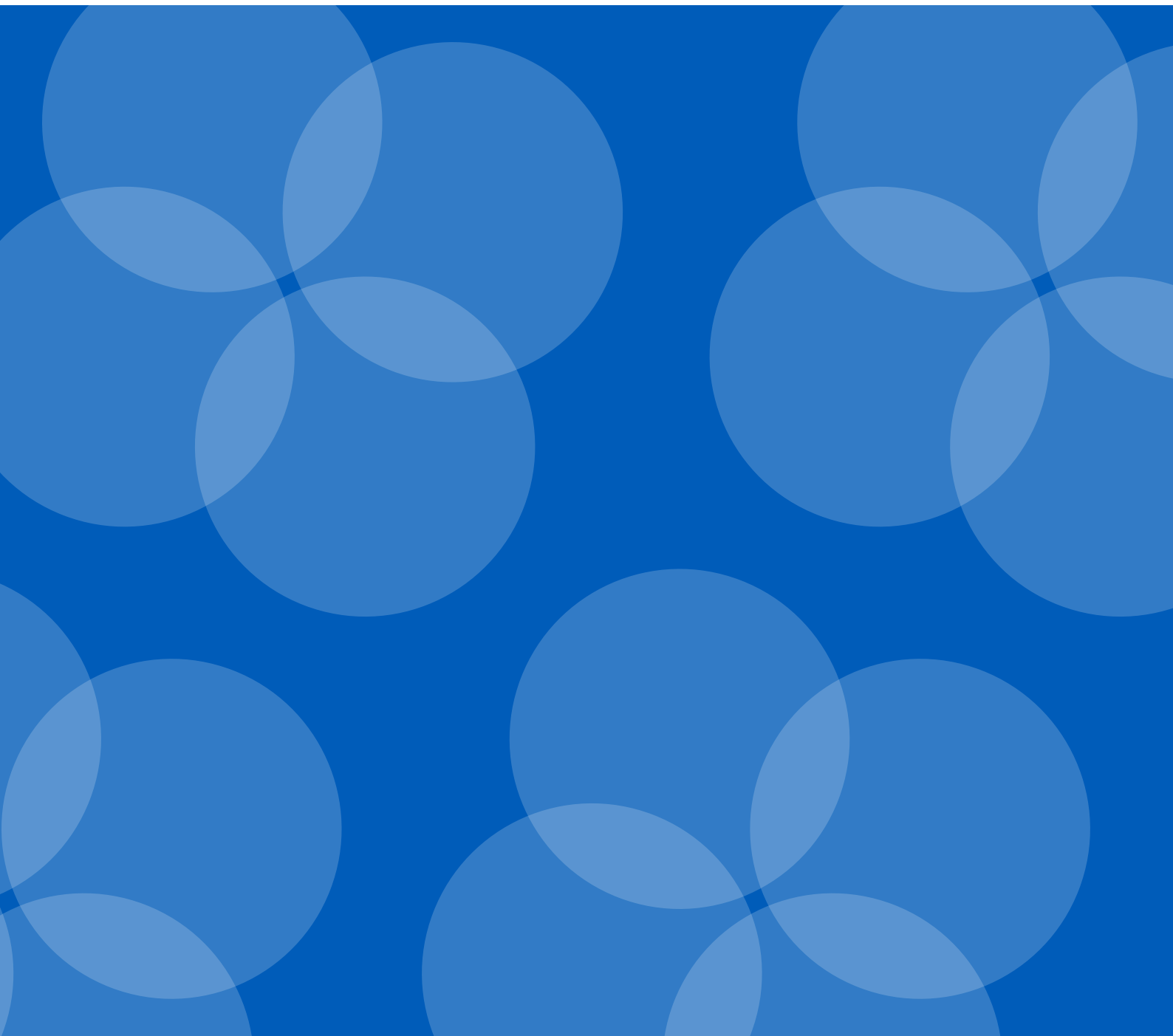


Att svara eller inte svara

Svenska elevers motivation att genomföra PISA-provet



Att svara eller inte svara

Svenska elevers motivation att genomföra PISA-provet

Beställningsuppgifter:

Fritzes kundservice

106 47 Stockholm

Telefon: 08-690 95 76

Telfax: 08-690 95 50

E-post: skolverket@fritzes.se

Beställningsnummer: 15:1443

ISSN: 1652-2508

ISBN: 978-91-7559-153-7

Grafisk formgivning: AB Typoform

Tryck: Elanders Sverige AB, 2015

Upplaga: 500 ex

Förord

PISA resultaten från 2012 visade på en kraftig resultatnedgång i alla de tre kunskapsområden, läsförståelse, matematik och naturvetenskap, som ingår i PISA. En av de tänkbara delförklaringar som diskuterades var den kontext i vilken PISA-undersökningen genomförs. PISA-resultatet har inte någon inverkan på elevernas betyg och är därför ur elevsynvinkel ett så kallat *low-stakes* prov, alltså ett prov utan direkta konsekvenser. Det är rimligt att anta att elever är mer motiverade och anstränger sig mer på prov som är betygsgrundande. Skolverket bedömde därför att motivationsaspekter i samband med internationella kunskapsprov var något som behöver studeras mer ingående.

I denna rapport redovisas två studier på temat elevers motivation, ansträngning och uthållighet att göra kunskapsprovet i PISA. Anita Wester på Skolverkets enhet för resultatutvärdering har skrivit den inledande texten med Skolverkets bedömning.

Analysen i den första studien *Svenska elevers rapporterade motivation och ansträngning i PISA, över tid och i jämförelse med andra länder* har genomförts av Hanna Eklöf vid Institutionen för tillämpad utbildningsvetenskap, Umeå Universitet.

Den andra studien *Har svenska elevers uthållighet och engagemang förändrats över tid? En studie av elevers svarsmonster på PISA:s kunskapsprov* har skrivits av Matilda Nilsson på enheten för resultatutvärdering, Skolverket.

Stockholm i mars 2015

Anna Ekström
Generaldirektör

Anita Wester
Senior rådgivare

Innehåll

Förord 3

Skolverkets bedömning 6

STUDIE I. Svenska elevers rapporterade motivation och ansträngning i PISA, över tid och i jämförelse med andra länder 11

Sammanfattning 12

Inledning 13

Provmotivationsbegreppet 13

Kan provmotivation vara viktigt att ta hänsyn till i PISA? 14

Sverige i PISA 15

Syfte 15

Genomförande 16

Resultat 20

Har det skett förändringar över tid för ansträngnings-termometern, i Sverige och jämfört med andra länder? 34

Finns det ett samband mellan rapporterad provmotivation och provprestation? 37

Kan provmotivation förklara nedgången i de svenska resultaten? 41

Sammanfattande diskussion 49

Begränsningar och fortsatta studier 51

Avslutande kommentar 52

Referenser 53

Bilaga 54

**STUDIE II. Har svenska elevers uthållighet
och engagemang förändrats över tid?
En studie om elevers svarsmönster på
PISAs kunskapsprov 61**

Sammanfattning 62

Inledning 64

Syfte 65

Genomförande 66

PISAs upplägg och kunskapsprovets design 67

Analysens idé och utgångspunkter 68

Avgränsningar och val av mätningar att jämföra 73

Metoder som används: två olika tillvägagångssätt 75

Resultat 77

Elevernas svarsmönster på matematikuppgifter 77

Elevernas svarsmönster på trenduppgifter i läsförståelse 83

Elevernas svarsmönster uppdelat efter
flervalsfrågor och frågor med öppna svar 87

Sammanfattande diskussion 90

Förändringar i elevernas uthållighet mellan åren 90

Tydligt stöd för hypotesen om sjunkande kunskapsnivå i PISA 91

Visst stöd för hypotesen om minskat provengagemang
mellan PISA 2003 och PISA 2009 91

Några avslutande ord 92

Referenser 94

Bilaga 95

Skolverkets bedömning

Bakgrund

År 2012 deltog Sverige för femte gången i OECD:s internationella studie PISA (Programme for International Student Assessment). PISA genomfördes för första gången år 2000 och har sedan dess återkommit vart tredje år – 2003, 2006, 2009 och 2012. PISA undersöker 15-åringars kunskaper i matematik, läsförståelse och naturvetenskap, där vart och ett av ämnena är huvudämne vid varje mättillfälle. Elevernas kunskaper testas genom ett prov. Genom enkäter fångas elevernas inställning till skolan och undervisningen.

I de första omgångarna av PISA visade svenska elever mycket goda kunskaper i samtliga tre kunskapsområden med resultat som låg över OECD-genomsnittet. Under hela decenniet har resultaten sedan successivt sjunkit för att 2012 ligga under OECD-genomsnittet i de tre kunskapsområdena. Sverige är också det land som haft den sämsta resultatutvecklingen av samtliga deltagande länder. I Skolverkets rapport, *PISA 2012* (2013) presenteras ett antal hypoteser som tänkbara delar av förklaringen till de försämrade resultaten. En av dessa förklaringshypoteser handlar om att eleverna kanske inte anstränger sig på PISA-provet i samma utsträckning som i tidigare PISA-studier, dvs. motivationen att göra sitt bästa på provet kan alltså ha minskat över tid.

Mot bakgrund av resultatförsämringen, som noterats inte bara i PISA, utan också tidigare kunnat ses i PIRLS 2011 (Skolverket, 2012a), TIMSS 2011, årskurs 8, (Skolverket, 2012b.) och som också helt nyligen konstaterats i den fördjupade analysen av PIAAC¹ (SNS, 2014), har Skolverket bland annat initierat arbetet med de två studier som här presenteras. Studierna kompletterar varandra på så sätt att analysen i den ena baseras på elevernas svar på frågor om ansträngning och motivation att göra PISA-provet, medan analysen i den andra bygger på elevernas faktiska svarsmonster i PISA-proven.

Studie I. Svenska elevers svar på frågor om motivation och ansträngning i PISA, över tid och i jämförelse med andra länder

Syfte

Syftet med studien är att undersöka svenska elevers svar på frågor om ansträngning och motivation att göra sitt bästa på PISA-provet 2012, samt att studera rapporterad ansträngning över tid, i jämförelse med andra länder och i relation till resultat på PISA-provet. Ansträngning och motivation att genomföra PISA-provet har mätts genom två olika mått: ”ansträngningstermometern”, som består av en bild på två termometrar med en skala, 1–10, där eleverna fått ange dels hur mycket de ansträngde sig på PISA-provet och dels hur mycket de skulle ha ansträngt sig om resultaten på provet skulle ha påverkat betygen. Ansträngningstermometern är ett gemensamt mått för alla deltagande nationer i PISA 2003, 2006 och 2012. Provmotivationsskalan däremot, baseras på sex frågor om hur motiverade eleverna kände sig att genomföra PISA-provet och lades till som ett nationellt tillägg i svenska PISA 2012. Såväl ansträngningstermometern som

1 SNS (2014). *Lära för livet? Om skolans och arbetslivets avtryck i vuxnas färdigheter.*

provmotivationsskalan är generella, och således inte direkt inriktade på något enskilt ämne. Båda baseras på elevernas egen bedömning av sin ansträngning och motivation.

Resultat

Svenska elever rapporterar en förhållandevis låg grad av ansträngning i PISA, internationellt och i jämförelse med länder som brukar anses jämförbara med Sverige. Eleverna rapporterar också en låg grad av provmotivation i jämförelse med samma svenska population elever som svarade på liknande frågor i samband med ett nationellt prov. Vidare rapporterar svenska elever en förhållandevis stor skillnad mellan ansträngning på PISA-provet och hur mycket de skulle ha ansträngt sig om provet räknats till betyget. Resultaten visar också att svenska elevernas grad av rapporterad ansträngning är något lägre nu än i tidigare PISA-undersökningar.

Det finns ett samband mellan rapporterad ansträngning och prestation i Sverige, ett samband som inte är försumbart och som är starkare än i de flesta andra länder. Vi ser dock samband av samma storlek i flera av de andra, oss näraliggande, länderna. Effekten av ansträngning på prestation är signifikant även om man kontrollerar för andra variabler. Detta gäller såväl för Sverige som för de flesta av de länder vi jämfört oss med.

Att svenska elever rapporterar en förhållandevis låg grad av motivation och ansträngning, att skattningen av ansträngning har gått ner något, och att skattad motivation och ansträngning har en effekt på prestation är ett intressant och viktigt resultat i sig. Det finns alltså ett flertal indikationer på att elevernas provmotivation och ansträngning i provsituationen är en viktig faktor att ta hänsyn till och fortsätta att följa.

Sammanfattningsvis kan det dock konstateras att nedgången i rapporterad ansträngning är så pass liten – och effekten av ansträngning är heller inte särskilt stark – att det verkar osannolikt att det är denna nedgång som starkt bidragit till de svenska elevernas svaga prestation i PISA 2012.

Enligt den statistiska analysen skulle minskad ansträngning kunna förklara en liten del av resultatförsämringen (exempelvis 4–5 poäng² av 31 poäng i matematik i PISA mellan 2003 och 2012), men den största delen förklaras sannolikt av andra faktorer, såsom att eleverna inte har tillräckliga kunskaper om det som mäts i PISA.

Studie II. Har svenska elevers uthållighet och engagemang förändrats över tid? En studie av elevers svarsmönster på PISA:s kunskapsprov

Syfte

Syftet med studien är att undersöka om delar av resultatnedgången i PISA 2012 kan sägas bero på att elever i den senaste undersökningen inte har haft uthållighet och engagemang för att besvara provfrågorna i samma utsträckning som tidigare, och i förlängningen om resultaten från PISA 2012 kan anses vara tillförlitliga. Tyngdpunkten ligger på matematikdelen, men också läsuppgifter ingår i analysen.

2 Medelvärdet är 500 och standardavvikelsen 100.

Vi undersöker förändringar i elevernas svarsmönster för provuppgifterna när de kommer först i provhäftet (position 1) och när de kommer sist i provhäftet (position 4) på två sätt: (1) dels för varje position mellan två PISA-mätningar och (2) dels för skillnaden mellan dessa två positioner inom en mätning (elevernas uthållighet). De svenska elevernas svarsmönster tolkas med det genomsnittliga OECD-mönstret som referens, för att kunna bedöma om de svenska elevernas mönster framstår som avvikande från OECD-elevernas. Däremot görs inga jämförelser med enskilda länder i denna analys. Två hypoteser har formulerats: (1) Resultatnedgången i PISA beror på en sjunkande kunskapsnivå hos de svenska eleverna. (2) Resultatnedgången i PISA beror på ett minskat engagemang bland eleverna för att skriva kunskapsprovet. I denna analys kan förändringar i elevernas svarsmönster inte kopplas direkt till resultatet på PISA-provet. Därför är det inte möjligt att ge ett svar på hur stor del av nedgången i resultat uttryckt i antal eller andel poäng som kan förklaras av elevers svikande engagemang.

Resultat

Elevernas uthållighet för matematikuppgifter tycks vara minst i PISA 2003, och högst i PISA 2006, medan uthålligheten i PISA 2012 hamnar någonstans däremellan. Elevernas uthållighet för uppgifter i läsförståelse är överlag oförändrad i Sverige. Undantaget är att uthålligheten enligt andelen elever som besvarar läsuppgifterna korrekt minskar mellan PISA 2003 och PISA 2009.

De svenska elevernas svarsmönster tyder sammantaget på att det främst är kunskapsnivån som har sjunkit fram till det senaste PISA-provet 2012. Stöd för denna hypotes ser vi i många av de jämförelser som görs, men särskilt när vi betraktar förändringen av andelen elever som besvarar uppgifterna korrekt i matematik, men också i läsförståelse. Det finns en tydlig, jämnstor nedgång i andelen elever som besvarar uppgifterna korrekt oavsett om uppgifterna kommer tidigt eller sent i provhäftet mellan de flesta jämförda PISA-mätningarna.

Det finns även förändringar i andelen elever som lämnar uppgifter obesvarade som stödjer denna hypotes. Dessa förändringar är dock oftast något mindre och något mer svårbedömda. Mellan PISA 2003 och PISA 2009, samt i viss mån också mellan PISA 2009 och PISA 2012 har det skett en jämn uppgång i andelen elever som lämnar läsuppgifterna obesvarade. Analysen ger också indikationer på att andelen elever som hoppar över uppgifter i matematik ökar mellan PISA 2006 och 2012. När uppgifterna kommer tidigt i provhäftet är uppgången mellan dessa år visserligen inte statistiskt säkerställd, men uthållighet är oförändrad mellan mätningarna. Det ger ytterligare, om än något svagare, stöd åt denna hypotes.

Det finns även vissa tecken som tyder på att provengagemanget bland svenska elever kan ha minskat något. Det är främst i jämförelsen av andelen elever som besvarar läsuppgifter korrekt mellan PISA 2003 och 2009 som ser vi ett svarsmönster som ger stöd för hypotesen om ett minskande engagemang hos eleverna. Mellan dessa år är andelen elever som besvarar uppgifterna korrekt oförändrad när läsuppgifterna ligger först i provhäftet, men minskar när uppgifterna ligger sist, vilket alltså innebär att uthålligheten minskar från PISA 2003 till 2009. Det bör dock noteras att en liknande, om än något mindre, förändring även framkommer i OECD i genomsnitt.

Även om det inte är möjligt att ge ett svar uttryckt i antal eller andel poäng av resultatnedgången, så är den samlade bedömningen att denna eventuella minskning i engagemang inte skulle kunna förklara någon större del av den nedgång i resultat som svenska elever har uppvisat under det senaste decenniet.

Den samlade bilden

De båda studierna utgör goda komplement till varandra, inte minst därför att Studie I baseras på elevernas självskattade motivation och ansträngning medan Studie II fångar elevernas faktiska beteende. Även om resultaten från de båda analyserna ska tolkas med försiktighet förstärker de varandra. Ingen av dem stödjer tanken om att det är bristande motivation och/eller ansträngning under provgenomförandet som till större delen skulle kunna förklara den kraftiga nedgången i PISA-resultaten. Studie II visar att något entydigt mönster med minskande uthållighet över tid inte kan konstateras.

Visserligen visar studie I att skillnaden i motivation mellan att göra ett high-stakes prov³ och ett low-stakes prov är stor hos svenska elever. I internationell jämförelse är det hos de svenska eleverna som den största skillnaden återfinns. Motivationen hos svenska elever är alltså inte på topp vid PISA-genomförandet, men motivationen hos elever i andra deltagande länder är inte heller maximal, även om den är något högre än svenska elevers. Att motivationen inte är maximal i något land är ett resultat som för övrigt förefaller tämligen rimligt med tanke på att eleverna inte får någon konkret ”vinst” av sin insats på PISA-provet. Varken eleverna eller lärarna får heller veta de enskilda elevernas resultat.

Resultaten av de två studierna sammantagna visar att det knappast är förändringar i motivationen, engagemanget eller ansträngningsnivån som på ett avgörande sätt försämrat resultaten, utan att den största delen sannolikt förklaras av andra faktorer, såsom att eleverna inte har tillräckliga kunskaper om det som mäts i PISA.

Det är dock viktigt att ha i åtanke att det inte finns några enkla samband eller enkla förklaringar till vare sig de svenska elevernas fallande resultat, eller till hur ansträngning och prestation samspelar. Genom fördjupade studier med olika perspektiv kan bitar i ett pussel läggas som vartefter ger ökad förståelse av elevers kunskapsutveckling, liksom av samspelet mellan socioemotionella och kognitiva variabler, men då det handlar om komplexa fenomen blir de enkla svaren endast grova beskrivningar av hur verkligheten fungerar.

Även om dessa analyser har haft ambitionen att fördjupa kunskaperna om svenska elevers svarsbeteende och skattade motivation och ansträngning i PISA, över tid och i jämförelse med andra länder, finns det fortfarande många aspekter kvar att belysa. Två sådana är skillnader mellan pojkar och flickor och vad de kan bero på, liksom mer fördjupade studier av de elever som inte alls svarar på frågorna om ansträngning och motivation. En annan aspekt är eventuella skillnader mellan skolor när det gäller skattad ansträngning. Eftersom PISA är ett prov utan betydelse för eleverna är det tänkbart att hur skolan kommunicerar vikten av en bra prestation på PISA till eleverna kan ha betydelse för deras motivation och ansträngning.

3 Det finns inget enkelt och tydligt svenskt begrepp för detta, att provet får viktiga konsekvenser för eleven. Därför används det engelska uttrycket.

Studier av elevers svarsbeteende och skattade ansträngning och motivation skulle också kunna kompletteras med mer småskaliga studier som exempelvis observationer eller intervjuer med elever som deltar i studier av typen PISA, för att få en ökad förståelse av hur proven uppfattas av eleverna, men också av vad de egentligen svarar på när de svarar på frågor om ansträngningstermometern och andra frågor relaterade till motivation, engagemang och uthållighet.

STUDIE I

**Svenska elevers rapporterade
motivation och ansträngning i PISA,
över tid och i jämförelse
med andra länder**



Sammanfattning

Syftet med föreliggande studie har varit att studera svenska elevers grad av rapporterad motivation och ansträngning i PISA, över tid och i jämförelse med ett urval andra länder. Analysen avsåg bidra med kunskap om huruvida provmotivation tycks vara en viktig aspekt att ta hänsyn till, liksom om huruvida de svenska elevernas sjunkande prestation i PISA kan förklaras av att de anstränger sig mindre, snarare än av andra faktorer såsom att kunskapsnivåerna har sjunkit.

I sammanfattning indikerar resultaten från denna studie att provmotivation verkar vara en viktig faktor att hålla under uppsikt i kontexten av PISA men samtidigt att de sjunkande resultaten i PISA inte i någon högre utsträckning kan förklaras av att eleverna anstränger sig mindre.

Följande resultat ligger till grund för slutsatsen att provmotivation kan vara en viktig faktor att ta hänsyn till i kontexten av PISA:

- Svenska elever rapporterar en förhållandevis låg grad av ansträngning i PISA, internationellt och i jämförelse med länder som brukar anses jämförbara med Sverige. Eleverna rapporterar också en låg grad av provmotivation i jämförelse med samma population elever som svarade på liknande frågor i samband med ett nationellt prov.
- Svenska elever rapporterar en förhållandevis stor skillnad mellan ansträngning på PISA-provet och hur mycket de skulle ha ansträngt sig om provet räknats till betyget
- De svenska elevernas grad av rapporterad ansträngning är något lägre nu än i tidigare PISA-undersökningar
- Det finns ett samband mellan rapporterad ansträngning och prestation i Sverige, ett samband som inte är försumbart och som är starkare än i de flesta andra länder. Samband av samma storlek ses dock också i flera av de andra, oss näraliggande, länderna.
- Det finns en signifikant effekt av skattad ansträngning på prestation även om man kontrollerar för andra variabler. Detta gäller såväl för Sverige som för de flesta av de länder vi jämfört oss med

Även om inga av dessa punkter enskilt signalerar kris avseende de svenska elevernas motivation och ansträngning så indikerar de tillsammans att provmotivation är en faktor som inte bör ignoreras.

Dock finner analyserna lite stöd för att brist på motivation är en avgörande variabel i det avseendet att lägre ansträngning skulle kunna förklara nedgången i de svenska resultaten:

- Nedgången i skattad ansträngning är så pass liten, och effekten av skattad ansträngning på prestation är inte så stark, att det verkar osannolikt att det är detta som förklarar de svenska elevernas kraftigt försämrade prestation i PISA 2012.
- Enligt den statistiska analysen skulle lägre ansträngning kunna förklara en liten del (4–5 poäng) av resultatförsämringen, men den största delen förklaras sannolikt av andra faktorer, såsom att eleverna inte har tillräckliga kunskaper om det som mäts i PISA.

Inledning

Internationella jämförande studier av elevers kunskaper har under de senaste årtiondena vuxit i omfattning, och de har ibland haft en betydande inverkan på utbildningssystem världen över. Den studie som har fått det största genomslaget på senare tid är sannolikt PISA. Föreliggande studie fokuserar svenska elevers rapporterade motivation och ansträngning i PISA. För att ge en bakgrund till analysen presenteras här närmast definitioner och antaganden när det gäller motivation i provsituationer, hur denna problematik relaterar till PISA, och varför det kan vara relevant att studera detta ur ett svenskt perspektiv.

Provmotivationsbegreppet

Motivation är en viktig aspekt när det gäller individers utveckling, lärande och prestationer i olika sammanhang. Motiverade individer tenderar att vara mer uthålliga, målmedvetna, aktiva, och prestera bättre. (se Pintrich & Shunk, 2002; Wigfield & Eccles, 2000). Motivation kan definieras på olika sätt, och mätas på olika nivåer. *Provmotivation* mäts på situationsspecifik nivå. Denna typ av motivation är ett tillstånd som är knutet till ett givet prov, och kan definieras som «villigheten att arbeta med provuppgifter och att lägga ner ansträngning och uthållighet i detta arbete».

Det finns några antaganden som visar på vikten av att ta hänsyn till provmotivation i vissa provsammanhang. Ett sådant antagande är att det behövs ett visst mått av motivation för att eleven ska kunna demonstrera sina kunskaper på ett bra sätt. För en bra prestation på ett prov behövs förstås primärt tillräckliga kunskaper. Har man inte kunskap om det som provet mäter hjälper det inte att vara väl motiverad. Utöver tillräckliga kunskaper behövs dock också en tillräcklig grad av motivation att anstränga sig för att visa dessa kunskaper. Ett annat antagande är att problemet med låg motivation är störst i provsituationer där resultatet inte har några konsekvenser för provtagaren (så kallade low-stakes prov). En hög grad av motivation följer ofta av att man sätter något slags värde på den uppgift som ska göras, att uppgiften eller konsekvensen av uppgiften betyder något för individen, direkt eller indirekt. Det kan då också antas att när provresultatet inte har några konsekvenser, kommer vissa elever att värdera uppgiften lågt och därmed inte vara motiverade att göra sitt bästa. Resultatet kan bli att elever presterar under sin faktiska förmåga. Att inte ta hänsyn till motivation i provsituationen kan således medföra att resultaten tolkas felaktigt, att det som tas för ett rent mått på kunskap egentligen även är ett mått på något annat (till exempel, motivation).

Kan provmotivation vara viktigt att ta hänsyn till i PISA?

På politisk nivå är studier som PISA att betrakta som *high-stakes*. Resultaten används ofta av politiker och beslutsfattare som indikatorer på utbildningens kvalitet och har ibland varit bidragande orsaker till reformer av utbildningssystem. Eftersom det mediala och politiska intresset för resultaten från internationella jämförande studier ofta är stort, bidrar rapporteringen av dessa studier också till att forma allmänhetens bild av utbildningens tillstånd i det egna landet. Från elevernas perspektiv är dock PISA ett exempel på en mätning där det saknas incitament att göra sitt bästa och det därför kan vara viktigt att ta hänsyn till elevers provmotivation vid tolkningen av resultaten. Eleven måste ägna timmar åt ett prov som för vissa kan vara ganska krävande, medan resultatet av elevens insats inte har någon praktisk betydelse för eleven personligen. Resultatet räknas inte till betyget och ingen återkoppling sker på individuell nivå, resultaten redovisas endast på landsnivå långt efter att studien är genomförd. En bra prestation leder därmed inte till någon form av uppskattning eller belöning, och en dålig prestation ger inte några negativa konsekvenser. Detta skulle kunna leda till att eleven sätter lågt värde på en bra prestation i PISA, anstränger sig mindre och presterar sämre än han eller hon hade gjort om provets resultat haft konsekvenser.

Å andra sidan skulle det faktum att eleverna representerar sitt land i en internationell studie kunna vara tillräckligt för att väcka ett intresse och skapa motivation. Att provet inte har några konsekvenser kan också göra provsituationen mindre stressande än vid andra prov, vilket kan vara positivt för prestationen, även om man inte är maximalt motiverad. Man ska inte heller underskatta elevers intresse av att göra bra ifrån sig bara för sin egen skull, för att de "alltid gör sitt bästa", eller för att "hjälpa till med studien". Vissa elever kan tycka det är intressant och meningsfullt att få testa sig själva utan just den betygsstress som omgärdar flera andra prov.

Alla dessa argument skulle kunna vara sanna, och det finns sannolikt elever i alla ovanstående kategorier vid en given provsituation: de som tycker det känns meningslöst såväl som de som tycker det känns spännande och intressant. Ganska få studier av provmotivation har genomförts i den faktiska kontexten av internationella studier som TIMSS och PISA, men de studier som ändå har gjorts i dessa och andra kontexter har generellt visat att elever rapporterar lägre motivation på low-stakes prov jämfört med high-stakes prov, och att provmotivationen tenderar att ha en viss betydelse för prestation (se exempelvis Eklöf & Knekta, 2014, eller Wise & DeMars, 2005 för en översikt).

Skulle elever i vissa länder vara väl motiverade att göra sitt bästa medan elever i andra länder saknar motivation att anstränga sig skulle jämförelser av resultat i termer av kunskapsnivåer kunna bli missvisande. I diskussionerna kring PISA förs det också ibland fram teser om att elever i vissa länder sätter ett större värde på en bra prestation i PISA än elever i andra länder, beroende på landets provkultur och kultur i stort, eller hur provet framställs för eleverna. Frågan om provmotivation i PISA blir således en fråga om huruvida provresultaten är tillförlitliga: om ett lands resultat faktiskt speglar elevernas kunskaper, och om jämförelser mellan länder faktiskt är jämförelser av kunskapsnivåer och inte av motivationsnivåer.

Sverige i PISA

Resultaten från de senaste omgångarna av PISA har varit nedslående för Sveriges del. Sedan den första PISA-undersökningen år 2000 har de svenska resultaten sjunkit kontinuerligt, och Sverige är nu det land i PISA-studien där resultaten försämrats allra mest (Skolverket, 2013). Skulle Sveriges sjunkande resultat kunna bero på att eleverna blivit mindre motiverade att anstränga sig och göra sitt bästa på ett prov som PISA-provet och att svenska elever är mindre motiverade att göra sitt bästa än elever i andra länder? Prov- och bedömningsklimatet i den svenska skolan har förändrats avsevärt under det senaste decenniet, bland annat på så sätt att Sverige idag har fler nationella prov i fler ämnen och fler årskurser än vi haft någon gång tidigare och eleverna får betyg från tidigare årskurs. I årskurs nio skriver eleverna nationella prov i fem ämnen. De nationella proven används av lärarna som ett stöd vid betygsättningen och upplevs som viktiga för eleverna. Det är tänkbart att prov som PISA av eleven uppfattas som mindre viktigt att lägga sin tid och kraft på i det ökade utbudet av nationella prov som de svenska niondeklassarna gör idag.

Mot denna bakgrund är syftet med föreliggande studie att titta närmare på svenska elevers grad av rapporterad motivation och ansträngning i PISA, över tid och i jämförelse med ett urval av andra länder.

Syfte

Syftet med denna studie är att undersöka svenska PISA-elevers rapporterade grad av ansträngning och motivation att göra sitt bästa i PISA, att studera rapporterad ansträngning på PISA över tid, i jämförelse med andra länder och i relation till prestation på PISA-provet. Provmotivation har mätts genom två olika mått: ”ansträngningstermometern” som är ett gemensamt mått för alla deltagande nationer i PISA sedan 2003, liksom genom en provmotivationsskala, som lades till som ett nationellt tillägg i svenska PISA 2012. Båda dessa mått bygger på självskattning, att eleven själv får ta ställning till ett antal påståenden som handlar om motivation och ansträngning.

Studien utgår från följande frågeställningar och de redovisas också i denna ordning i separata avsnitt i det som följer:

1. Tycks svenska elever vara motiverade att göra sitt bästa i PISA 2012? Under denna frågeställning presenteras beskrivande resultat av vad eleverna svarat på provmotivationsskalan och ansträngningstermometern.
2. Finns det skillnader mellan länder avseende självskattad ansträngning på ansträngningstermometern i PISA 2012? Här tittar vi på resultatet från ansträngningstermometern i ett internationellt perspektiv och för ett urval av länder
3. Har det skett förändringar över tid avseende resultat på ansträngningstermometern, i Sverige och i jämförelse med ett urval av andra länder?
4. Finns det något samband mellan rapporterad provmotivation och provprestation för de svenska eleverna, respektive i ett urval av länder?
5. Kan provmotivation förklara nedgången i de svenska resultaten? Finns någon effekt av skattad ansträngning när man tar hänsyn till andra attityd/motivationsvariabler?

Innan resultatredovisningen presenteras redogörs nedan mycket kortfattat för de analyser och de avgränsningar som gjorts liksom de instrument som använts.

Genomförande

Metod och instrument

Data från framför allt PISA 2003 och 2012, men i viss mån även från PISA 2006, har använts i analyserna. Alla data utom de nationella tilläggen har laddats ner från <http://www.oecd.org/pisa/pisaproducts/> där de finns fritt tillgängliga. Data från elevenkäten, ansträngningstermometern och de nationella tilläggen med provmotivationsfrågor har använts tillsammans med provresultat. Bearbetning och analys har gjorts i programmen SPSS och IEA IDB Analyser. Beskrivande statistik, signifikanstester, korrelations- och regressionsanalyser liksom analyser av skalornas homogenitet och dimensionalitet har gjorts. Om inte annat anges har alla analyser gjorts med viktade värden och med alla fem plausibla värden som beroendevariabler (varför viktade värden bör användas och varför man i PISA arbetar med resultatdata i form av flera ”plausibla värden” finns beskrivet i OECD, 2013a).

Fokus kommer att ligga på analyser av ansträngningstermometern, då denna kan användas för att titta på skillnader över tid och i relation till andra länder, medan resultaten från provmotivationsskalor bara gäller för Sverige och bara för 2012. I första hand redovisas resultat för Sverige, men i vissa av analyserna med ansträngningstermometern inkluderas ett urval av andra länder. Detta urval innehåller de nordiska länderna samt övriga länder kring Östersjön (utom Ryssland). Dessutom har Nederländerna, Australien och Nya Zeeland inkluderats i länderurvalet, då detta är länder som vi ibland jämför oss med i dessa sammanhang.

Fokus kommer vidare primärt att ligga på matematik, då matematik var huvudområde såväl 2003 som 2012, och mer tillförlitliga jämförelser därför kan göras över tid. Resultaten är således mer stabila när de relateras till matematik än då de relateras till de andra två ämnesområdena i PISA: läsförståelse och naturvetenskap. De motivationsskalor som ingår i elevenkäten för PISA 2012, och som också används som variabler i vissa av analyserna, handlar också främst om elevers attityder till och motivation att lära matematik. Läsförståelse och naturvetenskap kommer dock också att inkluderas i ett flertal av analyserna, för att få en uppfattning om huruvida ansträngning/motivation har olika effekt på resultatet i olika ämnen.

Ansträngningstermometern


Ansträngningstermometern användes för första gången i PISA 2000 i tre av de deltagande länderna: Norge, Tyskland och Australien. År 2003 användes termometern för första gången i alla länder, och sedan dess har den använts även 2006 och 2012. Ansträngningstermometern ligger sist i provhäftet och innehåller två frågor som eleven får besvara på en tiogradig skala (en termometer), (se Figur 1). I Figur 1 finns också en tredje skala (kolumnen till vänster), där siffran 10 redan är markerad för att illustrera en påhittad situation (vilken som helst) som är väldigt viktig för eleven personligen, och där denne därför skulle göra sitt allra bästa. På skalan i mittenkolumnen ombeds eleven sedan markera hur mycket denne ansträngde sig på PISA, *om eleven jämför sin ansträngning på PISA med denna påhittade situation*. På skalan i den högra kolumnen får eleven ange hur mycket den skulle ha ansträngt sig på PISA om provresultatet hade räknats till betyget. Det kan noteras att skattningen på termometern handlar om PISA-provet totalt, det vill säga såväl läsförståelse, matematik och naturvetenskap.

Även om ansträngningstermometern använts i flera PISA-omgångar har ganska få analyser gjorts just med fokus på elevernas självskattade ansträngning, och vi vet således inte så mycket om detta ännu. Med utgång från data från PISA 2000 och 2003 (läsförståelse) har det dock skrivits en avhandling med fördjupade analyser av elevernas ansträngning och med ett särskilt fokus på deras ”relativa ansträngning” (differensen mellan hur mycket de ansträngde sig på PISA och hur mycket de säger att de skulle ha ansträngt sig om resultatet räknats till betyget) (se Butler, 2008).

I sin avhandling drar Butler slutsatsen att varierande grad av motivation bland de deltagande länderna inte är något stort hot mot PISA-resultaten då skillnaderna mellan länder inte var så stora som man hade kunnat befara. Hon kom också fram till att det internationellt fanns ett samband mellan ansträngning och prestation, och att detta samband inte var försumbart men heller inte kunde förklara särskilt mycket av skillnaderna mellan länder i ett internationellt perspektiv (Butler, 2008). Utifrån dessa analyser drar också OECD slutsatsen att ”Reassuringly, students’ self-reports on this subject suggests that the effort they invest in PISA is fairly stable across countries” (OECD, 2007, p. 52). Då detta är den hittills enda större studie av ansträngningstermometern som tycks vara gjord kommer resultaten från den aktuella studien att relateras till Butlers studie där det är lämpligt.

Figur 1. Ansträngningstermometern.

Försök föreställa dig en verklig situation (i skolan eller i något annat sammanhang) som är väldigt viktig för dig personligen, då du skulle göra ditt allra bästa och anstränga dig maximalt för att prestera bra.

	I en sådan situation skulle du markera det högsta värdet på "ansträngningstermometern", som visas nedan:	I jämförelse med den situation du just tänkt dig, hur mycket ansträngde du dig med det just avslutade provet?	Hur mycket skulle du ha ansträngt dig om resultatet på provet skulle ha påverkat dina betyg?
	<input checked="" type="checkbox"/> 10	<input type="checkbox"/> 10	<input type="checkbox"/> 10
	<input type="checkbox"/> 9	<input type="checkbox"/> 9	<input type="checkbox"/> 9
	<input type="checkbox"/> 8	<input type="checkbox"/> 8	<input type="checkbox"/> 8
	<input type="checkbox"/> 7	<input type="checkbox"/> 7	<input type="checkbox"/> 7
	<input type="checkbox"/> 6	<input type="checkbox"/> 6	<input type="checkbox"/> 6
	<input type="checkbox"/> 5	<input type="checkbox"/> 5	<input type="checkbox"/> 5
	<input type="checkbox"/> 4	<input type="checkbox"/> 4	<input type="checkbox"/> 4
	<input type="checkbox"/> 3	<input type="checkbox"/> 3	<input type="checkbox"/> 3
	<input type="checkbox"/> 2	<input type="checkbox"/> 2	<input type="checkbox"/> 2
	<input type="checkbox"/> 1	<input type="checkbox"/> 1	<input type="checkbox"/> 1

Provmotivationsskalan

Provmotivationsskalan består av sex påståenden (se Tabell 1, nedan) Varje påstående besvarades på en fyrgradig skala (från ”håller absolut med” till ”håller absolut inte med”). Provmotivationsfrågorna låg i slutet av elevenkäten, som eleverna besvarar efter att de gjort provet. En liknande skala har använts i ett flertal tidigare studier i den svenska kontexten (TIMSS Advanced 2008, TIMSS 2003, Nationella prov, Högskoleprovet). Skalan innehåller dels frågor som avser mäta hur stort värde eleverna sätter på provet (om en bra prestation på provet känns viktig), dels frågor som handlar om hur motiverade eleverna var att anstränga sig liksom hur mycket de ansträngde sig. Just den aktuella skalan har använts i flera tidigare PISA-omgångar i Norge, och har där fungerat väl. Analyser av skalans egenskaper i svenska PISA 2012 tyder också på att den utgör ett homogent och endimensionellt mått.

Ansträngningstermometern och provmotivationsskalan är delvis överlappande (i alla fall i teorin) då de båda handlar om ansträngning. Båda dessa mått bygger också på självskattning och vi har därmed bara tillgång till elevernas egna utsagor om hur mycket de säger att de ansträngde sig eller hur motiverade de var. Vi kan inte veta hur mycket de ansträngde sig egentligen, eller hur motiverade de egentligen var, utan får lita på att eleverna ger en sanningsenlig bild av hur de själva kände och agerade. Detta är problem som alla självskattningsskalor brottas med, men självskattning används i många studier och det finns gott stöd för att resultat från självskattningsskalor kan vara användbara, men man bör förstås vara medveten om deras begränsningar.

Provmotivationsskalan handlar mer om hur eleverna kände för provet än vad ansträngningstermometern gör, då den bara frågar just efter ansträngning och inte direkt efter motivation. Då de besvaras vid olika tillfällen är det också tänkbart att de påverkas av olika ”fel”. Ansträngningstermometern besvaras direkt efter provet, och här kan det finnas en större risk att eleven svarar på, inte bara hur mycket den ansträngde sig, utan också hur den tyckte att det gick på provet den nyss gjorde, hur svårt provet upplevdes, vilken tänkt situation eleven relaterat sin skattade ansträngning till (se Figur 1), etc. Provmotivationsskalan besvaras i slutet av elevenkäten, och här finns en risk att eleverna är trötta på att svara på frågor och därför inte svarar alls/inte ger helt sanningsenliga svar. En elev som är omotiverad att anstränga sig på provet är kanske inte heller motiverad att svara på frågor om sin motivation. Det finns också en risk att de svarar på sin motivation för hela dagen, både för provet och för enkäten. Utgångspunkten är ändå att dessa båda mått kan ge en inblick i hur eleverna uppfattade PISA, hur motiverade de var att göra sitt bästa och hur mycket de ansträngde sig på provet.

Resultat

Tycks svenska elever vara motiverade att göra sitt bästa i PISA 2012?

Provmotivationsskalan

Hur stor andel elever som höll med respektive inte höll med om respektive påstående i provmotivationsskalan redovisas i Tabell 1, tillsammans med medelvärdet för varje fråga. Observera att i denna tabell gäller att ju lägre medelvärde, desto mer håller man med om påståendet (då 1 = håller absolut med, 4 = håller absolut inte med).

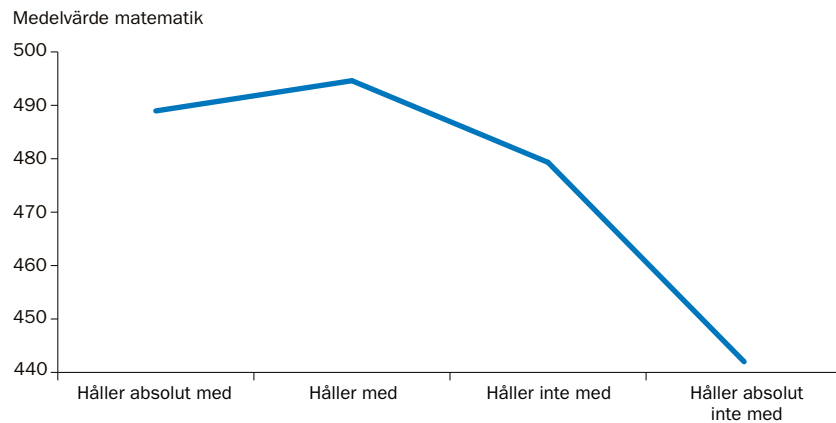
Tabell 1. Procentandel av de svenska eleverna i PISA 2012 som angett respektive svarsalternativ på frågorna i provmotivationsskalan, tillsammans med medelvärden på uppgiftsnivå.

	1) Håller absolut med (%)	2) Håller med (%)	1) + 2) totalt (%)	3) Håller inte med (%)	4) Håller absolut inte med (%)	3) + 4) totalt (%)	M	SD
1. Jag kände mig motiverad att göra mitt bästa på PISA-provet	16,1	46,8	62,9	27,1	10,1	37,2	2.31	.86
2. Jag gjorde en bra insats på PISA-provet	15,5	58,5	74,0	19,8	6,2	26,0	2.17	.76
3. Det var viktigt för mig att göra mitt bästa på PISA-provet	11,2	34,7	45,9	41,8	12,3	54,1	2.55	.85
4. Jag arbetade med uppgifterna i provet utan att ge upp även om vissa uppgifter kändes svåra	15,1	47,1	62,2	29,3	8,5	37,8	2.31	.83
5. Det betydde mycket för mig att göra bra ifrån mig på PISA-provet	8,6	28,0	36,6	46,7	16,6	63,3	2.71	.84
6. Jag gjorde mitt bästa på PISA-provet	22,0	44,3	66,3	24,5	9,3	33,8	2.21	.89

Slår man ihop andelarna för de två positiva svarsalternativen respektive de två negativa svarsalternativen (de blå kolumnerna i Tabell 1) framgår att en majoritet av eleverna uppger att de kände sig motiverade att göra sitt bästa på PISA-provet (fråga 1), att de gjorde sitt bästa på provet (fråga 6) och arbetade med uppgifterna i provet även om de kändes svåra (fråga 4). Tre fjärdedelar av eleverna håller med om att de gjorde en bra insats på provet (fråga 2). Mindre än hälften av eleverna uppger att det kändes viktigt för dem att göra bra ifrån sig på PISA (fråga 3), och en dryg tredjedel av eleverna uppgav att det betydde mycket för dem att göra bra ifrån sig i PISA (fråga 5). Det bör noteras att dessa procentandelar gäller bland de elever som faktiskt svarade på frågorna. För varje fråga var det runt tio procent bortfall.

Som ett exempel på hur olika svar på motivationsfrågorna är relaterade till provresultat redovisas i Figur 2 medelresultat i matematik per skattat värde för fråga 1, ”Jag kände mig motiverad att göra mitt bästa på PISA-provet.

Figur 2. Resultat på PISAs matematikdel per skalsteg på frågan ”Jag kände mig motiverad att göra mitt bästa på PISA-provet.



Frågorna 2), 4), och 6) i Tabell 1 ovan uppvisar likartade mönster som det i Figur 1, det tycks alltså som att en högre grad av rapporterad provmotivation är relaterat till högre prestation, med en liten dipp för dem som ”absolut håller med om” påståendet. De som inte alls svarade på frågan presterade som gruppen som inte alls höll med, eller sämre. För de två frågor som handlar om huruvida provet uppfattades som viktigt ser mönstret lite annorlunda ut: de elever som ”absolut höll med” om att det var viktigt respektive betydde mycket att göra sitt bästa på provet presterade avsevärt sämre än de elever som ”höll med” om påståendena. Att eleverna rapporterar att de tycker provet är jätteviktigt är således i sig ingen garanti för att eleverna presterar väl, däremot har tidigare studier visat att om eleverna rapporterar att provet är viktigt är eleverna också mer motiverade att anstränga sig.

Jämfört med liknande studier som gjorts i TIMSS-kontexten, där liknande om än inte identiska frågor använts, så rapporterar de svenska eleverna i PISA 2012 en högre grad av motivation än elever i tredje året på gymnasiet som deltog i TIMSS Advanced 2008 (se Eklöf, Japelj Pavecic, & Gronmo, 2014). I TIMSS Advanced 2008 var det exempelvis mindre än 35 procent av de svenska eleverna som uppgav att de kände sig motiverade att göra sitt bästa eller att de arbetade med alla uppgifter i provet utan att ge upp. Gruppen elever som

deltog i TIMSS Advanced 2008 skiljer sig dock på många sätt från elevgruppen i PISA, vilket gör direkta jämförelser mindre lämpliga. Eleverna i TIMSS Advanced är tre år äldre än eleverna i PISA och förmodligen betydligt mer studietaktiska. Jämfört med elever som deltog i TIMSS 2003 tycks PISA-eleverna rapportera något lägre grad av motivation (Eklöf, 2006), men upp-lägget för att mäta provmotivation var något annorlunda i TIMSS 2003. Jämfört med liknande frågor som använts i samma elevkohort men i kontex-ten av nationella prov, uppger en mindre andel av eleverna i PISA 2012 att de kände sig motiverade: I en studie av Eklöf och Knekta (2014) som inkluderade ett stickprov elever i årskurs 9 som skrev nationellt prov i naturvetenskap våren 2012, angav mellan 80 och 90 procent av eleverna att de kände sig motiverade att göra sitt bästa på provet, att de arbetade med alla uppgifter utan att ge upp, och att det var viktigt för dem att få ett bra resultat på provet.

Sammanfattningsvis indikerar de svenska elevernas självskattningar på prov-motivationsskalan i PISA 2012 att endast en liten andel elever struntade helt i att göra sitt bästa på PISA-provet, men att ganska många ändå tycks ha haft ganska ljum inställning till provet, och de rapporterar en lägre grad av motiva-tion än årskurs 9-elever från samma kohort som skrev ett nationellt prov.

Även om jämförelser med tidigare studier i TIMSS-kontexten respektive den svenska nationella prov-kontexten ger en viss indikation på om eleverna i PISA kan anses rapportera en hög eller en låg grad av motivation, är det ändå svårt att dra några egentliga slutsatser då vi i dagsläget inte har några andra PISA-stickprov att jämföra med för just provmotivationsskalan. Den ansträngnings-skala som används i PISA internationellt kan då ge en bättre indikation i det avseendet. Nedan presenteras därför inledningsvis svenska resultat för ansträng-ningstermometern i PISA 2012, och i nästa avsnitt görs sedan en jämförelse med andra länder avseende skattningar på ansträngningstermometern.

Ansträngningstermometern – deskriptiva resultat PISA 2012

Ansträngningstermometern ligger sist i provhäftet i PISA, och eleverna besvarar den direkt efter provet, före elevenkäten. I databaserna rapporteras två frågor: dels hur mycket eleverna ansträngde sig på PISA-provet jämfört med en maxi-mal prestation (se Figur 1, mittenkolumnen i figuren), dels hur mycket de skulle ha ansträngt sig om provresultatet räknats för deras betyg (frågan till höger i Figur 1). Det redovisas även en differens mellan dessa två värden som ett slags mått på relativ ansträngning. Nedanstående handlar framför allt om den ena frågan: hur mycket eleverna uppger att de ansträngde sig på PISA-provet, men analyser har även gjorts för differensvariabeln.

I PISA 2012 har Sverige ett medelvärde på 7,03 på den tiogradiga skalan avseende hur mycket de ansträngde sig på PISA-provet. 8,4 procent av de elever som besvarade frågan markerade en tia på ansträngningstermometern, det vill säga maximal ansträngning. Det vanligaste värdet att ange är en åtta, nästan en fjärdedel av de svenska eleverna har markerat en åtta på ansträngningstermome-tern. Det näst vanligaste värdet är en sju, som markeras av ungefär 20 procent av eleverna. Det tredje vanligaste värdet är en nia, som har markerats av ungefär 15 procent av eleverna. Vidare är det ungefär 20 procent av eleverna som anger en femma eller mindre i ansträngning. I tabell 2 nedan redovisas hur stor andel av eleverna som markerat respektive skalsteg på ansträngningsskalan, där "1"

står för minimal ansträngning och ”10” står för maximal ansträngning. Vidare redovisas medelresultat i matematik, läsförståelse och naturvetenskap för respektive skalsteg på ansträngningsskalan. Exempel: 11,6 % av eleverna markerade en sexa i ansträngning på ansträngningstermometern. Dessa elever hade ett genomsnittligt resultat på matematikdelen i PISA på 469 poäng.

Tabell 2. Andel elever som angivit respektive värde (1–10) på ansträngningstermometern (AT) i PISA 2012, tillsammans med genomsnittligt provresultat för varje skalsteg.

Skattat värde på AT	% av svenska elever	Resultat matematik 2012	Resultat läsförståelse 2012	Resultat naturvetenskap 2012
1	2,0	387	359	382
2	2,0	431	420	432
3	3,1	446	437	446
4	4,7	443	444	450
5	8,4	448	450	452
6	11,6	469	478	478
7	20,4	490	500	500
8	23,6	502	512	511
9	15,7	511	525	521
10	8,4	498	507	505

Resultaten ser väldigt lika ut för matematik, läsförståelse och naturvetenskap, med högre genomsnittlig provpoäng för i princip varje skalsteg mellan 1–9 på ansträngningsskalan. De som markerat en tia på ansträngningsskalan, det vill säga maximal ansträngning, presterar i genomsnitt lite sämre än de som markerat en nia eller en åtta. Gruppen elever som markerade en tia på ansträngningstermometern var en ganska heterogen grupp avseende prestation i PISA. Vissa hade väldigt låga provpoäng medan andra hade väldigt höga provpoäng, och denna kategori kan sannolikt innehålla såväl elever som upplevde provet som väldigt svårt och som därför verkligen ansträngde sig men helt enkelt inte förmådde prestera bättre, som elever med höga kunskapsnivåer och en hög grad av ansträngning, och därmed en hög provpoäng. Men i genomsnitt har denna grupp alltså något lägre provpoäng än de elever som markerat nio eller åtta i ansträngning.

I tabellen ovan presenteras resultaten för de elever som faktiskt svarat något på ansträngningsskalan. Drygt sju procent av de svenska eleverna svarade inte på frågan om hur mycket de ansträngde sig på PISA-provet. Medelvärdet för matematik för dessa elever är 410 poäng, och liknande medelvärden erhöles för läsförståelse och naturvetenskap. De som inte svarar alls på frågan om hur mycket de ansträngde sig presterar således ganska dåligt i PISA, i genomsnitt någonstans mellan de som angav en etta och de som angav en tvåa i ansträngning. Det är möjligt att denna grupp, de som underlät att besvara frågan i slutet på provhäftet, kanske hör till de mindre provmotiverade eleverna. En annan förklaring skulle kunna vara att denna grupp är svagpresterande elever som inte hann besvara alla frågor i provhäftet, och därmed inte hann fram till frågan om hur mycket de ansträngde sig på provet.

Att en klar majoritet av eleverna markerar en 7:a eller högre på ansträngnings-skalan kan inte sägas tyda på en allmän brist på självskattad ansträngning bland de svenska eleverna. Samtidigt är det ändå en andel elever som rapporterar en låg grad av ansträngning, och de som rapporterar en låg grad av ansträngning har ett sämre resultat i PISA. Ansträngningsskalan och provmotivationsskalan förmedlar liknande bilder: en majoritet av de svenska eleverna rapporterar en rimlig (se Butler, 2008) grad av motivation och ansträngning, men mellan en fjärdedel och en tredjedel av eleverna uppger en ganska låg grad av motivation och ansträngning. Liksom för provmotivationsskalan är det lite svårt att uttala sig om resultaten på ansträngningsskalan representerar höga eller låga skattningar om man inte har något att jämföra med. En jämförelse som kan göras är med hur eleverna svarat på den andra frågan i ansträngningstermometern: hur mycket de skulle ha ansträngt sig på PISA-provet om det hade räknats till betyget. En sådan jämförelse kan ge en indikation på hur eleverna värderar PISA-provet jämfört med hur de värderar betygsgrundande prov. Medelvärdet på frågan om hur mycket eleverna hade ansträngt sig på PISA-provet om det hade räknats till betyget är för Sverige 9,51 på den 10-gradiga skalan. Svenska elever rapporterar alltså att de skulle ha ansträngt sig mycket om provet hade räknats till betyget. Differensen mellan skattad ansträngning på PISA-provet och uppskattad ansträngning om provet hade räknats till betyget blir 2,47 skalsteg. De svenska eleverna uppger i genomsnitt således att de skulle ha ansträngt sig ”ett par snäpp till” om provet hade räknats till betyget. Fördelningar för differensvariabeln för de svenska eleverna, liksom genomsnittlig matematikprestation för respektive värde på differensvariabeln redovisas i Bilaga 1, figur D1 respektive D2.

Ovan har det redovisats beskrivande resultat för svenska elevers skattade ansträngning och motivation i PISA 2012. Jämförelser med andra länder som deltog i PISA och jämförelser över tid kan sätta in ovanstående resultat i en större kontext. I avsnittet som följer analyseras eventuella skillnader mellan länder när det gäller skattad ansträngning på PISA-provet.

Finns det skillnader mellan länder avseende skattad ansträngning på ansträngningstermometern i PISA 2012?

I figur 3 redovisas medelvärden för skattad ansträngning i PISA 2012 för alla länder som deltog i PISA, i figur 4 redovisas medelvärden på landsnivå för uppskattad ansträngning om PISA-provet hade räknats till betyget, och i figur 5 redovisas differensen mellan dessa båda skattningar.

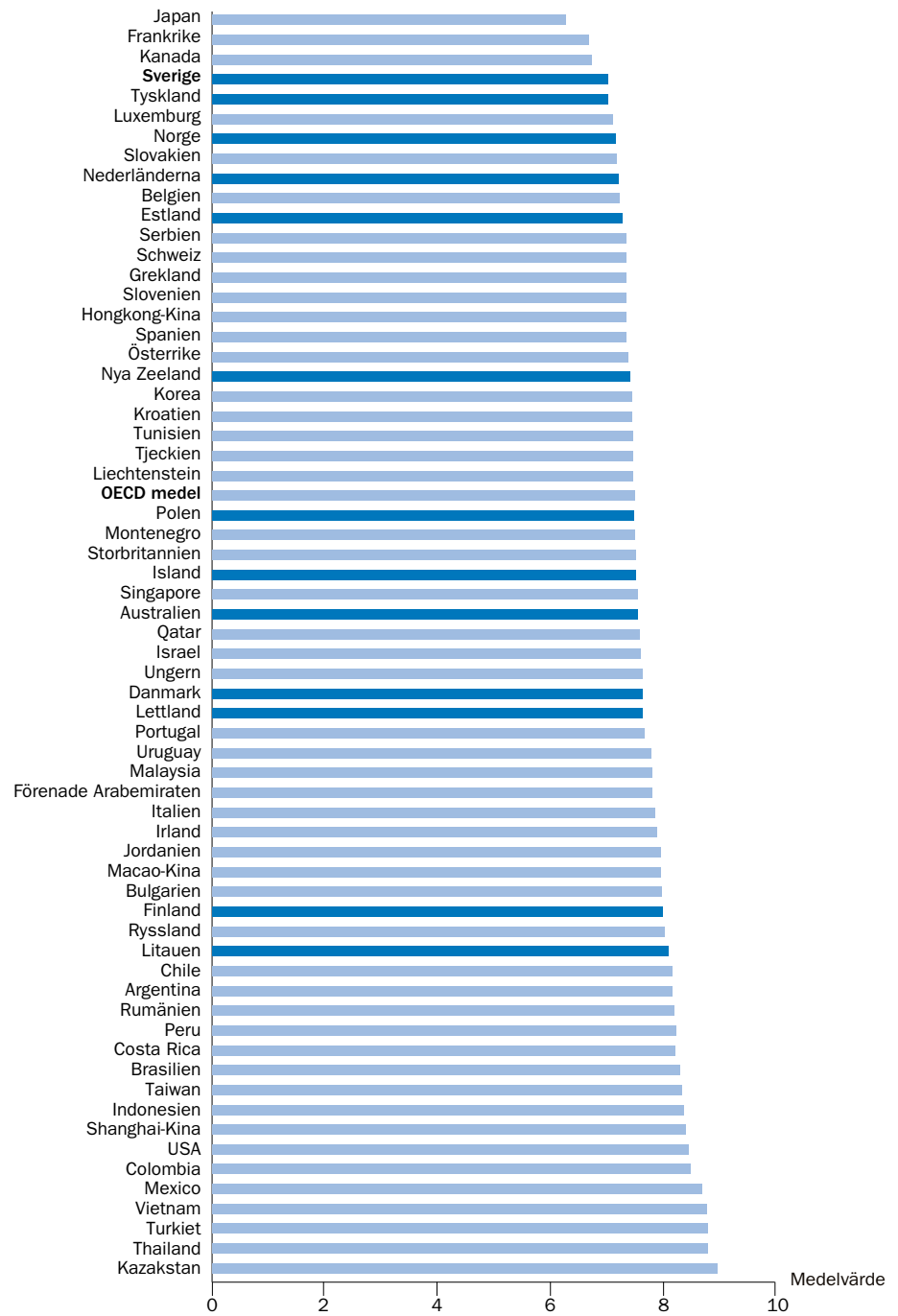
Som framgår av dessa tre figurer så är Sverige ett av de länder som rapporterar den lägsta graden av ansträngning i PISA 2012, den högsta graden av ansträngning om provet hade räknats till betyget, liksom den största differensen mellan dessa två skattningar.

Skillnaderna mellan länder i skattad ansträngning är dock ganska små i de flesta fall. Även om skillnaderna mellan Sverige och de andra länderna i många fall är statistiskt signifikanta är de, om man tittar på storleksskillnader med hjälp av mått på effektstorlek (standardiserade medelvärdeskillnader), ganska små. Majoriteten av länderna har ett medelvärde på skattad ansträngning i PISA 2012 som ligger mellan 7 och 8 på den tiogradiga skalan (se figur 3). De länder som rapporterar jämförelsevis mycket hög grad av ansträngning (högre än 8,50 i

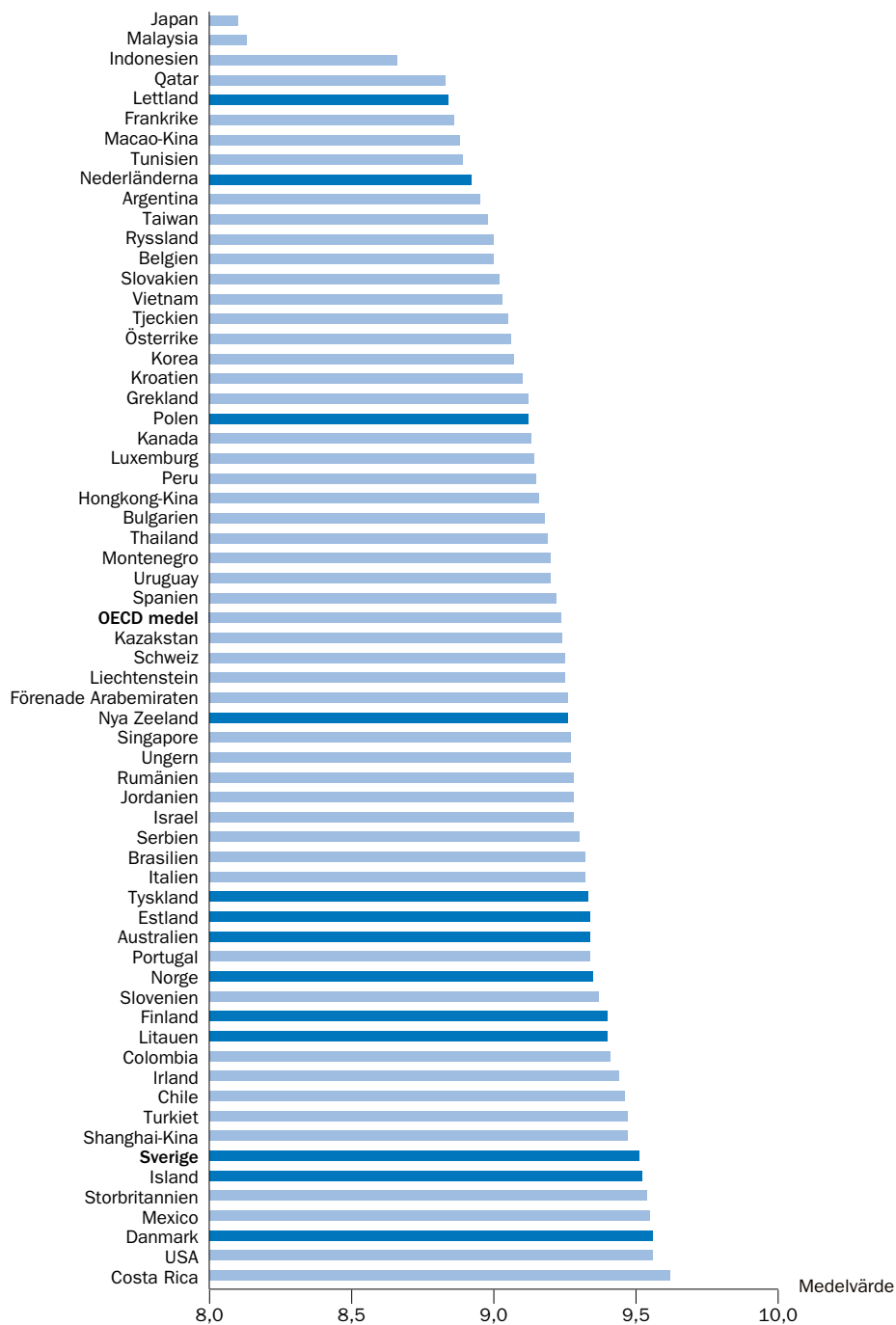
genomsnitt) är alla lågpresterande länder, liksom flera av de länder som rapporterar en medelansträngning på mer än 8,0. I den senare kategorin finns dock också exempelvis Shanghai och Taipei, som är bäst presterande respektive bland de bäst presterande i PISA 2012. Det land som rapporterar den jämförelsevis lägsta ansträngningen, Japan, är ett av de högst presterande länderna. De japanska eleverna ligger också lägst när det gäller skattning av hur mycket de skulle ha ansträngt sig om provet hade räknats till betyget, så det är möjligt att japanska elever i allmänhet tenderar att ge låga skattningar på en skala som den aktuella.

Man kan möjligen, med vissa undantag, se ett mönster som kan vara kulturellt betingat eller relaterat till typ av utbildningssystem/utbildningsnivå, där elever i vissa kulturer behandlar svarsskalor av denna typ på ett sätt (ger lägre, kanske mer realistiska, skattningar) medan elever i andra kulturer eller elever i utvecklingsländer behandlar denna typ av skalor på ett annat sätt (ger högre, kanske socialt önskvärda skattningar). Tidigare forskning har visat att elever från ekonomiskt mindre utvecklade länder tenderar att svara såsom de tror att man borde svara snarare än vad de egentligen tycker (King, Murray, Salomon, & Tandon, 2004). I andra länder kan det vara oacceptabelt att säga att man struntade i att anstränga sig på ett prov. Såsom påpekas av Butler (2008) kan det också vara så att i väldigt lässvaga länder kan innebörden i frågorna i ansträngningstermometern vara svår att förstå, och därmed att svara på. Det förutsätts också att eleven förstår innebörden av en termometer, och/eller principen för skattning på en skala från 1 till 10. Även om det kan tyckas självklart för vissa behöver det inte vara det för alla. Det kan ju också vara så att väldigt högpresterande elever i vissa länder uppfattar provet som lätt och inte behöver anstränga sig särskilt mycket för att ändå få ett bra resultat, medan lågpresterande elever uppfattar provet som mycket svårt och att det därför krävde mycket ansträngning, utan att det resulterar i en hög provpoäng. Med det sagt är det förstås ändå av intresse att se hur elever i andra länder svarade på frågan om ansträngning, och detta beskrivs i figurerna nedan.

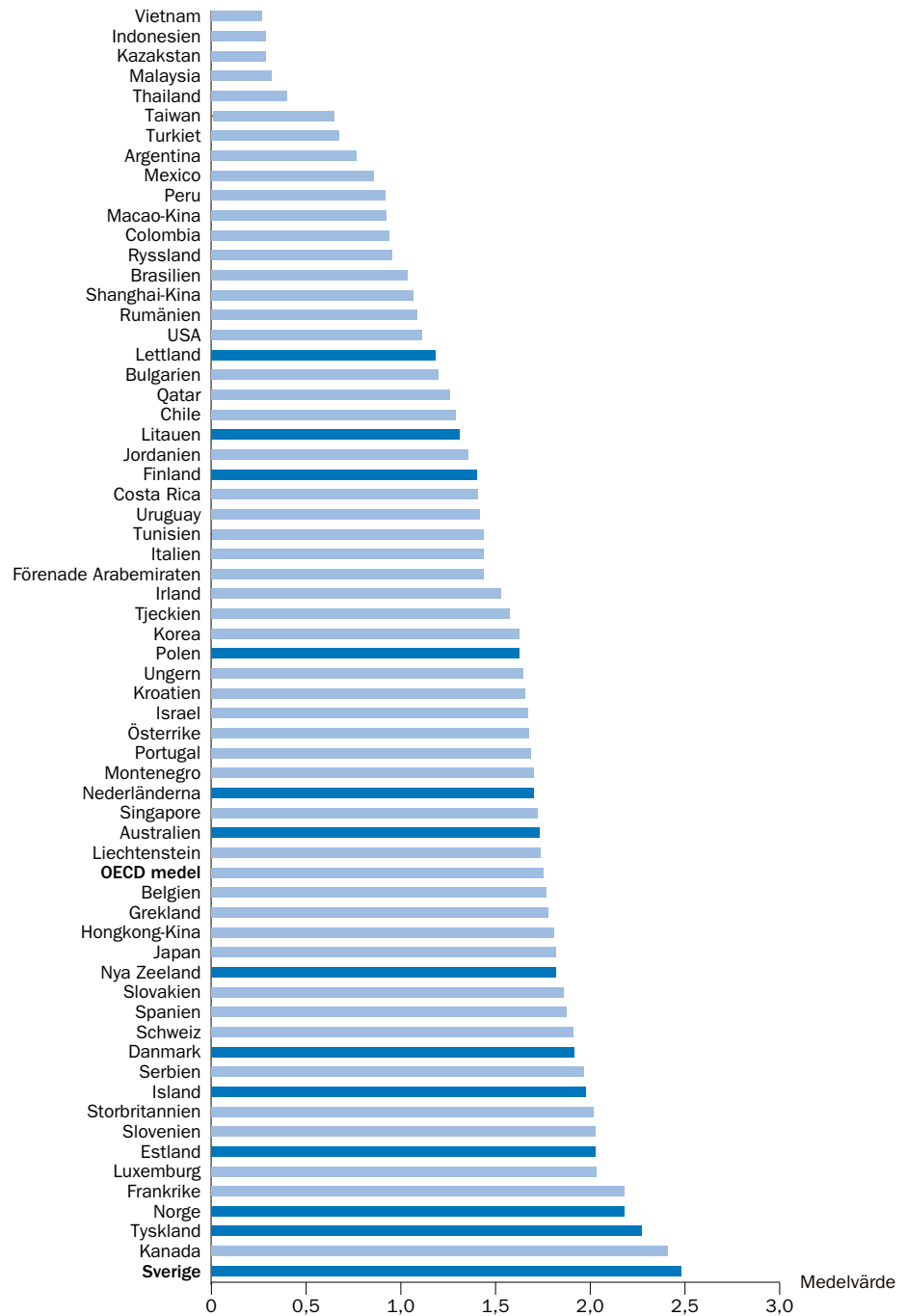
Figur 3. Medelvärde per land för skattad ansträngning i PISA 2012. Sverige har, tillsammans med andra länder som ingår i urvalet i denna studie, mörkblå staplar.



Figur 4. Medelvärde per land för uppskattad ansträngning om provresultatet i PISA hade räknats till betyget. Sverige har, tillsammans med andra länder som ingår i urvalet i denna studie, mörkblå staplar.



Figur 5. Skillnad mellan genomsnittlig skattad ansträngning på PISA och genomsnittlig uppskattad ansträngning om provresultatet hade räknats till betyget, per deltagande land. Sverige har, tillsammans med andra länder som ingår i urvalet i denna studie, mörkblå staplar.



I figurerna presenteras medelvärdet för OECD-länderna, då detta är gängse standard vid rapportering av PISA-resultat. Medelvärdet för skattad ansträngning på PISA 2012 var för OECD-länderna 7,49, medan det internationella medelvärdet (samtliga länder, OECD- såväl som icke OECD-länder) är 7,75. Ett något lägre medelvärde för OECD-länderna, således. För skattningen av ansträngning om PISA-provet hade räknats till betyget är OECD-medelvärdet 9,24 medan det internationella medelvärdet är 9,20, i princip samma. Det internationella medelvärdet för differensvariabeln är 1,44 medan OECD-medelvärdet för samma variabel är 1,75. Eleverna i OECD-länderna gör således i genomsnitt större skillnad mellan PISA-provet och betygsgrundande prov.

Det är möjligt, och troligt, att ansträngningstermometern inte fungerar riktigt likadant och inte mäter riktigt samma sak överallt, varför man bör vara försiktig vid jämförelser mellan länder. Det gör också att det blir ganska missvisande att exempelvis jämföra Sveriges resultat med ett internationellt medelvärde, där dock jämförelser avseende differensvariabeln kan anses vara något mindre påverkade av svarsstilar etc. (se Butler, 2008). Jämförelser över alla 65 länder är över huvud taget ganska svårtolkade. I denna rapport har därför ett antal länder, som ligger geografiskt och/eller kulturellt nära Sverige eller på annat sätt brukar vara intressanta för oss att jämföra oss med, valts ut för mer ingående jämförande analyser. Dessa är de nordiska länderna (Sverige, Norge, Finland, Island, Danmark) och andra länder runt Östersjön (Estland, Lettland, Litauen, Polen, Tyskland), samt Nederländerna, Australien och Nya Zeeland.

I tabell 3 nedan redovisas genomsnittliga skattningar på ansträngningstermometern för respektive land i urvalet, hur mycket landets medelvärde skiljer sig från Sveriges uttryckt i standardiserade medelvärdeskillnader (d)⁴; uppskattad ansträngning om provet räknats till betyget; liksom för differensen mellan dessa två skattningar, för urvalet av länder. Dessutom presenteras landets resultat i matematik i PISA 2012.

4 Den standardiserade medelvärdeskillnaden har beräknats genom formeln för Cohen's d , och Sverige har jämförts parvis med de andra respektive länderna på följande sätt: $d = \text{Sveriges medelvärde } (M_1) \text{ minus *landets* medelvärde } (M_2) \text{ dividerat med den poolade standardavvikelsen för de båda ländernas medelvärden: } \sqrt{(\sigma^2 + \sigma^2) / 2}$

Tabell 3. Genomsnittlig skattad ansträngning i PISA 2012, skillnad (*d*) jämfört med Sverige, genomsnittlig skattad ansträngning om provet räknats till betyget, differensen mellan skattad ansträngning och ansträngning om betyg, skillnad (*d*) jämfört med Sverige och matematikresultat för landet i PISA. Urval av länder.

Land	Skattad ansträngning i PISA	<i>d</i>	Skattad ansträngning om betyg	Differens/Relativ ansträngning	<i>d</i>	Matematikresultat
Sverige	7.03		9.51	2.47		478
Tyskland	7.05	.01	9.33	2.27	.10	514
Norge	7.16	.06	9.35	2.18	.14	489
Nederländerna	7.22	.10	8.92	1.70	.40	523
Estland	7.31	.14	9.34	2.03	.22	521
Nya Zeeland	7.44	.21	9.26	1.82	.33	500
Polen	7.50	.23	9.12	1.62	.42	518
Island	7.55	.25	9.52	1.97	.23	493
Australien	7.56	.27	9.34	1.78	.37	504
Danmark	7.64	.32	9.56	1.92	.28	500
Lettland	7.66	.35	8.84	1.18	.67	491
Finland	8.00	.52	9.40	1.40	.58	519
Litauen	8.10	.59	9.40	1.30	.62	479

De svenska eleverna rapporterar lägre skattad ansträngning på PISA än de andra länderna i urvalet, och en större differens mellan skattad ansträngning på PISA och uppskattad ansträngning om det hade räknats till betyget. Just den förhållandevis stora differensen mellan de två skattningarna skulle möjligen kunna tolkas som ett resultat av den nationella prov-period de svenska eleverna kan ha befunnit sig i: att de vid samma tid gör flera prov som kan vara viktigare för betygen än PISA, och att de därför gör en större skillnad mellan de olika typerna av prov. Å andra sidan har de svenska eleverna även i tidigare PISA-test gjort jämförelsevis stor skillnad mellan PISA och betygsgrundande prov, mer om detta i kommande avsnitt.

Tyskland, Norge, Estland och Nederländerna ligger tillsammans med Sverige ganska lågt på skattad ansträngning i PISA. Island, Danmark, Nya Zeeland, Australien, Polen och Lettland ligger ungefär ett halvt skalsteg över Sverige i genomsnitt, medan Finland och Litauen i genomsnitt ligger ungefär ett skalsteg över Sverige. För de nordiska länderna gäller att i synnerhet de finska eleverna, men också de danska och de isländska, rapporterar en högre grad av ansträngning på PISA än vad de svenska eleverna gör. Alla de nordiska länderna ligger högt i skattad ansträngning om provresultatet hade räknats till betyget, medan Nederländerna och Lettland ligger lite lägre.

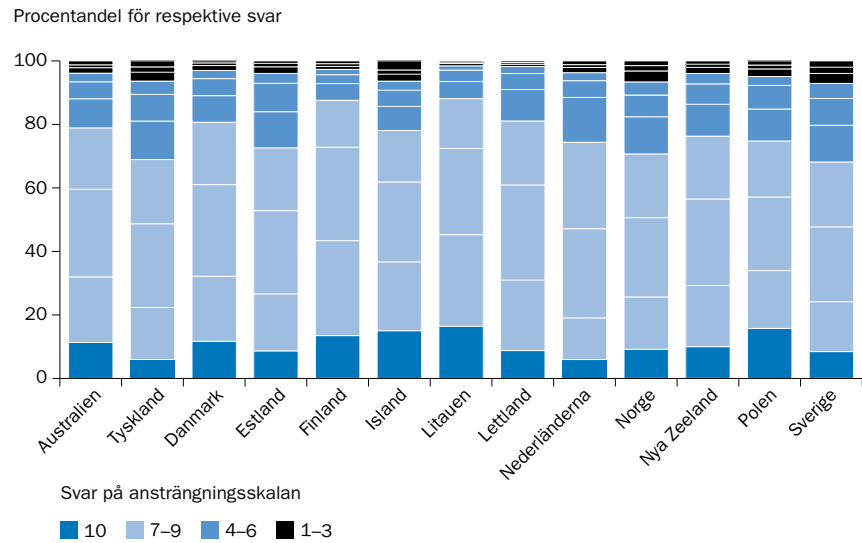
Måttet på effektstorlek är som nämnts Cohen's *d*. Värdet på Cohen's *d* runt 0,20 brukar betraktas som en liten skillnad i praktiken, värdet runt 0,50 brukar betraktas som en medelstor skillnad och värdet större än 0,80 brukar betraktas som en stor skillnad. Som Hattie (2009) har konstaterat kan ett värde som är litet i ett sammanhang betraktas som stort i ett annat, beroende på vad som studeras. Med tanke på att den totala variationen mellan länder är ganska begränsad för skattad ansträngning (de flesta har ett medelvärde mellan 7 och 8

och en standardavvikelse runt 2) kan ett helt skalstegs skillnad, som i jämförelsen mellan Sverige och Finland nog betraktas som ganska stor, även om d-måttet indikerar att den är medelstor.

För differensvariabeln är d-värdena större än för skattad ansträngning när Sverige jämförs med i synnerhet de icke-nordiska länderna. Detta ger ytterligare stöd åt resultatet att svenska elever tycks göra större skillnad på PISA-provet och betygsgrundande prov än elever i flera andra länder.

Fördelning av svar för de olika skalstegen på ansträngningstermometern för länderna i urvalet presenteras i figur 6. Här redovisas hur stor andel av eleverna som angett respektive värde på ansträngningstermometern (min ansträngning = 1, max ansträngning = 10). Eftersom svaren är snett fördelade (få elever skattar en väldigt låg ansträngning) är för överskådlighetens skull värden mellan 1–3 svartfärgade medan värden 4–6, 7–9 respektive 10 har olika nyanser av blått. I figur 7 redovisas sedan det genomsnittliga matematikresultatet för respektive värde på ansträngningstermometern, för respektive land.

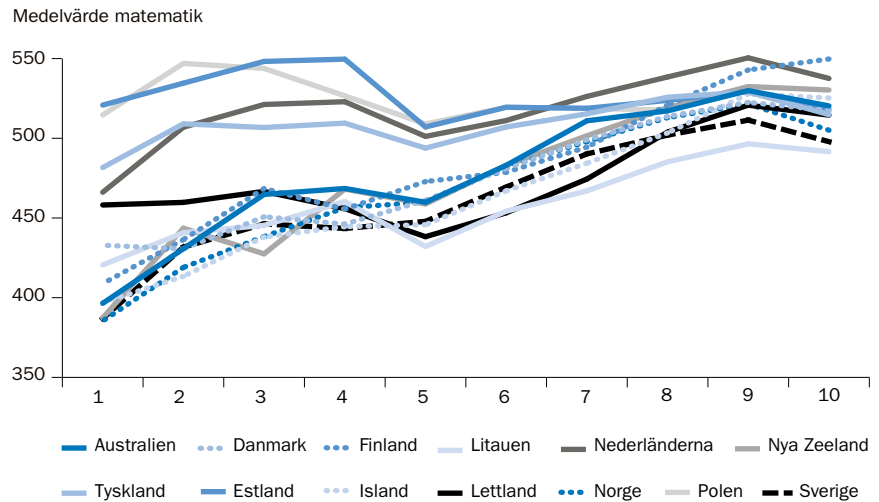
Figur 6. Fördelning av svar på ansträngningstermometern för ett urval av länder.



Även om Sverige i genomsnitt rapporterar lägre ansträngning än övriga länder i urvalet avviker det svenska skattningsmönstret inte dramatiskt från de andra ländernas. När det gäller andelen som inte svarat på frågan så varierar det från endast cirka 3 procent för Finland och Island, till ungefär 10 procent för Australien och Nederländerna. Danmark, Sverige, Norge och Polen ligger mellan sex och åtta procent.

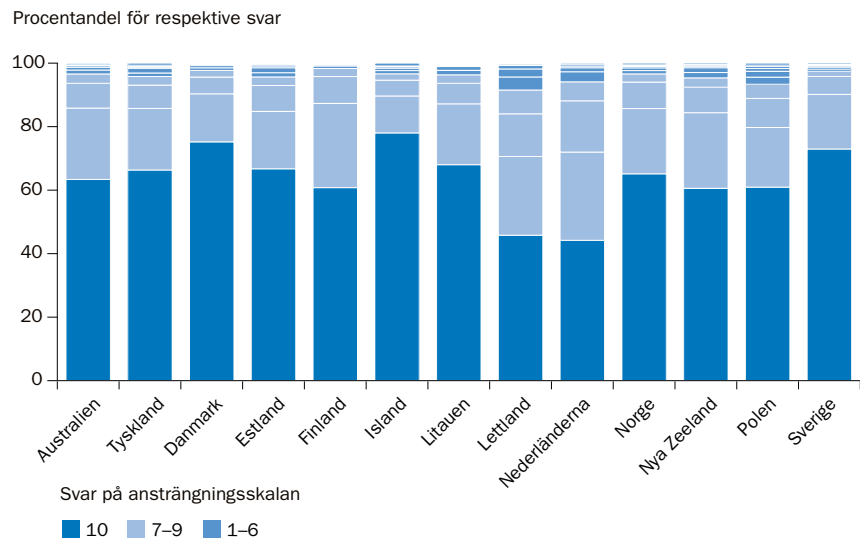
I figur 7 ser det för Polen och Estland (linjerna ”högst upp” i figuren) inte ut som att det finns någon samvariation mellan hur mycket eleverna rapporterar att de anstränger sig och hur bra de presterar i PISA. Elever som rapporterar låg grad av ansträngning presterar minst lika bra som de som rapporterar en hög grad av ansträngning. För Tyskland och Nederländerna ser vi liknande mönster, medan det för de nordiska länderna, liksom för Australien, Nya Zeeland och Lettland ganska tydligt är så att en högre grad av skattad ansträngning hänger ihop med en högre prestation – upp till en skattad nia på ansträngningstermometern. För alla länder utom Finland är det en lägre genomsnittlig prestation för gruppen som angett en tia på ansträngningstermometern.

Figur 7. Matematikresultat per svar på ansträngningsskalan för urvalet av länder i PISA 2012. Sverige har en svart streckad linje, övriga nordiska länder har prickade linjer och de andra länderna i urvalet har heldragna linjer.



I figur 8 redovisas fördelning av svar på frågan om hur mycket man skulle ha ansträngt sig på PISA-provet om det räknats till betyget. Staplarna i figuren representerar samma urval av länder som ovan. Eftersom svaren är väldigt snett fördelade är för överskådlighetens skull värden mellan 1–6, 7–9 respektive 10 givna olika nyanser av blått. I alla länder utom Lettland och Nederländerna är det överlägset vanligast att eleverna markerat en 10:a på hur mycket de skulle ha ansträngt sig om PISA-provet räknats till betyget. I alla länder är det också en ganska stor andel som markerat en 9:a, medan övriga skalsteg har ganska små andelar.

Figur 8. Fördelning av svar på frågan om hur mycket man skulle ha ansträngt sig om PISA-provet räknats till betyget. Ett urval av länder.



Hittills har det redovisats resultat för skattad provmotivation och skattad ansträngning för Sverige i PISA 2012. Det har vidare presenterats resultat för ansträngningstermometern internationellt, och för ett urval av länder i PISA 2012. Ansträngningstermometern har dock använts även i tidigare PISA-cykler, så det är möjligt att studera eventuella förändringar över tid. Detta görs i det avsnitt som följer.

Har det skett förändringar över tid för ansträngningstermometern, i Sverige och jämfört med andra länder?

I detta avsnitt tittar vi först på förändringar över tid i PISA vad gäller svenska elevers rapporterade ansträngning på PISA-provet. En stor minskning i ansträngning skulle kunna vara en förklaring till Sveriges sjunkande resultat, om skattad ansträngning skulle ha en effekt på prestation som inte kan förklaras av andra variabler. Därefter tittar vi på urvalet av länder och ser om det skett någon förändring över tid i dessa länder. Om andra länder skulle ha stabila ansträngningsskattningar över tid medan Sveriges skulle sjunka, skulle detta kunna förklara Sveriges fallande ranking – återigen förutsatt att skattad ansträngning har effekt på prestation i Sverige och även i andra länder, samt att skalan för ansträngning kan antas vara konstant över tid.

Jämfört med PISA 2003 och 2006 har de svenska elevernas rapporterade ansträngning sjunkit något. Sverige hade i PISA 2003 ett medelvärde på ansträngningstermometern på 7,38. I PISA 2006 var medelvärdet 7,37, och i PISA 2012 var det 7,03. En minskning med 0,35 skalsteg på den tiogradiga skalan från 2003 till 2012 (med matematik som huvudämne båda dessa år), och 0,34 skalsteg från 2006, medan det i princip inte var någon skillnad mellan 2003 och 2006 (då matematik respektive naturvetenskap var huvudämne). 2009 användes inte ansträngningstermometern. Nedgången från föregående år till 2012 är statistiskt signifikant enligt t-test, men effektstorleken är ganska liten ($d = 0,18$ vilket brukar räknas som en ganska svag effekt). Skattad ansträngning om PISA-provet hade räknats till betyget ligger tämligen stilla mellan åren (9,46, 9,57, respektive 9,51 på den tiogradiga skalan), medan differensen mellan de två skattningarna följaktligen har ökat till 2012.

Fördelningarna av svar på frågan om hur mycket eleven ansträngde sig på PISA har förändrat sig något från 2003 till 2012, som en följd av att eleverna rapporterar något lägre ansträngning, men mönstret är det samma. En mindre andel elever markerar exempelvis en tia i ansträngning 2012 jämfört med 2003 (8,4 % jämfört med 11,6 %). Såväl 2012 som 2003 (och 2006) presterar dock eleverna som markerat en tia sämre än de som markerat en nia. Noterbart är att andelen elever som inte svarat alls på frågan om hur mycket de ansträngde sig har ökat mellan 2003/2006 och 2012. I PISA 2003, liksom i PISA 2006, var det drygt tre procent av eleverna som inte svarade på frågan, i PISA 2012 var motsvarande andel drygt sju procent. Liksom 2012 presterade de elever som lämnat frågan blank i PISA 2003 (100 stycken) i paritet med dem som angivit en etta eller en tvåa i ansträngning (412 poäng i genomsnitt) medan de som hörde till kategorin ”övriga ogiltiga svar” (53 elever) hade ett genomsnittligt resultat på 487 poäng. I urvalet av länder redovisar även Australien och Nederländerna en ökad andel elever som inte besvarat frågan om ansträngning,

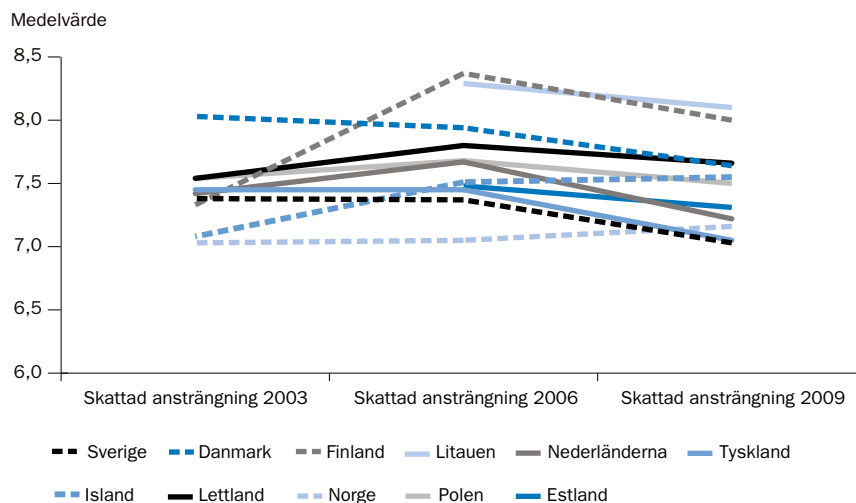
medan Tyskland och Lettland redovisar en minskad andel. Övriga länder har ungefär lika stort bortfall nu som tidigare. Jämfört med exempelvis Indonesien, som har 53 procents bortfall på ansträngningstermometern, är siffrorna ganska små för Sverige och det övriga urvalet av länder, icke desto mindre kan det vara värt att uppmärksamma att andelen som inte svarar på frågan har ökat i Sverige sedan PISA 2006, då detta möjligen säger något om eleveras motivation. Som nämnts tidigare, om man inte är motiverad att anstränga sig och göra sitt bästa på provet kanske man inte heller orkar bemöda sig med att svara på frågor om hur mycket man ansträngde sig. Alternativt kan det större bortfallet förklaras av en större andel svagpresterande elever som inte hinner svara på alla frågor i provhäftet, och därför inte hinner fram till frågan om ansträngning. Ansträngningstermometern ligger allra sist i provhäftet.

I PISA 2012 var det som nämnts tre länder som rapporterar en lägre nivå av ansträngning än Sverige. I PISA 2003 var det åtta länder (om man utgår från de länder som även deltog i PISA 2012) som rapporterar en lägre grad av ansträngning (Belgien, Finland, Luxemburg, Österrike, Island, Norge, Frankrike, Japan). Utgår man från alla länder, även de som sedan inte deltog i PISA 2012, var det 11 länder som rapporterade en lägre grad av ansträngning än Sverige i PISA 2003. I PISA 2006 var det sju länder som rapporterade en lägre grad av ansträngning. Sverige har alltså ”tappat ett antal placeringar” sett till skattad ansträngning, även om minskningen i absoluta termer inte är särskilt stor. Sverige rapporterade förvisso den största differensen mellan skattad och uppskattad ansträngning i PISA 2012, men 2006 hade Sverige den näst största differensen (efter Norge), så svenska elever har redan tidigare ”gjort skillnad på prov och prov”, och i alla PISA-omgångar där ansträngningstermometern varit med har Sverige legat ganska lågt i genomsnittlig rapporterad ansträngning.

Samtidigt är Sverige det land som visar den största resultatförsämringen av alla länder sedan 2003, -31 poäng för såväl matematik som läsförståelse, -21 poäng för naturvetenskap (resultaten är mer osäkra för läsförståelse och naturvetenskap). Från 2006 har Sveriges medelpoäng sjunkit med 24 poäng för matematik och läsförståelse, och 18 poäng för naturvetenskap. Även för andra länder kan man se lägre skattningar på ansträngningstermometern kombinerat med lägre prestation (Danmark är ett exempel), men bilden är, förstås, komplex. På Island rapporterar eleverna högre ansträngning 2012 än de gjorde 2003 och 2006 medan matematikresultaten gått ner ganska kraftigt. I Finland rapporterar eleverna högre ansträngning 2012 än de gjorde 2003 (dock lägre än de gjorde 2006) och även här har resultaten gått ner. I Tyskland å andra sidan, rapporterar eleverna lägre ansträngning 2012 jämfört med 2003 medan resultaten har gått upp. I Australien och Nya Zeeland har skattning av ansträngning i genomsnitt inte förändrats men prestationen har gått ner. Ser man till alla länder som deltog både i PISA 2003 och PISA 2012 ses heller inga tydliga trender. Internationellt är typvärdet, det vanligaste värdet, på skattad ansträngning på PISA-provet en åtta såväl 2003, 2006 som 2012. Typvärdet på uppskattad ansträngning om resultaten räknades till betyget är 10 såväl 2003, 2006 som 2012.

Hur medelvärdet för skattad ansträngning på PISA förändrats över tid för de länder som Sverige primärt jämförs med i denna rapport redovisas i Figur 9.

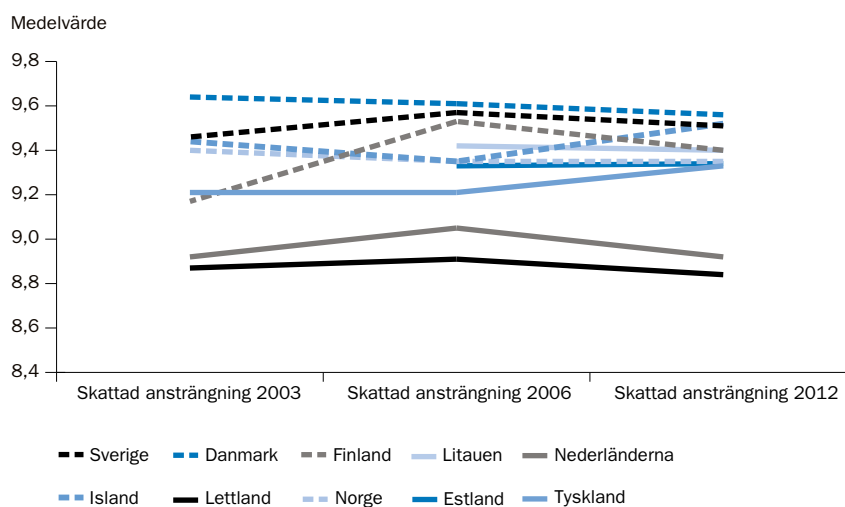
Figur 9. Trender i skattning av ansträngning på PISA-provet för Sverige + urval av länder. De nordiska länderna har streckade linjer i figuren medan de övriga länderna i urvalet har heldragna linjer.



När det gäller skattad ansträngning visar Island en uppgång över åren, Danmark visar en nedgång. Sverige och Tyskland ligger ganska stilla mellan 2003 och 2006 och går sedan ned till 2012. Norge ligger ganska stilla, men visar på en liten uppgång till 2012. Finland visar upp ett lite udda mönster, med en stor uppgång mellan 2003 och 2006 och sedan en nedgång till 2012. Estland och Litauen var inte med i PISA 2003 varför deras linjer är kortare.

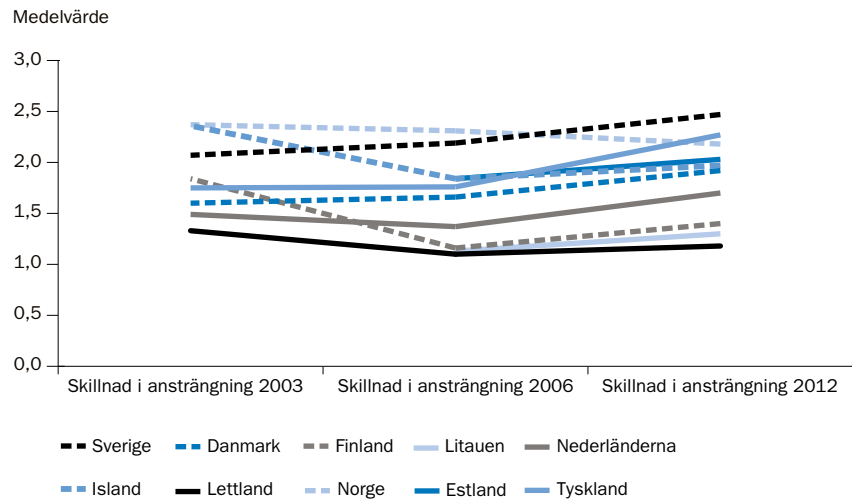
För uppskattad ansträngning om provet hade räknats till betyget ligger skattningarna ganska stilla mellan åren (se Figur 10 nedan, notera att skalan är ganska hoptryckt, det som kan se ut som tydliga upp- eller nedgångar representerar ganska små skillnader). Skillnader i differenser mellan de båda skattningarna (Figur 11) beror således främst på skillnader i skattad ansträngning på PISA-provet.

Figur 10. Trender i skattad ansträngning om PISA-provet räknats till betyget.



Lettland och Nederländerna utmärker sig en aning då de i medeltal anger en lägre grad av ansträngning om provet varit betygsgrundande jämfört med de andra länderna i urvalet. Finland uppvisar ett liknande mönster som i föregående figur, en ganska stor uppgång mellan 2003 och 2006 och sedan en nedgång till 2012. I övrigt har inga stora förändringar skett om länderna jämförs med sig själva.

Figur 11. Trender i differens mellan skattad ansträngning på PISA och uppskattad ansträngning om provet hade räknats till betyget.



Detta avsnitt fokuserar på förändringar över tid när det gäller skattad ansträngning på PISA, och uppskattad ansträngning om PISA-provet hade räknats till betyget. Vi har kunnat se att de svenska elevernas rapporterade ansträngning i PISA har gått ner något om man jämför 2012 med 2003 och 2006. Den uppskattade ansträngningen om provet hade räknats till betyget ligger ganska stilla över åren, och differensen mellan de två skattningarna har ökat.

I det som följer ska en central fråga behandlas: hur ser relationen mellan skattad ansträngning och prestation ut, och kan den lägre graden av skattad ansträngning förklara den försämrade prestationen i PISA? Detta avhandlas i de två följande avsnitten, och vi börjar med att titta på samband i form av korrelationer mellan skattad ansträngning och prestation, för Sverige och i ett internationellt perspektiv.

Finns det ett samband mellan rapporterad provmotivation och provprestation?

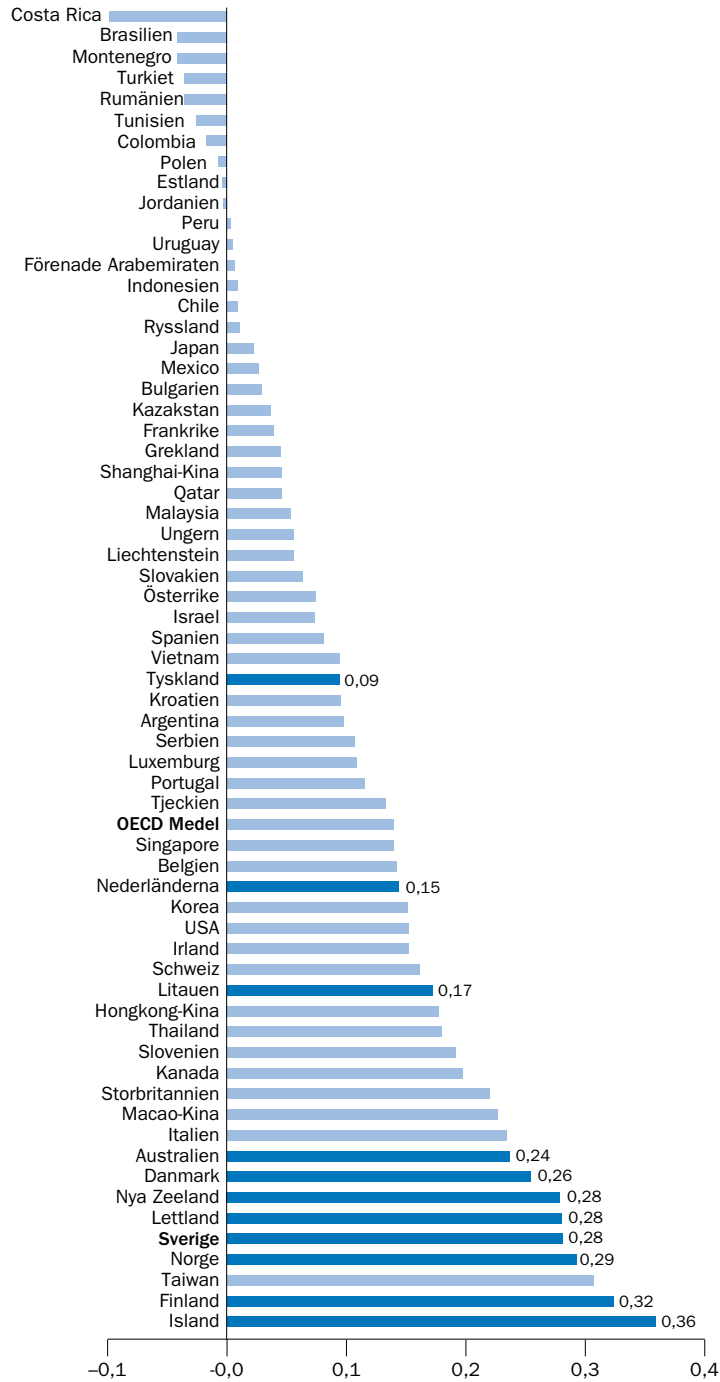
I absoluta tal är förändringen i skattad ansträngning mellan de olika PISA-omgångarna inte så stor för Sveriges del (eller för övriga länderna som Sverige jämförs med här), men det är en nedgång, och om det skulle vara så att ansträngning har starka samband med prestation skulle denna nedgång kunna vara viktig att ta hänsyn till vid tolkning av förändringar i resultatet. I det som följer närmast redovisas analyser av samband (korrelationer) mellan skattad ansträngning och prestation.

Vid en jämförelse som inkluderar alla länder som deltog i PISA 2012 (se Figur 12) blir det tydligt att sambanden ser olika ut för olika länder. I några länder finns i princip inget samband mellan rapporterad ansträngning och prestation i PISA, och i några länder är sambandet till och med negativt. I de flesta länder är sambandet positivt, men styrkan i sambandet varierar mellan länder, möjligen som en effekt av att frågan fungerar på lite olika sätt i olika länder. I Figur 11 redovisas korrelationer (uttryckt som "r") mellan skattad ansträngning på PISA och prestation på PISAs matematikdel år 2012. Sambanden är likartade för läsförståelse och naturvetenskap. Det genomsnittliga sambandet mellan skattad ansträngning och prestation är för OECD-länderna $r = .14$, något högre än det internationella medelvärdet som är $r = .10$. Motsvarande OECD-medelvärden avseende samband med läsning och naturvetenskap är $r = .17$ respektive $r = .15$.

Av alla deltagande länder redovisar Island det starkaste sambandet (uttryckt som "r", värdet ovanför respektive land i urvalet), följt av Finland. I Sverige är sambandet mellan skattad ansträngning och matematikprestation $r = .28$, och samma eller liknande styrka på sambandet återfinns i Norge, Danmark, Lettland, Nya Zeeland och Australien. I Tyskland, Holland och Litauen är sambandet positivt men ganska svagt. I Estland och Polen finns det inget samband mellan skattad ansträngning och prestation. För Sverige har sambandet mellan ansträngning och prestation varit ganska stabilt över åren och för de tre olika ämnena. Korrelationskoefficienten varierar från $.25$ (Naturvetenskap år 2003) till $.32$ (Läsförståelse år 2006).

Sambandet mellan skattad ansträngning och prestation tycks vara extra starkt i de nordiska länderna. Av de åtta länderna med det starkaste sambandet är fem nordiska. Detta kan dock inte tolkas som att det bara är i de nordiska länderna som ansträngning har någon betydelse, men kanske säger det något om hur man i olika länder tolkar och därmed svarar på ansträngningsfrågan, hur svars-skalan tolkas, att det är ok i vissa länder att säga att man inte anstränger sig, att eleverna gör mer skillnad mellan prov och prov i vissa länder – eller så har det andra anledningar som vi inte kan veta något om utifrån tillgängliga data.

Figur 12. Samband mellan skattad ansträngning och matematikprestation i PISA 2012, alla länder. Sverige har, tillsammans med andra länder som ingår i urvalet i denna studie, mörkblå staplar. Sambandets styrka (r) för länderna i urvalet redovisas ovanför respektive stapel. Ju högre värde desto starkare samband, värden mellan -1 – +1 är möjliga.



I majoriteten av länder finns det alltså positiva samband mellan skattad ansträngning och prestation i PISA, men i inget land är sambandet jättestarkt. Korrelationer i storleken $.20 - .30$ kan räknas som ganska modesta, dock inte försumbara, samband. Av Figur 12 (korrelationsfiguren) framgår att det blir ganska missvisande att prata om samband mellan skattad ansträngning och prestation på aggregerad nivå, eller att jämföra ett enskilt land med ett internationellt medelvärde, då det säger ganska lite om hur det ser ut i andra enskilda länder. Det internationella medelvärdet för sambandet mellan ansträngning och prestation är $.10$, men skulle detta användas för att exempelvis hävda att ansträngning har ett positivt men endast svagt samband med prestation i PISA ger det inte någon särskilt bra bild av hur det ser ut i de nordiska länderna, eller i Costa Rica, Brasilien eller Montenegro.

Det finns heller ingen tydlig koppling på landsnivå mellan nivå på skattad ansträngning och samband mellan ansträngning och prestation. Finland och Taiwan är länder som rapporterar förhållandevis höga nivåer av ansträngning, medan Sverige och Norge rapporterar förhållandevis låga nivåer av ansträngning, och i alla dessa länder är sambandet mellan skattad ansträngning och prestation jämförelsevis starkt. Exempel på det motsatta finns också. I Ryssland och Shanghai rapporterar eleverna hög grad av ansträngning, medan eleverna i Japan rapporterar den lägsta graden av ansträngning, och i dessa länder är sambandet mellan ansträngning och prestation i princip 0.

I exempelvis Polen och Estland, länder som lyft sig rejält i PISA prestationsmässigt, finns inget samband alls mellan skattad ansträngning och prestation, vare sig nu eller i tidigare PISA-undersökningar (Estland deltog dock inte i PISA 2003).

För att exemplifiera att de varierande sambanden mellan ansträngning och prestation kan bero på att självskattningsskalor fungerar olika i olika länder gjordes motsvarande sambandsanalys som i Figur 12 för den PISA-skala som mäter självuppfattning i matematik ("Jag lär mig saker snabbt i matematik", "Matematik är ett av mina bästa ämnen" etc.) och matematikprestation. Resultaten redovisas i Bilaga 1, Figur A1. Analysen med denna skala visar på liknande sätt att det är ganska stora skillnader mellan länder avseende sambandets styrka, liksom att det i princip är samma länder som uppvisar de starkaste sambanden mellan skattad självuppfattning och prestation som mellan skattad ansträngning och prestation: alla nordiska länder ligger topp åtta i båda figurerna.

När det gäller den svenska provmotivationsskalan (se Tabell 1) så är elevernas skattning på denna skala också positivt relaterad till prestation i PISA ($r = .21$). Även de enskilda frågorna i skalan är positivt korrelerade till prestation. För de två frågorna som handlar om en bra prestation på PISA uppfattas som viktigt är sambandet dock svagt ($r < .10$) då sambanden är kurvilinearare än linjära, medan sambanden är lite starkare för de frågor som handlar om motivation att anstränga sig ($r = .15 - r = .25$).

Sammanfattningsvis finns alltså ett positivt och signifikant samband mellan provmotivation och prestation i Sverige, och sambandet är av samma styrka som setts i ett flertal tidigare studier. I de övriga nordiska länderna liksom i ett flertal av de länder som "liknar oss" och som vi brukar vilja jämföra oss med i de internationella studierna finns också tydliga samband mellan rapporterad ansträngning och prestation, medan det i ett internationellt perspektiv finns exempel

på såväl länder där det inte finns något samband alls, länder där sambandet är mycket svagt, och länder där sambandet är negativt.

I avsnittet som följer närmast presenteras resultat från ett antal analyser där effekten av skattad ansträngning på prestation studerats lite mer i detalj: om effekten är konstant över åren, om andra motivationsvariabler som mäts i PISA är viktigare än ansträngning, och om en lägre grad av ansträngning nu jämfört med tidigare kan förklara den svenska nedgången i PISA

Kan provmotivation förklara nedgången i de svenska resultaten?

Hittills har vi konstaterat att svenska elever rapporterar en jämförelsevis låg grad av skattad ansträngning på PISA-provet år 2012. Vi har också kunnat se att de svenska elevernas rapporterade grad av ansträngning har minskat något sedan 2006, medan det inte var någon skillnad i rapporterad ansträngning mellan 2003 och 2006. Vidare verkar det finnas positiva samband mellan rapporterad ansträngning, provmotivation och provprestation så att en högre rapporterad motivation och ansträngning hänger ihop med högre prestation på provet. Dessa samband har varit ganska konstanta över åren för Sveriges del.

Nedan ska vi först titta lite närmare på hur mycket av elevernas prestation som skulle kunna förklaras av ansträngning/motivation och om de sjunkande resultaten kan bero på att eleverna anstränger sig mindre nu än tidigare.

Men, även om man skulle finna en signifikant effekt av ansträngning på prestation, behöver inte detta innebära att effekten kvarstår när andra variabler läggs in i analysen. Det skulle exempelvis kunna vara så att elevernas rapporterade ansträngning egentligen kan förklaras av hur bra de tycker att de är i matematik, hur uthålliga de är mer generellt när det gäller skolarbete, eller hur intresserade de är av matematik. Skulle det vara så kommer en eventuell effekt av ansträngning att försvinna eller minska när man tar hänsyn till de andra aspekterna. Därför har vi också haft ambitionen att analysera om elevernas rapporterade grad av ansträngning/provmotivation kan förklara något av variationen i matematikresultaten i PISA 2012 utöver det som förklaras av andra variabler, såsom inre och yttre motivation, självuppfattning i matematik, självtillit när det gäller att klara matematiska uppgifter, matematikängslan, socioekonomisk bakgrund etc.

Kan brist på ansträngning vara förklaringen till de svenska elevernas sjunkande resultat?

I regressionsanalysen uppskattades hur mycket av de svenska elevernas prestation i PISA som förklaras av deras (självrapporterade) grad av ansträngning och motivation, och hur mycket av nedgången i prestation i de olika ämnena som skulle kunna förklaras av nedgång i ansträngning. Oberoende variabel i denna analys var rapporterad ansträngning på PISA-provet (ansträngningstermometern). Beroende variabel var resultatet i matematik, naturvetenskap respektive läsförståelse. Samma analys gjordes med provmotivationsskalan, men här finns bara resultat från 2012 eftersom provmotivationsskalan inte använts tidigare. Resultaten från analysen presenteras i Bilaga 1, Tabell B1.

Regressionsanalysens B- och β -koefficienter

En regressionsanalys genererar ett antal värden som indikerar hur stor effekt en eller flera variabler har på andra variabler. Här vill vi veta hur stor effekt ansträngningstermometern (och senare även andra variabler) har på prestation i PISA. Några av dessa värden är B-koefficienten, β -koefficienten och R^2 , andelen förklarad varians. B-koefficienten, eller b-värdet, kan här tolkas som att en ökning med ett steg på ansträngningstermometern ger en ökning i prestation motsvarande b-värdet. β -koefficienten, eller β -värdet, är ett standardiserat mått på effekten. β -värdet är detsamma som korrelationskoefficienten om man bara har en oberoende variabel. R^2 anger hur stor del av variationen i prestation som kan förklaras av skattad ansträngning (eller andra oberoende variabler).

För Sveriges del resulterade analysen med ansträngningstermometern som prediktor av matematikprestation i PISA 2012 i ett b-värde på 12,20. Liknande värden erhöles för de andra ämnena och för övriga år (se Tabell B1 i Bilaga 1). Statistiskt ger alltså en ökning med ett skalsteg på ansträngningstermometern en ökning med 12,20 poäng på PISA-provet, en signifikant effekt. (Exempel: En ökning från 7 till 8 på ansträngningstermometern skulle enligt regressionsanalysen ge 12,20 poäng mer på provet, exempelvis 490 poäng i stället för 478). Detta stämmer också ganska väl överens med de inledande deskriptiva resultaten över medelresultat per skalsteg på ansträngningsskalan: ju högre skattad ansträngning, desto högre prestation. Vidare förklarar, enligt regressionsanalysen, ansträngning ungefär 8 procent av variationen i elevernas prestation i matematik. Effekten är tämligen konstant inom respektive ämne om vi jämför resultat från PISA 2003 och PISA 2006 med PISA 2012. Effekten är också i princip densamma mellan ämnena, även om en något starkare effekt ses när det gäller läsförståelse.

Effekten av ansträngning på prestation i de olika ämnena i PISA är alltså ungefär lika stark nu som den var 2003, men eleverna skattar en lägre grad av ansträngning nu jämfört med då. Kan detta vara orsaken till de svenska elevernas sämre resultat? Enligt regressionsanalysen kan det vara en bidragande orsak, men knappast den huvudsakliga förklaringen. Som vi sett i tidigare avsnitt har de svenska elevernas rapporterade grad av ansträngning i genomsnitt gått ner med 0,35 skalsteg på ansträngningsskalan sedan 2003. Ett skalstegs minskning (exempelvis från ett värde av åtta till ett värde av sju) skulle innebära 12,20 PISA-poäng mindre. 0,35 skalstegs förändring motsvarar mindre än så, mellan 4–5 PISA-poäng (om man tar ett genomsnitt av effekten 2003 respektive 2012). Sveriges resultat har dock sjunkit betydligt mer än så, i matematik med 31 PISA-poäng sedan 2003. Några av dessa poäng (eller ~ 15 % av nedgången) skulle alltså kunna förklaras med lägre grad av ansträngning.

Vi tittade också på effekter av elevers minskade ansträngning i relation till PISA 2006, då det är mellan 2006 och 2012 som vi ser en nedgång i elevernas skattade ansträngning. Mellan 2003 och 2006 var det ingen nedgång i skattad ansträngning (dock var det olika huvudämnen dessa år vilket skulle kunna spela en viss roll för elevernas skattade ansträngning), men en liten nedgång i prestation, i samtliga ämnen. Nedgången i prestation mellan 2006 och 2012 är 24 poäng för matematik och läsförståelse och 18 poäng för naturvetenskap. 4–5 poäng av dessa skulle således, i teorin, kunna förklaras av lägre ansträngning hos de svenska eleverna. Notera att resultatattenden för i synnerhet naturvetenskap är något osäker då naturvetenskap bara varit huvudämne vid ett tillfälle.

När det gäller provmotivations-skalan, så är även den en signifikant prediktor av prestation. Enligt regressionsanalysen ger en ökning med ett skalsteg på provmotivations-skalan en ökning med 26 poäng på PISA-provet. Ett högre värde än för ansträngningstermometern, men då har provmotivations-skalan inte lika många skalsteg och en jämförelse är därmed inte meningsfull. Provmotivations-skalan är även en signifikant prediktor av resultatet i svenska och naturvetenskap (se Tabell B1 i Bilaga 1).

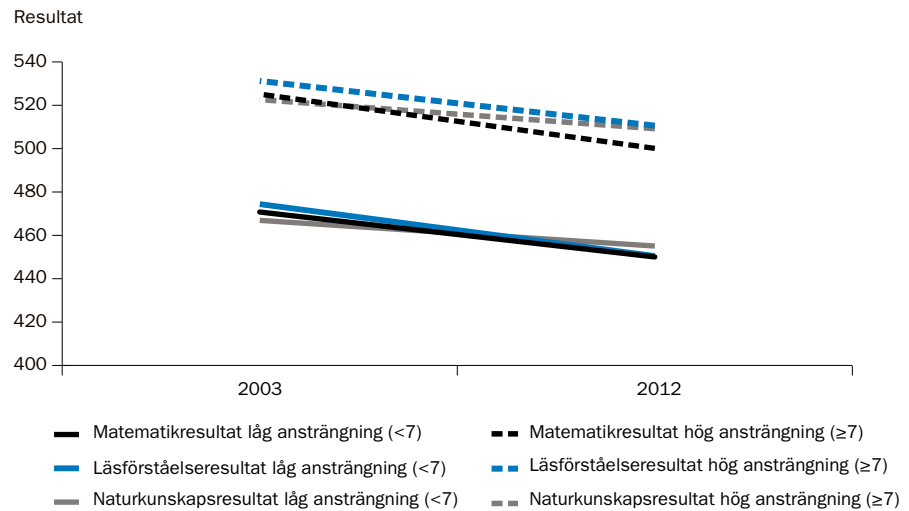
Om vi skulle korrigera för den lägre genomsnittliga graden av självskattad ansträngning (mätt med ansträngningstermometern) i PISA 2012 och skatta elevernas resultat, skulle de svenska eleverna ha haft ungefär 4,5 poängs högre resultat i matematik, 483 poäng i stället för 478. Regressionen säger dock egentligen ingenting om kausalitet, orsak-verkan (om det är den högre graden av skattad ansträngning som faktiskt orsakar den bättre prestationen), och vi vet inte hur det verkligen skulle ha sett ut om svenska elever skulle ha rapporterat samma grad, eller en högre grad, av ansträngning nu jämfört med tidigare. Det finns flera exempel på länder, exempelvis Finland och Island, där det finns lika starka samband mellan ansträngning och prestation som i Sverige, och där eleverna nu rapporterar en högre grad av ansträngning än de gjorde i PISA 2003, medan prestationen ändå har gått ner. Det är således vanskligt att dra alltför långtgående slutsatser enbart utifrån regressionsanalysen.

Konstateras kan dock att elevernas rapporterade grad av ansträngning tycks ha en signifikant ”effekt” på prestation, att en viss del av det svenska resultatet kan förklaras av ansträngning, men att den minskade graden av ansträngning knappast kan vara en huvudsaklig förklaring till att svenska elever presterar sämre nu än de gjorde 2003. Sannolikt förklaras den största delen av den svenska nedgången helt enkelt av att andra faktorer, till exempel att eleverna har sämre kunskaper om det som mäts i PISA, vilket i sin tur kan bero på många olika saker.

Resultatförändring för elever som rapporterar en hög respektive en låg grad av ansträngning

I PISA 2003 var det 25 procent av eleverna som rapporterade en ansträngning på PISA motsvarande en sexa på ansträngningsskalan eller mindre. I PISA 2012 hade denna andel ökat till 32 procent. Skulle det kunna vara så att de elever som rapporterar låg ansträngning försämrat sina resultat mer än de som rapporterat hög ansträngning och skulle en sådan förskjutning därmed kunna vara ytterligare en förklaring till den allmänna nedgången? Korrelations- och regressionsanalyserna ovan ger inga indikationer på detta, men för att åskådliggöra det hela lite tydligare redovisas i Figur 13 genomsnittliga resultat för åren 2003 och 2012 för matematik, läsförståelse och naturvetenskap, uppdelat på huruvida eleverna rapporterat en låg (mindre än sju på ansträngningsskalan) eller en hög (sju eller högre på ansträngningsskalan) grad av ansträngning på PISA-provet.

Figur 13. Genomsnittligt PISA-resultat för gruppen elever som rapporterar en låg (sex eller lägre på ansträngningstermometern) respektive en hög (sju eller högre på ansträngningstermometern) grad av ansträngning. Uppdelat på ämne och för PISA 2003 respektive 2012.



Nedgången i resultat är i princip lika stor för de som rapporterar en hög ansträngning som för de som rapporterar en låg grad av ansträngning på PISA. I matematik är nedgången till och med lite större för dem som rapporterar en hög grad av ansträngning. Det är alltså inte så att de som rapporterar en hög upplevd ansträngning på provet presterar lika bra nu som tidigare och att det är de elever som har en självskattad låg motivation som står för resultatförsämringen. Resultatet är säkrast för matematik, då det var huvudområde såväl 2003 som 2012, men trenderna ser likadana ut för alla ämnen.

Har ansträngning någon effekt på prestation i urvalet av länder

Regressionsanalysen med alla länder i urvalet visar att alla nordiska länder har b-värden mellan 11,95 (Danmark) och 16,65 (Finland). Australien, Lettland och Nya Zeeland har värden som liknar Sveriges. Tyskland, Nederländerna och Litauen har lägre, men fortfarande signifikanta värden. Polen och Estland har icke signifikanta värden (b-koefficienterna i princip 0, och den förklarade variansen är 0). I Polen och Estland har alltså skattad ansträngning ingen effekt på prestation, medan effekten är starkast i de nordiska länderna tillsammans med Australien, Lettland och Nya Zeeland. Den förklarade variansen i övriga länder i urvalet, uttryckt som R^2 , är .01, .02 respektive .03 i Tyskland, Nederländerna respektive Litauen och mellan .06 och .13 i de nordiska länderna, Australien, Lettland och Nya Zeeland. Annorlunda uttryckt, enligt den enkla linjära regressionen förklaras mellan 0 procent (i Polen och Estland) och 13 procent (på Island) av variationen i matematikresultatet av elevernas skattade ansträngning på PISA-provet. För Sveriges del var denna siffra 8 procent.

Samtliga resultat som presenterats så här långt gäller för hela stickproven, utan att de delats upp på något sätt. Vissa tidigare studier har visat på att det finns skillnader mellan pojkar och flickor när det gäller motivation att göra sitt bästa på prov som saknar konsekvenser för eleverna: att flickor tenderar att

rapportera en högre grad av motivation och ansträngning, medan sambandet mellan motivation och prestation ibland har varit starkare för pojkar. Det finns dock också studier som inte visat på några större könsskillnader. För att undersöka hur det sett ut i PISA för i synnerhet Sveriges del gjordes motsvarande analyser för vissa av de aspekter som redovisats ovan fast uppdelat på pojkar och flickor. Översiktliga resultat av denna analys presenteras här närmast.

Är det någon skillnad mellan pojkar och flickor?

I PISA 2012 hade de svenska flickorna ett medelvärde på ansträngningsskalan på 7,26 på den tiogradiga skalan, medan det för pojkarna var 6,80. Flickor rapporterar alltså en högre grad av ansträngning än vad pojkarna gör. Sambandet mellan skattad ansträngning och prestation är också något starkare för flickornas del, i synnerhet i matematik, medan det för läsförståelse är i princip samma.

Flickorna rapporterar i genomsnitt också en något högre grad av provmotivation, mätt med provmotivationskalan, jämfört med pojkarna. Medelvärdet för hela gruppen på hela skalan är 2,37, för flickorna är det 2,35 och för pojkarna är den 2,40 (notera att ett lägre värde här står för en högre grad av motivation). Det är framför allt på frågorna ”Jag kände mig motiverad att göra mitt bästa...” och ”Jag gjorde mitt bästa...” som flickorna ger högre skattningar medan skillnaderna är ganska små på övriga frågor. Överlag är skillnaderna mellan pojkar och flickor dock inte särskilt stora. Även här är sambandet mellan skattning på provmotivationskalan och prestation något starkare för flickorna.

Analysen av könsskillnader när det gäller skattad ansträngning i urvalet av länder ger samma bild som för Sverige: flickor skattar en högre grad av ansträngning än vad pojkar gör. Den minsta skillnaden mellan könen finns i Nederländerna och på Nya Zeeland, där det bara skiljer någon tiondel. Ett flertal länder har differenser av ungefär samma storlek som Sverige, runt ett halvt skalsteg, medan det i exempelvis Polen skiljer nästan ett helt skalsteg.

En analys av trender i skattad ansträngning för pojkar respektive flickor för Sveriges del visar att nedgången är större för flickorna. Såväl 2003 som 2006 var könsskillnaderna större när det gäller skattad ansträngning. Minskningen i skattad ansträngning för flickornas del är 0,46 respektive 0,43 skalsteg på ansträngningsskalan när 2012 jämförs med 2003 respektive 2006. Minskningen för pojkarnas del är 0,25 respektive 0,26 skalsteg. Flickorna rapporterar alltså en relativt sett lägre ansträngning nu än tidigare men fortfarande rapporterar de en högre ansträngning än pojkarna.

Effekten av ansträngning på prestation är ganska likartad för pojkar och flickor, över de olika åren och de olika ämnena. Dock skiljer det någon (i läsförståelse) eller några (i matematik) poängenheter, där b-värdena är högre för flickor. Effekten av ansträngning på prestation är något starkare för flickor än för pojkar, men på det stora hela är det inga stora skillnader. Det finns exempel på länder där könsskillnaderna är betydligt större. I Tyskland till exempel är skattad ansträngning signifikant relaterad till prestation för flickorna, medan det för pojkarna nästan inte finns något samband alls.

Uppdelat på kön skulle ytterligare någon poäng av förändringen i flickornas resultat kunna förklaras av minskad ansträngning, och någon färre poäng av förändringen av pojkarnas resultat. Man hade därför kunnat tänka sig att pojkarnas

resultat skulle ha försämrats något mindre än flickornas. I praktiken är det dock tvärtom: pojkarna har försämrat sina resultat mest. I exempelvis matematik har pojkarna presterat bättre än flickorna i PISA 2003 och PISA 2006 (trots att de även då rapporterade en lägre grad av ansträngning än flickorna). I PISA 2012 är det ingen skillnad mellan pojkar och flickor i matematikprestation. I läsförståelse har alltid flickorna presterat bättre, men i PISA 2012 är skillnaden till flickornas fördel större än den varit tidigare. Dessa resultat visar också på att det sannolikt inte finns några enkla samband eller enkla förklaringar till de svenska elevernas fallande resultat, eller till hur ansträngning och prestation interagerar utan relationer mellan icke-kognitiva och kognitiva variabler är komplexa fenomen.

I denna studie har bara översiktliga analyser uppdelat på kön gjorts. Ytterligare fördjupade analyser skulle behöva göras för att om möjligt nå en bättre förståelse av dynamiken när det gäller kön, ansträngning och prestation. Resultaten är i linje med tidigare studier när det gäller att flickor rapporterar en högre grad av ansträngning och motivation än pojkar, men visar till skillnad från tidigare studier inte att sambanden mellan ansträngning/motivation och prestation är starkare för pojkar.

Har provmotivation någon effekt på svenska elevers provprestation när man kontrollerar för andra bakgrundsvariabler som normalt har samband med prestation?

Som nämndes i inledningen så kan motivation mätas på olika nivåer. Provmotivation mäts på situationsspecifik nivå, i form av individens motivation att göra sitt bästa på ett givet prov, eller en given uppgift. Det vanligaste är kanske att attityder och motivation mäts på domänspecifik nivå (motivation att lära sig matematik) eller mer generell nivå (motivation att lära sig och bli bättre i skolan mer allmänt). Det är inte orimligt att tänka sig att det finns samband mellan domänspecifik och situationsspecifik motivation, och därför kan det vara meningsfullt att analysera dem tillsammans för att se om de "egentligen tycks mäta samma sak". I PISAs elevenkäter finns ett flertal frågor som på ett eller annat sätt anknyter till domänspecifik motivation. Man har i PISA bildat skalor, index, av dessa frågor som mått på aspekter som självuppfattning i matematik, inre och yttre motivation att lära matematik, självtillit, matematikängslan, uthållighet, hur man förklarar misslyckanden etc (OECD, 2014b). Resultaten för de svenska eleverna på dessa skalor är rapporterade i den nationella rapporten för PISA (Skolverket, 2013), liksom i mer detalj i OECDs rapport "PISA 2012 Results: Ready to learn. Students' engagement, drive, and self-beliefs" (OECD, 2014b) varför de inte behandlas närmare här. Som en bakgrund kan det bara nämnas att de svenska eleverna rapporterar en högre grad av positiv självuppfattning, högre inre och yttre motivation, högre självtillit, men också högre matematikängslan nu jämfört med 2003. Jämfört med OECD-medelvärdet avviker Sverige inte på något extremt sätt när det gäller nivåer på dessa skalor, svenska elever ligger dock jämförelsevis lågt när det gäller matematikängslan.

Följande index användes som variabler i regressionsanalysen tillsammans med ansträngningstermometern:

- *Självuppfattning i matematik* – huruvida man tycker att man är bra i matematik, att man lär sig lätt etc.
- *Självförtroende* – huruvida man känner sig säker på att man kan lösa olika matematiska uppgifter
- *Inre motivation* – intresse för matematik
- *Yttre motivation* – värdet av matematik
- *Matematikängslan*
- *Uthållighet* – ansträngning
- *Attribuering* – hur man förklarar misslyckande i matematik (inre/yttre orsaker)
- *Provmotivationsfrågorna* (6 st, för analys av enbart det svenska stickprovet).

I analysen användes även PISAs Index för ekonomisk, social och kulturell status (ESCS) för att ta hänsyn till socioekonomisk bakgrund.

Samtliga dessa variabler har signifikanta och tydliga samband med matematikprestation när de analyseras var för sig. Sambanden är i många fall ganska starka, detta gäller i synnerhet självuppfattning och självförtroende, och för samtliga variabler utom provängslan positiva. Mönstren ser i princip likadana ut i de nordiska länderna, medan sambanden ser något annorlunda ut i vissa av de andra länderna. I exempelvis Tyskland och Nederländerna är det ett svagare samband mellan självuppfattning och prestation, men ett starkare samband mellan självförtroende och prestation.

I just denna regressionsanalys användes bara delar av stickproven i de olika länderna. Detta beror på att PISA år 2012 använde tre olika enkätversioner, vilket innebär att alla elever inte svarat på alla enkätfrågor. I stället för att skatta svar för de elever (en tredjedel av det totala stickprovet) som inte besvarat vissa frågor användes här enbart de elever som besvarat elevenkät version B, där samtliga ovanstående index finns med. Det kan noteras att de svenska eleverna som besvarat version B har ett något högre medelvärde på ansträngningstermometern jämfört med den totala gruppen. Då syftet med just denna analys inte primärt är att studera grad av ansträngning, utan att mer allmänt se om ansträngning har betydelse för prestation även om man tar hänsyn till en rad andra aspekter, har analysen gjorts utan vidare korrigeringar.

Resultaten presenteras mer i detalj i Bilaga 1 (Tabell C1), men sammanfattningsvis kan det konstateras att effekten av rapporterad ansträngning på prestation kvarstår även om man tar hänsyn till en rad andra variabler och deras effekt på prestation. För Sveriges del händer inte så mycket med ansträngningsskalan, effekten på prestation minskar något, från ungefär 12 poäng per ökad enhet på ansträngningsskalan till drygt 8 poäng (notera att detta då egentligen bara gäller för de elever som besvarat enkätversion B), medan effekten för vissa av de andra variablerna (attribuering av misslyckanden, uthållighet) försvinner helt när variabler som självuppfattning och självförtroende läggs in i modellen. Ansträngning har ungefär lika stor effekt på prestation som socioekonomisk bakgrund och matematikängslan (för matematikängslan är effekten negativ: ju mer ängslan, desto sämre prestation). I den avhandling kring ansträngning i PISA 2003 som nämnts tidigare kom Butler fram till ungefär samma resultat (Butler, 2008).

En ambition med denna regressionsanalys var att se om det skulle gå att ”förklara bort” ansträngning med att denna aspekt överlappar mycket med andra aspekter som har med domänspecifik motivation att göra, men ansträngning har alltså fortfarande en tydlig effekt på prestation och tycks mäta något som inte täcks av de befintliga skalorna i PISA-enkäten.

När motsvarande multipla regression gjordes för det svenska stickprovet med provmotivationsskalan i stället för ansträngningstermometern var även provmotivationsskalan en signifikant prediktor också när andra variabler inkluderades i analysen

Motsvarande analys gjordes också för de andra länderna i länderurvalet. För samtliga övriga länder i urvalet förutom Polen och Estland gäller också att skattad ansträngning på ansträngningstermometern är en signifikant prediktor av prestation, även om man kontrollerar för en rad andra motivationsvariabler, liksom för socioekonomisk bakgrund (mätt via indexet för ekonomisk, kulturell och social status) (se Tabell C1, Bilaga 1).

Nedan diskuteras avslutningsvis resultaten från de analyser som presenterats i denna studie. Tyder resultaten på att provmotivation är en viktig faktor att ha med i beräkningen när vi tolkar resultaten från PISA, eller kan vi kanske bortse från provmotivationsfaktorn?

Sammanfattande diskussion

En majoritet av de svenska eleverna säger sig ha varit ganska motiverade och instämmer i att de gjorde sitt bästa på PISA-provet. En majoritet av eleverna markerar också ganska höga skattningar på den tiogradiga ansträngningsskalan. Samtidigt är det ganska stora andelar av eleverna, om än en tydlig minoritet, som rapporterar en ganska låg grad av motivation, och en låg grad av ansträngning.

Ska detta resultat tolkas som bra eller dåligt? Är eleverna överlag motiverade eller inte? Är det positivt att en majoritet av eleverna rapporterar att de var motiverade att göra sitt bästa, eller är det ett problem att en ganska stor andel rapporterar att de inte kände sig så motiverade, att de inte gjorde sitt bästa? Det optimala vore naturligtvis att alla elever skulle hålla med om att de kände sig motiverade att göra sitt bästa på provet och att de uppgav maximal ansträngning. Men, PISA är ett prov utan konsekvenser för eleverna, individuella resultat blir aldrig kända för vare sig elever eller lärare och betyget påverkas inte. Det verkar därför inte rimligt att förvänta sig att alla elever skulle uppfatta provet som oerhört viktigt eller att alla skulle anstränga sig maximalt. Det prov där 100 procent av provtagarna rapporterar maximal motivation och ansträngning finns sannolikt inte. Även i kontexten av exempelvis Högskoleprovet, ett prov där resultatet kan ha stora konsekvenser för provtagaren, och där liknande motivationsfrågor använts i en enkätstudie, finns det en viss andel av provtagarna som uppger att de inte kände sig motiverade att göra sitt bästa, även om andelen är betydligt lägre än i PISA (se Eklöf & Knekta, 2014).

Vid en internationell jämförelse med alla deltagande länder i PISA 2012 visar det sig att Sverige är ett av de länder som rapporterar den lägsta graden av ansträngning i PISA 2012, den högsta graden av ansträngning om provet hade räknats till betyget, liksom den största differensen mellan dessa två skattningar. Detta skulle kunna vara allvarligt, men även om Sverige ligger lågt i ett internationellt perspektiv så avviker vi inte på något extremt sätt, det är ganska många länder som ligger i samma spann som Sverige. Liknande resultat i tidigare PISA-omgångar har av OECD tolkats som att motivation inte är något större problem eftersom alla länder i medeltal redovisar liknande grad av ansträngning (OECD, 2007). Det finns dock faktiskt en del skillnader mellan länder i skattad ansträngning, och framför allt skiljer det en del i hur sambanden mellan skattad ansträngning och prestation ser ut.

Det har vid ett flertal tillfällen i denna rapport diskuterats varför det kan vara svårt, och kanske inte heller lämpligt, att jämföra ett enskilt lands resultat på en variabel som rapporterad ansträngning med alla andra deltagande länder, eller för den delen med det internationella medelvärdet. Bara för att det är möjligt att göra jämförelser betyder inte alltid att det är lämpligt att göra alla jämförelser som är möjliga, så att säga. Ibland kan det snarare ge missvisande resultat.

Till de jämförande analyserna i denna rapport gjordes därför ett urval av länder, bestående av de nordiska länderna, länder runt Östersjön, Nederländerna, Australien och Nya Zeeland. Även då vi jämför oss med detta urval av länder rapporterar Sverige den lägsta graden av ansträngning, fast även här är skillnaderna mycket små i jämförelse med länder som exempelvis Norge och Tyskland. Jämfört med till exempel Finland ligger dock Sverige betydligt lägre i skattad ansträngning.

Analysen av skattad ansträngning över tid visar att det skett en nedgång för de svenska eleverna med ungefär 0,35 skalsteg på den tiogradiga ansträngningsskalan sedan PISA 2003. Även om detta inte är en väldigt stor nedgång kan den ändå vara viktig att fundera kring. Möjligen säger den något om de förändringar som ägt rum i det svenska prov- och bedömningssystemet det senaste decenniet. I tidigare PISA-undersökningar hade reformerna i det svenska nationella provsystemet ännu inte genomförts, eleverna hade således inte den mängd nationella prov som de har i dag, och som de hade 2012. Det är möjligt att det ökande antalet prov som har betydelse för eleverna är en bidragande orsak till att de rapporterar en något lägre grad av ansträngning på PISA, som år 2012 sammanföll med den nationella prov-perioden och som kanske då stack ut som ett mindre viktigt prov, i alla fall för vissa av eleverna. Skulle denna hypotes ha någon relevans ter det sig dock än viktigare att fortsätta följa elevernas rapporterade grad av ansträngning och motivation i PISA då det är möjligt att ansträngningen i PISA kommer att sjunka ytterligare något. Elever i kommande PISA-omgångar kommer att ha än större erfarenhet av ”andra prov” än vad eleverna i tidigare PISA-omgångar har haft. Med mer erfarenhet av prov och betyg kan eleverna antas bli mer medvetna om var det är värt att anstränga sig eller inte. Det gäller för skolor, lärare och andra att verkligen kommunicera vikten av att eleverna anstränger sig på studier som PISA.

För samtliga år som ansträngning mätts i PISA ses en signifikant effekt av ansträngning på prestation, i alla tre ämnen som undersöks i PISA. Effekten är ungefär lika stark för alla ämnen, och alla år. Effekten av ansträngning på de svenska elevernas prestation är dock inte så stark, och nedgången i skattad ansträngning inte så stor, att det finns underlag för att dra slutsatsen att lägre ansträngning är orsaken till de svenska elevernas lägre prestation. Enligt den statistiska analysen ger ett stegs ökning på ansträngningstermometern en prestationsökning med dryga 10 poäng. Förändringen i skattad ansträngning är dock betydligt mindre än ett helt skalsteg, ja, mindre än ett halvt skalsteg, och med denna logik skulle den lägre graden av ansträngning kunna förklara ett fåtal poäng av den nedgång som skett sedan 2006, medan större delen av nedgången har andra förklaringar.

Det har vid ett par tillfällen i denna rapport gjorts jämförelser med samma population elever i en nationell provkontext, där eleverna som tog ett nationellt prov rapporterade högre grad av motivation och ansträngning än eleverna som deltog i PISA. Kan man utgå från att eleverna presterade mer ”på toppen av sin förmåga” på det nationella provet än på PISA? Ja, kanske, men som vanligt är verkligheten komplex och det finns som regel flera orsaker till en given prestation i en given situation, motivation är en av dem.

Det är också i en mer generell bemärkelse svårt att avgöra vad som är en ”sann” prestationen. Tillståndet i den svenska skolan tolkas ofta via betygsmedelvärden, nationella provresultat och resultat i internationella jämförande studier. Skulle man exempelvis jämföra resultat på nationella prov med resultat på PISA är det dock inte självklart att en given individ bör få samma resultat på de båda proven. Vare sig nationella prov, betyg, eller PISA-resultat redovisar sanningen, utan de representerar olika mått av olika karaktär med olika problem och möjligheter. Elever med gedigna kunskaper i exempelvis matematik bör rimligen kunna göra bra ifrån sig såväl på ett nationellt prov som på PISAs matematikdel, men om resultaten skulle skilja sig åt är det inte med nödvändig-

het så att ett resultat ”är mer rätt” än ett annat. Detta är ytterligare en anledning till att försöka mäta även socioemotionella förmågor i provsituationen, såsom motivation och provängslan, för att bättre kunna bilda sig en uppfattning om hur olika variabler spelar in i olika provsituationer.

Över huvud taget finns inga enkla samband eller enkla förklaringar till vare sig de svenska elevernas fallande resultat, eller till hur ansträngning och prestation interagerar. Genom fördjupade studier av olika aspekter utifrån olika perspektiv kan bitar i ett pussel läggas som vartefter ger ökad förståelse av elevers kunskapsutveckling, liksom av interaktionen mellan icke-kognitiva och kognitiva variabler, men då det handlar om komplexa fenomen blir de enkla svaren sällan exakta representationer av hur verkligheten fungerar.

Vad avser effekter av ansträngning/provmotivation skulle man kunna fundera på om det förutom den direkta effekt på provprestation som har belysts i denna studie även skulle kunna finnas en indirekt effekt, så tillvida att brist på motivation att anstränga sig i PISA också skulle kunna vara en indikation på hur mycket man orkar anstränga sig i skolan mer generellt, med läxor, uppgifter etc. Det är tänkbart att elevernas skattade ansträngning på PISA säger något om elevernas inställning utöver PISA-provet: att det kan finnas en indirekt effekt som går från skattad ansträngning på PISA via en mer generell minskad ansträngning i skolan som i sin tur leder till sämre resultat i skolan. En tänkvärd hypotes som skulle kunna prövas.

Begränsningar och fortsatta studier

Uppskattningen av hur mycket av elevernas prestation som kan förklaras av ansträngning är baserad på en statistisk modell som tittar på linjära samband, och som dessutom bygger på självskattning. Självskattning som metod har sina begränsningar: det förutsätter för det första att alla förstår frågan på det sätt det var tänkt, och för de andra att de svarar på just den frågan och inte väger in andra aspekter, för de tredje att alla svarar ärligt och med (den upplevda) verkligheten överensstämmande.

Även om denna analys har haft ambitionen att fördjupa kunskaperna om svenska elevers skattade motivation och ansträngning i PISA, över tid och i jämförelse med andra länder, så finns det fortfarande många aspekter kvar att fördjupa sig i ytterligare.

Två sådana aspekter är skillnader mellan pojkar och flickor och vad de kan bero på, liksom mer fördjupade studier av de elever som inte alls svarar på frågorna om ansträngning och motivation.

En annan aspekt är eventuella skillnader mellan skolor när det gäller skattad ansträngning. Eftersom PISA som bekant är ett prov utan betydelse för eleverna är det tänkbart att hur skolan kommunicerar vikten av en bra prestation till eleverna kan ha betydelse för deras motivation och ansträngning. Att eleverna är samplade på skolnivå gör också att det finns ett beroende mellan elever, varför det rent statistiskt också är korrekt att analysera variabler uppdelat på olika nivåer. Preliminära flernivåanalyser indikerar dock att den allra största delen av variationen (94 %) ligger på individnivå.

Studier med ansträngningstermometern och provmotivationskalan skulle också kunna kompletteras med mer småskaliga studier som exempelvis observationer eller intervjuer med elever som deltar i studier av typen PISA, för att få en ökad förståelse av hur proven uppfattas av eleverna, men också av vad de egentligen svarar på när de svarar på frågor av typen ansträngningstermometern.

Avslutande kommentar

Det finns ett flertal indikationer på att elevernas provmotivation och ansträngning i provsituationen är en viktig faktor att ta hänsyn till och fortsätta följa. De svenska elevernas skattade ansträngning har förändrats något över tid, vi ligger förhållandevis lågt i rapporterad ansträngning jämfört med länder som vi brukar jämföra oss med, skattad ansträngning och motivation har en effekt på provprestation, och denna effekt kvarstår även om man tar hänsyn till andra attityd/motivationsvariabler.

Samtidigt är nedgången i rapporterad ansträngning så pass liten – och effekten av ansträngning är heller inte särskilt stark – att det verkar osannolikt att det är denna nedgång som starkt bidragit till de svenska elevernas svaga prestation i PISA 2012. Enligt den statistiska analysen skulle ansträngning kunna förklara några poäng av resultatförsämringen (exempelvis 4–5 poäng av 31 poäng i matematik i PISA mellan 2003 och 2012), men den största delen förklaras sannolikt snarare av andra faktorer, såsom att eleverna inte har tillräckliga kunskaper om det som mäts i PISA.

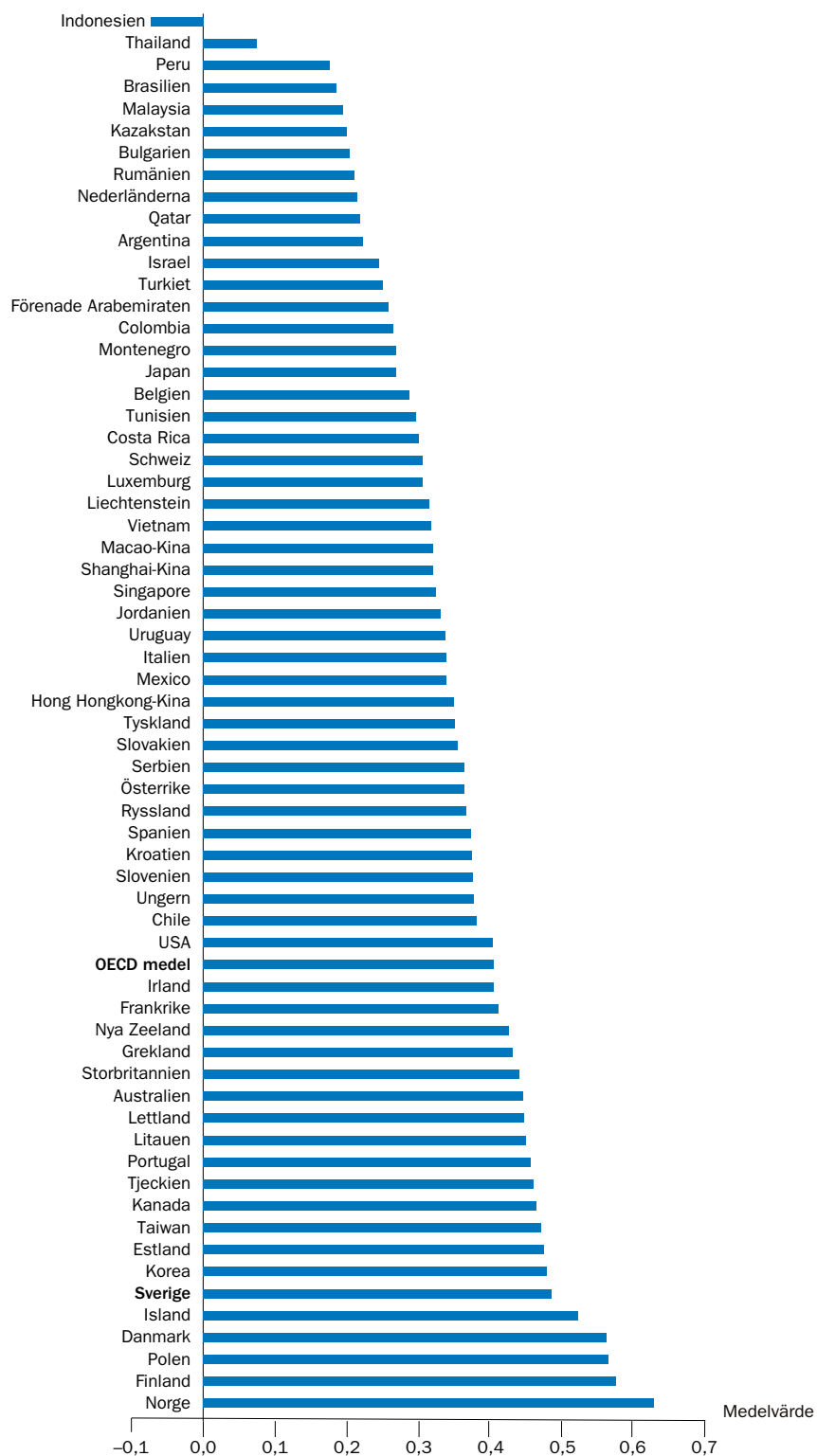
Referenser

- Butler, J. C. (2008). *Interest and effort in large-scale assessment: the influence of student motivational variables on the validity of reading achievement outcomes*. PhD thesis, University of Melbourne, Faculty of Education.
- Eklöf, H. (2006). *Motivational beliefs in the TIMSS 2003 context: Theory, measurement and relation to test performance*. Doctoral Dissertation, Umeå University, Department of Educational Measurement.
- Eklöf, H., Japelj, B., & Grønmo, L. S. (2014). A cross-national comparison of reported effort and mathematics performance in TIMSS Advanced 2008. *Applied Measurement in Education, 27*, 31–45.
- Eklöf, H., & Knekta, E. (2014). *Different stakes, different motivation? Studies of test-taking motivation in different assessment contexts*. Paper presented at the annual meeting of the American Educational Research Association (AERA), Philadelphia, April 2014 (In symposium: The Impact of Test-Taking Motivation and Test Consequences on the Validity of Test Score Inferences).
- Hattie, J.A. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- King, G., Murray, C., Salomon, J., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98*(1), 191–207.
- OECD (2013a). *PISA 2012 results: What Students Know and Can Do: Student Performance in Mathematics, Reading and Science* (Volume I). Paris: OECD.
- OECD (2013b). *PISA 2012 results: Ready to Learn: Student Engagement, Attitudes and Motivation* (Volume III). Paris: OECD.
- OECD (2007). *PISA 2006: Science Competencies for Tomorrow's World* (Volume 1). Paris: OECD.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research and applications* (2nd Ed.). New Jersey, NJ: Merrill Prentice Hall.
- Skolverket. (2013). *PISA 2012. 15-åringars kunskaper i matematik, läsförståelse och naturvetenskap*. Stockholm: Skolverket
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81
- Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment, 10*(1), 1–17. doi:10.1207/s15326977ea1001_1

Bilaga

A. Korrelationer mellan skalan som mäter självuppfattning i matematik och matematikprestation

Figur A1. Samband mellan skalan som mäter självuppfattning i PISA och matematikprestation, alla länder.



B. Regressionsanalys – effekter av rapporterad ansträngning och motivation på provprestation

Tabell B1. Effekt av rapporterad ansträngning och motivation på provprestation i PISA 2003, 2006 respektive 2012

Prov	Effekt på provpoäng per skalsteg på anstr-termometer	Standardiserad effekt	Nedgång skattad ansträngning	Nedgång förklarad av ansträngning (poäng)	Nedgång prestation PISA (poäng)
	B	β	2003-2012/ 2006-2012	2003-2012/ 2006-2012	2003-2012/ 2006-2012
Ma 2003	13.65	.29			
Ma 2006	12.59	.28			
Ma 2012	12.20	.28	-0.35/ -0.34/	4,5/ 4,2/	31/ 24/
Läs 2003	13.80	.29			
Läs 2006	15.72	.32			
Läs 2012	15.18	.31	-0.35/ -0.34/	5/ 5,2/	31/ 24/
NO 2003	13.26	.25			
NO 2006	13.10	.28			
NO 2012	13.28	.29	-0.35/ -0.34/	4,6/ 4,5/	21/ 18/
	B Provmotivationsskalan	β			
Ma 2012	26.08	.20			
Läs 2012	29.27	.20			
NO 2012	27.32	.19			

C. Multipel regressionsanalys

– effekter av rapporterad ansträngning på provprestation när även andra variabler inkluderas i modellen

Samtliga variabler lades in i en regressionsmodell som oberoende variabler, och de fem plausibla värdena i matematik lades in som beroende variabel. Hur mycket respektive variabel bidrar till att förklara resultaten efter att man tagit hänsyn till de andra variablerna framgår i Tabell C1, där resultaten från den multipla regressionen för urvalet av 13 länder, stickproven som besvarat elev-enkät version B presenteras. Notera att här har använts de färdiga indexen som finns tillgängliga i databasen, de är transformerade av PISA till ett medelvärde på 0 (OECD-medelvärde) och en standardavvikelse på 1. Ansträngnings-termometern har därför också skalats om, så att den har ett medelvärde på 0 och en standardavvikelse på 1.

Tabell C1. Effekter (b + beta samt standardfel och t-värden) av ett antal variabler på matematikprestation i PISA 2012. Ett urval av länder.

Land	EqVar	b	beta	b.se	beta.se	b.t	beta.t
Australien	(CONSTANT)	497,78					
	Matematikångslan	-8,67	-0,09	2,35	0,02	-3,69	-3,74
	ESCS (SES)	23,63	0,20	1,80	0,02	13,10	13,15
	Attribuering	0,80	0,01	1,86	0,02	0,43	0,43
	Yttre motivation	-1,39	-0,01	2,10	0,02	-0,66	-0,67
	Inre motivation	-11,46	-0,13	2,36	0,03	-4,86	-4,95
	Självförtroende	39,23	0,44	2,08	0,02	18,83	19,13
	Uthållighet	-5,53	-0,06	1,96	0,02	-2,82	-2,78
	Självuppfattning	21,26	0,21	3,09	0,03	6,89	6,93
	Ansträngning	15,60	0,16	1,49	0,02	10,49	9,75
Danmark	(CONSTANT)	493,76					
	Matematikångslan	-8,08	-0,10	3,53	0,04	-2,29	-2,30
	ESCS (SES)	22,60	0,23	2,27	0,02	9,98	10,29
	Attribuering	1,18	0,01	2,39	0,03	0,49	0,49
	Yttre motivation	-0,85	-0,01	2,14	0,02	-0,40	-0,40
	Inre motivation	-12,49	-0,15	2,32	0,03	-5,39	-5,28
	Självförtroende	26,22	0,29	3,01	0,03	8,72	8,57
	Uthållighet	-4,57	-0,05	2,56	0,03	-1,78	-1,78
	Självuppfattning	29,81	0,37	3,54	0,04	8,42	8,21
	Ansträngning	18,71	0,21	2,18	0,02	8,59	9,01
Estland	(CONSTANT)	520,01					
	Matematikångslan	-18,53	-0,22	3,78	0,04	-4,90	-4,99
	ESCS (SES)	18,79	0,20	2,32	0,02	8,10	8,15
	Attribuering	4,48	0,05	3,45	0,03	1,30	1,32
	Yttre motivation	-2,69	-0,03	2,68	0,03	-1,00	-1,00
	Inre motivation	-12,46	-0,14	3,23	0,04	-3,86	-3,86
	Självförtroende	30,08	0,33	2,77	0,03	10,84	11,28
	Uthållighet	-5,44	-0,06	1,91	0,02	-2,86	-2,88
	Självuppfattning	21,33	0,24	4,02	0,05	5,30	5,23
	Ansträngning	4,98	0,06	2,47	0,03	2,02	2,00

Land	EqVar	b	beta	b.se	beta.se	b.t	beta.t
Finland	(CONSTANT)						
	Matematikångslan	-7,23	-0,08	2,74	0,03	-2,63	-2,64
	ESCS (SES)	14,43	0,14	2,06	0,02	7,01	7,17
	Attribuering	1,25	0,01	1,55	0,02	0,80	0,80
	Yttre motivation	0,33	0,00	2,39	0,03	0,14	0,14
	Inre motivation	-12,67	-0,15	2,47	0,03	-5,14	-5,16
	Självförtroende	19,60	0,22	2,81	0,03	6,98	7,07
	Uthållighet	1,28	0,01	2,09	0,02	0,61	0,61
	Självuppfattning	33,50	0,43	2,35	0,03	14,27	13,93
	Ansträngning	18,09	0,18	1,96	0,02	9,23	9,45
Tyskland	(CONSTANT)						
	Matematikångslan	-18,44	-0,22	3,32	0,04	-5,56	-5,55
	ESCS (SES)	25,77	0,26	2,24	0,02	11,53	12,26
	Attribuering	12,52	0,13	2,77	0,03	4,53	4,47
	Yttre motivation	-5,74	-0,07	3,09	0,04	-1,86	-1,86
	Inre motivation	-3,85	-0,05	4,04	0,05	-0,95	-0,95
	Självförtroende	38,04	0,39	3,10	0,03	12,26	13,48
	Uthållighet	-4,53	-0,05	3,11	0,03	-1,45	-1,44
	Självuppfattning	10,39	0,13	4,49	0,05	2,32	2,30
	Ansträngning	11,21	0,12	2,25	0,02	4,99	4,85
Island	(CONSTANT)						
	Matematikångslan	-10,94	-0,12	3,28	0,04	-3,34	-3,33
	ESCS (SES)	10,60	0,10	2,86	0,02	3,71	3,81
	Attribuering	3,48	0,04	2,35	0,03	1,48	1,48
	Yttre motivation	7,34	0,08	2,87	0,03	2,56	2,55
	Inre motivation	-11,94	-0,13	3,36	0,04	-3,56	-3,59
	Självförtroende	23,76	0,29	3,11	0,04	7,64	7,67
	Uthållighet	-5,09	-0,06	2,92	0,03	-1,74	-1,75
	Självuppfattning	27,50	0,32	4,07	0,05	6,76	6,84
	Ansträngning	19,76	0,24	2,28	0,03	8,65	8,20
Lettland	(CONSTANT)						
	Matematikångslan	-14,00	-0,13	4,47	0,04	-3,13	-3,19
	ESCS (SES)	22,89	0,25	2,59	0,03	8,85	9,59
	Attribuering	0,16	0,00	3,32	0,03	0,05	0,05
	Yttre motivation	-8,25	-0,09	3,03	0,03	-2,72	-2,74
	Inre motivation	-10,64	-0,11	4,08	0,04	-2,61	-2,63
	Självförtroende	24,71	0,25	4,69	0,05	5,27	5,45
	Uthållighet	-1,71	-0,02	2,57	0,03	-0,67	-0,67
	Självuppfattning	28,69	0,29	4,93	0,05	5,82	5,76
	Ansträngning	20,71	0,22	2,93	0,03	7,07	6,82

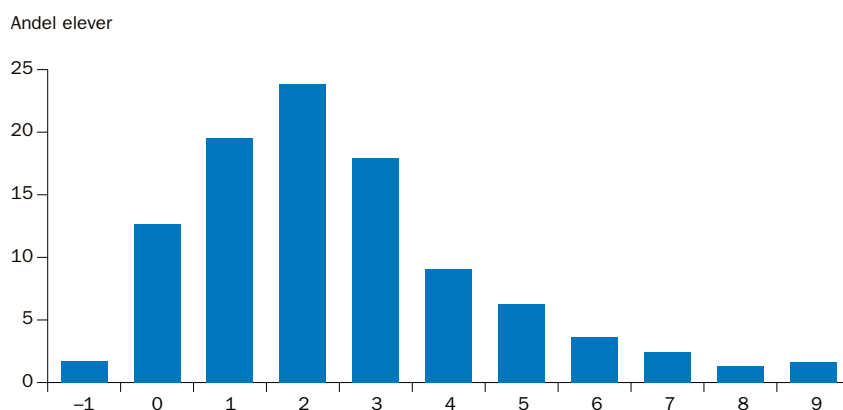
Land	EqVar	b	beta	b.se	beta.se	b.t	beta.t
Litauen	(CONSTANT)						
	Matematikängslan	-18,65	-0,21	2,49	0,03	-7,49	-7,54
	ESCS (SES)	23,84	0,25	2,31	0,02	10,33	10,61
	Attribuering	-0,01	0,00	2,02	0,02	-0,01	0,00
	Yttre motivation	0,67	0,01	1,92	0,02	0,35	0,35
	Inre motivation	-17,36	-0,21	2,37	0,03	-7,33	-7,61
	Självförtroende	31,19	0,34	2,83	0,03	11,02	12,75
	Uthållighet	-5,89	-0,06	2,92	0,03	-2,01	-1,98
	Självuppfattning	21,94	0,25	3,26	0,04	6,73	6,77
	Ansträngning	12,67	0,12	2,58	0,02	4,92	5,07
Nederländerna	(CONSTANT)	529,20		3,63		145,94	
	Matematikängslan	-12,75	-0,13	3,71	0,04	-3,44	-3,44
	ESCS (SES)	28,59	0,27	3,39	0,03	8,44	8,86
	Attribuering	-1,94	-0,02	3,89	0,04	-0,50	-0,50
	Yttre motivation	0,59	0,01	3,51	0,04	0,17	0,17
	Inre motivation	-4,81	-0,05	4,46	0,05	-1,08	-1,09
	Självförtroende	38,17	0,41	3,81	0,03	10,03	11,81
	Uthållighet	0,88	0,01	2,98	0,03	0,29	0,30
	Självuppfattning	-8,32	-0,09	3,88	0,04	-2,14	-2,17
	Ansträngning	8,46	0,09	2,43	0,03	3,48	3,46
Norge	(CONSTANT)	491,71		2,79		176,52	
	Matematikängslan	-9,85	-0,11	3,47	0,04	-2,84	-2,86
	ESCS (SES)	15,80	0,13	3,13	0,02	5,04	5,12
	Attribuering	1,81	0,02	2,68	0,03	0,67	0,67
	Yttre motivation	0,37	0,00	2,27	0,02	0,16	0,16
	Inre motivation	-10,16	-0,12	3,06	0,03	-3,32	-3,32
	Självförtroende	25,01	0,31	3,04	0,04	8,22	8,69
	Uthållighet	-1,29	-0,02	2,49	0,03	-0,52	-0,52
	Självuppfattning	32,97	0,40	3,10	0,04	10,63	10,53
	Ansträngning	15,73	0,19	1,94	0,03	8,11	7,48
Polen	(CONSTANT)	520,89		2,86		181,92	
	Matematikängslan	-19,20	-0,20	3,07	0,03	-6,25	-6,29
	ESCS (SES)	19,80	0,20	2,38	0,02	8,31	9,06
	Attribuering	-2,30	-0,03	2,00	0,02	-1,15	-1,14
	Yttre motivation	1,84	0,02	3,33	0,03	0,55	0,56
	Inre motivation	-15,67	-0,16	3,83	0,04	-4,10	-4,14
	Självförtroende	34,09	0,38	2,64	0,03	12,93	14,76
	Uthållighet	-1,92	-0,02	1,96	0,02	-0,98	-0,98
	Självuppfattning	23,08	0,25	3,37	0,04	6,85	6,73
	Ansträngning	0,87	0,01	1,78	0,02	0,49	0,49

Land	EqVar	b	beta	b.se	beta.se	b.t	beta.t
Sverige	(CONSTANT)	476,11		2,66		179,03	
	Matematikångslan	-17,28	-0,18	2,77	0,03	-6,23	-6,29
	ESCS (SES)	18,57	0,16	3,13	0,03	5,93	6,11
	Attribuering	-0,99	-0,01	2,86	0,03	-0,34	-0,35
	Yttre motivation	-7,91	-0,09	2,77	0,03	-2,85	-2,88
	Inre motivation	-8,55	-0,10	3,04	0,04	-2,82	-2,79
	Självförtroende	22,70	0,25	3,65	0,04	6,21	6,37
	Uthållighet	0,20	0,00	2,40	0,03	0,08	0,09
	Självuppfattning	29,31	0,32	3,97	0,04	7,39	7,17
	Ansträngning	16,03	0,19	1,79	0,02	8,95	8,56

D. Analyser av differensvariabeln (skillnaden mellan skattad ansträngning på PISA och uppskattad ansträngning om provresultatet hade räknats till betyget)

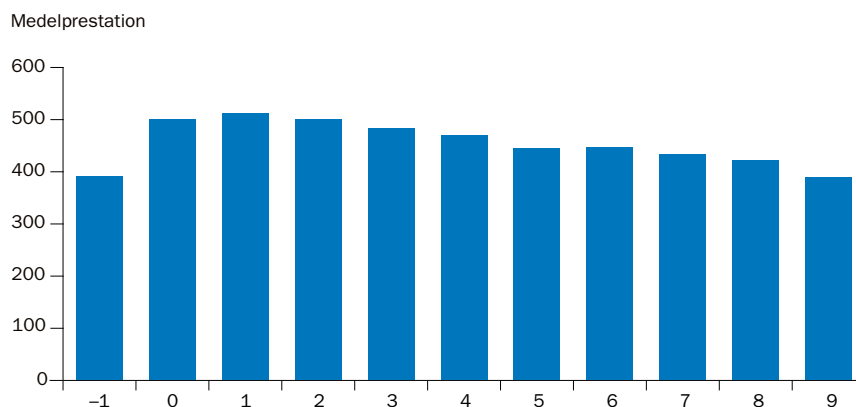
I figurerna, som gäller svenska data för PISA 2012, är samtliga kategorier av "realistiska skattare" (Butler, 2008: elever som skattat högre ansträngning på PISA jämfört med om provet räknats till betyget) sammanslagna till en kategori (-1) då grupperna var så väldigt små. Gruppen med "2" på differensskalan är de elever som har två skalstegs skillnad för skattad ansträngning och ansträngning om provet räknats till betyget (i typfallet, en åtta på den första och en tia på den andra). De som hör till kategorin "9" i figuren nedan är de studenter som har nio skalstegs skillnad: en etta i ansträngning på PISA men en tia om provet hade räknats till betyget. Denna grupp innehåller också väldigt få elever.

Figur D1. Andel elever med respektive värde på differensvariabeln (-1 = alla elever som rapporterar mer ansträngning på PISA än på betygsprov, 0 = elever som rapporterar lika mycket ansträngning på PISA som ett vanligt prov, 1 = elever som rapporterar ett steg mindre ansträngning på PISA än betygsprov 9 = elever som rapporterar minimal ansträngning på PISA men maximal ansträngning om det hade räknats till betyget).



Eleverna som markerat att de ansträngt sig ett steg mindre på PISA än de skulle ha gjort om provet räknats till betyget är den högst presterande gruppen i PISA, följt av de elever som har ett värde på 0 eller 2 i figuren ovan (se Figur D2 nedan).

Figur D2. Genomsnittlig prestation per "skalsteg" på differensskalan, Sverige, Matematik, PISA 2012



STUDIE II

Har svenska elevers uthållighet och engagemang förändrats över tid?

En studie om elevers svarmönster på PISAs kunskapsprov



Sammanfattning

Sverige har deltagit i PISA sedan den första mätningen år 2000. PISA undersöker 15-åringars läsförståelse, kunnande i matematik och naturvetenskap samt i vissa mätningar även olika former av problemlösningsförmåga.⁵ I de första omgångarna presterade de svenska eleverna på en relativt hög nivå inom samtliga undersökta områden, men i den senaste studien har Sverige istället placerat sig under OECD-ländernas genomsnitt. En hypotes om denna resultatnedgång är att det är elevernas engagemang att skriva kunskapsprovet i PISA som har minskat, snarare än att elevernas kunskaper blivit sämre. I den här studien undersöker vi om delar av resultatnedgången i PISA kan sägas bero på att elever i den senaste undersökningen inte haft uthållighet och engagemang för att besvara provfrågorna i samma utsträckning som tidigare.

Denna fördjupningsanalys gör inte anspråk på att ge ett helt uttömmande svar på frågan om det är engagemanget eller kunskapsnivån hos eleverna som sjunker vilket resulterat i ett försämrat resultat i PISA. Däremot är det en pusselbit bland flera om elevers motivation och ansträngning. I analysen framkommer ingenting som tyder på att resultatnedgången som PISA visat inte är en rättvis bild av elevernas kunskaper inom de områden som PISA mäter.

I denna analys undersöks om det tycks ha skett en förändring i hur de svenska eleverna besvarar uppgifterna i det senaste kunskapsprovet i PISA 2012 jämfört med tidigare studier, med fokus på matematik och läsförståelse. Med inspiration från en irländsk förlaga används PISA-provets utformning med en roterande design av provuppgifter för att jämföra elevernas svarsmönster. Dessa analyseras utifrån två möjliga hypoteser. Den första hypotesen säger att resultatnedgången i PISA beror på en sjunkande kunskapsnivå hos eleverna medan den andra hypotesen säger att resultatnedgången beror på ett minskat engagemang bland eleverna för att skriva kunskapsprovet. Båda dessa scenarier kan förekomma samtidigt samt även förstärka varandra, och det kan vara svårt att särskilja de två helt. För att ändå försöka undersöka vilken av de två förklaringsgrunderna som är mest trolig belyser vi elevernas svarsmönster på flera sätt i rapporten. I denna analys kan förändringar i elevernas svarsmönster inte kopplas direkt till resultatet på PISA-provet. Därför är det inte möjligt att ge ett svar på hur stor del av nedgången i resultat uttryckt i antal eller andel poäng som kan förklaras av elevers svikande engagemang.

Tanken bakom den första hypotesen, om en sjunkande kunskapsnivå, är att eleverna får svårare att besvara uppgifterna korrekt i lika stor utsträckning som tidigare oavsett om frågorna kommer tidigt eller sent i provet. Det innebär att det inte sker någon förändring i elevernas uthållighet mellan åren. Med uthållighet menas här huruvida eleverna blir tröttare mot slutet av provet och provtiden. Om uthålligheten avtar innebär det att andelen elever som besvarar uppgifterna korrekt mot slutet av ett prov är lägre än andelen elever som besvarar dem korrekt i början av samma prov. På motsvarande sätt förväntas andelen elever som lämnar en viss uppgift obesvarad vara lägre för dem som skriver uppgiften tidigt, mot dem som skriver den sent.

5 Problemlösning testades i PISA 2003 och digital problemlösning testades i PISA 2012

Den andra hypotesen, om ett minskat engagemang att skriva kunskapstestet, vilar istället på idén att elever inledningsvis besvarar de provuppgifter de kan, precis som elever med samma kunskapsnivå gjort tidigare år, men i högre grad än tidigare inte anstränger sig hela vägen igenom. Eleverna klarar sig lika bra i början av provet vilket tyder på att de har samma kunskaper. I slutet av provet presterar de däremot sämre än motsvarande elever tidigare år, vilket är en indikation på att de ger upp när det blir jobbigt, tråkigt eller när de blir trötta. Det innebär att elevernas uthållighet minskar mellan åren.

De svenska elevernas svarsmonster pekar mot att det är kunskapsnivån som har sjunkit under 00-talet. Denna slutsats grundar sig på att de mönster som stödjer denna hypotes är entydiga; det finns en tydlig nedgång i andelen elever som besvarar uppgifterna korrekt i såväl matematik som läsförståelse oavsett om uppgifterna kommer tidigt eller sent i provhäftet mellan de flesta jämförda PISA-mätningarna. Även förändringar i svarsmonster bland andelen elever som lämnar uppgifter obesvarade stödjer denna hypotes. Dessa förändringar är dock oftast något mindre och något mer svårbedömda.

Det finns även vissa tecken som tyder på att provengagemanget bland svenska elever kan ha minskat något. Dessa resultat är dock inte entydiga. Även om det inte är möjligt att ge ett svar uttryckt i antal eller andel poäng av resultatnedgången, så är den samlade bedömningen att denna eventuella minskning i engagemang inte skulle kunna förklara någon större del av den nedgång i resultat som svenska elever har uppvisat under det senaste decenniet.

Inledning

Sverige har sedan första mätningen år 2000 deltagit i OECD-projektet PISA (Programme for International Student Assessment). PISA undersöker 15-åringars läsförståelse, kunnande i matematik och naturvetenskap samt i vissa mätningar även olika former av problemlösningsförmåga.⁶

Den senaste PISA-undersökningen, PISA 2012, bekräftar och förstärker den bild som PISA tidigare visat och som bekräftas av andra internationella kunskapsundersökningar som PIRLS och TIMSS. Svenska grundskoleelevers kunskaper i läsförståelse, matematik och naturvetenskap har försämrats sedan millennieskiftet.⁷ Totalt sett är Sverige det land som haft den största resultatnedgången av alla länder som deltar i PISA. (Skolverket 2013a) För att kunna vända denna utveckling är en ökad förståelse för möjliga orsaker bakom resultatnedgången väsentlig. Som redan framhållits i tidigare rapporter från Skolverket går det inte att peka ut en isolerad förklaringsfaktor till den utveckling som skett vad gäller svenska elevers resultat under de senaste decennierna. Det är många faktorer som spelar in och det är svårt att fastställa ett orsakssamband mellan enskilda förändringar och resultatutvecklingen (Skolverket 2013b).

I Skolverkets rapport om PISA 2012 (Skolverket 2013a) diskuteras möjliga hypoteser om vad som kan ligga bakom den försämrade resultatutvecklingen, och det betonas att ytterligare analyser behövs innan det går att få en klarare bild till varför utvecklingen ser ut som den gör. En aspekt som diskuteras handlar om PISA-resultatens tillförlitlighet. Den tar sin utgångspunkt i att de elever som skriver PISA-provet är anonyma och att eleverna inte får någon återkoppling på sina resultat. Provet kan alltså inte räknas in som betygs- eller bedömningsunderlag. Denna typ av prov kallas ofta för low-stakes prov. I motsats till detta står så kallade high-stakes prov som är betydelsefulla för eleven. PISA:s kunskapsprov skrivs under vårterminen och de flesta 15-åringar går då i grundskolans årskurs 9. Det sammanfaller med en tid då eleverna också skriver ett flertal nationella prov. De nationella proven i årskurs 9 kan, till skillnad från PISA-provet, påverka elevens slutbetyg.

I Sverige har antalet nationella prov ökat från att ha varit tre stycken i årskurs 9 år 2003 till att vara fyra stycken år 2012.⁸ Å ena sidan kan man hävda att fler nationella prov kan ha lett till en ökad provvana bland eleverna, och att elevernas provängslan bör vara mindre på PISA-provet än andra prov eftersom det inte påverkar elevernas betyg. Det skulle kunna bidra till en prestation som bättre speglar elevernas kunskaper. Å andra sidan kan man istället mena att ett ökat antal prov under denna period gjort eleverna trötta på att skriva prov. Eftersom det för den enskilde eleven heller inte gör någon skillnad om de anstränger sig eller inte kan det leda till att eleverna inte gör sitt bästa på PISA-provet. Elevernas prestation skulle då sämre än tidigare spegla elevernas

6 Problemlösning testades i PISA 2003 och digital problemlösning testades i PISA 2012

7 Med undantag för naturvetenskap i TIMSS årskurs 4 mellan mätningen 2007 och 2011

8 2003 skrev elever i åk 9 nationella prov i svenska, engelska och matematik. När PISA 2012 genomfördes hade ett nationellt prov i no-ämnena tillkommit (varje elev gör ett prov i något av ämnena biologi, fysik eller kemi). År 2013 infördes nationella prov även i so-ämnena. På samma sätt som för no-ämnena gör varje elev ett prov i ett av so-ämnena. Numera är det alltså fem nationella prov för eleverna i åk9, men när PISA 2012 genomfördes skrev eleverna fyra nationella prov.

kunskaper och därmed sänka resultatens relevans för att användas som kunskapsmått. Om denna förklaring skulle stämma utgör den dock troligtvis endast en delförklaring till de sjunkande resultaten.

I PISA 2009 konstaterades att Irland var ett av de länder som tappat mest i läsförståelse av alla deltagande länder, precis som Sverige gjort i matematik i PISA 2012. I en artikel från Irland skriven av Cosgrove och Cartwright (2014) undersöker författarna två möjliga, eventuellt överlappande, förklaringar till den irländska resultatutvecklingen baserat på elevernas svarsmönster i PISA. Grundidén för denna fördjupningsstudie bygger på dessa två hypoteser, men tillvägagångssättet är vidareutvecklat.

Syfte

I denna studie undersöker vi om det tycks ha skett en förändring i hur eleverna besvarar uppgifterna i det senaste kunskapsprovet i PISA år 2012 jämfört med tidigare studier, med fokus på matematik och läsförståelse. Vidare undersöks om sådana eventuella förändringar främst signalerar en nedgång i kunskap eller om resultatnedgången åtminstone delvis kan härledas till ett hos eleverna bristande engagemang att skriva PISAs kunskapsprov. Syftet är att undersöka om elevernas engagemang för att skriva kunskapsprovet i PISA tycks ha förändrats mellan åren, och i förlängingen om de svenska elevernas resultat i PISA är tillförlitliga. Denna analys görs för att förstå mer om elevers uthållighet och engagemang i de provsituationer som de internationella storskaliga studierna som svenska elever deltar i grundar sig på.

Två hypoteser utgör grunden för analysen av elevernas svarsmönster. Den första hypotesen går ut på att resultatnedgången i PISA beror på en sjunkande kunskapsnivå hos eleverna. Den andra hypotesen är att resultatnedgången beror på ett minskat engagemang bland eleverna för att skriva kunskapsprovet. Båda dessa fenomen kan förekomma samtidigt och även förstärka varandra. Det kan också vara svårt att särskilja de två helt. För att ändå undersöka vilken av de två förklaringsgrunderna som är mest trolig belyser vi elevernas svarsmönster från flera håll. I denna analys kan förändringar i elevernas svarsmönster inte kopplas direkt till resultatet på PISA-provet. Därför är det inte möjligt att ge ett svar på hur stor del av nedgången i resultat uttryckt i antal eller andel poäng som kan förklaras av elevers svikande engagemang.

I nästa avsnitt presenteras en kort genomgång av PISAs design med fokus på kunskapsprovet och de egenskaper hos detta som vår analys använder sig av, samt en överblick över analysens idé, upplägg och metod. Därefter presenteras och analyseras resultaten. Avslutningsvis följer en sammanfattande diskussion.

Genomförande

I denna studie analyseras elevernas svarsmönster utifrån hur stor andel av eleverna som besvarar uppgifterna korrekt och hur stor andel av dem som lämnar frågorna obesvarade beroende på när i provet frågan kommer. Vi undersöker om dessa mönster förändras mellan åren. Eftersom matematik är huvudämne i PISA 2012 kommer vi att fokusera på matematikuppgifter, men också komplettera med läsuppgifter, då vi vill kunna dra så breda slutsatser som möjligt och inte endast fokusera ett specifikt ämnesområde. Matematikuppgifterna bedöms vara mer lästunga än de matematikuppgifter som svenska elever generellt är vana vid (Skolverket 2009 och 2009). Det kan alltså finnas en effekt av elevens läsförmåga på elevens benägenhet att orka besvara matematikuppgifter. Det är därför av intresse att komplettera matematikuppgifterna med just läsuppgifter. Vi har valt att inte inkludera naturvetenskap i denna analys.

En annan fråga som lyfts när PISA:s kunskapsprov diskuteras är hur uppgifterna är utformade och hur vana de svenska eleverna är vid denna typ av uppgifter. En aspekt av uppgifternas utformning är på vilket sätt eleverna förväntas ge sina svar. Vi undersöker därför även hur elevernas svarsmönster förändras när vi delar upp uppgifterna efter de två uppgiftstyperna flervalsfrågor och frågor där eleverna själva formulerar svaret.

I denna fördjupningsstudie fokuserar vi på förändringar i svar bland svenska elever över tid. För att fånga in eventuella allmänna trender och få en uppfattning om vad som kan påstås vara avvikande används genomsnittliga svarsandelar från eleverna i deltagande OECD-länder som referenspunkter.^{9,10} Ibland kommer dessa kort refereras till som OECD-genomsnittet. Det ska dock inte ses som en analys av OECD-länderna utan är ett stöd i tolkningen av de svenska elevernas svar. Inga jämförelser görs med enskilda länder i denna analys.

I detta avsnitt ger vi först en kort introduktion till upplägget på det kunskapsprov som eleverna gör i PISA och hur detta upplägg möjliggör vår analys. Efter det följer en diskussion om de avgränsningar som gjorts. Avsnittet avslutas med en närmare genomgång av de två olika tillvägagångssätt som används för att analysera elevernas svarsmönster.

9 De länder som bidragit till referensvärdena kan skilja sig åt mellan cykler eftersom vissa länder endast deltar i vissa cykler.

10 De svenska elevernas svarsandelar är viktade med vikten *Total Student Weight*. Svarsandelarna för OECD-länderna är viktade med vikten *Senate Weight*. Den sistnämnda gör så att samtliga länder bidrar lika mycket till medelvärdet. Det innebär att länder med väldigt många 15-åringar inte bidrar mer till OECD-medelvärdet än länder med färre 15-åringar.

PISAs upplägg och kunskapsprovets design

Det är i huvudsak tre kompetensområden som testas i PISA: matematik, naturvetenskap och läsförståelse. Varje gång studien genomförs, alltså vid varje mättillfälle, är ett ämne huvudämne och de andra två är biämnen. Varje kompetensområde är huvudämne i var tredje mätning, och återkommer alltså som huvudämne vart nionde år. En lista över vilka år PISA har genomförts och vilket ämne som varit huvudämne vid varje mätning presenteras i Tabell 1. All dokumentation om PISA, hur PISA genomförs och databaser från samtliga PISA-mätningar finns på OECDs hemsida.¹¹

Tabell 1. Huvudämne i PISAs kunskapsprov per mätning.

År/Mätning	Huvudämne i kunskapsprovet
2000	Läsförståelse
2003	Matematik
2006	Naturvetenskap
2009	Läsförståelse
2012	Matematik

I PISA gör varje elev endast ett urval av alla provfrågor, vilket innebär att uppgifterna i olika elevers provhäften skiljer sig åt. Vid varje PISA-mätning finns det både nya provfrågor och sådana som varit med vid minst en tidigare mätning. De provfrågor som återkommer vid fler än en mätning kallas för trenduppgifter.

Alla provfrågor delas in i block. En provfråga tillhör endast ett block och varje block innehåller cirka 12–14 provfrågor. Ett block utgörs endast av uppgifter inom samma ämnesområde. När ett ämne är huvudämne finns det vanligtvis 7 block som innehåller frågor på ämnet. Antalet block per biämne varierar mellan mätningarna med 2–4 block.

Det finns 13 olika provhäften i varje mätning.¹² Ett provhäfte innehåller 4 block, och varje block kan alltså placeras på fyra olika sätt i ett provhäfte. Det innebär att varje block förekommer på en av fyra olika positioner i provhäftet, se Tabell 2. Varje provhäfte består av block med uppgifter från minst två olika ämnesområden. Sedan PISA 2003 förekommer ett block en gång på varje position och två olika block förekommer inte tillsammans i mer än ett provhäfte. Detta kallas för balanserad roterande design.

¹¹ <http://www.oecd.org/pisa/pisaproducts/>

¹² Det finns även provhäften med lite lättare uppgifter, men dessa används inte i Sverige. I mätningen 2012 gjorde vissa elever testet på dator. Från och med mätningen år 2015 och framåt kommer alla elever göra uppgifterna på dator. I denna analys används endast elevsvar från tryckta provhäften.

Tabell 2. Exempel på blockens placering i de olika provhäftena. Exemplet taget från PISA 2012.

Matematikblocken är markerade med fet stil när de förekommer i position 1 och i position 4.

Provhäftets nummer	Block			
	Position 1	Position 2	Position 3	Position 4
B1	PM5	PS3	PM6A	PS2
B2	PS3	PR3	PM7A	PR2
B3	PR3	PM6A	PS1	PM3
B4	PM6A	PM7A	PR1	PM4
B5	PM7A	PS1	PM1	PM5
B6	PM1	PM2	PR2	PM6A
B7	PM2	PS2	PM3	PM7A
B8	PS2	PR2	PM4	PS1
B9	PR2	PM3	PM5	PR1
B10	PM3	PM4	PS3	PM1
B11	PM4	PM5	PR3	PM2
B12	PS1	PR1	PM2	PS3
B13	PR1	PM1	PS2	PR3

P = Provet ges på papper, M = Matematik (Mathematics), R = Läsförståelse (Reading), S = Naturvetenskap (Science)

Alla provfrågor delas in i **block**. En provfråga tillhör endast ett block och varje block innehåller cirka 12–14 provfrågor. Alla provfrågor i ett block tillhör samma ämnesområde.

Det finns 13 olika **provhäften** i varje mätning i PISA. Ett provhäfte innehåller 4 block. Varje provhäfte består av block med uppgifter från minst två ämnesområden.

Ett block kan placeras på fyra olika sätt i ett provhäfte. Det finns alltså fyra så kallade **positioner** för blocken i provhäftena.

Sedan PISA 2003 förekommer varje block en gång på varje position och två olika block förekommer inte tillsammans i mer än ett provhäfte. Detta kallas för en **balanserad roterande design**.

Analysens idé och utgångspunkter

I denna studie används det faktum att den roterande designen gör att vi kan observera hur olika elever besvarar samma uppgift både när den ligger tidigt och när den ligger sent i ett provhäfte både inom ett år och mellan två år. Det innebär att vi kan undersöka om elevers beteende när de skriver provet har förändrats mellan de år PISA genomförts och om en sådan förändring ger stöd åt hypotesen att engagemanget att skriva PISAs kunskapsprov minskat över åren, eller om mönstret snarare tyder på att det är kunskapsnivån som sjunkit.

Vilka svarsmönster stödjer de respektive hypoteserna?

De svarsmönster som stödjer respektive hypotes presenteras i Tabell 3a, och det svarsmönster som visar på att elevernas uthållighet avtar inom en mätning presenteras i Tabell 3b. Denna analys omfattar elevernas svarsmönster för korrekta svar och obesvarade uppgifter. På vissa uppgifter i kunskapsprovet kan man få en eller två poäng för delvis korrekta och helt korrekta svar. Dessa andelar läggs samman och kategoriseras som rätt svar. De uppgifter som lämnas obesvarade av elever kan vara sådana som antingen är överhoppade mitt i ett provhäfte eller sådana som eleven inte hinner till i slutet på provet. Dessa två kategorier av obesvarade uppgifter analyseras tillsammans.

Tabell 3a. Svarsmönster enligt de två hypoteserna.

Uppställning av de förändringar i svarsmönster som ger stöd för hypotesen om en sjunkande kunskapsnivå respektive minskat engagemang att skriva PISAs kunskapsprov.

		Förändring mellan år 1 och år 2	
		P1	P4
Sjunkande kunskapsnivå	Andel korrekta svar	Minskande	Minskande
	Andel obesvarade	Ökande	Ökande
Sjunkande provengagemang	Andel korrekta svar	Oförändrad	Minskande
	Andel obesvarade	Oförändrad	Ökande

Tabell 3b. Svarsmönster för avtagande uthållighet inom en mätning.

Uppställning av de svarsmönster som visar att elevernas uthållighet avtar under provsituationen. Om elevernas uthållighet minskar mellan åren tyder det på att elevernas engagemang har sjunkit.

		Förändring mellan P1 och P4	
		Inom år 1	Inom år 2
Avtagande uthållighet i provsituationen	Andel korrekta svar	Minskande	Minskande
	Andel obesvarade	Ökande	Ökande

På de följande sidorna illustreras konceptet med jämförelsen av svarsmönster och de tre kombinationer av de två hypoteserna som är möjliga. Detta görs med hjälp av figurerna 1–3. Den genomsnittliga andelen elever som besvarar uppgifterna korrekt används som exempel. I detalj innebär detta att vi jämför elevernas svarsmönster för uppgifter i position 1 (provhäftets första frågor) och motsvarande uppgifter i position 4 (provhäftets sista frågor) på två sätt: (1) dels för varje position mellan två mätningar och (2) dels för skillnaden mellan dessa två positioner inom en mätning. Punkterna markerar den genomsnittliga andelen elever som besvarar uppgifterna korrekt per år och position. Den streckade linjen illustrerar jämförelsen mellan de två positionerna inom varje år, det vill säga om elevernas uthållighet avtar i respektive mätning. Skillnaden mellan två punkter i samma position illustrerar skillnaden mellan två mätningar. En sammanlagd bedömning av de två typerna av förändringar görs för att bedöma om stöd ges till någon av hypoteserna.

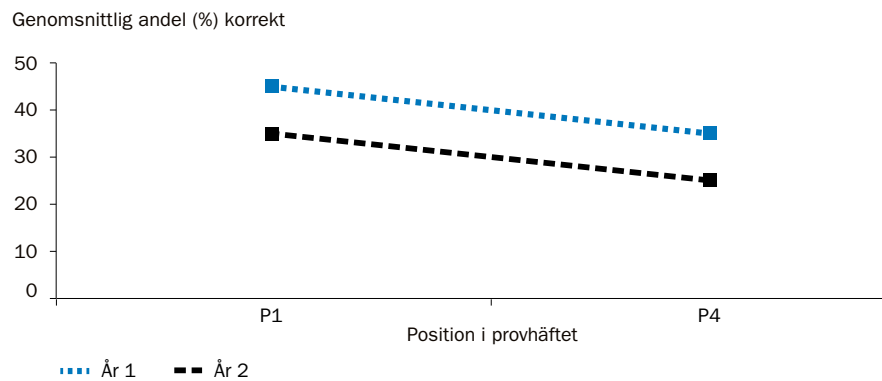
Det som menas med elevernas uthållighet är den effekt som kommer sig av att man mot slutet av längre perioder av koncentration och ansträngning kan börja tappa fokus och bli trött. Det vi förväntar oss i vilken provsituation som helst är att andelen elever som besvarar uppgifterna korrekt mot slutet av ett prov är lägre än andelen elever som besvarar dem korrekt i början av samma prov. På motsvarande sätt förväntas andelen elever som lämnar en viss uppgift obesvarad vara lägre för dem som skriver uppgiften tidigt, mot dem som skriver den sent.

Svarsmönster som ger stöd för hypotesen om sjunkande kunskapsnivå

Figur 1 visar ett exempel på hur elevers genomsnittliga andelar korrekta svar i position 1 och position 4 för två olika mätomgångar kan se ut då det är kunskapsnivån hos eleverna som lett till att resultaten sjunkit.

Figur 1. Generellt sjunkande kunskapsnivå.

Exempel på ett möjligt svarsmönster för genomsnittliga andelar elever som besvarat uppgifterna korrekt för de två positionerna vid två olika PISA-mätningar. Linjerna illustrerar jämförelsen mellan de två positionerna inom respektive år, det vill säga att elevernas uthållighet avtar mot slutet av provet.



Tanken bakom hypotesen om en sjunkande kunskapsnivå är att eleverna får svårare att besvara uppgifterna korrekt i lika stor utsträckning som tidigare, oavsett om frågorna kommer tidigt eller sent i provet. De försöker antingen svara på frågan, men svarar rätt i lägre utsträckning även tidigt i provet, eller struntar i att lämna svar på fler uppgifter än elever i tidigare mätningar.

Det innebär att det inte skett någon förändring i uthållighet mellan åren, vilket visas av att de två streckade linjerna är parallella, men att andelen elever som besvarar uppgifterna korrekta minskar mellan åren, och att andelen elever som lämnar uppgifterna obesvarade ökar, både i position 1 och position 4.

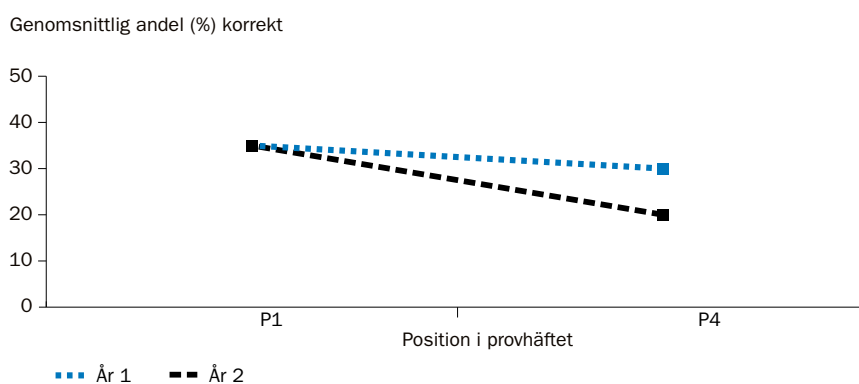
Ett sådant svarsmönster skulle också kunna uppkomma om en högre andel elever i en senare mätning bryr sig betydligt mindre om att försöka besvara uppgifterna, men att samtliga håller en jämn, låg nivå och inte blir mindre intresserade eller slutar besvara uppgifterna ju längre provet fortgår. Ett sådant scenario bedöms som mindre troligt, och om de i så fall bryr sig mindre för att de har en lägre kunskap och väljer att inte anstränga sig eller om de bryr sig mindre av någon annan anledning kan inte denna analys besvara.

Svarsmönster som ger stöd för hypotesen om minskat engagemang

Hypotesen om ett minskat engagemang att skriva kunskapsprovet vilar istället på idén att eleverna i en mätning inledningsvis skriver provet som tidigare år och besvarar de uppgifter de faktiskt kan, men i högre grad än tidigare inte anstränger sig hela vägen igenom. Detta svarsmönster illustreras av Figur 2. Eleverna klarar sig lika bra i början av provet som elever gjort i tidigare mätningar, vilket visar att de har samma egentliga förmåga. I slutet av provet presterar de däremot sämre än vad motsvarande elever gjort tidigare år, vilket är en indikation på att de ger upp när det blir jobbigt, tråkigt eller när de blir trötta.

Figur 2. Minskat engagemang.

Exempel på ett möjligt svarsmönster för genomsnittliga andelar elever som besvarat uppgifterna korrekt för de två positionerna vid två olika PISA-mätningar. Linjerna illustrerar jämförelsen mellan de två positionerna inom respektive år, det vill säga att elevernas uthållighet avtar mot slutet av provet.



Om stöd ska ges för denna hypotes kan man förvänta sig att andelen elever som besvarar uppgifterna korrekt och andelen elever som lämnar uppgifterna obesvarade är stabila i position 1, medan andelen elever som besvarar uppgifterna korrekt minskar och andelen elever som lämnar uppgifterna obesvarade ökar i position 4 mellan åren. Förändringen i uthållighet hos eleverna visas av att de streckade linjerna i Figur 2 inte längre är parallella, utan istället visar ett ökande ”glapp”. Det innebär att uthålligheten mellan de två åren ska vara signifikant skilda från varandra.

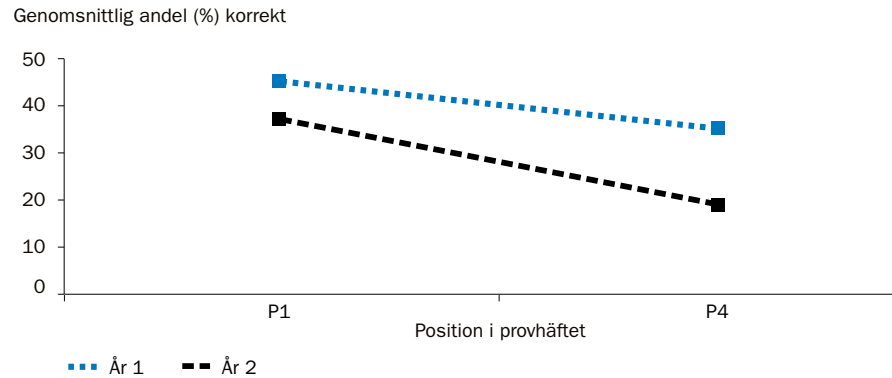
Huruvida ett eventuellt minskande engagemang har att göra med att provet inte betyder eller bidrar med något till eleverna, alltså provets low-stakes-karaktär, om det är en allmän inställning att akademiskt arbete och framgång är mindre viktigt för elever idag eller om det grundar sig på något annat kan inte besvaras i denna studie.

Sjunkande kunskapsnivå och minskat engagemang kan förekomma samtidigt

Dessa två scenarion kan förekomma samtidigt. De kan också förstärka varandra. En sådan situation illustreras i Figur 3. Figuren visar att det finns en skillnad mellan åren både i position 1 och i position 4, men med en lägre uthållighet år två. Det visar sig genom att skillnaden mellan de genomsnittliga andelarna elever som besvarar uppgiften korrekt i position 4 mellan åren är större än motsvarande skillnad i position 1.

Figur 3. Sjunkande kunskapsnivå i kombination med sjunkande engagemang.

Exempel på ett möjligt svarsmönster för genomsnittliga andelar elever som besvarat uppgifterna korrekt för de två positionerna vid två olika PISA-mätningar. Linjerna illustrerar jämförelsen mellan de två positionerna inom respektive år, det vill säga att elevernas uthållighet avtar mot slutet av provet.



En lägre uthållighet i en senare mätning kan bero på att elever som är mindre motiverade och engagerade inte orkar eller bryr sig om att ta sig igenom hela provet med fokus i samma utsträckning, men också att de kan bli jämförelsevis tröttare om provet upplevs som svårare än vad det upplevts tidigare år på grund av sämre förkunskaper. Det skulle då kunna påverka elevens möjligheter att besvara uppgifter korrekt genom hela provet, eller reflekteras i att fler uppgifter lämnas obesvarade mot slutet. På motsvarande sätt kan den som har ett lägre engagemang när de skriver provet välja bort svårare uppgifter snarare än att besvara dem korrekt i lägre utsträckning, kanske redan från provets början.

Det kan därför vara svårt att särskilja de två hypoteserna helt och tolkningar och slutsatser måste dras med viss försiktighet. Det är ändå rimligt att vissa skillnader i elevernas svarsmönster kan tillskrivas respektive hypotes. Vi belyser hypoteserna från flera olika håll för att öka möjligheterna att kunna dra slutsatser om elevernas ändrade beteende.

Flervalsfrågor eller frågor med öppna svar

– spelar det någon roll?

I PISA kan uppgifterna kategoriseras efter vilken sorts svar eleven förväntas lämna på uppgiften. Detta kallar vi för olika frågetyper. Samtliga analyser som görs av elevernas svarsmönster kommer också att göras uppdelat per frågetyp, för att få ytterligare information om elevernas svarsmönster och deras eventuella engagemang att skriva PISA:s kunskapsprov.

De exakta benämningarna eller definitionerna av de olika frågetyperna skiljer sig åt något mellan mättillfällena, men två kategorier av frågetyper är tydliga för samtliga år. Den ena kategorin innehåller frågor där eleven väljer mellan fasta svarsalternativ, så kallade flervalsfrågor. Den andra kategorin innehåller frågor där eleverna förväntas formulera sina egna svar, med eller utan förklaring eller motivering till hur de kommit fram till svaret. Denna sista grupp av frågor kallas öppna frågor.

Denna del av analysen försöker alltså säga något ytterligare om elevernas engagemang baserat på eventuella skillnader i svarsmönster mellan frågetyperna. Vi tänker oss att tröskeln att besvara en flervalsfråga med färdiga svarsalterna-

tiv är lägre för eleverna än tröskeln att börja formulera ett svar på egen hand. Rätt svar på en öppen fråga signalerar därför i högre grad att eleven faktiskt har kunskapen som frågan avser testa än vad rätt svar på en flervalsfråga gör. Därför kan vi också förvänta oss att det finns en viss skillnad i elevernas svarsfrekvenser beroende på frågetyp.

Om stöd ska ges för hypotesen att eleverna inte har samma engagemang för och därmed inte anstränger sig lika mycket att skriva kunskapsprovet vid senare års undersökningar kan man förvänta sig att eleverna i lägre grad besvarar öppna frågor korrekt vid senare mätningar än flervalsfrågor. Det finns alltid en chans för eleverna att gissa sig till rätt svar på en provfråga, men den chansen är mindre för en öppen fråga än en flervalsfråga. Om eleverna i lägre grad än tidigare kan det rätta svaret, och istället gissar mer på flervalsfrågor kommer andelen elever som besvarar flervalsuppgifterna korrekt också att minska, men i något mindre utsträckning än motsvarande andel för öppna frågor.

Om eleverna lämnar jämförelsevis fler öppna frågor än flervalsfrågor obesvarade i senare studier, kan det tyda på att de inte längre anstränger sig i lika stor utsträckning som tidigare. Det innebär att andelen elever som lämnar öppna frågor obesvarade i så fall ska öka över tid. Vi jämför därför framförallt om det finns skillnader i elevernas uthållighet mellan de två frågetyperna mellan två mätningar.

Avgränsningar och val av mätningar att jämföra

Vi har valt att prioritera ämnen där samma provuppgifter som finns med i PISA 2012 också förekommer i en undersökning så långt tillbaka i tiden som möjligt, eftersom vi vill jämföra med en omgång av PISA då det svenska resultatet var betydligt bättre än vad det är idag. Det måste också finnas trenduppgifter att jämföra mellan mätningarna, alternativt vara två mätningar med samma huvudämne.

Huvudfokus på elevers svarsmönster på matematikuppgifter

En genomgång av vilka provuppgifter som förekommer i fler än en undersökning och därmed är vad vi kallar för trenduppgifter har gjorts för åren 2003–2012. PISA 2000 är inte inkluderad i sammanställningen eftersom den studien inte använder en balanserad design.¹³

I denna rapport undersöks i huvudsak matematikuppgifter, eftersom matematik var huvudämne i PISA 2012. Elevernas svarsmönster i matematik i mätningen år 2012 jämförs med elevernas svarsmönster både i PISA 2003 och i PISA 2006. Eftersom matematik var huvudämne både år 2003 och år 2012 finns det fler matematikuppgifter totalt sett i dessa studier än i andra, och en jämförelse mellan dessa två år kan baseras på mer information än för de år då endast trenduppgifter kan jämföras.

Vi jämför också enbart trenduppgifter mellan år 2003 och år 2012, även om dessa uppgifter är utspridda på olika block i de två mätningarna. Slutligen jämförs trenduppgifter i matematik mellan mätningarna år 2006 och år 2012. Det är samma trenduppgifter som ingår i mätningen 2003, men i mätningarna år 2006 och år 2012 ligger de samlade i samma block.

13 I PISA 2000 var läsförståelse huvudområde, vilket innebär att även läsförståelse varit huvudområde i två PISA-mätningar. Däremot användes inte en balanserad design i PISA 2000 och jämförelser av positionseffekter blir därför inte möjliga.

Kompletterande analys av elevers svarsmönster på uppgifter i läsförståelse

Vi har också valt att analysera svarsmönster för läsuppgifter, eftersom vi inte primärt är intresserade av ett ämnesspecifikt resultat. Matematikuppgifterna i PISA bedöms vara mer lästunga än de matematikuppgifter som svenska elever är vana vid. Det kan därför finnas en effekt av läsförståelse på benägenheten att orka besvara matematikuppgifter i PISA. Analysen av läsuppgifter utgår endast från trenduppgifter.

För jämförelsen av uppgifter i läsförståelse används mätningarna år 2012, år 2009 och år 2003. I läsförståelse blev en ny uppsättning uppgifter trenduppgifter i och med att ämnet var huvudämne i mätningen år 2009. Det innebär att det för läsförståelse i PISA 2012 bara finns trenduppgifter från år 2009, och att en annan uppsättning uppgifter är desamma mellan mätningarna år 2009 och år 2003. Jämförelser av elevernas svarsmönster på trenduppgifter i läsförståelse görs därför dels mellan mätningarna år 2012 och år 2009, och dels mellan år 2009 och 2003. Notera att det inte finns några möjligheter att jämföra trenduppgifter direkt mellan 2012 och 2003 eftersom inga av provuppgifterna är desamma mellan dessa år.¹⁴

För naturvetenskap byttes en stor del av uppgifterna ut 2006. Det är få trenduppgifter från PISA 2003 i senare studier och de som finns kvar är inte samlade i samma block mellan åren. På grund av detta, och eftersom matematikuppgifterna i PISA bedömts vara mer lästunga än vad svenska elever är vana vid har vi valt att begränsa oss till att titta på uppgifter i matematik och läsförståelse, men inte i naturvetenskap.

En **generell jämförelse** av elevers svar på uppgifter i **matematik** mellan år **2012** och år **2003** görs. I båda dessa mätningar är matematik huvudämne, vilket innebär att finns det fler uppgifter i matematik att analysera.

Enbart **trenduppgifter i matematik** jämförs också mellan år **2012** och år **2003**. Uppgifterna ligger däremot inte samlade i samma block i båda mätningarna, vilket innebär att vi inte kan ta hänsyn till uppgifternas ordning i denna analys.

Samma **trenduppgifter i matematik** jämförs också mellan år **2012** och år **2006**. I dessa två mätningar ligger trenduppgifterna samlade i samma block i båda mätningarna.

I läsförståelse blev en ny uppsättning uppgifter trenduppgifter i och med att ämnet var huvudämne i mätningen år 2009. Därför jämförs **trenduppgifter i läsförståelse** mellan år **2012** och år **2009**.

I **läsförståelse** jämförs också den första uppsättningen **trenduppgifter** mellan år **2009** och år **2003**, eftersom vi vill kunna jämföra med mätningar så långt tillbaka i tiden som möjligt. Notera att det inte finns några möjligheter att jämföra trenduppgifter direkt mellan 2012 och 2003 eftersom inga av provuppgifterna är desamma mellan dessa år.

¹⁴ Blocken med läsförståelseuppgifter i mätningen år 2009 delas därför upp i två grupper, en som är jämförelsegrupp till elevernas svar i läsförståelse 2012, och en som är jämförelsegrupp för elevernas svar på dessa uppgifter från 2003.

Metoder som används: två olika tillvägagångssätt

I analysen av elevernas svarsmönster används två olika tillvägagångssätt. Båda sätten går ut på att undersöka hur elevernas svarsandelar på likvärdiga eller samma provuppgifter i genomsnitt har förändrats över tid, men de har olika utgångspunkter. Den ena ansatsen är mer generell och utgår från en allmän bild av elevernas svarsmönster. Den kan därför innehålla mer information och bli mindre känslig för avvikelser på enskilda frågor. Den andra ansatsen är snävare och är ett komplement till det generella tillvägagångssättet. Den baseras på direkta jämförelser mellan elevernas svarsandelar på enskilda frågor, och kan därför endast baseras på trenduppgifter. Sammantaget ger de en god bild av hur elevernas svarsmönster förändras inom provhäftena samt mellan studier.

Generell jämförelse av samtliga provuppgifter inom ett ämnesområde

Den mer generella analysen baseras på en linjär regressionsanalys. I denna ansats behövs en större mängd provuppgifter, och den lämpar sig därför bäst för år då ett ämne är huvudämne. Denna ansats kan alltså endast användas för jämförelsen mellan matematikuppgifter år 2003 och år 2012, då vi har tillgång till fler uppgifter än för andra år och för andra ämnen. Ansatsen är generell såtillvida att den inte kräver att samtliga provuppgifter för de två åren är desamma, utan kan ta hänsyn till jämförelser av två provuppgifter som anses vara likvärdiga. Det hanteras genom att i analysen ta hänsyn till vilken placering uppgiften har inom blocket (om den kommer på första, andra eller tredje plats, och så vidare, kallat sekvens) och till uppgifternas svårighetsnivå.¹⁵ Dessutom är drygt 40 procent av matematikuppgifterna i PISA 2012 trenduppgifter från mätningen år 2003, vilket ytterligare förstärker jämförbarheten mellan åren.

Analys av trenduppgifter inom ett ämnesområde

Det andra tillvägagångssättet i analysen utgår endast från de uppgifter som är trenduppgifter. I denna ansats görs en direkt jämförelse av elevernas svarsandelar på en viss fråga vid två olika tillfällen eller mellan två positioner i provhäftena vid samma tillfälle. Det finns färre trenduppgifter än det totala antalet uppgifter i ett ämne när det är huvudämne, vilket innebär att jämförelsen kan bli mer påverkad av elevernas svar på enskilda frågor än den generella analysen. Däremot baseras analysen enbart på de faktiska skillnaderna i andelen elever som hanterat samma fråga på ett visst sätt vid olika tidpunkter eller i olika positioner.

När trenduppgifterna ligger samlade i samma block vid båda tillfällena innebär det att frågorna ligger på samma plats i provhäftet, och eleverna möter samma uppgifter precis innan och precis efter uppgiften; två faktorer som gör att jämförbarheten ökar.

¹⁵ Information om hur uppgifternas svårighetsnivå beräknas finns i PISAs tekniska rapport (OECD 2014).

Trenduppgifterna i matematik i studien 2003 ligger däremot utspridda över flera olika block; hänsyn kan då inte tas till uppgiftens placering i provhäftet och de omgivande frågorna. Jämförelsen måste således tolkas mer försiktigt än den av trenduppgifter mellan 2006 och 2012. Eftersom vi i den generella analysen tar hänsyn till svårighetsgrad och plats i provhäftet för jämförelsen i matematik mellan dessa år ger de två tillvägagångssätten tillsammans en god bild av elevernas svarsmönster.

I den **generella analysen** som görs med hjälp av regressionsanalys undersöker vi *skillnader i elevernas genomsnittliga svarsandelar* mellan år och position i provhäftet. Om vi vill analysera elevernas uthållighet ett visst år baserat på andelen elever som besvarar uppgifterna korrekt tolkar vi regressionskoefficienten från en variabel som delar in elevernas svarsandelar utefter om uppgiften legat i position 1 (kodad som grupp 0) eller i position 4 (kodad som grupp 1). Denna koefficient ger oss skillnaden mellan de två grupperna och tolkas som "Skillnaden mellan andelen elever som besvarar uppgifterna korrekt när de ligger i position 1 jämfört med i position 4 är i genomsnitt XX procentenheter det aktuella året". Om denna skillnad inte är statistiskt signifikant kan vi således inte belägga att det finns en skillnad mellan elevernas genomsnittliga svarsandelar per position inom ett år.

I **analysen av trendfrågor** undersöker vi istället *genomsnittet av skillnaderna i elevernas svarsandelar* mellan år och position i provhäftet. För exemplet med uthållighet ett visst år baserat på andelen elever som besvarar uppgifterna korrekt beräknar vi först skillnaden mellan andelen elever som besvarat respektive trenduppgift när uppgiften ligger i position 1 och i position 4 det aktuella året. Vi subtraherar respektive andel i position 4 från position 1. När vi har dessa differenser beräknas medelvärdet av dem. Om detta medelvärde är negativt kan vi konstatera att det i genomsnitt skett en nedgång av andelen elever som besvarar uppgiften korrekt när uppgiften ligger sist i provhäftet, jämfört med när uppgiften ligger först i provhäftet. Vi testar om skillnaden är statistiskt signifikant med hjälp av ett så kallat parat t-test. Om denna skillnad inte är statistiskt signifikant kan vi således inte belägga att det i genomsnitt finns en skillnad mellan elevernas svarsandelar per position inom ett år.

Resultat

I de kommande avsnitten presenteras en sammanställning av de svarsmönster som framkommer i analysen. Resultatkapitlet är uppdelat i tre avsnitt. Först följer två avsnitt där elevernas svarsmönster redovisas uppdelat på ämnesområdena matematik och läsförståelse. Varje sådant delavsnitt avslutas med en sammanfattning av elevernas svarsmönster inom det ämnesområdet, samt vilken hypotes som svarsmönstren ger stöd åt. Därefter följer ett avsnitt i vilket analysen av elevernas svarsmönster uppdelat per frågetyp redovisas.

Alla skillnader som rapporteras i denna analys är sådana skillnader som är statistiskt signifikanta, och alltså inte kan påstås endast ha uppstått av den naturliga variation som finns vid urvalsundersökningar.¹⁶

Elevernas svarsmönster på matematikuppgifter

I följande avsnitt presenteras elevernas svarsmönster på matematikuppgifter. Först presenteras jämförelsen av PISA 2003 och PISA 2012, följt av jämförelsen av PISA 2006 och PISA 2012.

Överlag ger den generella analysen av samtliga matematikuppgifter för dessa år och analysen av enbart trenduppgifterna en överensstämmande bild av de svarsmönster och förändringar som undersöks. Resultaten från analyserna av matematikuppgifter finns i Tabell 4 och Tabell 5 i bilagan.

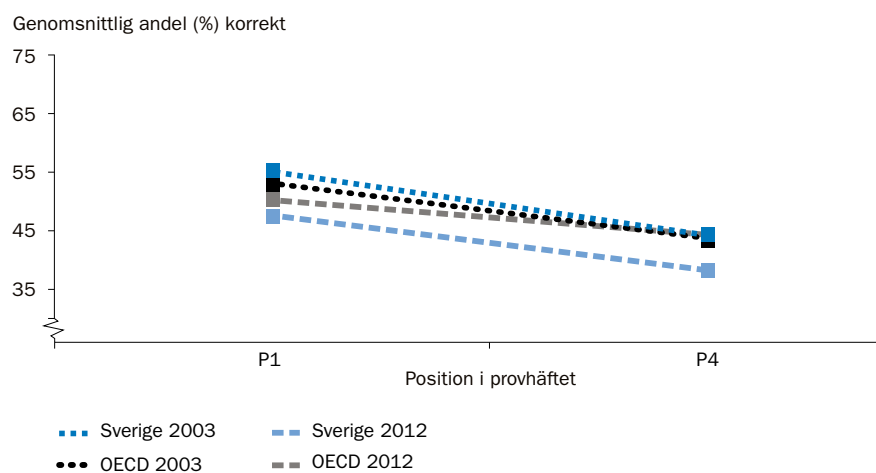
¹⁶ En signifikansnivå på 5 procent används. Tekniskt uttryck innebär det att vi påstår att medelvärdet är skilt från noll även efter att vi tagit hänsyn till den statistiska osäkerheten.

Elevernas svarsmönster för korrekt besvarade uppgifter i matematik i PISA 2003 och PISA 2012

Andelen svenska elever som besvarar matematikuppgifterna korrekt har minskat mellan 2003 och 2012 i båda positioner, enligt båda analysmetoderna. I Figur 4 visas dessa svarsmönster baserat på värdena från den generella analysen. I figuren ser vi att de två streckade linjerna som illustrerar elevernas uthållighet i princip är parallella för de svenska eleverna när vi undersöker andelen elever som besvarar uppgifterna korrekt. De liknar det scenario som beskrivs i Figur 1 om stöd för hypotesen om en generellt sjunkande kunskapsnivå.

Figur 4. Resultat från den generella analysen av uppgifter i matematik i PISA 2003 och PISA 2012: Korrekt besvarade uppgifter.

Genomsnittlig andel elever som besvarar matematikuppgifterna korrekt när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.



Andelen elever som besvarar uppgifterna korrekt när uppgifterna ligger först i provhäftet har enligt den generella analysen minskat med cirka 7 procentenheter mellan år 2003 och år 2012. Motsvarande förändring när provuppgifterna istället ligger sist i provhäftet, det vill säga i position 4, är för Sveriges del en nedgång med cirka 5 procentenheter. Nedgången i andelen svenska elever som besvarar uppgifterna korrekt mellan position 1 och position 4 enligt de två olika analysmetoderna är i PISA 2003 är runt 11–12 procentenheter, medan samma andel i PISA 2012 är cirka 9–10 procentenheter. Dessa procentenheter är inte signifikant skilda från varandra, vilket innebär att elevernas uthållighet inte förändrats mellan dessa år.

Enligt den generella analysen har andelen elever som besvarar uppgifterna korrekt i OECD som genomsnitt minskat när uppgifterna kommer först i provhäftet, men är oförändrad när uppgifterna kommer sist i provhäftet. Det innebär att uthålligheten förbättrats i genomsnitt i OECD mellan PISA 2003 och PISA 2012. Nedgången i andelen elever som besvarar uppgifterna korrekt mellan år 2003 och år 2012 är alltså större i Sverige än i OECD som genomsnitt.

Elevernas svarsmönster för obesvarade uppgifter i matematik i PISA 2003 och PISA 2012

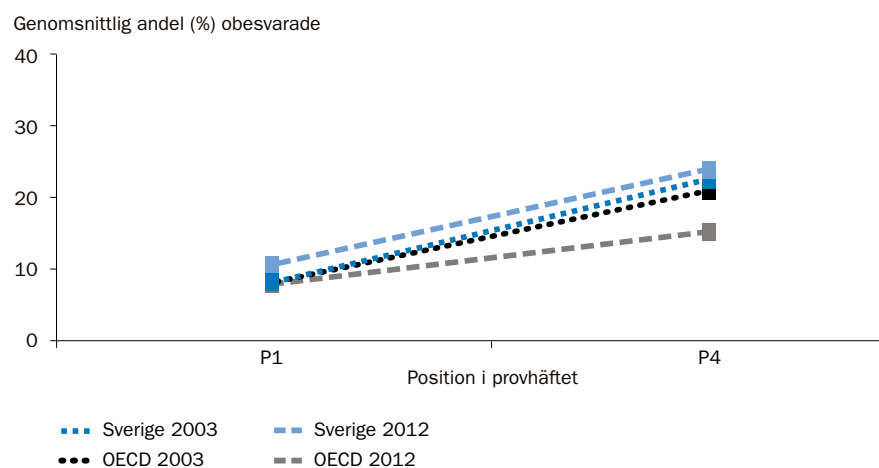
Svarsmönstret för andelen elever som lämnar matematikuppgifterna obesvarade är relativt lika mellan PISA 2003 och PISA 2012 i Sverige. Det syns också i Figur 5, där elevernas svarsmönster för överhoppade uppgifter enligt den generella analysen för år 2003 och år 2012 visas. Enligt den generella analysen är inga skillnader i elevernas svarsmönster för överhoppade uppgifter statistiskt säkerställda. Uthålligheten, alltså förändringen mellan då en uppgift besvaras tidigt i ett provhäfte jämfört med sent i ett provhäfte är enligt denna analys cirka 13–14 procentenheter.

Analysen av trenduppgifter påvisar en liten ökning av andelen elever som hoppar över uppgifter när de ligger först i provhäftet, men däremot en nedgång när de ligger sist i provhäftet. Det innebär att uthålligheten enligt denna analys har ökat något från PISA 2003 till PISA 2012.

I Figur 5 framkommer att uthålligheten utifrån obesvarade matematikuppgifter i PISA 2012 är mindre bland de svenska eleverna än för OECD-länderna i genomsnitt.

Figur 5. Resultat från den generella analysen av uppgifter i matematik i PISA 2003 och PISA 2012: Obesvarade uppgifter.

Genomsnittlig andel elever som hoppar över matematikuppgifterna när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.

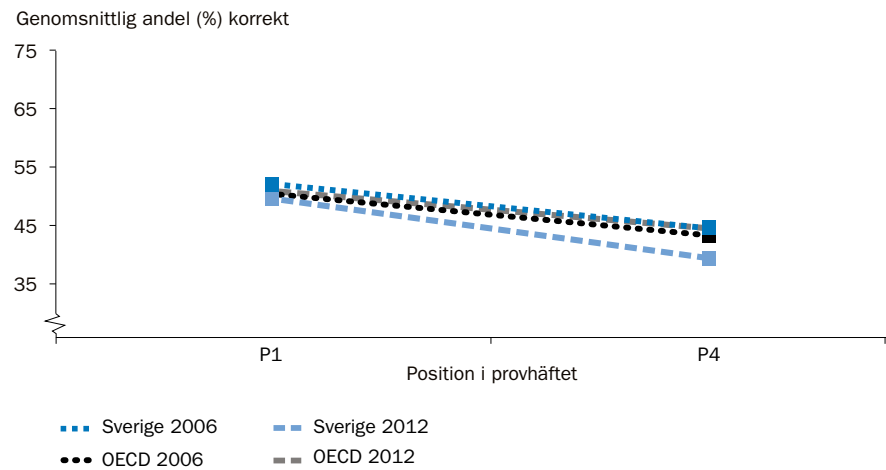


Elevernas svarsmönster för korrekt besvarade trenduppgifter i matematik i PISA 2006 och PISA 2012

Elevernas svarsmönster på matematikuppgifter i PISA 2012 jämförs också med PISA 2006. Denna jämförelse görs endast på trenduppgifter, vilka är färre än antalet matematikuppgifter som används i den generella analysen av svarsmönster mellan år 2003 och år 2012. Elevernas svarsmönster för korrekta svar visas i Figur 6.

Figur 6. Resultat från analysen av trenduppgifter i matematik i PISA 2006 och PISA 2012: Korrekt besvarade uppgifter.

Genomsnittlig andel elever som besvarar matematikuppgifterna korrekt när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.



När vi jämför trenduppgifter i matematik mellan PISA 2006 och PISA 2012 kan vi se en nedgång i andelen korrekt besvarade uppgifter för eleverna i båda positioner. Uthålligheten enligt de korrekt besvarade uppgifterna tycks vara lika omfattande i dessa två studier, och dessa mått inte statistiskt skilda från varandra. Detta mönster liknar det vi ser i jämförelsen mellan PISA 2003 och PISA 2012, och ger alltså stöd åt hypotesen om att det är kunskapsnivån som sjunkit fram till mätningen år 2012.

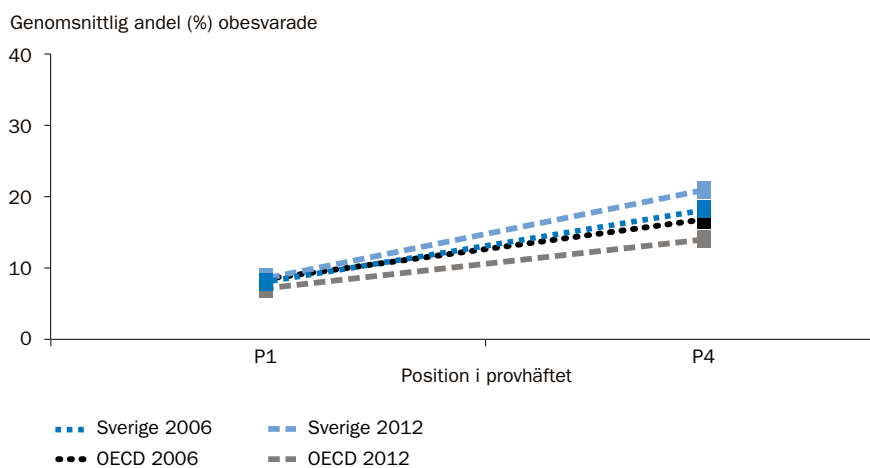
För position 1 har andelen elever som besvarar uppgifterna korrekt minskat med cirka 3 procentenheter, och motsvarande andel i position 4 är cirka 5 procentenheter. Skillnaden mellan position 1 och position 4 inom respektive mätning för båda dessa år är runt 8–10 procentenheter. Den svenska utvecklingen avviker något från den genomsnittliga utvecklingen i OECD-länderna, där någon förändring av andelen elever som besvarar uppgifterna korrekt knappast kan påstås ha skett mellan dessa två år.

Elevernas svarsmönster för obesvarade trenduppgifter i matematik i PISA 2006 och PISA 2012

I jämförelsen av andelen elever som lämnar uppgifterna obesvarade mellan PISA 2006 och PISA 2012 är skillnaden liten, både när uppgifterna kommer först i provhäftet respektive när de kommer sist i provhäftet. Det sker en liten ökning av andelen elever som lämnar uppgifterna obesvarade dessa två år. När uppgifterna kommer tidigt i provhäftet är uppgången mellan åren dock inte statistiskt säkerställd, medan motsvarande andel när uppgifterna kommer sist är cirka 3 procentenheter. Ett sådant mönster skulle kunna ge stöd åt hypotesen om sjunkande engagemang. Eftersom uthålligheten är oförändrad mellan åren ger dock den samlade bilden inget sådant stöd. Istället ges ett svagt stöd för hypotesen om sjunkande kunskapsnivå.

Figur 7. Resultat från analysen av trenduppgifter i matematik i PISA 2006 och PISA 2012: Obesvarade uppgifter.

Genomsnittlig andel elever som hoppar över matematikuppgifterna när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.



Förändringen av andelen elever som lämnar uppgifterna obesvarade i position 1 jämfört med position 4 är alltså ungefär lika stor i Sverige år 2012 som år 2006 (cirka 10–12 procentenheter), vilket framgår i Figur 7 genom att dessa två streckade linjer följer varandra tätt. Uthålligheten i mätningen år 2006 tycks dock vara större än i mätningen år 2003, vilket framkommer i Tabell 5.

Genomsnittet av andelen elever som lämnar uppgifterna obesvarade för OECD som helhet tycks snarare ha minskat något även mellan dessa år. Även för dessa år tycks uthålligheten utifrån obesvarade matematikuppgifter vara något mindre bland de svenska eleverna än för OECD-länderna i genomsnitt.

Sammanfattning av elevernas svarsmönster på matematikuppgifter

De svenska elevernas svarsmönster på matematikuppgifter stödjer hypotesen om att det är kunskapsnivån som sjunkit fram till PISA 2012.

Analysen har visat att andelen elever som besvarar uppgifter i matematik korrekt har sjunkit både när dessa uppgifter kommer först och när de kommer sist i provhäftet mellan såväl PISA 2003 och 2012, som mellan PISA 2006 och 2012. Den ger också indikationer på att andelen elever som hoppar över uppgifter ökar mellan PISA 2006 och PISA 2012, och att uthålligheten mätt som andelen elever som lämnar uppgifterna obesvarade, är oförändrad mellan dessa mätningar. Dessa förändringar är något mindre än de för andelen elever som besvarar uppgifterna korrekt.

Elevernas svarsmönster på trenduppgifter i läsförståelse

I följande avsnitt presenteras elevernas svarsmönster på uppgifter i läsförståelse. Först presenteras jämförelsen av PISA 2009 och PISA 2012, följt av jämförelsen av PISA 2003 och PISA 2009.¹⁷ Samtliga skillnader för elevernas svarsmönster för uppgifter i läsförståelse presenteras i Tabell 6 och Tabell 7 i bilagan.

Elevernas svarsmönster för korrekt besvarade trenduppgifter i läsförståelse i PISA 2009 och PISA 2012

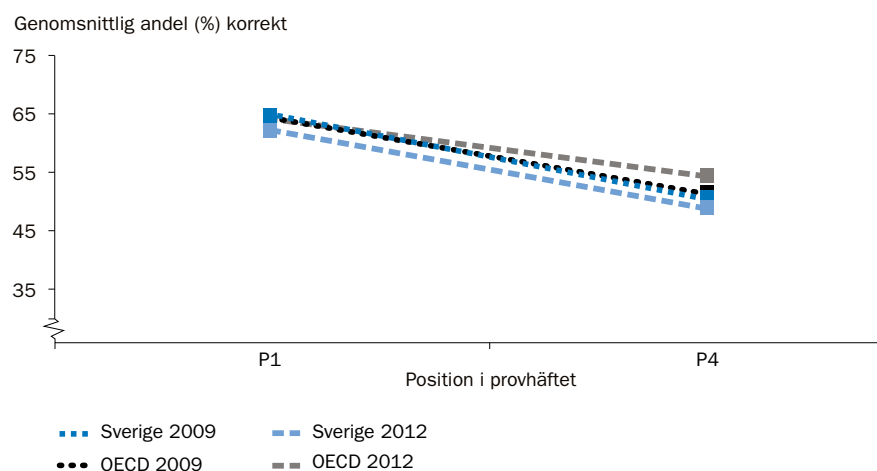
Elevernas svarsmönster för andelen elever som besvarar läsuppgifterna korrekt i PISA 2009 och PISA 2012 visas i Figur 8. Det finns ingen skillnad i uthållighet mätt i den genomsnittliga andelen elever som besvarar läsuppgifterna korrekt mellan dessa två år. I mätningen år 2012 är skillnaden i andelen elever som besvarar uppgifterna korrekt mellan positionerna cirka 13 procentenheter, och motsvarande siffra för mätningen år 2009 är cirka 14,5 procentenheter.

Andelen svenska elever som besvarar dessa uppgifter korrekt är i genomsnitt cirka 2–3 procentenheter lägre i mätningen år 2012 än år 2009 oavsett position. Detta sammantaget ger stöd åt hypotesen om en minskande kunskapsnivå.

Förändringarna i de svenska elevernas svarsmönster för andelen korrekta uppgifter i läsförståelse mellan år 2009 och år 2012 tycks avvika något från det mönster som ses för genomsnittet av eleverna i OECD. Där ser vi istället en ökning av andelen korrekt besvarade uppgifter på cirka 3 procentenheter när uppgifterna kommer sist i provhäftet, och ingen förändring när uppgifterna kommer först i provhäftet. Det betyder att uthålligheten uttryckt som andelen elever som besvarar läsuppgifterna korrekt är något större i OECD som genomsnitt år 2012 än år 2009. Det tyder på att varken kunskapsnivån eller engagemanget i OECD i genomsnitt har minskat.

Figur 8. Resultat från analysen av trenduppgifter i läsförståelse i PISA 2009 och PISA 2012: Korrekt besvarade uppgifter.

Genomsnittlig andel elever som besvarar trenduppgifterna i läsförståelse korrekt när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.



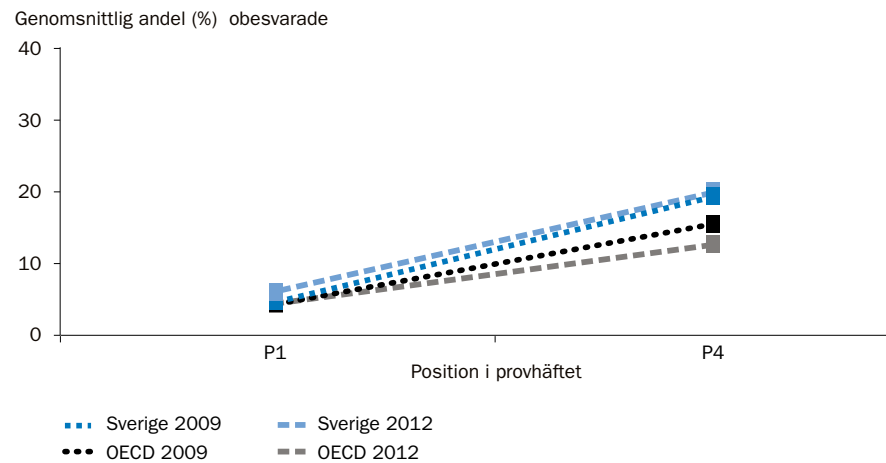
¹⁷ Kom ihåg att läsuppgifterna för 2009 delas upp i två grupper beroende på vilket år vi jämför med, eftersom alla trenduppgifter är olika i läsförståelse mellan 2006 och 2012.

Elevernas svarsmönster för obesvarade trenduppgifter i läsförståelse i PISA 2009 och PISA 2012

Andelen svenska elever som lämnar uppgifterna i läsförståelse obesvarade ökar något från år 2009 till år 2012, men främst i position 1. Dessa svarsmönster visas i Figur 9. Också dessa förändringar är små, endast runt en procentenhet. Uthålligheten för de svenska eleverna uttryckt som förändringen av andelen elever som lämnar läsuppgifterna obesvarade är cirka 14–15 procentenheter vid de båda mätillfällena, alltså ungefär som för andelen elever som besvarar uppgifterna korrekt. Sammantaget ger det visst stöd för hypotesen om sjunkande kunskapsnivå. De svenska elevernas uthållighet avviker något från uthålligheten i OECD i genomsnitt, som tycks vara något större, och också ökar något mellan mätningen år 2009 till mätningen år 2012.

Figur 9. Resultat från analysen av trenduppgifter i läsförståelse i PISA 2009 och PISA 2012: Obesvarade uppgifter.

Genomsnittlig andel elever som hoppar över trenduppgifter i läsförståelse när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.



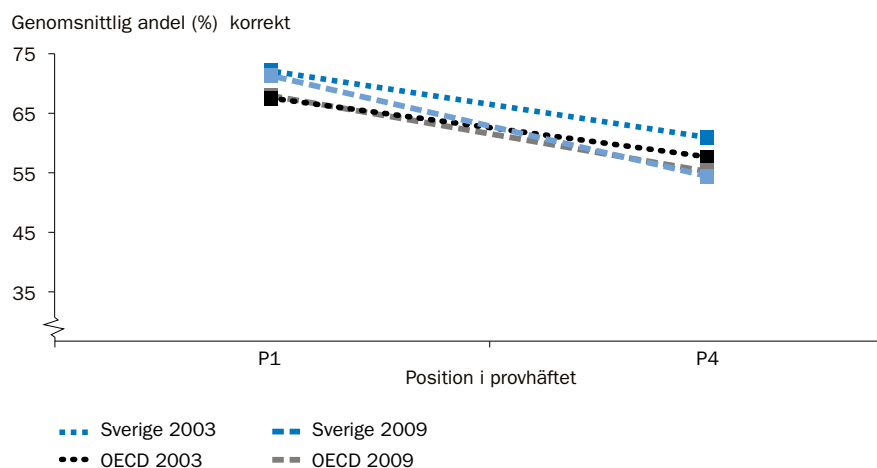
Elevernas svarsmönster för korrekt besvarade trenduppgifter i läsförståelse i PISA 2003 och PISA 2009

Förändringen av elevernas svarsmönster för de korrekt besvarade läsuppgifterna ser något annorlunda ut mellan dessa mätningar än i jämförelsen mellan år 2012 och år 2009, samt även jämfört med analyserna av matematikuppgifter. I Figur 10 visar de streckade linjerna på en minskande uthållighet mellan mätningen år 2003 och mätningen år 2009, men från samma nivå för den genomsnittliga andelen elever som besvarar uppgifterna korrekt i början av provhäftet. Det vill säga, andelen elever som besvarar uppgifterna korrekt är oförändrad när läsuppgifterna ligger först i provhäftet, men däremot minskar denna andel när uppgifterna ligger sist. Detta mönster liknar det som presenterats i Figur 2, nämligen scenariot om ett minskande engagemang att skriva kunskapsprovet. Skillnaden mellan position 1 och position 4 är cirka 11 procentenheter i mätningen år 2003, och cirka 17 procentenheter i mätningen år 2009.

Ett liknande mönster syns också för OECD-genomsnittet, men med en något mindre nedgång i elevernas uthållighet till mätningen 2009.

Figur 10. Resultat från analysen av trenduppgifter i läsförståelse i PISA 2003 och PISA 2009: Korrekt besvarade uppgifter.

Genomsnittlig andel elever som besvarar trenduppgifterna i läsförståelse korrekt när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.



Elevernas svarsmönster för obesvarade uppgifter i läsförståelse i PISA 2003 och PISA 2009

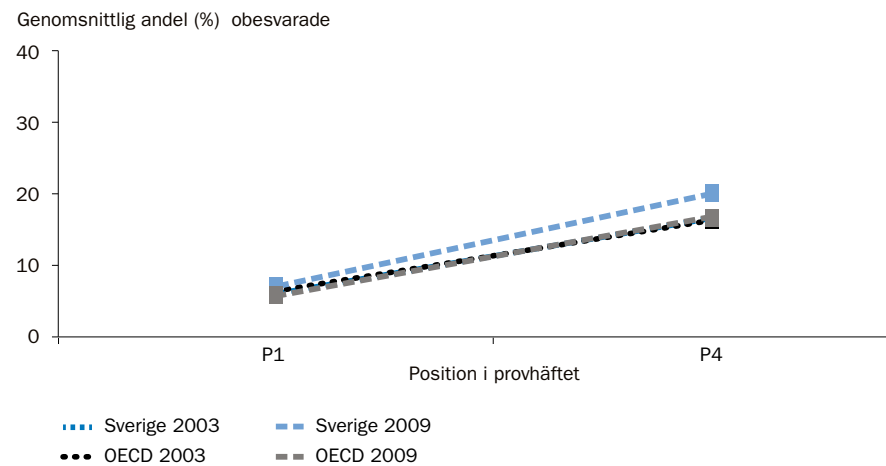
Andelen elever som lämnar uppgifterna i läsförståelse obesvarade tycks öka något även mellan PISA 2003 och PISA 2009, och här är skillnaderna statistiskt säkerställda både när uppgifterna kommer tidigt och sent i provhäftet. Dessa svarsmönster visas i Figur 11. Även dessa förändringar är relativt små; en ökning på runt 1 procentenhet när uppgifterna ligger först i provhäftet och runt 3 procentenheter när uppgifterna ligger sist.

Uthålligheten för de svenska eleverna uttryckt som förändringen av andelen elever som lämnar läsuppgifterna obesvarade är cirka 11–13 procentenheter vid de båda mättillfällena. Detta mönster ger stöd åt hypotesen om minskande kunskap.

I OECD i genomsnitt tycks överlag inga förändringar ha skett i elevernas svarsmönster för andelen obesvarade uppgifter mellan PISA 2003 och PISA 2009.

Figur 11. Resultat från analysen av trenduppgifter i läsförståelse i PISA 2003 och PISA 2009: Obesvarade uppgifter.

Genomsnittlig andel elever som hoppar över trenduppgifter i läsförståelse när uppgifterna kommer först respektive sist i provhäftet i respektive mätning.



Sammanfattning av elevernas svarsmönster på läsuppgifter

De svenska elevernas svarsmönster på uppgifter i läsförståelse stödjer främst hypotesen om sjunkande kunskapsnivå i de senare mätningarna. Andelen elever som besvarar läsuppgifterna korrekt minskar mellan mätningarna år PISA 2009 och år 2012. Dessa förändringar är i princip jämnstora oavsett om uppgifterna kommer först eller sist i provhäftet. Uthålligheten har därmed inte förändrats mellan åren. Samma mönster framkommer också enligt andelen elever som lämnar uppgifterna obesvarade mellan PISA 2003 och PISA 2009, samt i viss mån också mellan PISA 2009 och PISA 2012. Den genomsnittliga andelen elever som lämnar uppgifterna obesvarade ökar något oavsett om uppgifterna kommer först eller sist i provhäftet, tillsammans med en oförändrad uthållighet. Dessa förändringar är dock något mindre än motsvarande förändringar i analysen av matematikuppgifter. Detta mönster skiljer sig något från det vi ser i OECD som genomsnitt, där resultaten snarare tyder på en ökad kunskapsnivå eller ett ökat engagemang mellan dessa två år.

För läsförståelse framträder också en delvis annan bild av elevernas svarsmönster än den för matematikuppgifter. Förändringen av andelen elever som besvarar uppgifterna korrekt mellan PISA 2003 och 2009 tyder på ett minskande engagemang att skriva kunskapsprovet mellan dessa år. Mellan dessa mätningar är andelen elever som besvarar uppgifterna korrekt oförändrad när läsuppgifterna ligger först i provhäftet, medan den minskar när uppgifterna ligger sist. Det talar för att uthålligheten minskat från PISA 2003 till 2009. Det bör dock noteras att en liknande, om än något mindre, förändring även framkommer i OECD i genomsnitt.

Bilden är dock inte entydig. Förändringarna i andelen elever som lämnar uppgifterna obesvarade följer inte detta mönster. Uthålligheten enligt detta mått är oförändrat mellan såväl PISA 2003 och 2009, som mellan PISA 2009 och 2012 bland de svenska eleverna. För OECD-genomsnittet är uthålligheten oförändrad mellan år 2003 och år 2009, men ökar något mellan år 2009 och år 2012.

Elevernas svarsmönster uppdelat efter flervalsfrågor och frågor med öppna svar

I denna kompletterande analys delar vi upp elevernas svarsmönster utifrån om frågorna är flervalsfrågor eller öppna frågor.¹⁸ Vi tänker oss att tröskeln att besvara en flervalsfråga med färdiga svarsalternativ är lägre för eleverna än tröskeln att börja formulera ett svar på egen hand.

Om stöd ska ges för hypotesen att eleverna inte har samma engagemang för och därmed inte anstränger sig lika mycket att skriva kunskapsprovet vid senare års undersökningar kan man förvänta sig att eleverna jämförelsevis i lägre grad besvarar öppna frågor korrekt vid senare mätningar än flervalsfrågor. Motsvarande, om eleverna lämnar jämförelsevis fler öppna frågor än flervalsfrågor obesvarade i senare studier, kan det tyda på att de inte längre anstränger sig i lika stor utsträckning som tidigare. Detta innebär att vi framförallt tittar på om

¹⁸ Notera att när vi delar upp provfrågorna i två grupper kan den statistiska osäkerheten bli större i jämförelserna av elevernas genomsnittliga svarsandelar, eftersom vi nu har färre provfrågor i varje grupp.

det finns skillnader i elevernas uthållighet mellan de två frågetyperna mellan två mätningar. Resultaten av denna analys presenteras i tabellerna 8–11 i bilagan.

Elevernas svarsmönster för korrekt besvarade uppgifter förändras inte av frågetyp

Enligt den generella analysen av matematikuppgifter framkommer att frågetypen inte verkar spela någon roll för andelen elever som besvarar uppgifterna korrekt. När vi jämför samtliga provfrågor i matematik mellan PISA 2003 och PISA 2012 jämför vi de genomsnittliga svarsandelarna både med och utan hänsyn tagen till frågornas svårighetsnivå. Vid en jämförelse av elevernas svarsmönster utan hänsyn till frågans svårighetsnivå finns en signifikant skillnad i förändringen i andelen elever som besvarar uppgifterna korrekt beroende på frågetyp. När vi sedan också tar hänsyn till frågornas svårighetsnivå och i vilken ordning frågan kommer inom blocket försvinner denna skillnad. Det indikerar att frågetypen som sådan inte spelar roll för om eleverna besvarar uppgifterna korrekt. Istället tycks det framförallt vara i vilken ordning frågan kommer i provhäftet och dess svårighetsnivå som spelar roll.¹⁹

Den nedgång vi redovisat för andelen svenska elever som korrekt besvarar matematikuppgifterna, som ger stöd åt hypotesen om sjunkande kunskapsnivå, tycks inte förändras när elevernas svarsmönster delas upp efter de två frågetyperna. Skillnaden mellan åren tycks visserligen vara något mer stabil för de öppna frågorna, men uthålligheten skiljer sig inte åt mellan frågetyperna inom samma år för någon av de tre undersökta PISA-mätningarna.

För uppgifter i läsförståelse framkommer i princip inga skillnader alls i den genomsnittliga andelen elever som besvarar uppgifterna korrekt mellan flervalsfrågor och frågor med öppna svar. Den enda skillnad som framkommer är att den nedgång i uthållighet, enligt andelen elever som besvara uppgifterna korrekt, som noterats mellan PISA 2003 och PISA 2009, som alltså ger visst stöd åt hypotesen om ett sjunkande engagemang, främst beror på en nedgång bland frågor med öppna svar.

Frågetypen spelar viss roll för andelen elever som lämnar uppgifter obesvarade

När vi jämför elevernas uthållighet för andelen elever som lämnar frågor obesvarade är denna överlag lägre för öppna frågor än för flervalsfrågor, både i Sverige som i OECD som genomsnitt. Bland överhoppade matematikuppgifter i PISA 2003 och PISA 2006 syns dock inte ett sådant mönster. Detta indikerar att frågetypen tycks spela en viss roll för andelen elever som hoppar över uppgifter. Däremot tycks det generellt inte vara så att elevernas uthållighet minskar mer för öppna frågor än flervalsfrågor mellan åren, vilket inte ger stöd åt hypotesen om minskande engagemang.

Två saker kring elevernas svarsmönster för överhoppade matematikuppgifter kan dock kommenteras. Den första berör jämförelsen mellan PISA 2003 och

19 När vi i en regressionsmodell undersöker sambandet mellan frågornas svårighetsnivå och dess frågetyp för matematikfrågor i PISA 2003 och PISA 2012 tycks frågetypen inte påverka hur svår frågan upplevs i PISA 2003, medan ett sådant samband finns i PISA 2012. I den senare mätningen är öppna frågor svårare för eleverna än flervalsfrågor.

PISA 2012. Det framkommer inga skillnader mellan flervalfrågor och öppna frågor i den generella analysen mellan andelen elever som lämnar matematikuppgifterna obesvarade när dessa kommer först i provhäftet. Däremot tycks eleverna i något större utsträckning hoppa över öppna frågor än flervalfrågor när dessa kommer sist i provhäftet när vi enbart jämför trendfrågor. Det skulle kunna tyda på ett minskat engagemang hos eleverna i PISA 2012 jämfört med PISA 2003. Skillnaderna i uthållighet mellan dessa mätningar är dock inte signifikant skilda från varandra, vilket gör att de mönster som framkommer inte ger ett tydligt stöd åt hypotesen om sjunkande engagemang.²⁰

Det andra som bör kommenteras berör jämförelsen mellan PISA 2006 och PISA 2012. De svenska elevernas uthållighet minskar mellan dessa två mätningar för andelen elever som lämnar öppna frågor obesvarade, men inte för flervalfrågor. Det innebär att denna analys ger visst stöd för att elevernas engagemang mellan dessa år kan ha minskat. Det bör dock noteras att denna förändring inte får stöd i andelen korrekt besvarade uppgifter, och det är omöjligt att uteslutande avgöra om detta beror på minskat engagemang eller på om elever med lägre förmåga också blir tröttare än tidigare och därför i högre grad sorterar bort svåra uppgifter.

Sammanfattning av elevernas svarsmönster uppdelat efter frågetyp

När vi delar in elevernas svarsmönster i de två frågetyperna tycks resultaten överlag inte förändras. De skillnader som framträder ger sammantaget inget ytterligare stöd för hypotesen om att engagemanget att skriva PISAs kunskapsprov har minskat i de senare studierna.

De resultat vi redovisat för andelen svenska elever som korrekt besvarar uppgifterna tycks inte förändras när elevernas svarsmönster delas upp efter de två frågetyperna, varken för matematikuppgifter eller för uppgifter i läsförståelse.

För andelen elever som hoppar över uppgifter tycks frågetypen spela en viss roll. Däremot tycks det generellt inte vara så att elevernas uthållighet minskar mer för öppna frågor än flervalfrågor mellan åren. Analysen av elevernas svarsmönster för överhoppade frågor uppdelat per frågetyp ger därmed heller inte generellt något ytterligare stöd för att engagemanget har minskat mellan åren.

Det bör noteras att denna analys visar att elevernas uthållighet minskar för andelen elever som lämnar öppna matematikfrågor obesvarade mellan PISA 2006 och PISA 2012, medan motsvarande inte gäller för flervalfrågor i matematik. Det är dock svårt att uteslutande avgöra om dessa förändringar beror på minskat engagemang eller på om elever med lägre förmåga också blir tröttare än tidigare och därför i högre grad sorterar bort svåra uppgifter.

20 Det spelar roll här att öppna frågor har en högre svårighetsgrad än flervalfrågor i PISA 2012, medan ett sådant samband inte finns i PISA 2003. Se föregående fotnot. Dessutom tyder jämförelsen av elevernas svarsmönster på trenduppgifter för dessa år snarare på att skillnaden beror på att andelen elever i *lägre utsträckning än tidigare hoppar över flervalfrågor*, än att de i större utsträckning än tidigare hoppar över öppna frågor. Det tyder på att skillnaden mellan svarsandelarna för öppna frågor och flervalfrågor snarare handlar om frågornas svårighetsnivå 2012, än frågetypen. Sammantaget tycks förändringarna inte ge stöd för att engagemanget skulle minska mellan 2003 och 2012.

Sammanfattande diskussion

I den här studien undersöker vi svenska elevers svarsmönster på PISAs kunskapsprov och om dessa förändrats mellan de olika PISA-undersökningarna. Syftet är att undersöka om elevernas engagemang för att skriva kunskapsprovet i PISA förändrats mellan åren, och i förlängningen om de svenska elevernas resultat i PISA är tillförlitliga.

I detta avslutande kapitel presenteras den övergripande bilden som framträder av elevernas svarsmönster och stödet för de två hypoteser som prövas i studien sammanställs. Den första hypotesen säger att resultatnedgången i PISA beror på en sjunkande kunskapsnivå hos eleverna medan den andra hypotesen säger att resultatnedgången beror på ett minskat engagemang bland eleverna för att skriva kunskapsprovet.

Förändringar i elevernas uthållighet mellan åren

För samtliga mätningar som undersöks i denna analys ser vi att elevernas uthållighet avtar under provets gång. Det gäller för såväl Sverige som för OECD-genomsnittet. Det innebär att andelen elever som besvarar uppgifterna korrekt är högre om uppgifterna kommer tidigt jämfört med sent i provet, och omvänt att andelen elever som lämnar uppgifterna obesvarade är lägre när uppgifterna kommer tidigt jämfört med sent.

När vi jämför uthålligheten på matematikuppgifter mellan PISA 2003 och 2012 är den enligt den generella analysen oförändrad både utifrån andelen elever som besvarar uppgifterna korrekt och utifrån andelen elever som lämnar uppgifter obesvarade. Enligt analysen av enbart trenduppgifter tycks dock uthålligheten uttryckt som andelen elever som lämnar uppgifterna obesvarade ha ökat i PISA 2012. Uthålligheten i PISA 2006 är större än i PISA 2003, såväl när det gäller skillnaden i andel elever som har besvarat uppgifterna korrekt och som lämnat dem obesvarade. Däremot skiljer sig inte uthålligheten i PISA 2006 från 2012 i något avseende. Sammantaget tycks uthålligheten vara minst i PISA 2003, och högst i PISA 2006, medan uthålligheten i PISA 2012 hamnar någonstans däremellan.

Detta resultat skiljer sig något från mönstret bland OECD-länderna i genomsnitt. Där är uthålligheten i PISA 2012 större än i PISA 2003. Dessutom tycks uthålligheten utifrån matematikuppgifterna i PISA 2012 vara lägre bland de svenska eleverna än för OECD-länderna i genomsnitt.

Elevernas uthållighet i läsförståelse är överlag oförändrad i Sverige, både uttryckt som andelen elever som besvarar uppgifterna korrekt och som andelen elever som lämnar uppgifterna obesvarade. Bilden är dock inte entydig. Uthålligheten enligt andel elever som besvarar uppgifterna korrekt minskar mellan PISA 2003 och 2009. Denna nedgång indikerar att elevernas engagemang kan ha sjunkit mellan dessa år.

Denna avvikelse, med minskande uthållighet i andelen elever som besvarar uppgifter i läsförståelse korrekt mellan 2003 och 2009 framkommer också för OECD-genomsnittet.

För Sveriges del avviker dock elevernas svarsmönster i läsförståelse något från det som noteras i genomsnitt för alla elever i OECD när vi jämför PISA 2009 med PISA 2012. För OECD i genomsnitt ökar uthålligheten mellan

dessas år, medan uthålligheten i läsförståelse ligger på en relativt stabil nivå i Sverige mellan åren.

I analysen där vi delar upp elevernas svarsmönster per frågetyp framkommer ingen skillnad i elevernas uthållighet mellan öppna frågor och flervalsfrågor för andelen elever som besvarar uppgifterna korrekt i någon av de undersökta mätningarna.

Uthållighet för andelen elever som lämnar frågor obesvarade är överlag lägre för öppna frågor än för flervalsfrågor, både i Sverige som i OECD som genomsnitt. Bland överhoppade matematikuppgifter i PISA 2003 och PISA 2006 syns dock inte ett sådant mönster. Däremot tycks det generellt inte vara så att elevernas uthållighet minskar mer för öppna frågor än flervalsfrågor mellan åren.

Tydligt stöd för hypotesen om sjunkande kunskapsnivå i PISA

Stöd för hypotesen om sjunkande kunskapsnivå ser vi i många av de jämförelser som görs, men särskilt när vi betraktar förändringen av andelen elever som besvarar uppgifterna korrekt. För Sverige framkommer en tydlig nedgång i andelen elever som besvarar uppgifter i matematik korrekt både när dessa uppgifter kommer först och när de kommer sist i provhäftet mellan såväl PISA 2003 och 2012, som mellan PISA 2006 och 2012. Förändringarna i andelen korrekta svar är jämnstora oavsett placering av uppgifterna. Det ger ett tydligt stöd åt hypotesen om sjunkande kunskapsnivå i matematik. Analysen ger också indikationer på att andelen elever som hoppar över uppgifter ökar mellan PISA 2006 och 2012. När uppgifterna kommer tidigt i provhäftet är uppgången mellan åren visserligen inte statistiskt säkerställd, men uthållighet är oförändrad mellan mätningarna. Det ger ytterligare, om än något svagare, stöd åt denna hypotes.

Ser vi till uppgifter i läsförståelse stödjer de svenska elevernas svarsmönster främst hypotesen om sjunkande kunskapsnivå i de senare mätningarna. I jämförelsen mellan PISA 2009 och 2012 framkommer att det har skett en jämn nedgång i andelen elever som besvarar uppgifterna korrekt, om än i något mindre utsträckning än i jämförelsen av matematikuppgifter. Samma mönster framkommer också enligt andelen elever som lämnar uppgifterna obesvarade mellan PISA 2003 och PISA 2009, samt i viss mån också mellan PISA 2009 och PISA 2012. Den genomsnittliga andelen elever som lämnar uppgifterna obesvarade ökar något oavsett om uppgifterna kommer först eller sist i provhäftet, tillsammans med en oförändrad uthållighet. Dessa resultat står sig överlag när vi delar upp elevernas svarsmönster per frågetyp.

Visst stöd för hypotesen om minskat prov-engagemang mellan PISA 2003 och PISA 2009

I jämförelsen av andelen elever som besvarar läsuppgifter korrekt mellan PISA 2003 och 2009 ser vi ett svarsmönster som ger stöd för hypotesen om ett minskande engagemang hos eleverna. Mellan dessa år är andelen elever som besvarar uppgifterna korrekt oförändrad när läsuppgifterna ligger först i provhäftet, men minskar när uppgifterna ligger sist, vilket alltså innebär att uthålligheten minskat från PISA 2003 till 2009. Det bör dock noteras att en liknande, om än något mindre, förändring även framkommer i OECD i genomsnitt.

I analysen där vi delar upp eleverna svarsmönster efter flervalfrågor och frågor med öppna svar tycks de svenska elevernas uthållighet minskat uttryckt som andelen elever som lämnar öppna frågor obesvarade mellan PISA 2006 och PISA 2012. Motsvarande gäller inte för flervalfrågor. Det innebär att denna analys ger visst stöd för att elevernas engagemang mellan dessa år kan ha minskat.

Det bör dock noteras att det är svårt att uteslutande avgöra om dessa förändringar beror på minskat engagemang eller på om elever med lägre förmåga också blir tröttare än tidigare och därför i högre grad sorterar bort uppgifter som uppfattas som svårare.

Några avslutande ord

De svenska elevernas svarsmönster tyder främst på att det är kunskapsnivån som har sjunkit under 00-talet. Denna slutsats grundar sig på att de mönster som stödjer denna hypotes är entydiga; det finns en tydlig nedgång i andelen elever som besvarar uppgifterna korrekt i såväl matematik som läsförståelse oavsett om uppgifterna kommer tidigt eller sent i provhäftet mellan de flesta jämförda PISA-mätningarna. Även förändringar i svarsmönster bland andelen elever som lämnar uppgifter obesvarade stödjer denna hypotes. Dessa förändringar är dock oftast något mindre och något mer svårbedömda.

Det bör dock inte förbises att det även finns vissa mönster i svaren som stödjer hypotesen om ett sjunkande provengagemang. Detta gäller framförallt jämförelsen av svarsmönster för uppgifter i läsförståelse mellan PISA 2003 och 2009, baserat på nedgången i andelen elever som besvarar uppgifterna korrekt. En försvagande omständighet för detta är att andelen elever som lämnar läsuppgifterna obesvarade istället ger stöd åt hypotesen om en sjunkande kunskapsnivå, med en liten men jämn ökning oavsett om frågan kommer tidigt eller sent i provhäftet och med oförändrad uthållighet mellan åren.

Studien ger inget stöd för att uthålligheten har minskat mellan mätningarna vilket tyder på att elevernas engagemang knappast har förändrats, och därmed heller inte kunnat bidra till ett sjunkande resultat i någon större utsträckning. I denna analys kan förändringar i elevernas svarsmönster inte kopplas direkt till resultatet på PISA-provet. Det går därför inte att uppskatta hur stor en sådan del är uttryckt i antal eller andel poäng av resultatnedgången. Det tydligaste stöd som analysen visar för minskad uthållighet finns för korrekt besvarade uppgifter i läsförståelse. Detta resultat kan betyda att engagemanget särskilt sviktar för uppgifter i läsförståelse, eller så uppfattas läsuppgifterna som svårare än matematikuppgifterna när de kommer i slutet av ett provhäfte. Detta kan dock inte denna analys ge svar på.

De svenska elevernas uthållighet tycks generellt vara mindre än för eleverna i OECD i genomsnitt, vilket i sin tur tyder på att de svenska eleverna generellt skulle kunna vara mindre engagerade än OECD-genomsnittet. Eftersom analysen av elevernas svarsmönster främst tyder på att det är elevernas kunskapsnivå som har sjunkit ska detta resultat dock förstås i ljuset av att elevernas kunskapsnivå och deras uthållighet kan förstärka varandra på så sätt att elevernas uthållighet kan vara mindre då elevernas kunskapsnivå är lägre.

Sammantaget finns visst stöd för ett minskat provengagemang hos svenska elever. Dessa resultat är dock inte entydiga. Den samlade bedömningen av resultaten är att denna eventuella minskning i engagemang inte skulle kunna

förklara någon större del av den nedgång i resultat som svenska elever har uppvisat under det senaste decenniet. Som tidigare nämnts går det däremot utifrån denna analys inte att uppskatta hur stor, uttryckt i antal eller andel poäng av resultatnedgången, en sådan del är.

Denna fördjupningsanalys gör inte anspråk på att ge ett helt uttömmande svar frågan om det är engagemanget eller kunskapsnivån hos eleverna som sjunker vilket resulterat i ett försämrat resultat i PISA. Däremot är det en pusselbit bland flera om elevers motivation och ansträngning. Fler studier på detta och närliggande områden, som exempelvis jämförelser med liknande länder eller en komplettering med uppgifter i naturvetenskap som är huvudämne i PISA 2015, kan bidra med ytterligare kunskap. I analysen framkommer ingenting som tyder på att resultatnedgången som PISA visat inte är en rättvis bild av elevernas kunskaper inom de områden som PISA mäter.

Referenser

Cosgrove, J., Cartwright, F., (2014) Changes in achievement on PISA: the case of Ireland and implications for international assessment practice. *Large-scale Assessments in Education*, 2:2

OECD (2014) PISA 2012 *Technical Report*. OECD Publishing, Paris.

Skolverket (2006) *Med fokus på matematik och naturvetenskap. En analys av skillnader och likheter mellan internationella jämförande studier och nationella kursplaner*. Skolverkets aktuella analyser. Stockholm: Skolverket.

Skolverket (2013a) *PISA 2012: 15-åringars kunskaper i matematik, läsförståelse och naturvetenskap*. Rapport 398. Stockholm: Skolverket

Skolverket (2013b) *PISA 2012: 15-åringars kunskaper i matematik, läsförståelse och naturvetenskap. Resultaten i koncentrat*. Rapport 398. Stockholm: Skolverket

Skolverket (2013c). *Kommunernas resurstilldelning till grundskolor*. Rapport 391. Stockholm: Skolverket

Skolverket (2015). *Med fokus på matematik – Analys av samstämmighet mellan svenska styrdokument och den internationella studien PISA*. Stockholm: Skolverket

Bilaga

Tabell 4. Skattningar av elevernas svarsmönster i matematik, mellan positioner mellan år.

Sammanställning av förändringar i elevers svarsmönster på matematikuppgifter när blocket ligger i position 1 (P1) respektive i position 4 (P4) i provhäftet för studierna 2003, 2006 och 2012. Siffror i rött indikerar att förändringarna stödjer hypotesen om sjunkande kunskap. Inga förändringar i tabellen stödjer hypotesen om minskande engagemang.

Sverige		Förändring P1 mellan år		Förändring P4 mellan år	
		2003/2012	2006/2012	2003/2012	2006/2012
Andel korrekt	Generell analys, utan kontroll	-7,5		-5,9	
	Generell analys, kontroll†	-6,9		-5,3	
	Analys av trendfrågor	-4,7	-2,6	-2,8	-5,0
Andel obesvarat	Generell analys, utan kontroll	2,5		1,4	
	Generell analys, kontroll†	2,2		1,1	
	Analys av trendfrågor	1,5	0,6	-2,9	2,9
OECD		2003/2012	2006/2012	2003/2012	2006/2012
Andel korrekt	Generell analys, utan kontroll	-2,9		0,6	
	Generell analys, kontroll†	-2,3		1,1	
	Analys av trendfrågor	-1,0	0,5	2,7	1,2
Andel obesvarat	Generell analys, utan kontroll	-0,2		-5,7	
	Generell analys, kontroll†	-0,4		-5,9	
	Analys av trendfrågor	-0,7	-1,3	-8,0	-2,9

Om skillnaden i andelar för varje position mellan två år är signifikant skilda från varandra är dessa markerade med fet stil.

† Med kontroll för sekvens (uppgift 1–12), frågans svårighetsgrad och block.

Tabell 5. Skattningar av elevernas avtagande uthållighet på matematikuppgifter. Sammanställning av förändringar i elevers svarsmönster på matematikuppgifter mellan position 1 (P1) och position 4 (P4) inom samma år, för studierna 2003, 2006 och 2012. Inga förändringar i tabellen stödjer hypotesen om minskande engagemang.

Sverige		Förändring mellan P1 och P4 inom samma år		
		2003	2006	2012
Andel korrekt	Generell analys, utan kontroll	-11,0		-9,4
	Generell analys, kontroll†	-11,0		-9,4
	Analys av trendfrågor	-12,0^b	-7,7^c	-10,1
Andel obesvarat	Generell analys, utan kontroll	14,4		13,3
	Generell analys, kontroll†	14,4		13,3
	Analys av trendfrågor	16,7^{a, b}	10,0^c	12,3^c
OECD		2003	2006	2012
Andel korrekt	Generell analys, utan kontroll	-9,5		-6,0
	Generell analys, kontroll†	-9,5^a		-6,0^c
	Analys av trendfrågor	-10,2^{a, b}	-7,1^c	-6,4^c
Andel obesvarat	Generell analys, utan kontroll	12,8^a		7,2^c
	Generell analys, kontroll†	12,8^a		7,2^c
	Analys av trendfrågor	14,1^{a, b}	8,4^c	6,8^c

Om skillnaden i andelar mellan position 1 och position 4 inom ett år (uthålligheten) är signifikant skilda från varandra är dessa markerade med fet stil.

† Med kontroll för sekvens (uppgift 1–12), frågans svårighetsgrad och block.

a. Skild från 2012.

b. Skild från 2006.

c. Skild från 2003.

Tabell 6. Skattningar av elevernas svarsmönster i läsförståelse, mellan positioner mellan år.

Sammanställning av förändringar i elevers svarsmönster på läsuppgifter när blocket ligger i position 1 (P1) respektive i position 4 (P4) i provhäftet för studierna 2003, 2009 och 2012. Siffror i rött indikerar att förändringarna stödjer hypotesen om sjunkande kunskap. Siffror i blått indikerar att förändringarna stödjer hypotesen om minskande engagemang.

Analys av trendfrågor	Förändring P1 mellan år		Förändring P4 mellan år	
	2003/2009 ^a	2009 ^b /2012	2003/2009 ^a	2009 ^b /2012
Sverige				
Andel korrekt	-0,8	-2,7	-6,6	-1,6
Andel obesvarat	1,2	1,5	3,2	0,6
OECD				
Andel korrekt	0,4	-0,3	-2,6	3,0
Andel obesvarat	-0,5	0,1	0,3	-3,0

Om skillnaden i andelar för varje position mellan två år är signifikant skilda från varandra är dessa markerade med fet stil.

a. Jämför med år 2003 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012.

b. Jämför med år 2012 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012.

Tabell 7. Skattningar av elevernas avtagande uthållighet på läsuppgifter.

Sammanställning av förändringar i elevers svarsmönster på läsuppgifter mellan position 1 (P1) och position 4 (P4) inom samma år, för studierna 2003, 2009 och 2012. Siffror i blått indikerar att förändringarna stödjer hypotesen om minskande engagemang.

Analys av trendfrågor	Förändring mellan P1 och P4 inom samma år			
	2003	2009 ^a	2009 ^b	2012
Sverige				
Andel korrekt	-11,3^e	-17,1^f	-14,5	-13,4
Andel obesvarat	11,1	13,1	14,7	13,8
OECD				
Andel korrekt	-10,0^e	-13,0^f	-13,1^d	-9,8^e
Andel obesvarat	9,9	10,6	11,2^d	8,2^e

Om skillnaden i andelar mellan position 1 och position 4 inom ett år (uthålligheten) är signifikant skilda från varandra är dessa markerade med fet stil.

- Jämför med år 2003 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012.
- Jämför med år 2012 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012.
- Uthålligheten skild från 2012.
- Uthålligheten skild från 2009.
- Uthålligheten skild från 2003.

Tabell 8. Skattningar av elevernas svarsmönster i matematik, mellan positioner mellan år, uppdelat per frågetyp.

Sammanställning av förändringar i elevers svarsmönster på matematikuppgifter uppdelat per frågetyp när blocket ligger i position 1 (P1) respektive i position 4 (P4) i provhäftet för studierna 2003, 2006 och 2012. Siffror i rött indikerar att förändringarna stödjer hypotesen om sjunkande kunskap. Siffror i blått indikerar att förändringarna stödjer hypotesen om minskande engagemang.

		Skillnad i förändring mellan flervalsfrågor och frågor med öppna svarsalternativ							
		Förändring P1 mellan år				Förändring P4 mellan år			
		2003/2012		2006/2012		2003/2012		2006/2012	
		Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor
Sverige									
Andel korrekt	Generell analys, utan kontroll	2,1*	-13,9*			4,6*	-12,9*		
	Generell analys, kontroll†	-6,5	-7,1			-4,0	-6,1		
	Analys av trendfrågor	-3,6	-5,7	-0,9*	-4,1*	-0,2*	-5,3*	-3,2	-6,8
Andel obesvarat	Generell analys, utan kontroll	0,8	4,4			-3,9*	5,8*		
	Generell analys, kontroll†	3,2	2,4			-1,4	3,7		
	Analys av trendfrågor	1,0	1,9	0,6	0,6	-5,2*	-0,6*	-0,1*	5,7*
OECD		2003/2012		2006/2012		2003/2012		2006/2012	
Andel korrekt	Generell analys, utan kontroll	5,6*	-8,6*			9,2*	-5,3*		
	Generell analys, kontroll†	-2,5	-2,1			1,0	1,3		
	Analys av trendfrågor	-0,9	-1,1	0,4	0,6	3,2	2,3	1,8	0,5
Andel obesvarat	Generell analys, utan kontroll	-0,5	0,9			-7,1	-3,7		
	Generell analys, kontroll†	1,4	-0,8			-5,2	-5,4		
	Analys av trendfrågor	-0,3	-1,2	-0,3*	-2,3*	-7,8	-8,3	-2,8	-3,0

Om skillnaden i andelar mellan flervalsfrågor och öppna frågor för varje position mellan två år är signifikant skilda från varandra är dessa markerade med *.

Om skillnaden i andelar för varje position mellan två år är signifikant skilda från varandra är dessa markerade med fet stil.

† Med kontroll för sekvens (uppgift 1–12), frågans svårighetsgrad och block.

Tabell 9. Skattningar av elevernas avtagande uthållighet på matematikuppgifter, uppdelat per frågetyp.

Sammanställning av förändringar i elevers svarsmönster på matematikuppgifter uppdelat per frågetyp mellan position 1 (P1) och position 4 (P4) inom samma år, för studierna 2003, 2006 och 2012. Siffror i blått indikerar att förändringarna stödjer hypotesen om minskande engagemang.

Sverige		Skillnad i förändring mellan flervalsfrågor och frågor med öppna svarsalternativ					
		Förändring mellan P1 och P4 inom samma år					
		2003		2006		2012	
		Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor
Andel korrekt	Generell analys, utan kontroll	-11,5	-10,8			-9,0	-9,7
	Generell analys, kontroll†	-11,5	-10,8			-9,0	-9,7
	Analys av trendfrågor	-13,1^b	-10,9	-7,4^c	-7,9	-9,6	-10,6
Andel obesvarat	Generell analys, utan kontroll	14,0	14,6			9,3*	15,9*
	Generell analys, kontroll†	14,0	14,6			9,3*	15,9*
	Analys av trendfrågor	15,1^{a, b}	18,1^b	9,5^c	10,4^{a, c}	8,9^{*c}	15,6^{*b}
OECD		2003		2006		2012	
Andel korrekt	Generell analys, utan kontroll	-8,8	-9,9			-5,2	-6,6
	Generell analys, kontroll†	-8,8^a	-9,9^a			-5,2^c	-6,6^c
	Analys av trendfrågor	-10,0^a	-10,3^{a, b}	-7,4	-6,7^c	-5,9^c	-6,9^c
Andel obesvarat	Generell analys, utan kontroll	10,7^a	13,8			4,1 ^c	9,2
	Generell analys, kontroll†	10,7^a	13,8^a			4,1^{*c}	9,2^{*c}
	Analys av trendfrågor	11,7^{*a, b}	16,3^{*a, b}	6,7^{*a, c}	9,9^{*c}	4,2^{*b, c}	9,2^{*c}

Om skillnaden i andelar mellan flervalsfrågor och öppna frågor inom ett år (uthålligheten) är signifikant skilda från varandra är dessa markerade med *

Om skillnaden i andelar mellan position 1 och position 4 inom ett år (uthålligheten) är signifikant skilda från varandra är dessa markerade med fet stil.

a. Uthålligheten skild från 2012.

b. Uthålligheten skild från 2006.

c. Uthålligheten skild från 2003.

† Med kontroll för sekvens (uppgift 1–12), frågans svårighetsgrad och block.

Tabell 10. Skattningar av elevernas svarsmönster i läsförståelse, mellan positioner mellan år, uppdelat per frågetyp.

Sammanställning av förändringar i elevs svarsmönster på läsuppgifter uppdelat per frågetyp när blocket ligger i position 1 (P1) respektive i position 4 (P4) i provhäftet för studierna 2003, 2006 och 2012. Siffror i rött indikerar att förändringarna stödjer hypotesen om sjunkande kunskap. Siffror i blått indikerar att förändringarna stödjer hypotesen om minskande engagemang.

Skillnad i förändring mellan flervalsfrågor och frågor med öppna svarsalternativ								
Analys av trendfrågor	Förändring P1 mellan år				Förändring P4 mellan år			
	2003/2009 ^a		2009 ^b /2012		2003/2009 ^a		2009 ^b /2012	
Sverige	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor
Andel korrekt	-0,7	-0,9	-2,1	-3,2	-6,2	-6,9	-1,4	-1,8
Andel obesvarat	0,9	1,4	0,7*	2,1*	-1,4*	6,1*	0,3	0,8
OECD	2003/2009 ^a		2009 ^b /2012		2003/2009 ^a		2009 ^b /2012	
Andel korrekt	0,5	0,3	-0,2	-0,3	-3,1	-2,3	3,5	2,6
Andel obesvarat	-0,4	-0,5	0,1	0,1	-1,2*	1,2*	-2,7	-3,2

Om skillnaden i andelar mellan flervalsfrågor och öppna frågor för varje position mellan två år är signifikant skilda från varandra är dessa markerade med *.

Om skillnaden i andelar för varje position mellan två år är signifikant skilda från varandra är dessa markerade med fet stil.

a. Jämför med år 2003 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012.

b. Jämför med år 2012 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012.

Tabell 11. Skattningar av elevernas avtagande uthållighet på läsuppgifter, uppdelat per frågetyp.

Sammanställning av förändringar i elevers svarsmönster på läsuppgifter uppdelat per frågetyp mellan position 1 (P1) och position 4 (P4) inom samma år, för studierna 2003, 2009 och 2012. Siffror i blått indikerar att förändringarna stödjer hypotesen om minskande engagemang.

Skillnad i förändring mellan flervalsfrågor och frågor med öppna svarsalternativ								
Analys av trendfrågor	Förändring mellan P1 och P4 inom samma år							
	2003		2009 ^a		2009 ^b		2012	
	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor	Flervalsfrågor	Öppna frågor
Sverige								
Andel korrekt	-10,1	-12,1 ^d	-15,8	-17,9 ^e	-14,0	-14,9	-13,2	-13,5
Andel obesvarat	8,5*	12,7*	6,9*	17,0*	12,2*	16,8*	11,7*	15,5*
OECD								
Andel korrekt	-9,2 ^d	-10,5	-13,3 ^e	-12,8	-13,3 ^c	-12,9	-9,6 ^d	-10,0
Andel obesvarat	6,7*	11,8*	6,5*	13,2*	8,9* ^c	13,1* ^c	6,2* ^d	9,8* ^d

Om skillnaden i andelar mellan flervalsfrågor och öppna frågor inom ett år (uthålligheten) är signifikant skilda från varandra är dessa markerade med *.

Om skillnaden i andelar mellan position 1 och position 4 inom ett år (uthålligheten) är signifikant skilda från varandra är dessa markerade med fet stil.

- a. Jämför med år 2003 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012, och dessa två år jämförs således aldrig.
- b. Jämför med år 2012 – endast de trenduppgifterna inkluderade. Det är olika provfrågor som inkluderas för 2009 beroende på jämförelseår. Inga av uppgifterna i läsförståelse från proven 2003 finns med i proven 2012, och dessa två år jämförs således aldrig.
- c. Uthålligheten skild från 2012.
- d. Uthålligheten skild från 2009.
- e. Uthålligheten skild från 2003.

PISA resultaten från 2012 visade på en kraftig resultatnedgång i alla de tre kunskapsområden, läsförståelse, matematik och naturvetenskap, som ingår i PISA. En av de tänkbara delförklaringar som diskuterades var den kontext i vilken PISA-undersökningen genomförs. Mot bakgrund av denna resultatförsämring, har Skolverket bland annat initierat arbetet med de två studier som här presenteras. Studierna kompletterar varandra på så sätt att analysen i den ena baseras på elevernas svar på frågor om ansträngning och motivation att göra PISA-provet, medan analysen i den andra bygger på elevernas faktiska svarsmönster i PISA-proven. Även om resultaten från de båda analyserna ska tolkas med försiktighet förstärker de varandra. Oberoende av varandra visar analyserna att den kraftiga resultatnedgång som PISA visar inte kan förklaras av bristande motivation eller ansträngning hos eleverna.