtgående trend generellt sett. Med andra ord indikerar resultaten att de svenska eleverna presterar högre resultat på uppgifter som endast kräver faktakunskaper i ett skolämnesperspektiv samtidigt som de totala resultaten sjunker. En viktig fråga i detta sammanhang blir då hur denna trend kan förstås och förklaras. Existerar det en förändring av den naturvetenskapliga undervisningen mot en mer faktainriktad undervisning i Sverige och har elevernas förståelse av ämnesområdet successivt förändrats under 2000-talet? Vilken betydelse har de alltmer förekommande nationella proven och storskaliga testerna i detta sammanhang?

En helt annan typ av uppgifter har kategoriserats som *structure of science*. Denna typ av uppgifter fokuserar framförallt naturvetenskap som empiriskt ämnesområde, hur ny kunskap skapas i det vetenskapliga samhället och vad som utgör tillförlitlig kunskap.

Inom denna kategori finns också uppgifter som har för avsikt att mäta elevernas problemlösande förmåga, deras kreativa och kritiska tänkande i relation till ämnesområdet. I Tabell 2 beskrivs resultaten i tre påföljande PISA-mätningar från fyra exempel som kategoriserats tillhörande denna grupp av uppgifter.

Tabell 2: Elevernas prestationer uttryckt som procent korrekt svar på uppgifter i kategorin *structure of science*.

Item	Year	OECD	Den	Swe	Fin	Item	Year	OECD	Den	Swe	Fin
114Q03	2000	57.30	55.20	61.60	61.80	114Q05	2000	24.80	21.40	16.50	31.40
	2003	54.02	57.34	61.59	64.44		2003	22.26	16.57	24.04	30.77
	2006	53.94	55.78	54.83	66.63		2006	18.89	14.75	17.31	32.43
diff		-3.36*	-	-6.77*	+4.83*	diff		-5.91*	-6.65*	-	-
114Q04	2000	39.50	37.70	40.50	49.10	268Q02	2000	40.40	37.70	34.80	50.60
	2003	35.99	33.98	40.13	44.53		2003	37.08	30.52	26.71	51.75
	2006	34.47	34.62	29.91	47.64		2006	36.17	37.16	25.52	47.53
diff		-5.03*	-	-10.59*	-	diff		-4.23*	-	-9.28*	-

I tabell 2 är det möjligt upptäcka några olika trender beroende på vilket land och uppgift som presenteras. Exempelvis visar de svenska eleverna den största nedgången jämfört med de andra länderna och jämfört med OECDs genomsnitt på tre av uppgifterna. Den nedåtgående trenden i de svenska resultaten är dessutom signifikant i tre av uppgifterna och visar oförändrade resultat i den fjärde. Två av uppgifterna indikerar en kraftig nedåtgående trend på omkring tio procentenheter vilket nästan motsvarar en fjärdedel av eleverna. Även OECDs medelvärde indikerar en svag men signifikant nedåtgående trend under tidsperioden på samtliga uppgifter. De danska och finska resultaten visar inte motsvarande trender och tycks ha förbättrat resultaten på denna typ av uppgifter mellan mätningarna 2003 och 2006. Som tidigare nämnts är det inte möjligt att hitta några enkla orsaksamband utifrån denna typ av data men det är tydligt att resultaten på olika typer av uppgifter indikerar olika trender. Sammanfattningsvis kan man säga att analysen pekar på att det skett en förändring av svenska elevers prestationer som innebär att de presterar allt högre resultat på uppgifter som endast kräver att de kan återge naturvetenskapliga kunskapsprodukter och fakta medan uppgifter avsedda att undersöka elevernas problemlösningsförmåga och förståelse av naturvetenskapernas karaktär visar en relativt kraftig försämring under 2000-talet.

## Att använda klassrumsdata för att analysera hur elever förstår och löser PISA-uppgifter

En viktig frågeställning i relation till resultaten i de storskaliga studierna är hur elever förstår uppgifterna, hur de löser dem och om de brukar allvar och gör sitt bästa under denna typ av testsituationer. Inom detta område saknas i hög utsträckning forskningsstudier vilket bland annat kan bero på svårigheter att få tillgång till autentiska provsituationer utan att själva datainsamlingen påverkar resultaten. Om man vill undersöka hur elever resonerar och går tillväga när de löser denna typ av uppgifter kan man istället konstruera en situation där de får arbeta med PISA-uppgifter och dokumentera hur de går tillväga. Det är emellertid viktigt att konstatera att en sådan situation inte kan sägas motsvara själva provsituationen vilket inte heller var syftet i en studie av Margareta Serder och Anders Jakobsson (2013) som är uppbyggd kring en liknande design.

Omkring 70st niondeklassare fick i denna studie arbeta i smågrupper med PISA-uppgifter under ett antal NO-lektioner. Samtalen och elevernas tillvägagångssätt dokumenterades med hjälp av videoinspelningar och ljudupptagningar. Analyser av ett relativt stort antal samtal mellan eleverna pekar på några överraskande resultat. I ungefär hälften av grupperna förekom ett tydligt motstånd (resistance) mot att engagera sig i uppgifterna och flera av eleverna visade ett avståndstagande till innehållet. Från början var det svårt att förstå detta uppenbara motstånd eftersom det inte blev lika tydligt i alla grupper eller i alla uppgifter. Efterhand blev det i analysen alltmer tydligt att elevernas motstånd framförallt handlade om den naturvetenskapliga presentationen och den inbäddade meningen i uppgifterna som tycktes aktivera speciella *diskursiva identiteter* hos eleverna (Brown, Reveles & Kelly 2005). Ett exempel på detta var det uppenbara motstånd som eleverna visade angående de ”imaginära eller fiktiva elever” som agerar i uppgifterna. Dessa fiktiva elever använder sig nästan uteslutande av ett utpräglat akademiskt språkbruk, agerar som små vetenskapsmän och kvinnor och genomför självkonstruerade experiment och vetenskapliga undersökningar på sin fritid. I uppgifterna används exempelvis uttryck som ”din slutsats är förhastad” eller ”du måste försäkra dig om att alla andra faktorer som kan påverka växthuseffekten är konstanta” vilket tycktes väcka starkt motstånd hos flera av eleverna.

I analysen blev det också tydligt att motståndet mot de fiktiva elevernas språkbruk och agerande ofta medförde att eleverna i studien inte kom fram till själva kunskapsinnehållet eller problemlösningen utan fastnade i en förhandling om meningen med uppgiften. I studien (Serder & Jakobsson, 2013) menar författarna att utvärdera



elevernas kunskaper i naturvetenskap endast utifrån denna typ av uppgifter riskerar resultera i en ”mätning” av elevernas förmåga när det gäller att överbrygga kulturer och diskurser, det vill säga i vilken utsträckning eleverna förmår att passera gränslandet mellan en vardaglig och en vetenskaplig kultur eller diskurs. Enligt författarna utgör denna typ av kunskap eller kompetens i sig ett viktigt mål för undervisningen men resultaten kommuniceras vanligtvis inte utifrån en sådan slutsats. När OECD och PISA-organisationen presenterar resultaten görs istället anspråk på att testerna beskriver nivån på elevernas kunskaper och deras naturvetenskapliga allmänbildning. Författarna framhåller dessutom dilemmat att utvärdera elevers kunskaper med hjälp av uppgifter som beskriver konkreta och vardagliga situationer samtidigt som uppgiften formuleras utifrån ett vetenskapligt språkbruk.

I ytterligare en artikel (Serder & Jakobsson, kommande) analyserar författarna diskursiva betydelser av vissa specifika termer som används i uppgiftstexterna i PISA och hur eleverna förhandlar om dessa termers betydelse. Analyserna visar att ett antal av orden har speciella betydelser eller konnotationer beroende på i vilka sammanhang eller i vilka diskurser de används, vilket relativt ofta tycks skapa problem när eleverna försöker förstå vad som förväntas av dem. Exempelvis har ordet ”faktor” ett antal möjliga betydelser som får helt olika innebörder beroende på i vilken diskurs de används (t.ex. matematik, naturvetenskap, vardagliga). I analysen framkommer dessutom att eleverna använder dessa termer utifrån olika premisser i ett och samma samtal vilket ofta skapar förvirring och osäkerhet hur de löser uppgifterna. Betydelsen av orden blir därmed flertydiga och oklara för eleverna vilket också ofta innebär att de inte lyckas lösa uppgifterna eller ens påbörja själva problemlösningsprocessen. Några andra exempel på denna typ av diskursiva termer är ”mönster”, ”konstant” och ”referens” vilka får helt olika betydelser beroende på hur, var och när de används vilket tycks orsaka problem för eleverna. I uppgiftstexterna används däremot orden vanligtvis på ett oproblematiskt sätt utifrån att de endast har en innebörd och att det är möjligt att kommunicera denna enda betydelse (Se vidare Serder & Jakobsson, kommande).

## Diskussion

Det övergripande syftet med denna artikel har varit att diskutera möjligheter och problem när det gäller att återanvända PISA-data för att analysera kunskapstrender i internationella kunskapsmätningar

i naturvetenskap under 2000-talet. Som tidigare nämnts indikerar de svenska resultaten, i form av nationella medelvärden, en relativt kraftig nedgång när det gäller svenska elevers kunskaper under perioden 2000-2009 och vid några tillfällen uppvisar de svenska eleverna den största nedgången av samtliga deltagande länder (se t.ex. TIMSS 2007). Eftersom PISA-mätningarna endast beskriver nationella medelvärden för att kunna jämföra med andra länder saknas vanligtvis orsaksförklaringar eller nationella trendanalyser i rapporteringen från studierna. Flera tongivande forskare uttrycker en relativt skarp kritik mot denna procedur, kring de nationella medelvärdena och nationsrankingarna och ifrågasätter därmed studiernas övergripande validitet och reliabilitet. Utgångspunkten i denna artikel har istället varit att diskutera de svenska elevernas resultat utifrån forskningsstudier som använder alternativa eller komplementära metoder för att analysera PISA-data (t.ex. Davidsson, Karlsson & Oskarsson 2013; Jakobsson et al 2013, Serder & Jakobsson 2013). Dessa studier visar att grunddata från de storskaliga studierna kan användas för att relatera kunskaps-trender till faktorer som likvärdighet, kunskapssyn och elevers förståelse av ämnesinnehållet för att nå en djupare förståelse av trender i materialet. Syftet är med andra ord inte att beskriva enskilda elevers kunskapsnivå eller ens en allmän nationell kunskapsnivå utan mer att analysera underliggande strukturer och mönster i resultaten.

Exempelvis visar Davidsson, Karlsson och Orskarsson (2013) att det är möjligt att använda PISA-data för att påvisa att den nedåtgående trenden i de svenska resultaten i naturvetenskap och läsning kan förklaras utifrån att det framförallt är de lågpresterande och i viss mån de medelpresterande elevernas resultat som försämrats relativt kraftigt under 2000-talet. Analyser av materialet visar dessutom att skillnaden i de svenska resultaten uttryckt på skolnivå (mellanskolevarians) har fördubblats i naturvetenskap och nästan tredubblats i läsning under perioden 2000-2009. Genom att diskutera hur de svenska elevernas prestationer i PISA-naturvetenskap har förändrats över tid utifrån begreppet mellanskolevarians har det alltså blivit möjligt att visa resultaten i ett enskilt ämnesområde tycks vara relaterad till en allmänt förstärkt skolegregation i Sverige. Man skulle kunna uttrycka detta så som att det tycks existera ett samband mellan en alltmer förstärkt skolegregation och en ökad *kunskapssegregation* i naturvetenskap i den svenska grundskolan under perioden. Kunskapssegregation kan då sägas vara ett uttryck för hur elever på vissa skolor ämnesmässigt missgynnas av en generell skolegregation. En viktig fråga i detta sammanhang blir då om den svenska skolan kan sägas nå de mål om likvärdighet som finns uttryckta i skollagen och om den svenska skolan för närvarande kan erbjuda en likvärdig utbildning inom det

naturvetenskapliga ämnesområdet för alla elever. (För en vidare fördjupning och jämförelse se Davidsson, Karlsson & Oskarsson 2013).

Ett annat exempel hur man kan återanvända data är trendanalysen av elevernas prestationer på olika typer av PISA-uppgifter (Jakobsson et al 2013). I studien framkommer att de svenska eleverna generellt har förbättrat sina resultat på faktabaserade uppgifter och uppgifter som mäter enkla ämneskunskaper under perioden 2000-2009. Detta trots att de generella resultaten visar på en tydligt nedåtgående trend. Emellertid har elevernas resultat på uppgifter som avser att mäta problemlösning, kritiskt tänkande och frågeställningar som belyser naturvetenskapens karaktär sjunkit relativt kraftigt under 2000-talet. Är det möjligt att förstå resultaten som om det har skett en förändring av svenska elevers kunskapssyn och epistemologiska förståelse av skolans naturvetenskap? En viktig fråga i dessa sammanhang blir då om och i så fall på vilka sätt undervisningen i naturvetenskap har förändrats i Sverige? Har undervisningen utvecklats mot en mer faktabaserad undervisning på vissa skolor men inte på andra? Vilken roll spelar den mediala uppmärksamheten på de storskaliga studierna och de nationella proven i dessa sammanhang? Här återstår flera viktiga frågeställningar som kräver ytterligare uppmärksamhet och analys i framtiden (se även Lundqvist & Lidar i denna tidskrift).

I några studier (Serder & Jakobsson 2013; Serder & Jakobsson, kommande) visar författarna att relativt stor andel av eleverna i deras studie antar en s.k. *diskursiv identitet* när de löser PISA-uppgifter i smågrupper. Detta tar sig uttryck i ett explicit motstånd och avståndstagande mot de fiktiva elever som agerar i uppgifterna och dessa elevers användning av ett korrekt akademiskt och naturvetenskapligt språkbruk. Detta trots att uppgiften berör en vardaglig kontext. Motståndet medför att ungefär hälften av grupperna hindras eller störs i själva problemlösningsprocessen. I analyser av elevernas samtal framkommer dessutom att relativt många elever upplever termer som "faktor", "mönster", "konstant" och "referens" som problematiska eftersom orden har olika betydelser och konnotationer i olika diskurser och skolämnen.

Sammantaget utgör alla dessa analyser exempel på att det är möjligt att använda PISA-data på alternativa eller komplementära sätt för att upptäcka underliggande mönster och strukturer ur ett skolämneperspektiv som troligtvis inte blivit synliggjorda annars. Samtidigt är det viktigt att lyfta fram att OECD och PISA-organisationen utgår från antagandet att det är möjligt att "mäta" elevernas individuella kunskaper utifrån ett psykometriskt beräknat medelvärde och att dessa produkter kan säga något om elevernas ämneskunskaper och utbildningssystemens effektivitet (se t.ex. tekniska rapporter OECD

2006; 2012). I dessa sammanhang pekar ett relativt stort antal internationella forskningsartiklar på svagheter och tillkortakommanden när det gäller PISA-studiernas tillförlitlighet och de anspråk som OECD gör kring studiernas resultat. Framst är kritiken riktad just mot den psykometriska konstruktionen av de nationella medelvärdena, de nationella rankingarna och mot att studierna används som endimensionella utvärderingar av deltagarländernas skolsystem. Exempelvis framhåller Mortimore (2009) att det hade varit mer konstruktivt och värdefullt om PISA-studierna metodologi förändrades mot att undersöka mer longitudinella perspektiv, att utvärdera medborgarnas livslånga lärande, att analysera trender och likvärdighet, att tolka lärarnas uppfattningar och samla in fördjupande bakgrundsdata. I ett sådant perspektiv blir generella nationsrankingar meningslösa och resultatberskrivningen mer inriktad på att förstå och beskriva framgångsfaktorer och hinder i ländernas utbildningssystem när det gäller att skapa en likvärdig och framgångsrik skola för alla.

Ur ett sociokulturellt perspektiv är det uppenbart att ifrågasätta i vilken utsträckning kunskapsmätningar av detta slag förmår att beskriva elevernas kunskaper och kompetenser. Vid provtillfällena sitter eleverna isolerade från varandra och omvärlden och förväntas lösa uppgifter utan tillgång till de resurser de normalt använder i liknande sammanhang. Denna traditionella provsituation utgår med andra ord ifrån antagandet att det är möjligt att locka fram och rimligt beskriva människors flerdimensionella kunskaper och kompetenser utifrån ett endimensionellt tal (Popkewitz 2011). Man skulle kunna ifrågasätta om eleverna någonsin kommer lösa problem de ställs inför på detta sätt i resten av sina liv; utan tillgång till andra människor, internet, böcker, tidningar, minräknare, smartphones och dylikt. En viktig fråga blir då hur en utvärdering av elevers förmågor och kunskaper i olika typer av autentiska situationer skulle kunna formuleras och beskrivas. Hur skulle de iscensättas, hur skulle uppgifterna se ut, hur skulle produkterna bedömas och beskrivas och på vilka sätt skulle man kunna uttala sig om elevers styrkor och svagheter? Om man menar allvar med kritiken av de storskaliga studierna, dess produkter och användning bör man kanske också fundera på hur alternativ eller komplement till dessa skulle kunna formuleras.

**Kommentar:**

Denna forskning stöds av Vetenskapsrådet i Sverige (Dnr 721-2008-4717).

## Referenser

- Allerup, Peter (2007): Identification of group differences using PISA scales – considering effects of inhomogeneous items. I Stefan Hopmann; Gertrude Brinek & Michael Retzl. red *PISA According to PISA: Does PISA Keep what it Promises?* s 175–202. Wien: LIT.
- Bautier, Elisabeth & Rayou, Patrick (2007): What PISA really evaluates: literacy or students' universes of reference? *Journal of Educational Change* 8 (1), 359–364.
- Biesta, Gert (2010): *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. Boulder: Paradigm Publishers.
- Bonnet, Gerard (2002): Reflections in a critical eye: on the pitfalls of international assessment. *Assessment in Education* 9 (3), 387–399.
- Bottani, Noberto. & Vrignaud, Pierre (2005): *La France et les évaluations internationales. National Report to the HCEE*. Paris: HCEE.
- Brown, Bryan, Reveles, John & Kelly, Gregory (2005): Scientific literacy and discursive identity. A theoretical framework for understanding science learning. *Science Education* 89 (5), 779–802.
- Brunner, Martin; Artelt, Cordula; Strauss, Stefan & Baumert, Jürgen (2007): Coaching for the PISA test. *Learning and Instruction* 17 (2), 111–122.
- Bunar, Nihad (2010): Choosing for quality or inequality: current perspectives on the implementation of school choice policy in Sweden. *Journal of Education Policy* 25 (1), 1-18.
- Davidsson, Eva; Karlsson, Karl-Göran & Oskarsson, Magnus (2013): Trender och likvärdighet – svenska elevers resultat på PISA naturvetenskap i en internationell jämförelse. *Utbildning & Demokrati* (22) 3, 37–52.
- Fensham, Peter (2009): Real world contexts in PISA science: implications for context-based science education. *Journal of Research in Science Teaching* 46 (8), 884–896.
- Fertig, Michael (2004): What can we learn from international student performance studies – Some methodological remarks. *RWI: Discussion Paper*. No. 23. Essen: RWI.
- Goldstein, Harvey (2004): International comparisons of student attainment: some issues arising from the PISA study. *Assessment in Education* 11 (3), 319–330.
- Grek, Sotiria (2009): Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy* 24 (1), 23–37.

- Jakobsson, Anders; Mäkitalo, Åsa & Säljö, Roger (2009): Conceptions of knowledge in research on students understanding of the greenhouse effect: methodological positions and their consequences for representations of knowing, *Science Education* 93 (6), 978–995.
- Jakobsson, Anders; Davidsson, Eva; Karlsson, Karl-Göran. & Oskarsson, Magnus. (2013): Exploring epistemological trends in students' understanding of science from the perspective of large-scale studies. *ISRN Education*. 13 (1), 4-18. (Open access: Article ID 196014).
- Kallstenius, Jenny (2010): *De mångkulturella innerstadsskolorna : Om skolval, segregation och utbildningsstrategier i Stockholm*. Acta Universitatis Stockholmiensis.
- Kjaernsli, Marit & Lie, Svien (2011): Students' preference for science careers: international comparisons based on PISA 2006. *International Journal of Science Education* 33 (1), 121–144.
- Lavonen, Jari; Lie, Svein; Macdonald, Allyson; Oskarsson, Magnus; Reistrup, Claus & Sørensen, Helene (2009): Science education, the science curriculum and PISA. I Thomas Matti (red) *Northern Lights on PISA 2006* s 31-58. *Differences and Similarities in the Nordic Countries*. Copenhagen: Nordic Council of Ministers, TemaNord 2009:547.
- Mortimore, Peter (2009): *Alternative models for analysing and representing countries' performance in PISA*. Paper commissioned by Education International Research Institute. Brussels: 2009.
- Lietz, Peter (2009): A meta-analysis of gender differences in reading achievement at secondary school level. *Studies in Educational Evaluation* 32 (4), 317–344.
- Lundgren, Ulf P (2011): PISA as a political instrument – One history behind the formulating of the PISA programme. I Miguel Pereyra; Hans-Georg Kotthoff & Robert Cowen (red). *PISA under Examination – Changing Knowledge, Changing Tests, and Changing Schools* s 17–30. The Netherlands: Sense Publisher.
- Lundqvist, Eva & Lidar, Malena (2013): Nationella prov i NO och lärares val av undervisningsinnehåll. *Utbildning & Demokrati*. (22) 3, 85-106.
- OECD (1999): *Measuring student knowledge and skills: A new framework for assessment*. Paris: OECD Publications.
- OECD (2003): *PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD Publications.

- OECD (2004): *Learning for tomorrow's world*. First results from PISA 2003. Paris: OECD Publications.
- OECD (2006): *PISA 2006, Technical Report*. Paris: OECD Publications.
- OECD (2007): *Executive summary PISA 2006: Science competencies for Tomorrow's world*. Paris: OECD Publications.
- OECD (2010): *PISA results 2009. What Students Know and Can Do. Student Performance in Reading, Mathematics and Science (Volume I)*. Paris: OECD Publications.
- OECD (2012): *PISA 2009 Technical Report*, PISA, OECD Publishing.
- Olsen, Rolf (2005): An exploration of cluster structure in scientific literacy in PISA: Evidence for a Nordic dimension? *Nordic Studies of Science Education* 1 (1), 81–94.
- Olsen, Rolf & Lie, Svein (2011): Profiles of students' interest in science issues around the world: analysis of data from PISA 2006. *International Journal of Science Education* 33 (1), 121–144.
- Popkewitz, Thomas (2011): PISA – Numbers, standardizing conduct, and the alchemy of school subjects. I Miguel Pereyra; Hans-Georg Kotthoff & Robert Cowen (red). *PISA under examination – Changing Knowledge, Changing Tests, and Changing Schools* s 31-46. The Netherlands: Sense Publisher.
- Puchhammer, Michael (2007): Language-based item analysis — problems in intercultural comparisons. I Stefan Hopmann; Gertrude Brinek & Michael Retzl (red). *PISA According to PISA: Does PISA Keep what it Promises?* s 127–138. Wien: LIT, 2007.
- Rasch, Georg (1960): *Probabilistic Models for some Intelligence and Attainment Tests*. Chicago: University Press.
- Roberts, Douglas (1982): Developing the concepts of “curriculum emphases” in science education. *Science Education* 66 (2), 243–260.
- Roberts, Douglas (2007): Linné scientific literacy symposium opening re-remarks. I Cedric Linder; Leif Östman, & Per-Olof Wickman (red), *Promoting Scientific Literacy: Science Education Research in Transition* s 9–17. Proceedings of the Linnaeus Tercentenary Symposium, Uppsala University. Uppsala: Geotryckeriet.
- Serder, Margareta & Jakobsson, Anders (2013): “Why bother so incredibly much?” – Student perspectives on PISA science assignments. Accepted to be publish in *Cultural Studies in Science Education* 2013.

- Serder, Margareta & Jakobsson, Anders (kommande): Language games – the meaning potential of a science test.
- Schoultz, Jan; Säljö, Roger & Wyndhamn, Jan (2001): Conceptual knowledge in talk and text: What does it take to understand a science question? *Instructional Science* 29 (3), 213-236.
- Sjøberg, Svein (2007): PISA and “real life challenges”: mission impossible. I Stefan, Hopmann; Gertrude Brinek & Michael Retzl. (red). *PISA According to PISA: Does PISA Keep what it Promises?* s 203–224. Wien: LIT, 2007.
- Skolverket (2008): *TIMSS 2007. Svenska grundskoleelevers kunskaper i matematik och naturvetenskap i ett internationellt perspektiv*. Rapport 323, Stockholm: Skolverket.
- Skolverket (2010): *Rustad att möta framtiden? PISA 2009 om 15-åringars läsförståelse och kunskaper i matematik och naturvetenskap*. Internationell rapport 352. Fritzes, Stockholm.
- SOU 2008:109: *En hållbar lärarutbildning*. Statens offentliga utredningar. Ministry of Education in Sweden.
- TIMSS (2007): *International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eight Grades*. IEA, international.
- Trumberg, Anders (2011): *Den delade skolan: segregationsprocesser i det svenska skolsystemet*. Örebro: Örebro Studies in Human Geography 6.
- Wuttke, Joachim: (2006). Fehler, Verzerrungen, Unsicherheiten in der PISA Auswertung. I Thomas Jahnke, & Werner Meyerhöfers (red). *PISA & Co: Kritik eines Programms* s 101–154. Hildesheim/ Berlin: Franzbecker.
- Uljen, Michael (2007): The hidden curriculum of PISA – the promotion of neo-liberal policy by educational assessment. I Stefan Hopmann; Gertrude Brinek & Michael Retzl. (red). *PISA According to PISA: Does PISA Keep what it Promises?* s 295–304. Wien: LIT, 2007.
- Öst, John; Andersson, Eva & Malmberg, Bo (2010): *School choice and increasing performance difference: a counterfactual approach*. Stockholm Research Reports in Demography 2010:11. Stockholm University.