

Sambandet mellom læsførståelseresultat i PISA og nasjonella prøv: En fråga om motivation i prøv där konsekvenserna är små respektive stora

Stefan Johansson^{1*}, Linda Borger¹ & Rolf Strietholt²

¹Göteborgs universitet, Sverige; ²TU Dortmund, Tyskland

Sammanfattning

Denna studie undersöker om det finns belegg för hypotesen att vissa elever inte gör sitt bästa på prøv som inte har några konsekvenser för dem. Ett sådant prøv är PISA som 15-åringar runt om i världen genomför utan att få några svar på hur väl de har presterat. Trots att PISA-resultaten kan ha en betydande inverkan på utbildningsreformer både i Sverige och globalt, saknas kunskap om hur motivationen påverkar elevernas prestationer. Urvalet bestod av 15-åriga svenska elever som deltog i PISA-prøvet 2018. Eleverna rapporterade även sin ansträngning på prøvet. PISA-data kopplades samman med registerdata, vilket möjliggjorde analyser av sambandet mellom PISA-resultat och nasjonella prøvbetyg samt hur detta varierade beroende på elevers ansträngningsnivå. Huvudsaklig analysmetod var regressionsanalys. Resultaten visar att svenska elever i stor utsträckning ansträngde sig på PISA-prøvet, men att de skulle ha ansträngt sig ännu mer om betyget hade påverkats. Høgt skattad ansträngning var positivt relaterad till bättre PISA-resultat, även efter kontroll av det nasjonella prøvbetyget i læsførståelse. Dessutom framkom att mer motiverade elever oppvisade ett sterkere samband mellom sina PISA-resultat og nasjonella prøvbetyg. Resultaten tyder på att en mindre gruppe elever underpresterade relativt sine nasjonella prøvresultat. Studien betonar vikten av att undersøke og forstå faktorer som påverkar sambandet mellom prøv med små konsekvenser og prøv med store konsekvenser.

Nyckelord: PISA; læsførståelse; nasjonella prøv; sjølvskattning; prøvmotivation; anstrængning

Abstract

The Relationship Between Reading Achievement in PISA and National Tests: A Question of Test Motivation in Low-Stakes Versus High-Stakes Assessments

The following study examines whether there is evidence for the hypothesis that students do not do their best on low-stakes tests. One such test is PISA, which 15-year-olds around the world take without receiving any feedback about how they performed. Despite the significance of PISA results for educational reforms in Sweden and globally, there is limited knowledge about the impact of motivation on student performance. The sample comprised 15-year-old Swedish students who participated in the 2018 PISA assessment. The students also reported their effort on the PISA test. PISA data was linked with registry data, enabling analyses of the relationship between PISA results and national test grades, and how this varied depending on the students' effort levels. Regression

*Korrespondens: Stefan Johansson, e-post: stefan.johansson@gu.se

© 2024 Stefan Johansson, Linda Borger & Rolf Strietholt. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

Citation: S. Johansson, L. Borger & R. Strietholt. "Sambandet mellom læsførståelseresultat i PISA og nasjonella prøv: En fråga om motivation i prøv där konsekvenserna är små respektive stora" Nordic Journal of Literacy Research: Temanummer om statlig måtning av læs- og skriveførståelse, Vol. 10(2), 2024, pp. 1–19. <http://doi.org/10.23865/njlr.v10.5302>

analysis was the main method of analysis. While most students exerted significant effort on the PISA assessment, they indicated a greater willingness to exert even more effort if the results had consequences for their grades. Higher effort was positively associated with better PISA performance, even after controlling for national test scores in reading comprehension. Additionally, it was found that more motivated students showed a stronger correlation between their PISA results and national test grades. The results suggest that a smaller group of students underperformed relative to their national test scores. These results emphasize the need to understand factors contributing to differential performance between low- and high-stakes tests.

Keywords: *PISA; national tests; reading comprehension; self-reported effort; test-taking motivation*

Ansvarig redaktör: Michael Tengberg

Mottagen: December, 2022; Antagen: November, 2023; Publicerat: Mars, 2024

Inledning

Internationella storskaliga kunskapsmätningar har fått en allt viktigare roll för att generera empirisk kunskap om utbildning samt för beslutsfattande och reformer på olika nivåer i samhället, både nationellt och internationellt (t.ex. Grek, 2009; Lindblad et al., 2018). En av de mest framträdande och inflytelserika kunskapsmätningarna är Programme for International Student Assessment (PISA), som genomförs av Organisationen för ekonomiskt samarbete och utveckling (OECD). PISA-mätningen görs vart tredje år och undersöker 15-åriga elevers kunskaper inom tre ämnesområden: läsförståelse, matematik och naturvetenskap. Mot bakgrund av den betydelse som kunskapsprov som PISA har på utbildningsdebatten och politiken är det av största vikt att resultaten är giltiga. Parallellt med ökningen av antalet kunskapsmätningar har det emellertid funnits en växande kritik som uttrycker en oro över PISA:s validitet. Kritiken berör bland annat frågor om urvalets representativitet (Jerrim, 2021), resultatens stora inflytande på nationella utbildningssystem samt risken för likriktning av utbildningssystemen mellan olika länder (Grek, 2009; Meyer & Benavot, 2013; Meyer et al., 2018).

Det förekommer en betydande mängd tidigare forskning som har undersökt validitetsfrågor i PISA angående innehållet som testas. Men det är relativt få studier som fokuserar på huruvida elever är motiverade att göra sitt bästa på PISA-proven jämfört med prov som har större personliga konsekvenser (så kallade high-stakes-prov), som exempelvis nationella prov. Att genomföra ett PISA-prov är ett ganska omfattande arbete där två timmar ägnas åt att skriva kunskapsprov och ytterligare 45 minuter åt att besvara en bakgrundsenkät. För många elever torde detta vara ansträngande och det krävs sålunda ett visst mått av uthållighet för att genomföra PISA-provet med gott resultat. Eleverna får ingen återkoppling på sina resultat och resultatet rapporteras inte för enskilda elever eller skolor utan endast på landsnivå. Det är rimligt att anta att alla elever inte är motiverade att göra sitt bästa på PISA-provet på grund av avsaknaden av konsekvenser. Om en sådan hypotes vinner stöd, är det möjligt att

slutsatserna som dras om elevernas kunskaper i Sverige och andra länder inte är fullt korrekta eftersom vissa elever underpresterar på PISA (jfr Wise & DeMars, 2010).

Huruvida elever i andra länder underpresterar är inte väl känt, men PISA-provet är inte "high-stake" för elever i något land. Dock finns det skäl att tro att motivationen till att göra PISA varierar mellan länder. Baserat på data från PISA 2018 visar De Bortoli (2021) exempelvis att elever i Sverige indikerade en större vilja att lägga ner mer ansträngning i PISA-provet om resultatet hade bidragit till deras betyg i skolan. Andelen som skulle ha lagt ner mer ansträngning i PISA om resultatet haft betydelse för betyget var 9 procentenheter högre än OECD-genomsnittet och 20–30 procentenheter högre än i östasiatiska länder som Korea och Taiwan. Dessa resultat tyder på att graden av seriositet kan variera över länder och att detta i förlängningen skulle kunna påverka resultatjämförelserna. Det är dock inte klart om de elever som uppger att de kunde ha presterat bättre faktiskt kan det och inte heller klart hur PISA-resultaten påverkas, eftersom adekvata data för att undersöka detta har saknats. Sverige är emellertid ett av få och kanske enda land där det går att relatera PISA-resultat till prov med större konsekvenser. Detta innebär att vi kan undersöka om eleverna tenderar att underprestera på PISA-provet relativt de nationella proven.

Trots att PISA-resultatet har begränsade konsekvenser för individen, kan det få betydande konsekvenser på policynivå och ligga till grund för utbildningsreformer i Sverige och internationellt. Dock råder det fortfarande en brist på kunskap om hur elevernas motivation att göra sitt bästa på provet påverkar resultatet. Denna studie syftar till att fylla denna kunskapslucka genom att belysa sambandet mellan läspoäng i PISA 2018, där konsekvenserna för den enskilde eleven är begränsade, och nationella provbetyg i läsförståelse, där konsekvenserna är mer betydande, samt undersöka hur elevers motivation kan påverka detta samband.

Teoretiska utgångspunkter: provmotivation och förväntan-värde-teorin

Begreppet *provmotivation* [eng. *test-taking motivation*] definieras enligt Baumert och Demmrich (2001) som "the willingness to engage in working on test items and to invest effort and persistence in this undertaking" (s. 441). I enlighet med denna definition kommer provdeltagarens motivationsnivå att reglera den ansträngning som de är villiga att lägga ner i den specifika provsituationen. *Provansträngning* kan således förstås som den insats en provdeltagare gör för att uppnå högsta möjliga resultat på provet (Wise & DeMars, 2005, s. 2). Provmotivation kan även ses som ett underbegrepp till den mer övergripande termen *prestationsmotivation* [eng. *achievement motivation*] (Pintrich & Schunk, 2002). Medan prestationsmotivation kan mätas på olika nivåer, är provmotivation kopplad till en specifik provsituation och handlar om provdeltagarens motivation att prestera bra på just det provet (Eklöf, 2010).

En teoretisk grund för föreliggande studie är förväntan-värde-teorin (Atkinson, 1957; Feather, 1982; Pintrich, 1989; Pintrich & Schunk, 2002; Wigfield & Eccles, 2000) som används för att förstå sambandet mellan ansträngning och prestation i

PISA. Enligt en modell av Wigfield och Eccles (2000), som bygger på förväntan-värde-teorin, beror prestationsmotivation på elevens förväntningar på att lyckas med en viss uppgift, värdet som eleven sätter på uppgiften och interaktionen mellan dem. Tron på att lyckas, som bland annat handlar om elevens uppfattningar om sin kompetens, och vikten som läggs vid uppgiften, exempelvis i form av intresse, personligt värde och kostnad, antas alltså påverka prestation, ansträngning och uthållighet. Eftersom low-stakes-prov inte har några direkta personliga konsekvenser kopplade till prestationen, är det enligt förväntan-värde-teorin rimligt att anta att provdeltagares värdeuppfattning och därmed ansträngning generellt är lägre i sådana prov.

Teorin om förväntan-värde har framgångsrikt använts som ramverk i tidigare forskning för att förklara och förstå provmotivationsbegreppet. Exempelvis jämförde Eklöf och Knekta (2017) olika urval av provdeltagare som skrivit prov med stora respektive små konsekvenser utifrån förväntan-värde-teorin. Resultaten visar att elevers självrapporterade ansträngning var signifikant och positivt korrelerad med provresultaten i båda provsituationerna. Dock var sambandet lägre i low-stakes-prov, där korrelationen mellan självskattad ansträngning och prestation exempelvis var $r = 0,27$ i PISA 2012 ($n = 4\,129$) jämfört med $r = 0,48$ för ett urval av elever som gjorde 2012 års nationella prov i NO-ämnen i årskurs 9 ($n = 536$). Dessutom förblev den rapporterade nivån av provmotivation en signifikant prediktor för prestation, även efter kontroll av andra domänspecifika motivationsvariabler och bakgrundsegenskaper.

Påverkar provmotivation resultaten i prov med små konsekvenser?

Att med bestämdhet hävda att elever underpresterar i PISA-provet eller på andra prov där konsekvenserna är små är besvärligt om inte adekvata data finns att tillgå. Få studier har haft tillgång till provresultat med stora respektive små konsekvenser för en och samma individ. Tidigare forskning i kontexten av PISA drar emellertid slutsatsen att elevers motivation och ansträngning vid genomförande av provet är en viktig faktor att ta hänsyn till i tolkningen av resultaten (t.ex. Butler & Adams, 2007; Debeer et al., 2014; Pools & Monseur, 2021; Zamarro et al., 2019). Eklöf och Hopfenbeck (2019) som undersökt elevers självrapporterade ansträngning i PISA 2012 fann att eleverna skulle ha ansträngt sig mer för att prestera bra i PISA om deras betyg hade påverkats av PISA-resultaten. Intressant nog hade PISA:s låginsatskaraktär ett större inflytande på svenska elevers motivationsnivå än på elever i något annat land.

Eklöf och Knekta (2017) har jämfört olika mått på självrapporterad motivation från prov med små respektive stora konsekvenser. Ett övergripande resultat är att provdeltagare generellt rapporterade en lägre grad av motivation och ansträngning i prov med små personliga konsekvenser jämfört med prov som hade större konsekvenser, som exempelvis att resultatet räknades till betyget. Wise och DeMars (2005) redovisade effektstorlekar från tolv empiriska studier där provresultat för grupper av elever med hög respektive låg motivation jämfördes. Resultaten indikerar att högmotiverade elever presterade i genomsnitt mer än en halv standardavvikelse högre

resultat jämfört med elever med lägre provmotivation. Flera studier har dessutom funnit ett positivt samband mellan provdeltagares självrapporterade motivation och ansträngning och deras prestation, men korrelationen tenderar att vara svag till måttlig (t.ex. Eklöf & Knekta, 2017; Wise & DeMars, 2005). En översikt av Silm et al. (2020) visar att den genomsnittliga korrelationen mellan självskattad ansträngning och provprestation i de granskade studierna är $r = 0,33$.

Ett möjligt problem med att mäta motivation med självskattning är att motivation kan vara relaterad till förmåga (Wise & DeMars, 2005). Om så är fallet kan elever som uppger högre motivation prestera bättre på ett prov inte bara tack vare sin högre motivation utan också sin högre förmåga. Det är därför viktigt att statistiskt kontrollera för förmåga när självskattade mått på motivation används för att undersöka om motivationen är kopplad till provsituationen eller om den är av mer generell art.

Sammanfattningsvis kan vi konstatera att det finns tydliga prestationsskillnader mellan elever med olika grad av motivation, vilket gäller såväl prov med som utan konsekvenser för individen. Det är dock fortfarande oklart om elever generellt presterar bättre på prov med stora konsekvenser jämfört med prov där konsekvenserna är ringa eller inga. En anledning till detta är avsaknaden av lämpliga data.

Syfte

Syftet med denna studie är att undersöka om det finns belägg för hypotesen att vissa elever underpresterar på prov med små konsekvenser, som exempelvis PISA-provet. Nyligen konstaterade Skolverket (2022) att sambandet mellan PISA-resultat och nationella provbetyg var måttligt högt, cirka 0,6. Även om PISA-prov och nationella prov tjänar olika syften och därför inte förväntas ha en perfekt överensstämmelse, finns det likheter vad gäller innehåll, begrepp och kognitiva processer. Trots att PISA-provet inte direkt kopplar till länders kursplaner, är samstämmigheten mellan läsning i PISA och kursplanen i svenska god (Johansson et al., 2019). PISA-provet 2018 utfördes dock på dator, medan de nationella proven genomfördes på papper samma år.¹ Texterna i det nationella provet är generellt längre än i PISA som även har en högre andel flervalsfrågor. Det bör också noteras att PISA:s läsförståelsedel har en adaptiv testmodell sedan 2018, vilket innebär att elever delvis får olika uppgifter baserat på i vilken mån de svarat rätt på tidigare uppgifter.

Eftersom eleverna saknar tydliga incitament till att prestera sitt bästa på PISA-provet, kan sambandet med det nationella provbetyget i läsförståelse till viss del förklaras av ansträngningsnivån. Motivationen att göra ett prov torde generellt vara högre för det nationella provet, vars resultat kan påverka betygsättningen. Således är vår hypotes att sambandet mellan läsförmåga i PISA-provet och betyg i det nationella provet i läsning är olika för olika nivåer av ansträngning att göra PISA. Om en elev inte har gjort sitt bästa på PISA, förutser vi ett lägre samband.

¹ Digitala nationella prov införs successivt i Sverige från och med 2024.

Studien syftar till att besvara följande forskningsfrågor:

1. Skulle eleverna ha ansträngt sig mer om PISA-resultatet hade haft betydelse för betygen?
2. Hur ser sambandet ut mellan PISA-resultatet i läsförståelse och det nationella provbetyget i läsförståelse för elever med olika ansträngningsnivå?

Material och metod

Deltagare

I studien ingick 5 504 svenska elever i åldern 15 år (50,2 procent flickor) som skrev PISA-provet 2018. Av de deltagande eleverna gick majoriteten i årskurs 9 ($n = 5\,334$), medan 108 elever gick i årskurs 8 och 62 elever i gymnasieskolans årskurs 1. Urvalet i PISA görs i två steg: ett stratifierat urval av ett antal skolenheter och ett slumpmässigt urval av elever från dessa skolenheter.² I samband med den svenska PISA-mätningen 2018 samlades det för första gången in personnummer för de deltagande eleverna ($n = 5\,432$). Detta har möjliggjort att PISA:s elevunderlag kunnat sammankopplas med nationella registerdata som bland annat innehåller information om elevernas bakgrund, slutbetyg och provbetyg från de nationella proven i årskurs 9.

PISA-resultat i läsförståelse

Den beroende variabeln i denna studie är elevernas resultat på läsprovet i PISA. Eleverna som skriver PISA-provet genomför ett två timmar långt datorbaserat prov inom de tre ämnesområdena *läsförståelse*, *matematik* och *naturvetenskap*. Provuppgifterna består av både flervalsfrågor (cirka 2/3) och öppna frågor. Läsförståelse var huvudområde i PISA 2018, vilket innebär att det finns fler frågor inom detta område än inom de andra två. Detta gör också att resultaten kan presenteras mer i detalj för just läsförståelse. I PISA benämns läsförståelse som *reading literacy*, vilket ”handlar om elevers förmåga att förstå, använda, utvärdera, reflektera över och engagera sig i texter för att uppnå sina egna mål, utveckla sina kunskaper och sin potential och för att delta i samhället” (Skolverket, 2019, s. 10). Läsförståelseuppgifterna avser pröva följande tre läsprocesser: *inhämta information*, *förstå* samt *utvärdera och reflektera*.

Resultaten i PISA redovisas inte som råpoäng, eftersom en elev endast svarar på en delmängd av alla provfrågor. På detta sätt täcks större kunskapsområden in utan att proven blir för krävande för varje enskild elev. Sålunda svarar elever på olika block med provfrågor. I PISA 2018 var läsprovet adaptivt för första gången, vilket innebär att elever blev tilldelade olika svåra uppgifter beroende på hur de hade presterat på tidigare uppgifter. För att elevers provresultat ska kunna jämföras skattas varje elevs provpoäng enligt en statistisk procedur i form av tio plausibla (troliga eller rimliga) värden med ett medelvärde på 500 poäng och en standardavvikelse på 100 poäng.

² För mer information om urval och bortfall, se Skolverket, 2019.

De plausibla värdena bygger på en latent modell och baseras dels på elevens provresultat i PISA, dels på bakgrundsinformation, exempelvis deras könstillhörighet och föräldrarnas genomsnittliga utbildningsnivå. För att korrigera för osäkerheten i skattningen måste alla tio plausibla värden användas i analyserna för att standardfelen ska bli korrekta (Rutkowski et al., 2010). I dessa analyser har vi använt datorprogrammet Mplus som har integrerade funktioner för att stödja användningen av plausibla värden.

Frågor om elevers ansträngning och motivation i PISA

Information om elevers skattade provmotivation i PISA hämtades från den så kallade ansträngningstermometern (se Figur 1 nedan). PISA:s ansträngningstermometer ligger sist i provhäftet och består av en tiogradig skala där eleverna ska skatta hur mycket de ansträngde sig på PISA-provet (faktisk ansträngning). Därtill får eleverna en fråga om hur mycket de skulle ha ansträngt sig om resultatet hade haft betydelse för betyget (hypotetisk ansträngning). Frågorna ställs i förhållande till en tredje skala (se vänsterkolumnen i Figur 1). I den skalan är siffran 10 redan är markerad för att illustrera en verklig situation som är mycket betydelsefull för eleven och där eleven därför skulle ge sitt allra bästa. Deskriptiv statistik för ansträngningstermometern finns i Tabell 1 och 2 nedan.

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?
	☒ 10	☐ 10	☐ 10
	☐ 9	☐ 9	☐ 9
	☐ 8	☐ 8	☐ 8
	☐ 7	☐ 7	☐ 7
	☐ 6	☐ 6	☐ 6
	☐ 5	☐ 5	☐ 5
	☐ 4	☐ 4	☐ 4
	☐ 3	☐ 3	☐ 3
	☐ 2	☐ 2	☐ 2
	☐ 1	☐ 1	☐ 1

Figur 1. PISA:s ansträngningstermometer.

Källa: Skolverket, 2015, s. 18.

Det är värt att påpeka att ansträngningstermometern avser hela PISA-provet och således inte är ämnesspecifik. Detta innebär att den kan tolkas något olika beroende på vilket ämnesområde (läsning, matematik eller naturvetenskap) som eleven relaterade till mest när de svarade. Något som talar för att läsning sannolikt har varit framträdande är att det var flest läsuppgifter i PISA-provet 2018, eftersom läsförståelse var huvudämne då.

Provbetyg i läsförståelse

De nationella proven i årskurs 9 har större personliga konsekvenser för elever än PISA-provet, eftersom resultatet kan påverka slutbetyget. Syftet med de nationella proven är att stödja en likvärdig och rättvis bedömning och betygssättning. Provresultatet utgör en del av lärarnas bedömningsunderlag och testar specifika delar av kursplanen inom varje ämne (Skolverket, 2022). Förutom att läraren ska använda provresultaten som stöd för betygssättning, får eleven vanligtvis återkoppling på hur de presterade. Detta skiljer sig från PISA-provet där eleverna inte informeras om sina individuella resultat.

Det nationella provet i svenska i årskurs 9 består av tre delprov: delprov A (tala), delprov B (läsa) och delprov C (skriva). I denna studie används information om resultatet i delprov B (*NpLäs*), som prövar förmågan att läsa texter av varierande svårighetsgrad inom sakprosa och skönlitteratur. Läsprovet består av ett texthäfte med cirka 30 uppgifter som besvaras med både öppna frågor och flervalsfrågor. Uppgifterna avser pröva fyra läsförståelseprocesser, vilka är konceptuellt sett nära besläktade med de läsprocesser som PISA:s ramverk definierar (Tengberg & Skar, 2017):

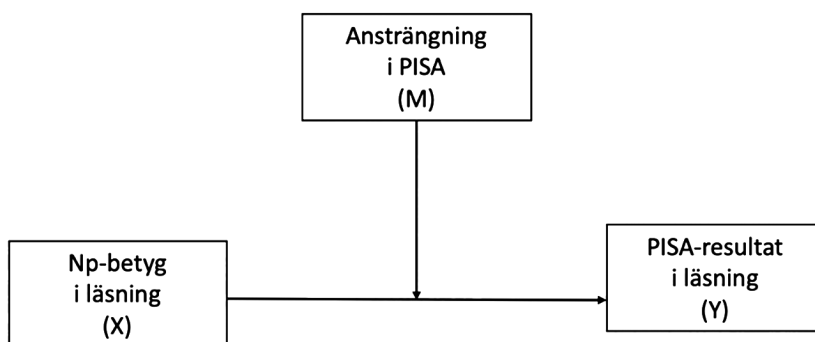
- *Hitta efterfrågad information*
- *Dra enkla slutsatser*
- *Sammanföra och tolka information och idéer samt reflektera*
- *Granska och värdera innehåll, språk och textuella drag.*

Det nationella provet i svenska för årskurs 9 utvecklas vid Institutionen för nordiska språk vid Uppsala universitet på uppdrag av Skolverket. Det är konstruerat utifrån kursplanerna för ämnena svenska och svenska som andraspråk. Provet är gemensamt för båda ämnena. Betygsskalan går från A till F, där A–E är godkända betyg. Betygsstegen A–F motsvaras numeriskt av värdena 20,0, 17,5, 15,0, 12,5, 10,0 och 0,0. Deskriptiv statistik för resultat på det nationella läsförståelseprovet i PISA 2018 ges i Tabell 2 nedan.

Analysmetoder

För att besvara första forskningsfrågan om eleverna skulle ha ansträngt sig mer om PISA-resultatet haft betydelse för betygen gjorde vi en jämförelse mellan rapporterad faktisk och hypotetisk ansträngning i PISA med hjälp av deskriptiv statistik. För att svara på den andra forskningsfrågan om hur sambandet mellan PISA-poängen i läsning och det nationella provbetyget i läsning ser ut för elever med olika

ansträngningsnivå använde vi regressionsanalys med moderering, där läsförståelseresultatet i PISA är beroendevariabel. Styrkan i ett samband mellan två variabler X och Y kan variera beroende på en tredje variabel M som då modererar sambandet mellan X och Y (Hayes, 2013). Moderering prövas statistiskt genom att lägga till en interaktion mellan X och M, där X, M och XM är oberoende variabler och Y är beroende variabel (se Figur 2 nedan).



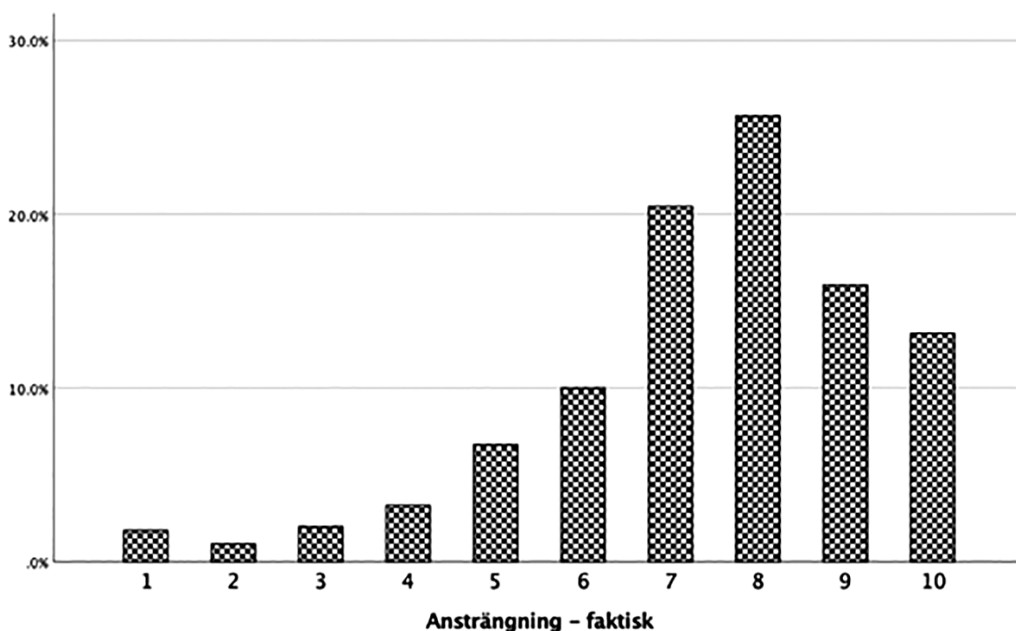
Figur 2. Konceptuell modell av moderering. Den modererande variabeln (M) påverkar styrkan av sambandet mellan X och Y.

För att testa antagandet om provmotivation modererar sambandet mellan PISA-poäng i läsning och det nationella läsprovsbetyget lade vi således till en interaktions-term mellan elevernas skattade faktiska ansträngning i PISA (*AnstrPISA*; M) och det nationella provbetyget i läsning (*NpLäs*; X). Variablerna centrerades först så att de fick medelvärde 0, för att underlätta tolkningen av regressionskoefficienterna. Därefter förde vi stegvis in provbetyget i läsning, ansträngningsvariabeln och korsprodukten mellan dessa.

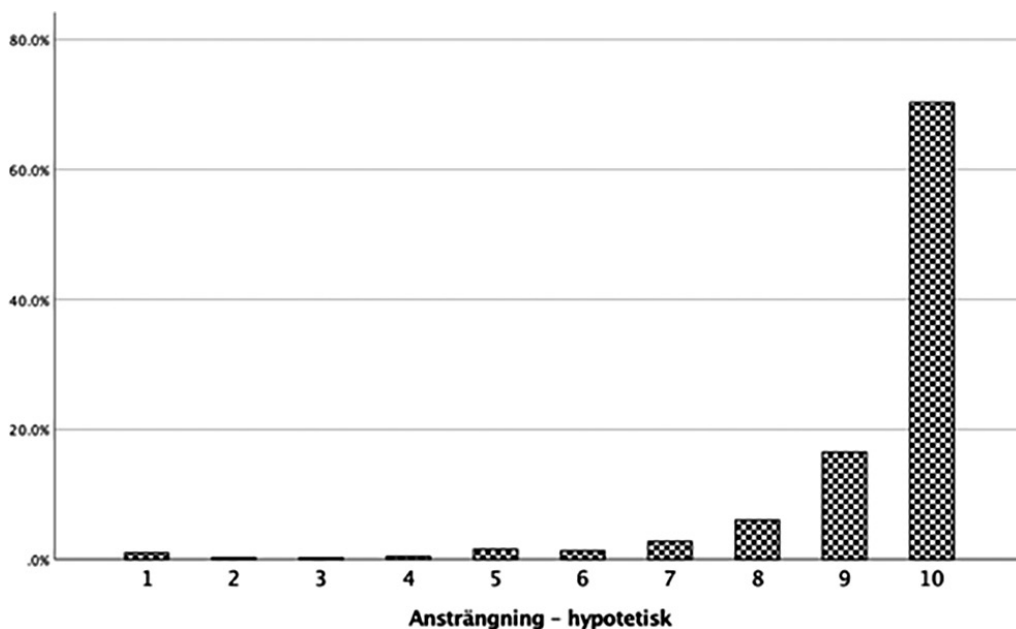
PISA-studien använder vikter för att kompensera för olika sannolikhet att komma med i urvalet. I regressionsanalysen använde vi den slutliga elevvikten (*W_FSTUWT*). Den hierarkiska struktur som finns i PISA-data med individ- och skolnivå tog vi hänsyn till genom att använda funktionen *complex option* i Mplus (Muthén & Muthén, 1998–2017). Databortfall hanterades med Full Information Maximum Likelihood Estimation (FIML) i Mplus. FIML använder all tillgänglig information för att beräkna modellens parametrar. De deskriptiva analyserna genomfördes med IBM SPSS Statistics (IBM Corp., 2022) och regressionsanalysen med Mplus (Muthén & Muthén, 1998–2017).

Resultat

Studiens första forskningsfråga belyser om eleverna är motiverade att göra sitt bästa på PISA: Skulle eleverna ha ansträngt sig mer om PISA haft betydelse för betygen? För att identifiera elever som eventuellt underpresterar på PISA har vi studerat skattningarna på ansträngningstermometern mer i detalj (se Figur 3–4 och



Figur 3. Procentuell svarsfrekvens på fråga 1 i Ansträngningstermometern i PISA 2018: ”I jämförelse med den situation du just tänkt dig, hur mycket ansträngde du dig med det just avslutade provet?”



Figur 4. Procentuell svarsfrekvens på fråga 2 i Ansträngningstermometern i PISA 2018: ”Hur mycket skulle du ha ansträngt dig om resultatet på provet skulle ha påverkat dina betyg?”

Tabell 1 nedan). Resultaten visar att eleverna lade ner relativt stor ansträngning på provet. Nästan 70 procent av eleverna har svarat '7' eller högre och totalt 25 procent '9' eller '10'. En jämförelse med hur mycket eleverna skulle ha ansträngt sig om PISA-resultatet haft betydelse för betyget visar att mycket tyder på att eleverna skulle ha ansträngt sig mer då (se Figur 4). Cirka 70 procent anger '10' och nästan 90 procent '9' eller '10', vilket signalerar att eleverna anser att de kunde ha gjort mer för att förbättra sin prestation. Medelvärdena visar också på en skillnad i elevernas ansträngning. För faktisk ansträngning ligger genomsnittet på 7,4, medan hypotetisk ansträngning ligger på 9,4 (se Tabell 1). Skillnaden mellan skattningarna undersöktes via ett T-test och den är statistiskt signifikant ($p < .001$).

Tabell 1. Deskriptiv statistik: Ansträngningstermometern, besvarad av svenska elever i PISA 2018.

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Ansträngning faktisk	4901	7,40	1,96	1	10
Ansträngning hypotetisk	4879	9,37	1,40	1	10

Baserat på denna självrapporterade information kan det utläsas att eleverna är relativt motiverade till att göra PISA-provet, men att de skulle ha ansträngt sig betydligt mer i medelvärde om provresultatet hade räknats till betyget.

Studiens andra forskningsfråga undersöker sambandet mellan PISA-resultatet i läsförståelse och det nationella provbetyget i läsförståelse för elever med olika ansträngningsnivå. För att besvara frågan har vi stegvis byggt upp en regressionsmodell. Deskriptiv statistik för de variabler som ingår i modellen presenteras i Tabell 2 nedan.

Tabell 2. Beskrivande statistik och korrelationer för de ingående variablerna i regressionsmodellen.

Variabel	<i>n</i>	<i>M (SD)</i>	PISA-poäng	<i>NpLäs</i>	<i>AnstrPISA</i>
PISA-poäng	5504	505,79 (107,54)	1		
<i>NpLäs</i>	4919	13,66 (5,02)	0.63	1	
<i>AnstrPISA</i>	4901	7,40 (1,96)	0.23	0.06	1

Not: Korrelationer, medelvärden och standardavvikelser beräknades med viktad data.

I ett första steg predicerade vi läsresultatet i PISA utifrån provbetyget i läsförståelse (*NpLäs*). Den standardiserade koefficienten var 0,62 (se Tabell 3, modell 1), vilket kan beskrivas som ett måttligt högt samband mellan de två läsproven (Skolverket, 2022). I ett andra steg lade vi endast till ansträngningsvariabeln faktisk ansträngning i PISA på en tiogradig skala (*AnstrPISA*), vilken har en positiv relation till PISA-resultatet ($\beta = 0,22$; se Tabell 3, modell 2). Resultatet visar att elever med högre skattad ansträngning presterar bättre på PISA, motsvarande en genomsnittlig ökning av tolv PISA-poäng per en enhets ökning på *AnstrPISA*-skalan.

Det skulle dock kunna vara så att det i regel är högpresterande elever som har högre provmotivation, exempelvis för att de känner att det går bra på provet och att deras motivation stärks av detta. I syfte att falsifiera denna hypotes förde vi in det nationella provbetyget i läsning (*NpLäs*) och den skattade ansträngningen (*AnstrPISA*) i samma modell, för att konstanthålla elevernas läsförmåga. Givet samma provbetyg har dock provansträngning fortfarande en signifikant relation till PISA-resultat ($\beta = 0,19$; se Tabell 3, modell 3). Även om elevernas nationella provbetyg i läsförståelse hålls på en likställd nivå, presterar de som rapporterar högre ansträngning i PISA bättre på PISA-provet.

Betyg i läsförståelse samt provmotivation är således viktiga för elevernas PISA-resultat. Men för att kunna besvara vår fråga om elever med lägre grad av ansträngning underpresterar på PISA-provet relativt nationella provbetyget måste vi undersöka sambandet mellan PISA-resultat och provbetyg i läsförståelse för elever med olika motivationsnivå. Hypotesen i denna studie är att om elever anstränger sig på PISA kommer sambandet mellan PISA-resultatet och provbetyget vara högre.

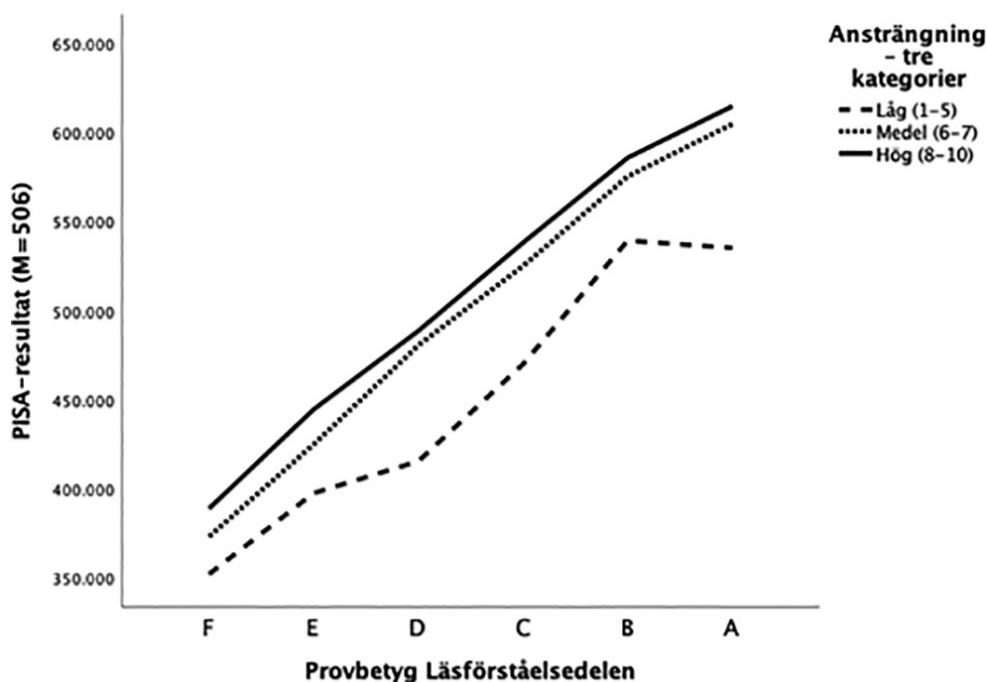
Tabell 3. Skattade regressionskoefficienter när utfallsvariabeln är PISA-resultat i läsförståelse. Resultat för modell 1, modell 2, modell 3 och modell 4.

		<i>b</i>	<i>SE b</i>	β
Modell 1	(Constant)	505,27***	2,45	
	<i>NpLäs</i>	13,19***	0,30	0,62
Modell 2	(Constant)	511,88***	3,06	
	<i>AnstrPISA</i>	12,00***	0,90	0,22
Modell 3	(Constant)	511,03***	2,23	
	<i>NpLäs</i>	12,78***	0,30	0,60
	<i>AnstrPISA</i>	10,23***	0,83	0,19
Modell 4	(Constant)	510,79***	2,24	
	<i>NpLäs</i>	12,81***	0,30	0,60
	<i>AnstrPISA</i>	10,48***	0,86	0,20
	<i>AnstrPISA</i> x <i>NpLäs</i>	0,38*	0,16	0,04

Not: *b* = ostandardiserad regressionskoefficient, β = standardiserad regressionskoefficient, *SE* = standardfel. *NpLäs* och *AnstrPISA* centerades vid sina medelvärden. $R^2 = 0,385$ för Modell 1, $R^2 = 0,050$ för Modell 2, $R^2 = 0,412$ för Modell 3, $R^2 = 0,413$ för Modell 4. * $p < .05$; *** = $p < .001$.

För att testa hypotesen går det att undersöka om ansträngning (*AnstrPISA*) modererar sambandet mellan PISA-poäng och nationellt provbetyg i läsförståelsedelen. Detta görs genom att lägga till en interaktionsterm mellan ansträngningsvariabeln (*AnstrPISA*) och provbetyget (*NpLäs*), vilken förs in i modellen som oberoendevariabel tillsammans med *AnstrPISA* och *NpLäs*. Resultatet i modell 4 (se Tabell 3) indikerar en signifikant positiv interaktion ($\beta = 0,04$, $p = .01$), vilket innebär att sambandet mellan provbetyg (*NpLäs*) och PISA-resultat i läsning är olika för olika grad av ansträngning i PISA (*AnstrPISA*). Den positiva koefficienten visar att det finns ett

något starkare samband mellan PISA-resultat och läsprovsbetyg när elevers ansträngning är högre. Sambandet mellan PISA-resultatet och provbetyget för en elev som har 1SD högre ansträngning skulle sålunda vara 0,64 (0,60+0,04). Motsvarande samband för en elev med 1SD lägre ansträngning på motivationsskalan är 0,56. För att illustrera interaktionen delade vi in *AnstrPISA* i tre kategorier; låg, medel och hög (se Figur 5).



Figur 5. Samband mellan PISA-resultat och provbetyg i läsförståelse för tre kategorier av ansträngning.

Figur 5 visar att sambandet mellan PISA-resultat och provbetyg är högt oberoende av elevernas motivationsnivå och att lutningen på linjen är mycket lik för elever som ansträngt sig på medel eller hög nivå. För elever som indikerar en lägre grad av ansträngning noterar vi en betydande diskrepans för elever som presterat 'A' på det nationella provet i läsförståelse och även för elever som presterat 'D'. Eftersom interaktionen är signifikant är sambandet mellan PISA-provet och nationella provet (*NpLäs*) något lägre för elever med lägre motivation (låg nivå). Det är dock särskilt de högst presterande i denna grupp som ger upphov till det lägre sambandet.

Diskussion och slutsatser

Den övergripande frågan för denna studie är om elever underpresterar i prov där resultatet inte har några direkta konsekvenser för individen (så kallade low-stakes-prov)

jämfört med prov som har mer betydande konsekvenser. PISA är ett prov som saknar direkta konsekvenser för den enskilde eleven. Individuella resultat blir aldrig kända för vare sig elever eller lärare och provet påverkar inte betyget. Det är därför rimligt att anta att alla elever inte gör sitt bästa. Vår hypotes var att elever med låg motivation att göra sitt bästa på PISA-provet skulle ha en svagare koppling mellan PISA-resultat och nationella provbetyg i läsförståelse jämfört med elever med hög motivation att göra PISA. Hypotesen bygger på antagandet att elever anstränger sig mer i det nationella provet eftersom det har större konsekvenser. Tidigare forskning indikerar också att elever troligtvis skulle anstränga sig mer i PISA om resultaten haft konsekvenser för betygen (Eklöf & Hopfenbeck, 2019; Eklöf & Knekta, 2017). Det är emellertid sällsynt att studier har tillgång till provresultat med små respektive stora konsekvenser för samma individ.

Elevers motivation att göra PISA-provet

Elevernas svar på den första frågan på ansträngningstermometern om hur mycket de anstränger sig på PISA-provet visar att de lade ner relativt stor ansträngning, trots att det är ett prov med små konsekvenser för dem personligen. Medelvärdet på den tiogradiga skalan är 7,4, vilket kan jämföras med OECD-snittet som ligger på 8,0. Detta resultat kan ses som betryggande ur ett validitetsperspektiv, eftersom det tyder på att de flesta svenska elever är motiverade att anstränga sig på PISA. Eftersom PISA-resultat kan ligga till grund för utbildningsreformer i Sverige och världen över, är det avgörande att poängen är trovärdiga och reflekterar elevernas faktiska förmåga (Wise & DeMars, 2010). Det finns dock en lägre andel elever som rapporterar en låg grad av ansträngning och dessa elever har överlag sämre resultat. Detta antyder att bristande ansträngning kan ha en negativ påverkan på elevernas prestationer.

Elevernas svar på den andra frågan på ansträngningstermometern tyder på att de skulle ha ansträngt sig ännu mer om resultatet hade haft betydelse för betyget. Medelvärdet var cirka 2 poäng högre. Skillnaden mellan faktisk och hypotetisk ansträngning på PISA-provet får emellertid ses som rimlig. Enligt förväntan-värde-teorin är det troligt att ett prov med mindre konsekvenser värderas lägre än ett prov med betydande personliga konsekvenser. Detta bekräftas även i tidigare forskning där provdeltagare generellt anger en lägre grad av motivation i prov med små personliga konsekvenser jämfört med prov som hade större konsekvenser (se t.ex. Eklöf & Knekta, 2017; Silm et al., 2020).

Intressant nog är Sverige ett av de länder i PISA-mätningen 2018 där en hög andel elever rapporterar att de skulle ha lagt ner mer ansträngning på PISA-provet om det hade räknats till betyget. Andelen av eleverna som har svarat så är 77 procent jämfört med OECD-snittet på 68 procent (De Bortoli, 2021). Det går därför att spekulera om att det svenska PISA-resultatet hade kunnat öka något relativt andra länder, men vi hade också behövt data från andra länder för att kunna dra slutsatser om det. Det är dock svårt att jämföra svar på subjektiva skattningsskalor mellan länder på grund av möjliga kulturella skillnader i svarsmönster samt skillnader i utbildningssystem

och utbildningsnivå. Elever i vissa kulturer tenderar att ge lägre, mer realistiska skattningar, medan elever i andra kulturer kan ge högre, möjligen mer socialt önskvärda skattningar (Butler & Adams, 2007; Skolverket, 2015).

Samband mellan ansträngning och provprestation i PISA

Vad gäller sambandet mellan rapporterad ansträngning och provprestation i PISA visar regressionsanalysen att skattad ansträngning har en positiv relation med PISA-resultatet i läsning med ett standardiserat betavärde på 0,22 ($p < .001$). Styrkan på relationen är i linje med tidigare forskning som rapporterar korrelationer mellan skattad ansträngning och prestation i low-stakes-prov (t.ex. Abdelfattah, 2010; Wise & DeMars, 2005). Det kan även jämföras med Eklöf och Knektas studie (2017) där korrelationen mellan elevers självrapporterade ansträngning och prestation i PISA 2012 var 0,27. När vi kontrollerar för elevernas nationella provbetyg i läsning, som är ett prov med mer betydande konsekvenser, och deras ansträngning i PISA i samma modell har provansträngning fortfarande en signifikant relation till PISA-resultatet ($\beta = 0,19$). Även när elevernas provbetyg i läsförståelse är likvärdiga, presterar de elever som rapporterat att de lagt ner mer ansträngning på PISA-provet alltså bättre. Samtidigt visar jämförelsen att PISA-resultaten varierar för elever som har samma betyg på det nationella provet, där eleverna med högre ansträngning har högre PISA-resultat.

För att testa antagandet att ansträngning modererar sambandet mellan PISA-poäng och betyg på det nationella provet i läsning, inkluderade vi en interaktionsterm mellan ansträngningsvariabeln och provbetyget. Resultaten visar på en statistiskt signifikant positiv interaktion, vilket indikerar att sambandet mellan PISA-poäng och läsprovsbetyg är något starkare när elevernas ansträngning är högre. Detta bekräftar vår hypotes att sambandet mellan läsförmåga i PISA, som är ett prov utan direkta konsekvenser, och betyget på det nationella provet, som har mer betydande konsekvenser, varierar beroende på elevernas ansträngningsnivå vid genomförandet av PISA-provet. Sambandet var svagast framför allt bland de elever som fick högsta betyg på det nationella provet i läsning samtidigt som de rapporterade låg ansträngning i PISA. Dessa resultat tyder på att det finns en grupp elever som underpresterar i PISA i förhållande till deras prestation på det nationella provet.

Studiens resultat visar således att provens olika karaktär och funktion i form av low- och high-stakes-prov har betydelse för elevernas prestation. För elever som är mer motiverade i PISA är sambandet mellan nationella provresultat och PISA-resultat högre. Även om korrelationen mellan PISA-poängen i läsförståelse och det nationella provbetyget i läsförståelse (cirka 0,6) kan beskrivas som måttlig (se t.ex. Skolverket, 2022), är storleken jämförbar med korrelationer som uppmäts i tidigare studier med liknande mått (se Hoge & Coladarsi, 1989; Südkamp et al., 2012). Det är dock viktigt att ha i åtanke att det nationella provet och PISA-provet inte mäter exakt samma aspekter, vilket innebär att vi inte bör förvänta oss perfekta samband mellan dem (jfr Skolverket, 2022). Båda proven avser att testa läsförståelse, men de skiljer sig åt vad

gäller innehåll, format och syften. Exempelvis har PISA:s läsförståelsedel, förutom en adaptiv provmodell, en roterande design som innebär att eleverna delvis svarar på olika frågor, medan det nationella provets resultat baseras på att alla elever svarar på samma frågor. I PISA-mätningen genereras därför plausibla värden med hjälp av en statistisk modell som baseras dels på elevens provresultat i PISA, dels på bakgrundsinformation om eleverna. En felspecificering i denna modell kan leda till en underskattad korrelation mellan PISA och de nationella proven (Hopfenbeck et al., 2018; Skolverket, 2022).

PISA-resultaten och de nationella provbetygen är dessutom behäftade med mätfel av olika slag, vilket påverkar tillförlitligheten som mått på elevers ”sanna” kunskaper (se t.ex. Tengberg & Skar, 2017). Betygen från de nationella proven, liksom PISA-resultaten, utgör båda tvärsnittsmått av elevers kunskaper vid en specifik provsituation. Provresultat från sådana enskilda mättillfällen är ofta påverkade av större slumpmässiga variationer. Slutligen kan skillnaden i skalor också resultera i en underskattning av korrelationen mellan mätningarna. PISA-poängen representerar en kontinuerlig skala, medan betygsskalan endast innehåller sex betygsteg.

Begränsningar och fortsatt forskning

En begränsning med denna studie är att enbart självskattade mått har använts. Detta innebär att vi förlitar oss på elevernas egna påståenden om sin ansträngning och motivation. Det finns emellertid flera nackdelar med självrapporteringsmetoder. Vi kan inte bedöma ärligheten hos eleverna när de ombads skatta sin ansträngning. Trots att graden av ansträngning var positivt korrelerad med provresultatet, kan vi inte avgöra i vilken utsträckning den skattade ansträngningen återspeglar elevens faktiska beteende i provsituationen. Det råder även osäkerhet kring jämförbarheten av svar som ges på subjektiva skattningsskalor, dels mellan elever inom samma land, dels mellan olika länder. Att jämföra den ”faktiska” och ”hypotetiska” ansträngningen kan också vara problematiskt. Det är möjligt att elever underskattar sin sanna ansträngning på PISA relativt den hypotetiska ansträngningen av olika skäl, exempelvis för att eventuella svårigheter i provet då kan ”skyllas på” bristande ansträngning snarare än bristande kunskap (OECD, 2019).

Det finns andra sätt att undersöka grad av ansträngning, exempelvis genom observationer av elevers faktiska beteende och hur de svarar på uppgifter. I digitala prov, som PISA numera är, samlas information om elevernas svarsbeteende, bland annat antal klick och den tid eleven lägger per uppgift. Genom att studera dessa data går det att lära sig mer om provdeltagarnas tankar och handlingar (se t.ex. Lundgren, 2023). I fortsatta studier skulle självskattning kunna kombineras med observerade mått (se t.ex. Hofverberg et al., 2022) för att få en bredare bild av elevers motivation och ansträngning, även om de senare måtten inte heller är fria från mätfel.

Tidigare studier med självskattningsmått indikerar att kvinnor generellt sett uppger en något högre grad av ansträngning än män. Skillnaden är särskilt tydlig i prov

med små konsekvenser och främst vid låga ansträngningsnivåer (Butler & Adams, 2007; DeMars et al., 2013). Butler och Adams (2007) har analyserat data från PISA 2000 och 2003 där de noterade att flickor i genomsnitt rapporterade en högre grad av ansträngning än pojkar. Borgonovi och Bieчек (2016) samt Balart och Oosterveen (2019) har undersökt elevers svarsmonster i olika PISA-omgångar. De har funnit att flickor var bättre på att bibehålla sin motivation under provets gång medan pojkar oftare tappade i uthållighet mot slutet av provet, vilket resulterade i en lägre andel korrekt besvarade uppgifter. Dessa resultat tyder på att provmotivation medierar sociala prestationsskillnader som könsskillnader. I svenska PISA-mätningen är könsskillnaderna i motivation för att göra provet inte särskilt stora, men flickor rapporterar i genomsnitt en något högre grad av ansträngning än pojkar (Skolverket, 2015). Tentativa analyser gjorda inom ramen för den här studien visar på liknande resultat, men hur kön, ansträngning och prestation samvarierar behöver dock undersökas med fördjupade analyser för att nå en bättre förståelse.

Författarbiografier

Stefan Johansson är docent i pedagogik vid Institutionen för pedagogik och specialpedagogik, Göteborgs universitet. Han forskningsintressen fokuserar bland annat på frågor om validitet i olika bedömningsformer och vilka faktorer som påverkar elevers kunskapsresultat i olika ämnen.

Linda Borger är universitetslektor i pedagogik vid Institutionen för pedagogik och specialpedagogik, Göteborgs universitet. Hennes forskningsintressen innefattar frågor om kunskapsbedömning och betyg i skolan, nationella och internationella storskaliga prov samt validitetsteori.

Rolf Strietholt är docent i pedagogik vid institutet för skolutvecklingsforskning, TU Dortmund, Tyskland. I sin forskning har han riktat fokus mot likvärdighetsfrågor inom utbildning och faktorer som påverkar elevresultat inom olika skolsystem. Ett särskilt intresse är internationella storskaliga studier.

Finansiering

Detta arbete finansierades av Riksbankens Jubileumsfond, projektnummer P20-0095.

Referenser

- Abdelfattah, F. (2010). The relationship between motivation and achievement in low-stakes examinations. *Social Behavior and Personality*, 38(2), 159–167. <https://doi.org/10.2224/sbp.2010.38.2.159>
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64(6), 359–372. <https://doi.org/10.1037/h0043445>
- Balart, P. & Oosterveen, M. (2019). Females show more sustained performance during test taking than males. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-11691-y>

- Baumert, J. & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16(3), 441–62. <https://doi.org/10.1007/BF03173192>
- Borgonovi, F. & Biecek, P. (2016). An international comparison of students' ability to endure fatigue and maintain motivation during a low-stakes test. *Learning and Individual Differences*, 49, 128–137. <https://doi.org/10.1016/j.lindif.2016.06.001>
- Butler, J. & Adams, R. J. (2007). The impact of differential investment of student effort on the outcomes of international studies. *Journal of Applied Measurement*, 8(3), 279–304.
- Debeer, D., Buchholz, J., Hartig, J. & Janssen, R. (2014). Student, school, and country differences in sustained test-taking effort in the 2009 PISA reading assessment. *Journal of Educational and Behavioral Statistics*, 39(6), 502–523. <https://doi.org/10.3102/1076998614558485>
- De Bortoli, L. (2021). Snapshots issue 16: How much effort are students putting into PISA? *Snapshots*, 16(16). <https://research.acer.edu.au/snapshots/vol16/iss16/1>
- DeMars, C., Bashkov, B. & Socha, A. (2013). The role of gender in test-taking motivation under low-stakes conditions. *Research & Practice in Assessment*, 8, 69–82.
- Eklöf, H. (2010). Skill and will: Test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy & Practice*, 17(4), 345–356. <https://doi.org/10.1080/0969594X.2010.516569>
- Eklöf, H. & Hopfenbeck, T. (2019). Self-reported effort and motivation in the PISA test. I B. Maddox (Red.), *International large-scale assessments in education: Insider research perspectives* (s. 121–136). Bloomsbury Academic.
- Eklöf, H. & Knekta, E. (2017). Using large-scale educational data to test motivation theories: A synthesis of findings from Swedish studies on test-taking motivation. *International Journal of Quantitative Research in Education*, 4(5), 52–71.
- Feather, N. T. (1982). Human values and the prediction of action: An expectancy-value analysis. I N. T. Feather (Red.), *Expectations and actions: Expectancy-value models in psychology* (s. 263–289). Lawrence Erlbaum.
- Grek, S. (2009). Governing by numbers: The PISA 'effect' in Europe. *Journal of Education Policy*, 24(1), 23–37. <https://doi.org/10.1080/02680930802412669>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. The Guilford Press.
- Hofverberg, A., Eklöf, H. & Lindfors, M. (2022). Who makes an effort? A person-centered examination of motivation and beliefs as predictors of students' effort and performance on the PISA 2015 science assessment. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.791599>
- Hoge, R. D. & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research*, 59(3), 297–313. <https://doi.org/10.3102/00346543059003297>
- Hopfenbeck, T., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J. & Baird, J.-A. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research*, 62(3), 333–353. <https://doi.org/10.1080/00313831.2016.1258726>
- IBM Corp. (2022). *IBM SPSS Statistics for Windows* (Version 29.0). IBM Corp.
- Jerrim, J. (2021). PISA 2018 in England, Northern Ireland, Scotland and Wales: Is the data really representative of all four corners of the UK? *Review of Education*, 9(3). <https://doi.org/10.1002/rev3.3270>
- Johansson, S., Klapp, A. & Rosén, M. (2019). *Läsförståelse i PISA 2018. Om relationen mellan läsförståelseuppgifterna i PISA och den svenska kursplanen*. Skolverket. <https://www.skolverket.se/publikationer>
- Lindblad, S., Pettersson D. & Popkewitz, T. S. (Red.). (2018). *Numbers, education and the making of society: International assessments and its expertise*. Routledge.
- Lundgren, E. (2023). *Exploring and modeling response process data from PISA: Inferences related to motivation and problem-solving* [Doktorsavhandling]. Umeå universitet.
- Meyer, H. D. & Benavot, A. (Red.). (2013). *PISA, power, and policy. The emergence of global educational governance*. Symposium Books.
- Meyer, H. D., Strietholt, R. & Epstein, D. Y. (2018). Three models of global education quality and the emerging democratic deficit in global education governance. I M. Akiba & G. K. LeTendre (Red.), *Routledge international handbook of teacher quality and policy* (s. 132–150). Routledge.
- Muthén, L. K. & Muthén, B. O. (1998–2017). *Mplus user's guide. Eighth Edition*. Muthén & Muthén.
- OECD. (2019). *PISA 2018 results (Volume I): What students know and can do*. OECD Publishing.

Sambandet mellan läsförståelseresultat i PISA och nationella prov

- Pintrich, P. R. (1989). The dynamic interplay of student motivation and cognition in the college classroom. I C. Ames & M. Maehr (Red.), *Advances in motivation and achievement: Vol. 6 Motivation enhancing environments* (s. 117–160). JAI Press.
- Pintrich, P. R. & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2 uppl.). Prentice Hall.
- Pools, E. & Monseur, C. (2021). Student test-taking effort in low-stakes assessments: Evidence from the English version of the PISA 2015 science test. *Large-scale Assessments in Education*, 9(10), 1–31. <https://doi.org/10.1186/s40536-021-00104-6>
- Rutkowski, L., Gonzalez, E., Joncas, M. & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142–151.
- Silm, G., Pedaste, M. & Täht, K. (2020). The relationship between performance and test-taking effort when measured with self-report or time-based instruments: A meta-analytic review. *Educational Research Review*, 31, 1–22. <https://doi.org/10.1016/j.edurev.2020.100335>
- Skolverket. (2015). *Att svara eller inte svara. Svenska elevers motivation att genomföra PISA-provet*. Skolverket. <https://www.skolverket.se/publikationer>
- Skolverket. (2019). *PISA 2018. 15-åringars kunskaper i läsförståelse, matematik och naturvetenskap*. Skolverket. <https://www.skolverket.se/publikationer>
- Skolverket. (2022). *PISA 2018 och betygen. Analys av sambanden mellan svenska betyg och resultat i PISA 2018*. Skolverket. <https://www.skolverket.se/publikationer>
- Südkamp, A., Kaiser, J. & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762. <https://doi.org/10.1037/a0027627>
- Tengberg, M. & Skar, G. (2017). Hur tillförlitligt är det nationella provet i läsning i åk 9? *Utbildning & Demokrati*. 26(2), 113–137.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wise, S. L. & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10(1), 1–17. https://doi.org/10.1207/s15326977ea1001_1
- Wise, S. L. & DeMars, C. E. (2010). Examinee noneffort and the validity of program assessment results. *Educational Assessment*, 15(1), 27–41. <https://doi.org/10.1080/10627191003673216>
- Zamarro, G., Hitt, C. & Mendez, I. (2019). When students don't care: Reexamining international differences in achievement and student effort. *Journal of Human Capital*, 13(4), 519–552. <https://doi.org/10.1086/705799>