

# **Ledarskapsmätning i skolan– en explorativ studie om svenska myndigheters mätningar av rektors pedagogiska ledarskap**

**Rebecka Persson**

Center for Educational Leadership & Excellence  
Stockholm School of Economics

PO Box 6501

11383 Stockholm

[Rebecka.Persson@hhs.se](mailto:Rebecka.Persson@hhs.se)

## **Sammanfattning**

Skalor från två svenska skolmyndigheter, som använts för att mäta rektorers pedagogiska ledarskap, undersöktes i syfte att utvärdera reliabilitet och validitet. Skalorna kommer från Skolverket och Skolinspektionen och har använts för självskattningar på Rektorsprogrammet respektive för lärares skattningar av rektorers ledarskap i Skolenkäten. Med data både från Rektorsprogrammet och från Skolenkäten, samt en ny datainsamling där alla skalor användes tillsammans, hittades stöd för att skalorna uppfyller konventionella krav om reliabilitet och validitet. Det innefattar att skalorna verkar mäta samma sak, har tillräckligt god precision och extern validitet i relation till läraromsättning. Skalorna bedömdes vara användbara både i praktisk verksamhet och i forskning, vilket dock villkoras av att hänsyn tas till att mätningar om rektors pedagogiska ledarskap beträffar två analysnivåer; lärare och skolor.

I Sverige har ledarskapsforskningen i skolan varit begränsad och beskrivits som ”viktig men negligerad” (Ärlestig m.fl. 2016). Ett hinder i utvecklingen mot mer ledarskapsforskning i svensk skola är datatillgång. Rektorer har hög arbetsbelastning och en trend inom många forskningsområden är människors minskade benägenhet att besvara enkäter. Två av våra svenska skolmyndigheter, Skolverket och Skolinspektionen, har samlat in data som handlar om rektorers ledarskap via enkätundersökningar, både med rektorers egna skattningar och lärares skattningar av rektorers ledarskap. Dessa data har inte validerats med konventionella vetenskapliga metoder och mätinstrumenten har inte jämförts med varandra. Denna studies syfte är därför både att jämföra de skalor som myndigheterna använt för att mäta pedagogiskt ledarskap hos rektorer samt att validera dessa med vetenskapliga metoder. Valideringen syftar både till att undersöka statistiska kriterier för mätningarnas kvalitet, och till skalornas användbarhet, dvs skalornas relation till andra faktorer/variabler än ledarskap. I detta fall kommer lärares skattningar om rektors ledarskap att sättas i relation till lärares benägenhet att vilja stanna i sin lärartjänst och i läraryrket. Data från de två myndigheterna kan på så sätt informera forskningen om de ledarskapsmätningar i skolan som pågått i mer än ett decennium, har den validitet som krävs för att komma forskningen till nytta. Sambandet mellan ledarskapskvalitet och medarbetares vilja att stanna kvar på arbetsplatsen är ett tämligen standardiserat utfall i organisationspsykologi (Carter et al., 2019), såväl som i internationell skolforskning (Balu et al., 2009; Ronfeldt et al., 2013). Om myndigheternas skattningar av rektorers ledarskap kan användas till detta är det möjligt att de även kan användas i andra forskningssammanhang, vilket skulle öppna stora möjligheter för ny samhällsrelevant forskning som kan leda till förbättringar i svenska skolor.

I forskningen om ledarskap i skolan är utbudet av ledarskapsteorier smalare än i den generella ledarskapsforskningen. De teoretiska ledarskapsbegrepp som används i ledarskapsforskning i skolans värld är ofta anpassade till skolkontexten. Det vanligast

förekommande begreppet är ”pedagogiskt ledarskap” (*instructional leadership*) som handlar om rektorers ledarskap kring undervisning och lärande (Daniëls et al., 2019). Rektorers pedagogiska ledarskap handlar om att definiera skolans mål, definiera den pedagogiska inriktningen och att utveckla skolans lärandeklimat (Hallinger, 2003). Andra centrala begrepp är ”situationellt ledarskap” som fokuserar inte bara på skolkontexten utan även på att skolor skiljer sig åt (Daniëls et al., 2019; Hallinger, 2011), samt transformativt ledarskap anpassat till skolkontexten (Leithwood & Jantzi, 1999). Även delat ledarskap (*shared leadership*) har studerats i en skolkontext (Wilhelmson & Döös, 2019). Denna typ av ledarskap innefattar inte bara rektor utan även andra personer i skolans ledningsgrupp. Ibland innefattas även lärare i det distribuerade ledarskapet på skolor, något som är vanligt förekommande framförallt på större skolor (Poekert et al., 2020).

Rektors ledarskap har påvisbar betydelse för skolan (Liebowitz & Porter, 2019). Böhlmark et al. (2012) fann kausalt plausibla samband mellan rektorsbyten och elevers skolprestationer samt lärares sjukskrivningsgrad och skolors personalomsättning. En viktig relaterad fråga är hur mycket pedagogiskt ledarskap som rektorer ägnar sig åt i relation till annat arbete. I Lindberg and Vanyushyn (2013) studie rapporterar svenska rektorer att nästan 70% av *uppgifterna relaterade till pedagogiskt ledarskap är nedprioriterade* av administrativt arbete och brådskande ärenden, och även *utfördes sämre* än andra uppgifter. Om rektorers ledarskap har effekt på elevresultat och skolans arbetsmiljö men till stor del måste väljas bort så är förståelsen för det pedagogiska ledarskapets vikt avgörande, då rektor ofta behöver prioritera bland uppgifter. En förutsättning för att utvärdera ledarskapsfaktorers betydelse är validerade mått, av det slag som utforskas av tex. Hellström and Hagquist (2019) om pedagogiskt ledarskap.

Sammanfattningsvis finns indikationer på att rektorers ledarskap har en viktig effekt på skolors arbetsmiljö och funktion, såväl som på elevers lärande. Dock saknas i stort sett

applicerad ledarskapsforskning bland skolor i Sverige, vilket leder till en brist på kunskap om hur ledarskap bäst bör utövas av rektor, samt även i vilken mån detta ledarskap kan förbättras av vidareutbildningsinsatser såsom tex Rektorsprogrammet. En förutsättning för sådan applicerad forskning är en stor datamängd av tillräckligt god kvalitet.

En stor potentiell resurs för applicerad ledarskapsforskning i svenska skolor är det enkätunderlag som samlats in av våra skolmyndigheter. Rektorsprogrammet som bedrivs av Skolverket har sedan 2009 samlat in enkätsvar om pedagogiskt ledarskap hos deltagare på Rektorsprogrammet och Skolinspektionen har sedan 2010 administrerat Personalenkäten till tusentals lärare varje år där de bland annat besvarat frågor om sina rektorers pedagogiska ledarskap. Dessa underlag har ännu inte utvärderats enligt konventionell psykometrisk metod och de olika skalorna för ledarskap har heller inte utvärderats i relation till varandra, vare sig inom eller mellan myndigheter. I denna studie undersöks de skalor som de två myndigheterna använt. Dessa har också jämförts med varandra, genom en nyinsamlad enkät där samtliga skalor använts tillsammans.

Undersökningen inkluderar fyra skalor. Den första är Skolverkets och har använts på Rektorsprogrammet. De tre övriga skalorna är Skolinspektionens och har använts i deras enkät riktad till lärare, kallad Personalenkäten. Två av dessa skalor handlar om rektors pedagogiska ledarskap respektive rektors förmåga att utveckla utbildningen. De har använts tillsammans i Personalenkäten fram till år 2021, för att sedan ersättas av en gemensam skala för pedagogiskt ledarskap. Denna nya skala är den fjärde skalan som behandlas i undersökningen. Skalornas reliabilitet, eller tillförlitlighet, och deras relationer till varandra undersöks. I Skolverkets skala undersöks invarians, eller gruppskillnader, mellan olika lärosäten, formella roller och kön. Också skalornas externa validitet, eller giltighet i relation till något annat relevant för lärare, undersöks. Detta görs med ett mått från TALIS (Teaching and Learning International Survey) som administreras av OECD (Organization for Economic

Cooperation and Development), som handlar om lärares benägenhet att vilja stanna på sin skola och i läraryrket, det vill säga läraromsättning (*teacher retention*). I en studie från 2021 var Sverige det land av 32 OECD-länder där lärares intentioner att lämna yrket var störst, med 41% lärare som övervägde att lämna yrket. Mexiko hade minst proportion lärare, 4,5%, som avsåg lämna yrket (Qin, 2021). Det finns viss påvisad relation mellan läraromsättning och rektors ledarskap (Casely-Hayford et al., 2022; Suárez & Wright, 2019).

Det är viktigt att data från myndigheter om pedagogiskt ledarskap valideras innan den används i forskning. Trots att frågorna som myndigheterna använt är inspirerade av forskning och har satts samman av experter på skolområdet inom myndigheterna, så är skalorna begränsade i sitt omfång och fullödlig validering, tex mot andra skalor så som ofta är brukligt i forskning, är inte möjlig. Att skalorna ändå jämförs och undersöks i termer av reliabilitet och viss extern validitet är viktigt eftersom användningen av skalorna kan vara omfattande. Till exempel så är resultaten från Skolinspektionens Personalenkät offentligt tillgängliga och kan användas i formella och informella bedömningar. Dessutom skulle den ansenliga mängden data som är insamlad kunna utgöra en viktig resurs för tex. longitudinell forskning om skola.

### ***Pedagogiskt ledarskap och psykologisk mätning***

Ett vanligt sätt att undersöka människors egenskaper är att ställa frågor om det som undersöks, så som Skolverket och Skolinspektionen gjort om rektors pedagogiska ledarskap. Det som man vill undersöka hos människor utgör sedan en fördelning, eller distribution, av svar på frågorna, vars medelvärde och variation blir föremål för statistisk analys (Edling & Hedström, 2003). När det man vill undersöka baseras på människors skattningar är mätningarna inte direkta (Furr, 2011; Rust & Golombok, 2014). Man brukar säga att begreppen som man mäter är "latenta". Därför behöver frågornas kvalitet utvärderas för att undersöka om det som man avsett mäta faktiskt fångas upp av frågorna, och i vilken mån det fångas upp.

## Metod

### Urval

Tre olika urval användes i studien. Ett urval kom från Rektorsprogrammet (  $n=3101$ ), där data samlats in av Skolverket mellan åren 2009 och 2020 och utgjorde självskattningar om pedagogiskt ledarskap av deltagare på Rektorsprogrammet. Urvalet bestod av  $n=1276$  rektorer,  $n=1205$  biträdande rektorer och  $n=620$  förskolechefer. Deltagarna var  $n=3023$  kvinnor och  $n=815$  män från sex olika lärosäten. Detta urval användes endast för invariansanalyser. Ytterligare ett urval fanns i data från Skolinspektionens Personalenkät,  $n=20070$ , från grundskolan höstterminen 2019. Ingen information om kön, ålder eller antal års yrkeserfarenhet fanns om urvalet. Urvalet från Skolinspektionen användes endast i en analys, för att undersöka samband på både lärar- och skolnivå i en flernivåanalys.

Urvalet som användes för majoriteten av analyserna i undersökningen utgjordes av ett för forskningsstudien egeninsamlat material, som var enkätsvar från upp till 718 lärare ur Lärarbarometern, varav  $n=638$  lämnat svar på hela enkäten. Av dessa var  $n=471$  kvinnor och  $n=167$  män, som i genomsnitt hade varit i läraryrket 22,5 år ( $SD=12,9$ ) och arbetat för sina nuvarande rektorer i genomsnitt 4,2 år ( $SD=3,6$ ).

### Skalor

Fyra olika skalor användes i studien. En skala kom från Skolverket, och de övriga tre från Skolinspektionen. I data från Skolverket, alltså Rektorsprogrammet, baseras data på en skala ensam, och i data från Skolinspektionen figurerar två av skalorna tillsammans. I studiens egeninsamlade data användes alla fyra skalor tillsammans, för att möjliggöra jämförelser.

Den första skalan bestod av fem indikatorer och har använts av Skolverket för självskattningar om pedagogiskt ledarskap för deltagare på Rektorsprogrammet. Skalan alternerades i studiens egen enkät för att lärare skulle kunna skatta sina rektorer pedagogiska ledarskap. Till exempel så formulerades indikatorn ” Det är viktigt för mig som ledare att

observera den pedagogiska verksamheten under ett par tillfällen varje termin” om till ” Det är viktigt för rektor som ledare att observera den pedagogiska verksamheten under ett par tillfällen varje termin”. Skalan från Skolverket har använts sedan 2009 och är fortfarande i bruk.

Resterande tre skalor kommer från Skolinspektionen. Två av skalorna har använts tillsammans i Personalkäten sedan 2012 fram till 2021, för att sedan ersättas av den fjärde och sista skalan. De två tidigare skalorna har tre indikatorer vardera och går under namnen ”Pedagogiskt ledarskap” och ”Utveckling av utbildningen”. Båda skalorna har använts i Personalkäten som är en del av Skolenkäten och administreras av Skolinspektionen. Personalkäten besvaras av Sveriges lärare årligen eller terminsvis. Båda skalorna handlar om rektors arbete och är utformade för att lärare ska skatta sina rektorers pedagogiska ledarskap.

Den sista skalan, Skolinspektionens nya skala som ersatt de två tidigare, har fyra indikatorer. Dessa berör inte rektor i direkt form, utan ”skolledningen” och skolan i en vidare bemärkelse. Förutom de fyra skalorna från de två skolmyndigheterna användes ett mått för läraromsättning, hämtat ur TALIS (Teaching and Learning International Survey) som administreras av OECD (Organization for Economic Cooperation and Development). Skalan har sju indikatorer, varav fyra handlar om lärares vilja att stanna i yrket och tre om lärares vilja att stanna på sin nuvarande skola. Skalan finns tillgänglig på engelska och översattes av författaren.

### **Statistiska analyser**

Data från Skolverkets Rektorsprogram analyserades för att undersöka skalans validitet i termer av gruppkillnader. Gruppkillnader mellan lärosätena där skalan använts, deltagares formella positioner och deltagares kön undersöktes i invarianstester i Mplus. Sedan användes data från Lärarbarometern för att undersöka de fyra olika skalorna separat. Respektive skalas

interna konsistens räknades fram och skalorna studerades var för sig i konfirmatoriska faktoranalyser (CFA) i Mplus. Skolinspektionens två skalor ”Pedagogiskt ledarskap” respektive ”Utveckling av utbildningen” undersöktes därutöver med exempeldata från Skolinspektionens Personalenkät. Med exempeldata gjordes en konfirmatorisk faktoranalys på två nivåer (MCFA).

Skalornas statistiska relationer till varandra undersöktes både i korrelationsanalys baserat på observerade variabler och på latent variabler i konfirmatorisk faktoranalys. Därefter relaterades ett antal kontrollvariabler till de olika skalorna, där skalorna var latent variabler. Sist undersöktes den externa validiteten hos måtten vidare i relation till läraromsättning (eng *teacher retention*), dvs en variabel om lärares självskattade benägenhet att stanna på sina nuvarande skolor och i läraryrket. Det vill säga, rektors pedagogiska ledarskap relaterades till läraromsättning.

## **Resultat**

### **Skolverkets mätningar av rektorers pedagogiska ledarskap från Rektorsprogrammet**

Skolverkets mätning av pedagogiskt ledarskap har skett inom ramen för Rektorsprogrammet, där rektorer själva skattat sitt pedagogiska ledarskap. Ett sätt att utvärdera om skalan fungerar är invarianstester. Med ”skala” menas här de frågor som använts, betraktade tillsammans. Med invarianstester försöker man förstå om det finns gruppkillnader i människors sätt att svara på frågorna, som också refereras till som ”indikatorer”. Analysen är stegvis, så att det först testas om de specifika frågorna har fungerat likadant i olika grupper. Om grupperna svarar på samma sätt på de olika frågorna i en skala så kan man gå vidare och jämföra grupperna i termer av skalan som helhet med alla frågor tillsammans, det vill säga man kan jämföra grupperna i termer av ”pedagogiskt ledarskap”. De fem frågor som använts för att mäta rektorers ledarskap på Rektorsprogrammet presenteras i Tabell 1.



**Tabell 1***Indikatorer Skolverkets skala om Pedagogiskt ledarskap*


---

F1	Det är viktigt för mig som ledare att observera den pedagogiska verksamheten under ett par tillfällen varje termin.
F2	Som ledare är jag välkommen att observera den pedagogiska verksamheten
F3	Jag prioriterar tid till att föra dialog med mina pedagoger om deras undervisning/verksamhet
F4	Pedagogerna önskar att jag har ett aktivt pedagogiskt ledarskap.
F5	Jag prioriterar tid för att stödja och utveckla mina pedagogers kompetens för uppdraget

---

***Invarianstester***

Skolverkets skala har administrerats till deltagare på Rektorsprogrammet på olika lärosäten och till personer med olika befattningar. Det vill säga, förutom rektorer utbildas även biträdande rektorer och förskolechefer på programmet. Ett sätt att förstå om en skala mäter samma sak i olika grupper är invarianstester. Att skalan fungera likadant på olika grupper är ett exempel på intern validitet. Om man t.ex. vill uttala sig om pedagogiskt ledarskap hos beslutsfattande personal i skolan behöver samma skala fungera på både rektorer och förskolechefer, annars kan man inte anta att samma företeelse mäts hos beslutsfattare med olika positioner, vilket är en förutsättning för att grupperna ska kunna jämföras. Att en skala fungerar likadant i olika grupper kallas för skalans invarians. Förutsatt att mätvariens kan etableras följer också en jämförelse av grupperns olika medelvärden på skalan för Pedagogiskt ledarskap. Skolverkets skala om Pedagogiskt ledarskap invarianstestades tre gånger, med befattning, lärosäte och kön som grupperande variabler. Resultaten redovisas i Tabell 2.

**Tabell 2***Invarianstester*


---

Typ av invarians	$\chi^2$	Df	Skillnad
<i>Befattning</i>			
Configural	42,637	12	
Metric	74,819	22	32,182 (10)
Scalar	128,740	30	53,921 (8)

*Lärosäte*

Configural	96,062	24	
Metric	158,351	49	62,289 (45)
Scalar	209,700	69	51,349(20)
<i>Kön</i>			
Configural	66,022	8	
Metric	70,430	13	4,408(5)
Scalar	92,223	17	21,793(4)

Inför invarianstesterna med Befattning som grupperande variabel testades en basmodell per befattning separat. De tre befattningarna i den grupperande variabeln var Rektor, n=1276, Biträdande rektor, n=1205, och Förskolechef, n=620. För att uppnå statistisk tillförlitlighet vad gäller modellernas anpassningsgrad (*eng model fit*) krävdes att två indikatorer i skalan tilläts korrelera i samtliga av de tre basmodellerna. I invarianstesterna betyder statistiskt signifikanta skillnader mellan modellernas  $\chi^2$ -värden att grupperna på något betydande sätt skiljer sig åt. Den första skillnaden mellan befattningar, som redovisas på raden "metric" i Tabell 2, är skillnaden mellan modellerna som heter "configural" och "metric". Dess två olika modeller är en estimering av all samvariation i data utan restriktioner, respektive en estimering av samvariationen i data när varje indikatorns bidrag till Pedagogiskt ledarskap tvingas vara likvärdigt i de olika grupperna. Skillnaden var inte statistiskt signifikant, vilket betyder att befattningsgrupperna inte skiljer sig åt i hur de har svarat på frågorna om Pedagogiskt ledarskap. Nästa skillnad, som redovisas på raden "scalar", är skillnaden mellan modellerna som heter "metric" och "scalar". I den senare modellen, den som heter "scalar", tvingades medelvärdena i de olika grupperna att vara likvärdiga. Eftersom skillnaden i medelvärden mellan den och den tidigare modellen inte var statistiskt signifikanta så skiljde sig grupperna inte åt i termer av sina medelvärden av Pedagogiskt ledarskap på ett statistiskt betydelsefullt sätt.

Också inför invarianstesterna med Lärosäte som grupperande variabel testades en basmodell för respektive av de sex olika lärosätena. Lärosätena var Uppsala, n=735,

Göteborg, n=716, Umeå, n=838, Mittuniversitetet, n=224, Stockholms universitet, n=775 och Karlstads universitet, n=548. Precis som när grupperingen av data handlade om olika befattningar så var det i samtliga sex basmodeller nödvändigt att låta två indikatorer korrelera med varandra för att nå tillräcklig anpassningsgrad. Varken skillnaden mellan modellerna ”configural” och ”metric” eller ”metric” och ”scalar” var statistiskt signifikant. Det betyder att deltagare på Rektorsprogrammet som besvarat skalan om Pedagogiskt ledarskap inte skiljde sig betydelsefullt mycket åt mellan olika lärosäten, varken i termer av själva mätningen eller i termer av de olika lärosätenas medelvärden.

En basmodell per kön estimerades, kvinnor, n=3023, män, n=815. Ingen statistisk skillnad kunde påvisas mellan de första två modellerna, men mellan modellerna ”metric” och ”scalar” skilde sig män och kvinnor åt på ett statistiskt urskiljbart sätt. Det betyder att män och kvinnor inte skiljde sig åt i hur de svarar på skalan eller själva mätningen, men att de skiljde sig åt i resultaten av mätningen. Skillnaden fanns i medelvärdena för kvinnor respektive män i den latent delen av den statistiska modellen. Medelvärdet i gruppen Män var lägre än i gruppen Kvinnor, dvs män skattade i lägre utsträckning att de ägnar sig åt pedagogiskt ledarskap. Också med kön som grupperande variabel var det nödvändigt att låta två av skalans indikatorer korrelera för att uppnå tillräckligt god statistisk anpassningsgrad.

Sammanfattningsvis visade invarianstesterna att Skolverkets skala om Pedagogiskt ledarskap var invariant över olika befattningar och över olika lärosäten, vilket betyder att mätningen verkar betyda samma sak när man tillfrågar olika befattningshavare eller deltagare vid olika lärosäten. Vidare skiljde sig män och kvinnor åt i sina skattningar av pedagogiskt ledarskap, där män i genomsnitt skattade sitt pedagogiska ledarskap lägre än kvinnor.

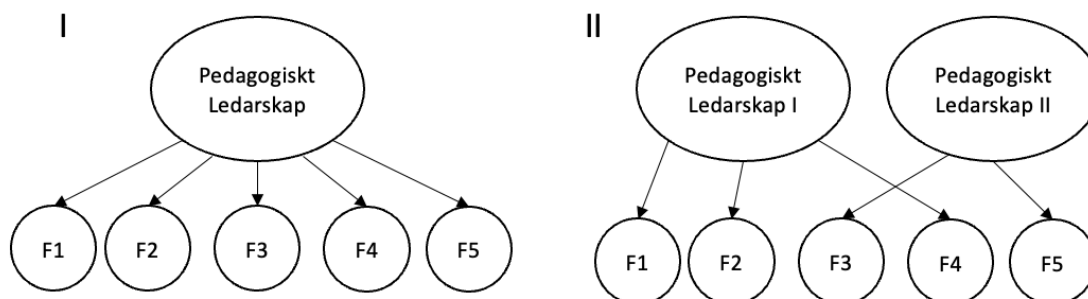
### ***Skalans innehåll***

Ett konsekvent mönster som framträdde i invarianstesterna och gruppjämförelserna var att två indikatorer korrelerade så mycket att deras samband måste tas med i beräkningen.

Det skulle kunna bero på att skalan om Pedagogiskt ledarskap innehåller två teman, snarare än ett tema, så att innehållet borde behandlas som två något olika aspekter av pedagogiskt ledarskap. För att testa detta jämförs en statistisk lösning där alla indikatorer samlades i ett övergripande innehållstema (I) och en statistisk lösning där indikatorerna samlades i grupper om två respektive tre indikatorer av två olika men korrelerande innehållsteman (II). Detta illustreras i Figur 1.

**Figur 1**

*Två olika statistiska lösningar för ett- respektive två innehållsteman i skalan för Pedagogiskt Ledarskap*



I samtliga invarianstester så behandlades skalan som en lösning med två innehållsteman, som till höger i Figur 1 (bild II). Invarianstesterna estimerades inte med två faktorer utan två indikatorer tilläts istället korrelera i modellerna, vilket är likvärdigt. Resultatet betyder att invarians i Skolverkets skala om pedagogiskt ledarskap förutsätts av att den behandlas som två olika men relaterade innehållsteman.

Eftersom det är indikator F3 och F5 som konsekvent samvarierar så måste deras relation tas hänsyn till. Dessa indikatorer har gemensamt att de handlar om att rektor *prioriterar tid för att utveckla sina pedagoger*. Indikator F1, F2, och F4 har någon form av observation gemensamt, dvs i vilken mån *rektor ägnar sig åt att observera den pedagogiska verksamheten*. Indikator F4 handlar inte explicit om observation men samvarierar med F1 och F2, vilka uttalat handlar om observation. I forskningen om rektorers ledarskap beskrivs

rektorers arbete med lärare ibland stegvis, där observation föregår intervention (Balu et al., 2009) Skolverkets skala skulle kunna sägas vara i enlighet med denna beskrivning.

### **Reliabilitet i skalorna**

Reliabilitet, som också kallas för intern konsistens, handlar om en skalas träffsäkerhet, eller hur väl de olika frågorna som utgör en skala hänger samman. Det är viktigt när man vill anta att de i tillräcklig mån handlar om samma sak. Med frågor eller påståenden som handlar om psykologiska begrepp vill man å ena sidan ha en hög reliabilitet, det konventionella värdet för tillräckligt hög reliabilitet är Cronbachs alfa,  $\alpha = .70$ . Å andra sidan är en alltför hög reliabilitet, nära värdet 1, inte heller önskvärd. Detta eftersom vi vet att psykologisk mätning alltid är förenat med felvarians, det vill säga innehåll om annat än det som skalan ska mäta (Borsboom, 2005; Furr, 2011). Med väldigt lite felvarians, och därmed väldigt hög reliabilitet, finns en risk att mätningen handlar om något som inte var avsett att mätas, tex socialt önskvärda värderingar. En god reliabilitet är därför gärna nära det konventionellt accepterade värdet  $\alpha = .70$ .

Ett sätt att undersöka reliabilitet förutom Cronbachs alfa,  $\alpha$ , som gäller en hel skalas observerade poäng, är att estimerar latent modeller för respektive skala med konfirmatorisk (*eng* "confirmatory") faktoranalys (CFA). I CFA kan respektive frågas relation till det som ska mätas undersökas. De fyra olika skalornas reliabilitet kommer därför att undersökas både med Cronbachs alfa och med CFA. Skolinspektionens skalor som använts 2010-2020 undersöktes både med enkätsvar från Lärarbarometern och från Skolinspektionens personalenkät från höstterminen 2019. Det senare gjordes för att studera reliabiliteten i skalorna både på lärar- och skolnivå parallellt. Skolinspektionens data innehåller, till skillnad från data insamlad via Lärarbarometern, svar från flera lärare per skola, vilket gör att också reliabilitet mellan skolor kan beräknas.

### ***Rektorsprogrammets skala om Pedagogiskt ledarskap***

För att kunna ställa frågor om rektors pedagogiska ledarskap till lärare i Lärarbarometern så formulerades Skolverkets frågor från Rektorsprogrammet om. De fem omformulerade frågorna som användes för att testa skalan presenteras i Tabell 3.

Reliabiliteten för skalan i Lärarbarometern, var  $\alpha=.614$ . I Tabell 3 rapporteras hur Cronbachs alfa skulle påverkas om respektive av de olika indikatorerna uteslöts, och i samtliga fall skulle  $\alpha$  sjunka, vilket betyder att frågorna tillsammans bidrar till skalans reliabilitet. Däremot understiger skalan den konventionella gränsen för god reliabilitet,  $\alpha=.70$ .

**Tabell 3**  
*Reliabilitet i Skolverkets skala om Pedagogiskt ledarskap*

	$\alpha$ utan indikator
F1 Det är viktigt för rektor som ledare att observera den pedagogiska verksamheten under ett par tillfällen varje termin.	.594
F2 Som ledare är rektor välkommen att observera den pedagogiska verksamheten	.606
F3 Rektor prioriterar tid till att föra dialog med mina pedagoger om deras undervisning/verksamhet	.468
F4 Jag som pedagog önskar att jag har ett aktivt pedagogiskt ledarskap.	.607
F5 Rektor prioriterar tid för att stödja och utveckla pedagogers kompetens för uppdraget	.457

För att undersöka om frågorna fungerar på samma sätt när de ställs till lärare som när man ställer dem till rektorer så estimerades en konfirmatorisk faktoranalys (CFA). Precis som när rektorers svar analyserades så behövde två av frågorna sammanlänkas med en korrelation i analysen för att modellen skulle uppvisa tillräckligt god anpassningsgrad. Återigen handlar det om indikator F3 och F5. Den konfirmatoriska faktoranalysen resulterade därför i två latent faktorer, en med indikator ett, två och fyra, samt en med indikator tre och fem, ( $\chi^2(4)=2.553$ , RMSEA= 0.00, CFI=1.00).

Resultatet indikerar att det både för rektorer och för lärare fanns två komponenter i Skolverkets skala om Pedagogiskt ledarskap som möjligen reflekterar samma sak- dvs ovan föreslagna ”rektors observationsarbete” på skolan samt ”uppföljande utvecklingsarbete”. Att dessa två innehållsteman framträder betyder att rektorer, både i självskattningar och i lärares skattningar av rektorer, varierar i sitt pedagogiska ledarskap så att en rektors observation av den pedagogiska verksamheten inte helt förutsäger eller hänger samman med utvecklingsarbetet som bedrivs av rektor. Dock finns ett samband mellan dessa innehållsteman vilket syns i korrelationen dem emellan,  $r= 0.211$ ,  $p<.001$ .

I undersökningen jämfördes frågorna som Skolverket använt för att mäta pedagogiskt ledarskap i två separata grupper, rektorer och lärare. Men dessa kunde inte länkas samman i data, vilket betyder att analyser av hur väl lärares skattningar stämde överens med specifika rektorers egna skattningar inte var möjlig. Istället handlade jämförelsen om skalans struktur när den används på rektorer respektive lärare. Som synes antar skalan samma struktur, det vill säga två innehållsteman.

### ***Skolinspektionens skalor***

Skolinspektionens har använt två olika skalor för mätning av pedagogiskt ledarskap mellan åren 2011 och 2020 som går under namnen ”Pedagogiskt ledarskap” respektive ”Utveckling av utbildningen”. Varje skala består av tre indikatorer som redovisas i Tabell 4, tillsammans med reliabilitetsindex. För ”Pedagogiskt ledarskap” var  $\alpha=.888$ , och för ”Utveckling av utbildningen” var  $\alpha =.885$ . Samtliga indikatorer bidrar till högre reliabilitet, , dvs.  $\alpha$  sjunker i samtliga fall om någon av indikatorerna utesluts.

**Tabell 4***Reliabilitet i Skolinspektionens skalor 2011-2020*

	$\alpha$ utan indikator
<b>Pedagogiskt ledarskap</b>	
F1 Rektor har god kunskap om det dagliga arbetet i skolan.	.850*
F2 Rektor arbetar för att främja elevernas kunskapsutveckling.	.845
F3 Rektor tar ansvar för det pedagogiska arbetet på skolan.	.825
<b>Utveckling av utbildningen</b>	
F1 Rektor ser till att utvärdering av elevernas kunskapsutveckling ligger till grund för utveckling av utbildningen.	.882
F2 Rektor ser till att arbetssätt förändras om eleverna inte uppnår kunskapskraven.	.827
F3 Rektor ser till att uppföljningen av mitt och mina kollegors arbete leder till relevanta förändringar i arbetssätt.	.799

De båda skalorna undersöktes tillsammans i en konfirmatorisk faktoranalys,  $\chi^2(8)=49.890$ ,  $p=0.001$ , RMSEA=0.088, CFI=0.986. De observerade variablernas laddningar på de latent faktorerna var alla signifikanta och var i storlekarna .786 till .916. Både det signifikanta värdet för  $\chi^2$  och RMSEA indikerade att modellen där var skala hade vars en latent faktor inte hade anpassningsgrad till data i önskvärd utsträckning. När påstående två i skalan om Pedagogiskt ledarskap och påstående ett i skalan om Utveckling av utbildningen fick korrelera så förbättrades modellens anpassningsgrad, men  $\chi^2$  förblev signifikant,  $\chi^2(7)=20.222$ ,  $p=0.005$ , CFI= 0.995, RMSEA= 0.053.

Korrelationen mellan de latent faktorerna för de två skalorna var  $r= 0.855$ ,  $p=0.001$ . Eftersom korrelationen var så stor testades också en konfirmatorisk faktoranalys med en gemensam latent faktor för alla sex påståenden. Åter blev samtliga relationer mellan observerade och latent variabler signifikanta med storlekar från .779 till .860, men flera index indikerade att data och statistisk modell inte hade önskvärd tydlig relation,  $\chi^2(9)=234.016$ ,  $p=.001$ , CFI= 0.922, RMSEA= 0.192. Inte heller när upp till fyra par observerade variabler tilläts korrelera uppnåddes konventionellt tillräckligt bra anpassningsgrad mellan data och statistisk modell.



För att vidare förstå hur de två skalorna hänger samman gjordes statistiska analyser på ett urval av Skolinspektionens data, från höstterminen 2019. Skolinspektionens data skiljer sig från data från Lärarbarometern, i det att den innehåller flera svar från lärare per skola. Detta innebär att svaren från lärare på samma skola är beroende av varandra, eftersom de är skattningar samma rektor. Den statistiska undersökningen av reliabilitet behöver därför ske både på individnivå och på skolnivå. Detta gjordes med en flernivåmodell (*eng.* multi-level CFA). Underlaget för denna analys utgjordes av 20 070 lärares enkätsvar från grundskolan, höstterminen 2019. Svaren var från 1268 skolor och det fanns i genomsnitt 16 svar per skola. När de båda skalorna samlades i ett innehållstema varierade indikatorernas laddning på de latent faktorerna mellan .779 och .853 på individnivå och mellan .963 och .998 på skolnivå. Modellens anpassningsgrad var  $\chi^2(16) = 1172.633$ ,  $p < .001$ , CFI=0.980, SRMR within= 0.028, SRMR between= 0.009. Det signifikanta  $\chi^2$ -värdet var väntat givet den stora datamängden. De latent faktorerna i analysen korrelerade med  $r = .873$  på individnivå och  $r = .976$  på skolnivå. Den mycket höga korrelationen mellan skalorna på skolnivå indikerade att innehållet i de båda skalorna nära perfekt överlappade när de behandlades på aggregerad nivå. På individnivå hittades modifieringsindex, dvs statistiska förslag som skulle förbättra modellen, beträffande korsladdningar för fem av de sex frågorna. Det betyder att indikatorerna inte kunde tillskrivas bara "Pedagogiskt ledarskap" eller "Utveckling av utbildningen". Istället verkade indikatorerna ha statistisk hemvist i båda skalorna. Både på individnivå och på skolnivå fanns alltså anledning att se skalorna som ett enda latent konstrukt i analysen.

Ett ytterligare mått på reliabilitet när data har en andra nivå, i detta fall skolnivå, är intraklasskorrelationer. Intraklasskorrelationer är ett mått på hur mycket som data inom skolor beror av varandra, eller hur mycket lärares skattningar beror på att de är på samma skola och alltså skattar samma rektor. Indikatorerna i Skolinspektionens data från höstterminen 2019

hade intraklasskorrelationer mellan ICC(1)=.212 och ICC(1)=.272 för ”Pedagogiskt ledarskap” och ICC(1)=.189 och ICC(1)=.201 för ”Utveckling av utbildningen”, när de estimerades i separata latent modeller. Korrelationerna inom skolor (utan felvarians), var ICC(1)=.364 respektive ICC(1)=.283. Intraklasskorrelationer kan tolkas i procent, vilket innebär att 36,4 % av variationen i den latent variabeln för ”Pedagogiskt ledarskap” kan tillskrivas variationen mellan skolor, snarare än variationen mellan lärare, liksom 28,3% av variationen i den latent variabeln för ”Utveckling av utbildningen”.

### ***Skolinspektionens nya skala***

Skolinspektionens nyare skala för mätning av pedagogiskt ledarskap har fyra indikatorer som redovisas i Tabell 5. Reliabilitet för skalan i sin helhet var  $\alpha = .841$ . Samtliga fyra indikatorer bidrog till skalans reliabilitet, vilket också visas i Tabell 5 där  $\alpha$  sjunker vid uteslutning av respektive indikator.

**Tabell 5**

*Reliabilitet i Skolinspektionens nya skala om Pedagogiskt ledarskap*

	$\alpha$ utan indikator
F1 I vilken utsträckning tycker du att skolledningen har tillräcklig kännedom om skolans verksamhet för att leda och samordna den?	.795
F2 I vilken utsträckning genomförs utvecklingsinsatser och förbättringsarbeten som du tycker behövs i skolan?	.768
F3 I vilken utsträckning är ni lärare delaktiga i skolans uppföljningsarbete (som del i det systematiska kvalitetsarbetet)?	.825
F4 Hur uppfattar du möjligheterna för er lärare att få stöd för att skapa studier om det skulle behövas?	.801

En konfirmatorisk faktoranalys visade att samtliga fyra indikatorer var signifikant relaterade till den latent variabeln med laddningar mellan .675 och .845. Modellen hade god anpassningsgrad,  $\chi^2(2)=3.687$ ,  $p = .158$ , CFI=998, RMSEA= .036.

### **Skalornas relationer till varandra**

För att bedöma huruvida skalorna mäter samma sak, det vill säga samma aspekter av pedagogiskt ledarskap, så undersöktes de statistiska relationerna mellan skalorna.

Undersökningen gjordes med två olika ansatser. I den första undersöktes alla frågor tillsammans, utan någon indelning i olika skalor, i en explorativ analys. Detta gjordes för att undersöka det gemensamma innehållet i frågorna, oberoende av vilken skala frågorna kommer ifrån. Samtidigt som explorativ analys kan vara informativt så är frågorna de facto använda i olika skalor vid myndigheterna, varför relationerna mellan dessa också undersöktes.

För att undersöka de olika skalornas relationer till varandra gjordes en principalkomponentsanalys (PCA) med samtliga indikatorer, följt av korrelationsanalyser på skalnivå, först med medelvärden för observerade indikatorer, sedan med latenta faktorer.

### ***Explorativ analys; PCA***

En PCA med två komponenter estimerades, efter det att två komponenter indikerats i en "scree plot". Den sammanlagda andelen förklarad varians i modellen var 81%. Samtliga indikatorer utom tre laddade tydligt på den första komponenten. De tre indikatorer som laddade på den andra komponenten kom alla från Skolverkets skala om pedagogiskt ledarskap, och var samma frågor som tenderade att bilda ett innehållstema om rektors observerande arbete. Resultatet indikerar att innehållet i dessa frågor skiljer sig, inte bara från övriga i skalan i vilken de använts, utan också från samtliga frågor i de tre skalor som Skolinspektionen använt. Resultatet visar även att när en explorativ och deskriptiv metod som PCA appliceras så är det inte uppenbart att frågorna tillhör olika skalor.

### ***Korrelationer mellan skalor***

Korrelationer mellan de fyra skalorna gjordes med respondenters medelvärden av observerade data. Resultaten rapporteras i Tabell 6. Som synes så är alla korrelationer stora, men en viss trend med något mindre korrelationer mellan Skolverkets skala om pedagogiskt ledarskap och övriga tre är urskiljbar. Korrelationerna mellan Skolverkets skala och övriga tre är ca  $r = .50$ , och korrelationerna mellan Skolinspektionens tre skalor överstiger alla  $r = .70$ . Korrelationerna estimerades också i en latent modell, dvs med variationen som är unik för var

skala borträknad, som också presenteras i Tabell 6. Även här var relationen mellan Skolverkets skala och övriga skalor mindre än relationerna mellan Skolinspektionens skalor.

**Tabell 6**  
*Korrelationer mellan skalorna*

	Observerade variabler				Latenta variabler			
	1	2	3	4	1	2	3	4
1.Skolverket Pedagogiskt ledarskap	1	.560	.517	.507	1	.777	.727	.712
2.Skolinspektionen Pedagogiskt ledarskap	.560	1	.766	.772	.777	1	.854	.889
3.Skolinspektionen Utveckling av utbildningen	.517	.766	1	.719	.727	.854	1	.823
4.Skolinspektionen Ny skala	.507	.772	.719	1	.712	.889	.823	1

*Not:* Antal observationer per korrelation varierade mellan  $n=675$  och  $n=643$ .

Sammanfattningsvis fanns inga statistiska belägg för att fyra olika skalor är urskiljbara när samtliga indikatorer studerades i PCA. Det styrker bilden av att skalorna utgör olika mått av samma företeelse; pedagogiskt ledarskap. Skalorna korrelerar även med varandra i hög grad. Det enda innehåll som kunde differentieras från övrigt datainnehåll fanns i Skolverkets skala. Detta indikerades av både PCA och korrelationsanalyserna.

### Kontrollvariabler

I datainsamlingen via Lärarbarometern, där frågor från alla fyra skalor ställdes till lärare, inkluderades sex kontrollvariabler. Att använda kontrollvariabler är ett sätt att testa om någon betydelsefull del av variationen i skalorna verkar bero på något annat än rektors pedagogiska ledarskap, tex hur många års erfarenhet i läraryrket man har eller andra omständigheter som kan skilja sig åt mellan lärare. Indikationer om att kontrollvariabler spelar en roll för lärares skattningar är på så vis en aspekt av den externa validiteten hos skalorna, det vill säga deras applicerbarhet på verksamheter som lärare finns i. Kontrollvariablerna i studien var lärares kön, yrkeserfarenhet i antal år, antal år på arbetsplatsen med nuvarande rektor (som man uppmanades använda som referens när man besvarade frågorna), antal biträdande rektorer på skolan, skolans organisationsform (fristående eller kommunal) samt en uppskattning av elevernas socioekonomiska status (SES).

Elevernas SES skattades genom lärarna fick besvara frågan ”I vilken mån skulle du säga att eleverna på skolan har en bakgrund som är fördelaktig för elevernas resultat? (bakgrund innefattar till exempel föräldrars utbildningsbakgrund)”.

Kontrollvariablerna relaterades till de fyra olika skalorna i separata latent modeller. Det vill säga, skala för skala var latent beroende variabel i regressionsmodeller där de observerade kontrollvariablerna var oberoende variabler. Resultaten presenteras i Tabell 7. För Skolverkets skala om Pedagogiskt ledarskap och kontrollvariablerna hade endast organisationsform (kommunal eller fristående skola) ett signifikant samband med skalan,  $\beta=.436, p=.001$ . Relationen var positiv och medelstor. Den positiva effekten indikerar att lärare som besvarat enkäten och arbetar vid fristående skolor skattar rektorers pedagogiska ledarskap högre än de som arbetar vid kommunala skolor.

**Tabell 7**

*Regressionsrelationer mellan myndigheternas fyra skalor om pedagogiskt ledarskap och kontrollvariabler*

	$\beta$	SD	$p$
<i>Skolverkets skala om Pedagogiskt ledarskap</i>			
Antal år i yrket	-0.006	0.003	0.077
Antal år på skolan	-0.007	0.013	0.602
Antal biträdande rektorer	-0.014	0.034	0.669
Elevernas SES	0.003	0.027	0.920
Fristående skola	0.436	0.136	0.001
Kön	0.128	0.104	0.219
<i>Skolinspektionens skala om Pedagogiskt ledarskap</i>			
Antal år i yrket	-0.007	0.003	0.035
Antal år på skolan	-0.008	0.013	0.547
Antal biträdande rektorer	-0.070	0.033	0.036
Elevernas SES	0.026	0.026	0.332
Fristående skola	0.464	0.134	0.001
Kön	0.183	0.101	0.069
<i>Skolinspektionens skala om Utveckling av utbildningen</i>			
Antal år i yrket	-0.004	0.003	0.310
Antal år på skolan	-0.015	0.013	0.249
Antal biträdande rektorer	-0.045	0.034	0.184
Elevernas SES	0.002	0.027	0.929

Fristående skola	0.352	0.136	0.010
Kön	0.125	0.102	0.221
<i>Skolinspektionens Nya skala om Pedagogiskt ledarskap</i>			
År som lärare	-0.008	0.004	0.067
År med nuvarande rektor	0.004	0.016	0.805
Antal biträdande rektorer på skola	-0.120	0.042	0.005
Elevernas socioekonomiska status	0.014	0.033	0.679
Fristående skola	0.421	0.173	0.015
Kön	0.092	0.128	0.471

Not: n=510

För relationerna mellan Skolinspektionens skala om Pedagogiskt ledarskap och kontrollvariablerna var tre av relationerna signifikanta. Två av dessa, ”Antal år i yrket”,  $\beta = -0.007$ ,  $p = 0.035$ , samt ”Antal biträdande rektorer”,  $\beta = -0.070$ ,  $p = 0.036$ , var små- eller nollrelationer mellan skalan och respektive kontrollvariabel, medan organisationsform, ”Fristående skola”, var signifikant och medelstort,  $\beta = 0.464$ ,  $p = 0.001$ .

För relationerna mellan Skolinspektionens skala Utveckling av utbildningen och kontrollvariablerna, var bara den mellan skalan och organisationsform, ”Fristående skola” statistiskt signifikant,  $\beta = 0.352$ ,  $p = 0.010$ .

Sist i Tabell 7 presenteras relationerna mellan Skolinspektionens nya skala om Pedagogiskt ledarskap och kontrollvariablerna. Två statistiskt signifikanta relationer återfanns, till ”Antal biträdande rektorer”  $\beta = -0.12$ ,  $p = 0.005$ , och ”Fristående skola”,  $\beta = 0.421$ ,  $p = 0.015$ .

Sammanfattningsvis var endast den kontrollvariabel som handlade om skolans organisationsform, alltså om skolan är fristående eller kommunal, återkommande och signifikant relaterad till lärares skattningar av rektorers pedagogiska ledarskap. Lärare som arbetar vid fristående skolor skattade alltså sina rektorers pedagogiska ledarskap högre än lärare vid kommunala skolor. Ett litet negativt samband hittades mellan Skolinspektionens

nya skala och antal biträdande rektorer på skolan. Det är möjligt att just denna skala skiljer sig från övriga på ett sätt som gör att biträdande rektorer är relevanta, eftersom frågorna är formulerade i termer av "ledningen" snarare än "rektor".

Att dessa kontrollvariabler påverkar lärares skattningar är av relevans för skalornas externa validitet. Vad gäller Skolinspektionens nya skala är det tänkbart att själva appliceringen inte gäller bara rektor utan hela ledningen på en skola, som även innefattar biträdande rektorer. Att kommunala respektive fristående skolor verkar ha ett systematiskt samband med lärares skattningar indikerar att skalorna har viss känslighet för lärares omständigheter som behöver beaktas när skalorna används.

### **Extern validitet**

Den externa validiteten testades vidare genom att relatera skalorna till en yttre faktor som anses relevant för begreppet pedagogiskt ledarskap, nämligen läraromsättning (Qin, 2021). I detta fall utgjordes måttet på läraromsättning av lärares benägenhet att vilja stanna på sin skola och i läraryrket. Återigen undersöktes relationerna mellan de fyra skalorna och den externa variabeln latenta regressionsmodeller, där det förväntades att lärares skattningar av rektorers pedagogiska ledarskap skulle ha en och positiv relation till lärares vilja att stanna, både i läraryrket och på sina skolor.

För att extern validitet ska anses vara indikerat får de statistiska relationerna inte vara större än relationerna mellan skalorna, eftersom det då är problematiskt att anta att skalan som handlar om lärares vilja att stanna mäter någonting annat än pedagogiskt ledarskap. Relationerna mellan de fyra skalorna om pedagogiskt ledarskap och läraromsättning estimerades i åtta separata modeller. Resultaten av analyserna mellan skalorna om pedagogiskt ledarskap och lärares benägenhet att vilja att stanna i yrket och på skolan presenteras i Tabell 8.

**Tabell 8***Korrelationer mellan myndigheternas skalor och läraromsättning, latent modeller*

	Vilja att stanna i yrket	Vilja att stanna på skolan
Skolverkets skala om pedagogiskt ledarskap	.277**	.463**
Skolinspektionens skalor i bifaktormodell		
Pedagogiskt ledarskap	.300**	.620**
Utveckling av utbildningen		
Hel skala	.308**	.618**
Skolinspektionens nya skala	.406**	.703**

*Not. \*= $p < .05$ , \*\*= $p < .001$ .*

Skolverkets skala om Pedagogiskt ledarskap hade en liten signifikant positiv relation till lärares Vilja att stanna i yrket,  $r=.277$ ,  $p<.001$ , ( $\chi^2(25)=173.205$ ,  $p<.001$ , CFI=.940, RMSEA= 0.091), med en modifiering av modellen som innebar att indikator ett och två tilläts att korrelera. Samma skalas relation till lärares Vilja att stanna på sina nuvarande skolor var  $r= .463$ ,  $p<.001$  ( $\chi^2(18)=169.734$ ,  $p=.001$ , CFI=.928, RMSEA=.108), även här med korrelation mellan indikator ett och två.

Skolinspektionens skala om Pedagogiskt ledarskap relaterades till Läraromsättning, först för Vilja att stanna i läraryrket och sedan för Vilja att stanna på nuvarande skola. För läraryrket var relationen till den latent variabeln medelstort:  $r=.30$ ,  $p<.001$ , ( $\chi^2(13)=29.690$ ,  $p=.005$ , CFI=.994, RMSEA=.044). För nuvarande skola var relationen stor  $r=.620$ ,  $p<.001$  ( $\chi^2(8)=24.297$ ,  $p=.002$ , CFI=, RMSEA=.055). Relationen mellan Läraromsättning och Skolinspektionens skala om Utveckling av utbildningen var också medelstor för yrket,  $r=.281$ ,  $p<.001$ ,  $\chi^2(13)=38.521$ ,  $p<.001$ , CFI=.991, RMSEA=.055, medan relationen för nuvarande anställning var stor,  $r= .533$ ,  $p<.001$  ( $\chi^2(8)=20.685$ ,  $p=.008$ , CFI=.995, RMSEA=.049).

Sist relaterades Skolinspektionens nya skala till Läraromsättning. För variabeln om Viljan att stanna i yrket var relationen medelstor,  $r= .406$ ,  $p<.001$ ,  $\chi^2(19)= 63.591$ ,  $p<.001$ , CFI=.984, RMSEA=.060, och för skolor var relationen stor,  $r= .703$ ,  $p<.001$ ,  $\chi^2(13)= 33.843$ ,



$p=.001$ , CFI=.991, RMSEA=.050. Båda relationerna är större samband än de mellan de andra tre skalorna och läraromsättning.

Sammanfattningsvis kan sägas att läraromsättning i undersökningen gav stöd åt de fyra skalornas externa validitet. Detta eftersom ett mönster av relationer var tydligt, där samtliga skalor hade större relation till lärares vilja att stanna på sin skola än till att stanna i yrket, samt att samtliga relationer var av betydande storlek.

## Sammanfattande diskussion

I denna studie har fyra skalor om pedagogiskt ledarskap från två olika svenska myndigheter undersökts. Skalorna ska mäta pedagogiskt ledarskap hos rektorer. En skala kom från Skolverket och har använts för rektorers självskattningar på Rektorsprogrammet, övriga tre skalor kom från Skolinspektionen och har använts för lärares skattningar av rektorer i en stor longitudinell datainsamling, där en skala nyligen har ersatt två andra. Data från Skolverkets Rektorsprogram användes i studien för att undersöka eventuella gruppskillnader mellan lärosäten, formella roller och kön, och data från Skolinspektionen användes för att undersöka reliabiliteten i skalor när de används både på individ- och skolnivå. Därtill har merparten av underökningens analyser gjorts på nyinsamlad data, där samtliga fyra skalor utgjort underlag för lärares skattningar av sina rektorer.

Invarianstester av Skolverkets skala om Pedagogiskt ledarskap för självskattning av rektorer under utbildning indikerade inga skillnader mellan lärosäten eller formella positioner (rektor, biträdande rektor och förskolechef), men dock att kvinnor tenderade att skatta sitt pedagogiska ledarskap något högre än män. Invarianstester av samma skala indikerade kontinuerligt att skalan utgörs av två latenta faktorer, eller något olika innehållsteman. Samma mönster med två innehållsteman återfanns när skalan använts för lärares skattningar av rektors pedagogiska ledarskap. Det är viktigt att indikationerna om att skalan innehåller två latenta faktorer bara möjliggör jämförelse av skalans struktur hos rektorers respektive lärares skattningar. Att tillfråga rektorer och lärare är två olika sätt att mäta pedagogiskt ledarskap. För en direkt jämförelse av de olika perspektiven som rektorer och lärare rimligen har, så är sammanlänkning av rektorer och lärare nödvändig, vilket inte varit möjligt i denna studie. Istället kunde bara skalans struktur jämföras i de olika grupperna. Resultatet indikerar att skalan fungerar likadant när den används för självskattning och för lärarskattning av pedagogiskt ledarskap hos rektorer.

Reliabiliteten hos skalorna undersöktes med beräkningar av Cronbach's alfa.

Skolverkets skala uppnådde inte konventionellt önskvärda värdet, vilket kan vara ytterligare en indikation på att skalan inte innehåller ett utan snarare två innehållsteman, som rör den tid som rektor lägger ner på observation och uppföljning av verksamheten, och ett innehållstema som rör själva utvecklingsarbetet. Sammanfattningsvis kan sägas att det är tillrådligt att använda Skolverkets skala på ett sätt som återspeglar två latenta faktorer snarare än en. Vid framtida datainsamling skulle med fördel ytterligare en fråga kunna introduceras för det innehållstema som i dagsläget bara har två indikatorer. Detta för att två frågor inte räcker för att en latent faktor ska vara statistiskt identifierbar.

Reliabilitetsindex för Skolinspektionens skalor "Pedagogiskt ledarskap" och "Utveckling av utbildningen" var höga, med Cronbach's alfa kring  $\alpha = .80$ , vilket är anmärkningsvärt med tanke på att respektive skala bara har tre indikatorer och reliabiliteten ökar med antal frågor. Även Skolinspektionens nya skala för Pedagogiskt ledarskap hade ett högt reliabilitetsindex trots sina relativt få fyra stycken items.

Det är möjligt att den så kallade halo-effekten har inflytande på dessa resultat, dvs att skrattare tenderar att bedöma ledaregenskaper på ett unisont sätt i olika frågor (Behrendt et al., 2017; Rosenzweig, 2007). När de två skalorna undersöktes med hjälp av Skolinspektionens data från höstterminen 2019 så framstod de två skalorna som i stort sett överlappande på både individ- och skolnivå. Vidare kunde 20-30 procent av variationen i Skolinspektionens data tillskrivas variation mellan skolor, snarare än variation mellan lärare, i skattningarna av rektorers arbete. Eftersom variationen mellan skolor är statistiskt sett stor, bör data hanteras på båda nivåer (Shrout & Fleiss, 1979). I praktiska termer, tex när rektorers arbete utvärderas med uppgifter från Skolinspektionens Personalenkät, är andelen medhåll mellan lärare (dvs inom skolor) kanske tvärt om relativt liten. Det vore därför väl värt att tex undersöka om det finns olika subgrupper av lärare, med olika benägenheter att bedöma sina

rektorer. Siffrorna med 20 eller 30 % kommer vidare från latent modeller, om observerade poäng används är medhålllet mellan lärare per skola lägre. När enskilda personer utvärderas måste ett mått för mätningens fel finnas. Det är tex praxis i rekrytering och urvalssituationer och i psykiatriska utredningar att alla testresultat utvärderas i relation till testers felmarginaler. Intraklasskorrelationerna i skolor är en typ av mått på felmarginalen i lärares skattningar av sina rektorer.

Vidare undersöktes skalornas relationer till varandra. I en explorativ analys kunde de fyra skalorna inte urskiljas från varandra, och svagt stöd hittades för mer än en latent faktor, eller ett innehållstema, i frågorna. Eftersom skalorna i praktiken är olika så jämfördes också deras korrelationssamband, där samtliga samband var stora och signifikanta, vilket stödjer att skalorna handlar om samma sak.

Av kontrollvariablerna i undersökningen var bara en relaterad till variationen i samtliga fyra skalor. Lärare som arbetade vid fristående skolor skattade sina rektorers pedagogiska ledarskap högre än vad lärare vid kommunala skolor. Det går inte att med basis i undersökningen att säga varför det förhåller sig så. En del av förklaringen skulle kunna ligga i att rektorer i friskolor i större utsträckning utvecklar strategier för att uppfattas som attraktiva arbetsgivare, både i rekryteringsprocessen och när det gäller att bibehålla personal (Minarik et al., 2003).

Bara i en skala hade kontrollvariabeln ”Antal biträdande rektorer på skolan” samband med lärares svar. Det gällde Skolinspektionens nya skala och sambandet var litet och negativt. Det betyder att ju fler biträdande rektorer på en skola, desto lägre har lärare skattat det pedagogiska ledarskapet. Skalan i fråga skiljer sig något från övriga skalor i sin utformning. Rektor omnämns inte i direkt form, utan istället ställs frågor om *skolledningens* förmåga att leda skolan och om lärares syn på aspekter av skolan generellt, utan specifik hänvisning till rektor. Därför kan skalan tänkas vara känslig för antal biträdande rektorer. Samtidigt som

detta tillför information som inte är tillgänglig med de andra skalorna, så villkoras jämförelsen av skalorna något. Det vill säga, objektet i skalorna är inte detsamma, men skalorna överlappar samtidigt så mycket att innehållet i dem tycks representera i det närmsta samma empiriska företeelse.

Skalornas externa validitet testades i relation till två variabler om Läraromsättning; lärares vilja att stanna i yrket och lärares vilja att stanna på sina nuvarande skolor. Samtliga av de fyra skalorna uppvisade medelstora samband med lärares vilja att stanna i yrket, och stora samband med lärares vilja att stanna på skolan. Störst samband hade Skolinspektionens nya skala till både lärares vilja att stanna i yrket och på sina nuvarande skolor.

Skalorna som undersökts kommer från två olika svenska skolmyndigheter där stora datamängder finns och ackumuleras. Både för praktisk skolverksamhet som kan komma att använda data, och för eventuell forskning, är skalvalidering ett grundläggande villkor. Undersökningen som gjorts indikerar att skalorna håller en sådan kvalitet att de kan vara informativa både i praktisk verksamhet och i forskning, men detta villkoras av datahantering som tar hänsyn till skalornas mätfel och de två olika analysnivåer som berörs; lärare och skolor. Det vill säga, att skalorna används med hänsyn till reliabilitet gällande både lärare och skolor, tex med mått om hur mycket lärare håller med varandra inom skolor. Fortsatt validering kan tex också vara i relation till fler begrepp än till läraromsättning, eller handla om att en myndighet byter ut indikatorer och relaterar dem till tidigare använda indikatorer för att mätningarna ska kunna anses valida över tid. Skalornas nuvarande form indikerar dock goda förutsättningar för att forskningen ska kunna använda dem till exempel i relation till läraromsättning, vilket öppnar för viktiga studier om ledarskap och den svenska skolans fungerande.

## Referenser

- Balu, R., Béteille, T., & Loeb, S. (2009). Examining teacher turnover: The role of school leadership. *Politique américaine*(3), 55-79.
- Behrendt, P., Matz, S., & Göritz, A. S. (2017). An integrative model of leadership behavior. *The leadership quarterly*, 28(1), 229-244.
- Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press.  
<https://doi.org/doi:10.1017/cbo9780511490026>
- Böhlmark, A., Grönqvist, E., & Vlachos, J. (2012). *Rektors betydelse för skola, elever och lärare*. Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU).
- Carter, S. P., Dudley, W., Lyle, D. S., & Smith, J. Z. (2019). Who's the boss? The effect of strong leadership on employee turnover. *Journal of Economic Behavior & Organization*, 159, 323-343.
- Casely-Hayford, J., Björklund, C., Bergström, G., Lindqvist, P., & Kwak, L. (2022). What makes teachers stay? A cross-sectional exploration of the individual and contextual factors associated with teacher retention in Sweden. *Teaching and Teacher Education*, 113, 103664.
- Daniëls, E., Honddeghem, A., & Dochy, F. (2019). A review on leadership and leadership development in educational settings. *Educational research review*, 27, 110-125.
- Edling, C., & Hedström, P. (2003). *Kvantitativa metoder: grundläggande analysmetoder för samhälls- och beteendevetare*. Studentlitteratur AB.
- Furr, M. (2011). *Scale construction and psychometrics for social and personality psychology*. SAGE Publications Ltd. <https://doi.org/doi:10.4135/9781446287866>
- Hallinger, P. (2003). Leading educational change: Reflections on the practice of instructional and transformational leadership. *Cambridge Journal of education*, 33(3), 329-352.
- Hallinger, P. (2011). Leadership for learning: Lessons from 40 years of empirical research. *Journal of Educational Administration*.
- Hellström, L., & Hagquist, C. (2019). School effectiveness in Sweden: psychometric properties of an instrument to measure pedagogical and social climate (PESOC) focusing on pedagogical leadership. *International Journal of Leadership in Education*, 1-21.
- Leithwood, K., & Jantzi, D. (1999). Transformational school leadership effects: A replication. *School effectiveness and school improvement*, 10(4), 451-479.
- Liebowitz, D. D., & Porter, L. (2019). The effect of principal behaviors on student, teacher, and school outcomes: A systematic review and meta-analysis of the empirical literature. *Review of educational research*, 89(5), 785-827.
- Lindberg, E., & Vanyushyn, V. (2013). School-Based Management with or without Instructional Leadership: Experience from Sweden. *Journal of Education and Learning*, 2(3), 39-50.
- Minarik, M. M., Thornton, B., & Perreault, G. (2003). Systems thinking can improve teacher retention. *The Clearing House*, 76(5), 230-234.
- Poekert, P. E., Swaffield, S., Demir, E. K., & A. Wright, S. (2020). Leadership for professional learning towards educational equity: A systematic literature review. *Professional Development in Education*, 46(4), 541-562.

- Qin, L. (2021). Country effects on teacher turnover intention: a multilevel, cross-national analysis. *Educational Research for Policy and Practice*, 20(1), 79-105.
- Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American educational research journal*, 50(1), 4-36.
- Rosenzweig, P. (2007). Misunderstanding the nature of company performance: The halo effect and other business delusions. *California Management Review*, 49(4), 6-20.
- Rust, J., & Golombok, S. (2014). *Modern psychometrics: The science of psychological assessment*. Routledge.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, 86(2), 420. <https://doi.org/doi.org/10.1037/0033-2909.86.2.420>
- Suárez, M. I., & Wright, K. B. (2019). Investigating school climate and school leadership factors that impact secondary STEM teacher retention. *Journal for STEM Education Research*, 2, 55-74.
- Wilhelmson, L., & Döös, M. (2019). Delat ledarskap i förskola och skola: om täta samarbeten som kräver och ger förutsättningar. In: Studentlitteratur AB.