



GÖTEBORGS UNIVERSITET
Utbildnings- och forskningsnämnden för lärarutbildning
Lärarprogrammet, examensarbete 10 poäng

Hållbar betygsättning?

- Påverkas betygsättningen på prov av att eleven är känd för läraren?

Gill Brink, Karin Svensson, Git Öjergren

”LAU660”

Handledare: Peter Esaiasson

Examinator: Birgitta Niklasson

Rapportnummer: HT06 2490-08

Abstract

Institution: Göteborg Universitet – Institutionen för pedagogik och didaktik
Författare: Gill Brink, Karin Svensson, Git Öjergren
Titel: Hållbar betygsättning – Påverkas betygsättningen på prov av att eleven är känd för läraren?
Arbetets art: Examensarbete 10 poäng, LAU 660
Handledare: Peter Esaiasson – Statsvetenskapliga Institutionen
Examinator: Birgitta Niklasson
Rapportnummer: HT06 2490-08
Antal sidor: 34

Rapporten har kommit till för att stilla vår nyfikenhet om det är så att lärare påverkas av yttre faktorer vid betygsättning av prov. Givet de betygskriterier som finns vet vi att betygen endast ska spegla elevens kunskaper i förhållande gällande betygskriterier. Trots att ämnet är känsligt så anser vi sannolikheten vara stor att lärare påverkas av yttre faktorer vid rättning av prov.

Inledningsvis berör vi lärarnas attityder, att tidigare forskning är begränsad och hur vi gått tillväga i vår undersökning. Syftet var att undersöka om lärare blir påverkade av att veta vem eleven är vid rättning av ett prov. I bakgrunden tar vi upp lagstiftning gällande betygsättning. Den tidigare forskningen berör lärares bedömningsarbete, dess orsak och verkan och hur man tar hänsyn till yttre faktorer. Betygshistoriken visar förändringarna som skett, vi tar även upp svårigheter och konsekvenser i förhållande till betyg. Metoden var experimentell design, fem lärare deltog samt 91 gymnasieelevers prov. Vi utformade ett prov av mycket enkel karaktär i ämnet Hållbar utveckling. Lärarna fick rätta provet två till tre gånger. Vid andra rättningstillfället var provet datautskrivet och utan elevens namn. I metoden går vi igenom genomförandet av undersökningen och de forskningsetiska överväganden som beaktades. Resultatet beskriver hur anonymitetseffekten inverkar på lärares sätt att betygsätta prov. Vi fann en signifikant skillnad på vissa frågor med större tolkningsutrymme. Medelvärdeskillnaden mellan första och andra rättningstillfället är 0,25 poäng. Det i sig är anmärkningsvärt men har dock ingen inverkan på betygsättningen på provet. Vi undersökte också huruvida likhetsfaktor, yrkeserfarenhet, kännedom om ämnet Hållbar utveckling och om rättningstid hade någon inverkan på betygsättningen. Vårt resonemang kring undersökningen och resultatet finns i diskussionen där vi även tar upp svagheter i vår undersökning samt ger förslag på framtida forskning.

Sökord: Bedömning, betygsättning, grading, pedagogik, psykologi, similarity

Innehållsförteckning

Inledning	3
Syfte och frågeställning	4
Bakgrund	5
Tidigare forskning	6
Forskning i början av förra seklet	9
Sammanfattning av tidigare forskning	9
Betygens historia	10
Absoluta betygssystemet	10
Relativa betygssystemet	10
Mål- och kunskapsrelaterade betygssystem	10
Orättvisa betyg	11
Prov	11
Svårigheter vid bedömning och betygsättning	11
Konsekvenser	12
Metod och praktiskt genomförande	12
Forskningsetik - Etiska övervägande	14
Resultat, se tabell 1:1	15
Analys av anonymitetseffekt, se tabell 2:1-2:3, 3:1-3:3	16
Fördjupning av undersökningen	20
Lika barn leka bäst?, se tabell 4:1-4:3	20
Trygg och säker i yrkesrollen?, se tabell 5:1-5:3	22
Hållbar betygsättning?, se tabell 6:1-6:3	24
Hare eller snigel?, se tabell 7:1-7-3, 8:1-8:2	26
Lärarnas egna kommentarer	29
Samtal med lärare 1	29
Samtal med lärare 2	29
Samtal med lärare 3	30
Samtal med lärare 4	30
Samtal med lärare 5	30
Samtal med utebliven lärare	30
Diskussion	31
Svagheter	33
Förslag på framtida forskning	34
Referenslista	35

Bilagor

Bilaga 1:1 – 1:2	Läroinstruktioner, provfrågor med facit
Bilaga 2:1 – 2:2	Försättsblad lärarenkät, lärarenkät
Bilaga 3:1 – 3:2	Försättsblad prov, prov
Bilaga 4:1	Kodbok

Inledning

Bedömning och betygsättning är en omfattande del av läraryrket, vi vill med den här rapporten berätta om svårigheter med bedömning av prestationer. Kunskapsbedömning är ett hett ämne som är värt att diskutera. Trots att ämnet är känsligt så anser vi att sannolikheten är stor att lärare påverkas av yttre faktorer vid rättning av prov. Det finns väldigt begränsat med forskning om prestationsbedömning i allmänhet. Vi har inte funnit någon svensk forskning på betygsättning av enskilda prov. Med risk för att trampa lärarprofessionen på tårna så ansåg vi att just en sådan undersökning var värdefull att utforska närmare. Är det möjligt att göra en likvärdig bedömning av elevers prestationer enligt skolverkets allmänna råd om betygsättning? Kan det finnas andra saker som lärare väger in vid sin bedömning?

Metoden som vi använt kan ge signaler om att även i de enklast utformade proven kan skillnader i poängsättning och sedan betygsättning finnas. Vår undersökning skulle kunna användas som en pilotstudie. Förhoppningen vi hyser är att man ska gå vidare med forskning i området. Den korta tid vi hade till vårt förfogande ledde till ett litet skrap på ytan av svårigheter med betygsättning.

Eftersom tidigare forskning hävdar att lärare tar hänsyn till personliga faktorer så bestämde vi oss för att göra en liten rundfrågning till personer i vår omgivning. Vi ställde frågan till ett trettiotal personer med olika bakgrund och i varierande åldrar, vissa var lärare andra elever eller före detta elever. De allra flesta av dem vi frågade hade en åsikt om bedömning och ansåg att den inte alltid var likvärdig. Endast en av de tillfrågade som dessutom var lärare var övertygad om sin objektivitet vid bedömning.

Historier om lärare har cirkulerat i alla tider, elever har sorterats in i kategorier som ”bråkstakar”, ”busfrön” och ”frökens fjäsk”. Lärare har påstått ha favoriter eller utsett syndabockar. Vi vill uppmärksamma problemet med att man aldrig går in utan viss förförståelse i nya situationer. Lärare ska, ”*Vid utvärdering, bedömning och betygsättning vara sakliga och rättvisa och därvid motstå otillbörlig påverkan*”.¹

Är betygen alltid likvärdiga och hur mycket påverkas läraren av eleverna? Ur likvärdighetssynpunkt är det viktigt att betygen uppfattas som korrekta. Det är nödvändigt för att betygssystemet skall uppfattas som legitimt och trovärdigt.² Men skulle det se annorlunda ut om eleven var okänd för läraren? Går det verkligen att koppla bort allt man som lärare känner till om eleven när man rättar ett prov? För att få svar på våra frågor, ville vi se hur lärare betygsätter ett enkelt utformat prov utan essäfrågor. Skulle vi i undersökningen finna stora skillnader vid betygsättningen är det ett anmärkningsvärt resultat och vad beror det i sådana fall på? Betygen kan vara livsavgörande för den enskilde eleven genom att de används som urvalsmetod vid ansökan om tillträde till högre studier.

Givet vår problemställning att undersöka om betygsättningen påverkas av att eleven är känd för läraren, med detta tar vi inte ställning huruvida det har en positiv eller negativ effekt för eleven. Vi hade från början för avsikt att göra en undersökning på flera gymnasieskolor i Göteborgsregionen. Gärna med minst 20 deltagande lärare. De skulle få rätta ett enskilt prov där hälften var de handskrivna originalproven från egna elever, den andra hälften skulle bestå av handskrivna prov från en klass okända elever. Antalet rättade prov skulle då uppgå till

¹ Lärarförbundet. (2004). *Lärarens handbok*

² Skolverket. (2005). *Bedömning och betygssättning*

drygt 500 st, vilket skulle ge en större datamängd att analysera. Resultatet skulle bli säkrare och få en högre validitet och eventuellt vara generaliserbarhet. För att göra undersökningen praktiskt hanterbar så blev vårt bidrag till forskningen att omfatta en analys av ett enskilt prov.

Vi valde att avgränsa oss till att analysera resultatet av rättningen av ett enskilt enkelt utformat prov. Eleven skulle vara känd respektive okänd för läraren. I det här fallet innebär det att läraren känner eleverna till namnen och har undervisat dem en längre tid. Okänd elev innebär att elevens namn är borttaget från provet och att svaren är överförda till datautskrift. I vår rapport kommer läsaren att stöta på ordet bedömning. Vi vill redogöra för vad det begreppet står för i det sammanhang som vi redovisar. Lärarna fick bedöma ett skriftligt prov fråga för fråga, dessutom skulle de betygsätta dem. Det var alltså endast det som stod på pappret som skulle bedömas och inte eleven efter kön, etniska eller sociala bakgrund. Skolverkets allmänna råd tar upp betygsättning generellt, för hela kurser och behandlar inte den enskilda betygsättningen på ett prov.

Om vi trots provets enkla utformning ändå finner stora skillnader i resultatet av rättningarna så anser vi att det finns behov av att undersöka problemet i större omfattning. Vi har inte funnit någon specifik forskning i det område som vi valt att bygga vår undersökning på. Den forskning som tidigare gjorts har varit av mer övergripande slag och inte så detaljerad att den beskrivit resultaten på enskilda prov. Däremot finns det tidigare forskning som pekar på att det finns stora olikheter i hur elever betygsätts. Lärare tar hänsyn till personliga egenskaper vid betygsättning. Det finns även samhällsvetenskaplig forskning som visar på att människor överlag tenderar att gynna dem som mest liknar en själv och alltså tillhör samma samhällsgrupp. De människor som är mer olika och främmande i utseende och samhällsklass tenderar att missgynnas, ja kanske till och med diskrimineras eller rent av få högre betyg för att kompensera detta, så kallad omvänd diskriminering.

I denna rapport presenteras resultatet av vår experimentliknande undersökning. Det blir därigenom möjligt att avläsa i vilken utsträckning och på vilket sätt som betygsättning och bedömning av ett prov kan variera. Resultatet visar på behovet av en medvetenhet i svårigheterna vid betygsättning. Vi ville göra en undersökning som ingen tidigare gjort och som fångar intresset hos fler personer än studenter på Pedagogien. Vi hoppas att studien bidrar till ett ökat intresse för bedömning och betygsättning. Den här rapporten är resultatet av den tanken.

Syfte och frågeställning

Vårt syfte var att undersöka om man som lärare blir påverkad av att man vet vem eleven är vid rättning av ett prov. Vi ville också ta reda på om bedömningen sedan blir den samma om eleven är anonym eller inte, för det kan ge både en positiv men även en negativ effekt för eleven. Tidigare forskning hävdar att lärare väger in andra aspekter vid bedömningen än det som Skolverket rekommenderar i sina allmänna råd.

För att tydliggöra vårt syfte kan detta konkretiseras i följande frågeställning:
Påverkas betygsättningen på prov av att eleven är känd för läraren?

Bakgrund

Lärare får ibland rapporter om elever som är nya i klassen där de beskrivs i negativa ordalag. Detta kan påverka läraren i sin bedömning av eleven. Det blir en slags förhandsinställning, läraren kommer att lägga märke till sådant som stödjer uppfattningen. Lärare kan också ha förutfattade meningar som grundar sig på otillräcklig och felaktig information och blir till fördomar. Det blir en slags schablonmässig bedömning och det kan röra sig om homosexuella, funktionshindrade, religiösa eller elever av annan etnicitet än läraren själv. Uppfattningarna blir stereotypa och generaliserande.³

Bedömningen av elevers prestationer får inte påverkas av den som bedömer. Betyget skall bara spegla elevers kunskaper i förhållande till kursmålen, och vara så entydigt som möjligt.⁴ I Prop. 1992/93:220 står det att bedömningen av en elevs kunskaper enligt riksdagsbeslut skall ske med utgångspunkt i kursplanernas mål och betygen skall relateras till de krav och kunskapsnivåer som uppnås.

Lärare har idag ett stort tolkningsutrymme när det gäller bedömning och betygsättning. Grunden måste alltid bygga på Skolverkets allmänna råd där läraren vid betygsättning skall:

- *utnyttja all tillgänglig information om elevens kunskaper i förhållande till kraven i kursplanen,*
- *beakta även sådana kunskaper som en elev tillägnat sig på annat sätt än genom den aktuella undervisningen,*
- *beakta såväl muntliga som skriftliga bevis på kunskaper och*
- *göra en allsidig bedömning av kunskaperna och därvid beakta hela kursen.*⁵

Världen uppfattas utifrån de attityder, värderingar och föreställningar som vi redan utvecklat, vilket gör att de tenderar att bli självförstärkande. Lärare är inget undantag utan kan som alla riskera att fastna i bestämda arbetsmetoder och uppfattningar.⁶ Men det är elevens kunskaper och förmåga som skall bedömas, inte elevens åsikter eller beteende.⁷ När vi gör bedömningar innebär det att vi värderar våra intryck i relation till någon eller något. De flesta lärarna har emellertid inte erhållit någon utbildning eller träning i att sätta betyg innan de lämnat lärarutbildningen och själva skall betygsätta sina elever.⁸ Betygsättning och utfärdande av betyg utgör myndighetsutövning gentemot eleven i vart fall när det gäller slutbetyg. Det är därför viktigt att bestämmelserna om bedömning och betyg tillämpas likvärdigt och rättssäkert.⁹

Vid en granskning som skolverket genomförde år 2000 visade det sig att fanns stora skillnader när det gäller förutsättningarna för en likvärdig bedömning av elevers kunskaper. Det fanns skillnader gällande lärarnas kunskaper kring kursplanernas innehåll t.ex. var uppfattningen och tankarna bakom betygsättningen i skolsystemet mycket varierande. Det fanns också stora skillnader i vad lärarna ansåg skulle ligga till grund för betygsättningen

³ Sjödén, S. (2001). *Psykologi för gymnasieskolan*

⁴ Korp, H. (2003). *Kunskapsbedömning – hur, var och varför*

⁵ Skolverkets allmänna råd. (2004). *Likvärdig bedömning och betygsättning*

⁶ Kveli, A-M. (1994). *Att vara lärare*

⁷ Skolverket. (2006).

⁸ Skolverket. (2002). *Att bedöma eller döma*

⁹ Skolverkets allmänna råd. (2004). *Likvärdig bedömning och betygsättning*

liksom skolläringens engagemang för att skapa likvärdiga förutsättningar.¹⁰ Mindre skolor tenderade att vara mer generösa vid betygsättningen än större skolor och friskolor. Det finns skolor som för att gynna sina elever medvetet agerar i strid mot de bestämmelser som har till syfte att säkra en likvärdig bedömning. ”Lärare lockas i konkurrens om individuella lönepotter att inför omgivningen//...//antingen dölja sina tillkortakommanden eller överdriva sina framgångar genom att sätta högre betyg än vad eleverna förtjänar. De uppmanas att i//...//lojalitetens namn utåt tåga om skolproblemet...”¹¹

Dessa rader är tagna ur vår referenslitteratur till LAU 650, i kursen ”Mänskliga möten”, 6 poäng, men vi finner ingen forskning i ämnet och får se det mer som en allmän skröna.

Tidigare forskning

Det vetenskapliga intresset för forskning om bedömning har varit svalt i de Nordiska länderna. De finns få empiriska studier att tillgå och de flesta av dem har varit av utvärderande karaktär.¹² Som vi tidigare nämnt så finns det inget stöd för att sätta betyg efter elevens ”personlighet” eller utifrån hur han eller hon betar sig om det inte finns angivet i målen för ämnet eller kursen.¹³ Trots det hävdar tidigare forskning att lärare påverkas. Ett experiment liknande vår undersökning kan bidra med ny kunskap om huruvida lärare tar hänsyn till vem eleven är när de betygsätter prov.

I den Norska rapporten: *Variationer i lärares bedömningsarbete - dess orsak och verkan*, beskrivs ett experiment i första årskursen på gymnasieskolor i Norge. I experimentet utgick forskarna från Correa & Gruver (1987) och deras teori om betyg som ett policyinstrument för läraren att försöka påverka eleven att använda mer tid till studier. En av undersökningsgrupperna gick ut på att lärare skulle sätta ganska låga betyg på eleverna, den andra gruppen gick ut på att eleverna skulle sätta press på lärarna att ge dem högre betyg. I gruppen av elever som blev hårdare bedömda av sina lärare ökade elevernas prestationer och de använde mer av sin fritid till studier. I gruppen där eleverna försökte påverka lärarna att höja deras betyg berodde det till stor del på lärarens bakgrund, karaktär och utbildning om de föll till föga och höjde betyget. Forskarna kom fram till att det i båda grupperna var svårt att hålla isär lärarens karaktär och hur denne bedömde eleven. Norska gymnasieelever som studerat ett år valde sedan inriktning såsom matte, samhällsvetenskap eller främmande språk. Enligt studien påverkades elevens val av de betyg de fick i årskurs ett. De elever som ökade prestationen när de fick ett lägre betyg verkade bli bättre förberedda för högre studier, de såg sitt betyg som ett ”medelbetyg”. De elever som inte lade ner extra energi när lägre betyg sattes valde inriktningar som inte var lika krävande som de högskoleförberedande, de såg sitt betyg som under medel. De elever som ”tjatade” sig till högre betyg gick också vidare till högre studier. Det fanns vissa svårigheter med studien eftersom lärares bakgrund ser så olika ut. Interaktionen mellan lärare och elev påverkas av lärarens karaktär såsom utbildning, erfarenhet och skicklighet.

I studien har man inte tagit hänsyn till socioekonomisk bakgrund, vilket har man gjort i tidigare studier. Antalet elever som deltog i experimentet var 1092, och de var ifrån 15 olika skolor. Betygsskalan gick ifrån 0-6. De flesta lärare hade varit i yrket 20-30 år. Det fanns stora

¹⁰ Ibid.,5

¹¹ Romhed, R. (2006). *Ord och bild*

¹² Korp, H. 2003:10

¹³ Skolverket. (2001). *Bedömning och betygsättning*

skillnader i betygsättningen mellan skolorna, den största var dock i matte, vilket forskarna tyckte var anmärkningsvärt, då matte är ett av de ämnen som skiljde sig minst i undervisningen. Beträffande lärare så fanns det få studier som relaterade till deras erfarenheter. De skolor där eleverna hade höga resultat i matematik hade lärare med högst kompetens inom området och tvärtom. Det kunde också vara så att de lärare som hade högst utbildning sökte sig till skolor med elever som presterade mera. Anmärkningsvärt var också att lärare med högre utbildning satte högre betyg och lärare med lägre utbildning men längre erfarenhet satte lägre betyg. Sammanfattningsvis menar forskarna att det finns två alternativa förklaringar, den ena var att lärare har olika erfarenheter av konsekvenserna av betygsättning, den andra var att olika lärare har olika förmåga att stå emot elever som försökte påverka dem att sätta högre betyg.¹⁴ Denna undersökning presenterar ett resultat som klart hävdar att lärare låter sig påverkas av elevers påtryckningar, nästa rapport hävdar att även elevers attityder spelar roll för betygsättningen.

Cross & Frary (1996) har noterat att lärare ofta värderar elever efter attityd, ansträngning och prestation. Det finns många svårigheter med att mäta prestationer som grundar sig i lärares utbildningsfilosofi och lärares svårighet i att agera både domare och försvarsadvokat. Även om det finns många orsaker till lärares svårigheter med betygsättning så måste den grundas på prestation, mognad, förmåga, ansträngning, kommunikationsförmåga och liknande faktorer. Trots dessa tydliga kriterier så har det framkommit att den gängse betygsättningen skiljer sig ifrån den rekommenderade. Lärare verkar mildra negativa sociala konsekvenser med att sätta högre betyg. Cross & Frary (1996) utökar antagandet om att mildra sociala konsekvenser till att gälla även lärarna. Antagandet går så långt så att det till och med omfattar betygsättning som lärarnas sätt att kontrollera elevernas uppförande.

Lärarna har haft svårare att beskriva och betygsätta elever med sämre prestationer, medan högpresterande elever blivit beskrivna i termer som: lysande, brilliant, uppför sig väl och är skötsam.

Denna undersökning omfattar betygsättningen av 310 middle- och highschoollärare i ett skolsystem, betyg sattes på 7367 middle- och highschool elever. Metoden som användes i undersökningen var enkäter där både elever och lärare skulle svara på huruvida de instämde i vissa påståenden. Följande påståenden är plockade ur undersökningen av lärarna: Hänsyn skall tas till studentens förmåga? (66 % instämde, eller tenderade att instämma). Värderar du ansträngning i termer som förmåga? (72 % Ja). Idealt skall det inte vara nödvändigt att ta hänsyn till förmåga när man sätter ett betyg. (49 % instämmer eller tenderar att instämma). Eleverna fick liknande frågor: Är det rättvist att ta med elevens förmåga när ett betyg sätts? (55 % Ja). Idealt sett skall det inte vara nödvändigt att ta hänsyn till förmåga när man sätter betyg. (62 % instämmer eller tenderar att instämma). Resultatet visade på att endast 49 % av lärarna ignorerade yttre påverkans faktorer vid betygsättning trots att alla lärare kände till de rekommenderade normerna. Även eleverna fick svara på frågor i undersökningen, 55 % av eleverna tyckte att det var okey att lärarna övervägde elevers förmåga när de bestämde den egna graden för betyg, medan 62 % tyckte att lärare skulle gå efter den fastställda normen i första hand. 81 % av lärarna och 70 % av eleverna ansåg att det skulle finnas rapporter liknande ordning och uppförande vid sidan av betyget. Trots att lärare svarade nej på frågor om ifall klasstillhörighet och etnicitet spelade roll i betygsättningen så visade resultatet helt annorlunda. 39 % av lärarna blev måttligt påverkade och 14 % blev mycket påverkade. 46 % ansåg att de litade på sin objektivitet.¹⁵ I denna undersökning så framgår det att många lärare

¹⁴ Bonesrönning, H. (1999). *The variation in teachers grading practises: causes and consequences*

¹⁵ Cross, L H and Frary, R B. (1996). *Hodgopodge Grading; Endorsed by Students and Teachers Alike*

har en medvetenhet om att elevers attityder påverkar dem. Om det nu är så att lärare anser sig vara opåverkade av sådant som klasstillhörighet, etnicitet och liknande så finns det forskning som pekar på motsatsen.

I samhällsvetenskapen har forskning bedrivits på hur likheter påverkat människor. Ett experiment utfördes för att undersöka om det spelade någon roll hur en kandidat i ett politiskt val såg ut för väljarna. Inom forskningen har det tidigare funnits en tendens att förbise betydelsen av det som inte "sägs" av politiker. Denna amerikanska undersökning som även hänvisade till tidigare forskning, menade att utseende, etnicitet och kön påverkade väljarna. Det visade sig att afroamerikaner hellre röstade på en kandidat som var afroamerikan än en nordamerikan. Väljarna hade också åsikter om att kvinnor var mer liberala och män var strängare och mer militäriska.

Belönade väljarna kandidater vars ansikte liknade ens eget? Forskning i socialpsykologi har sett storskaliga effekter av social påverkan av likheter. Enligt undersökningar är detta en mänsklig instinktiv handling. Detta gällde troligtvis även vid val av kandidat. Kandidater med mer attraktiva ansikten fick fler röster. Ett experiment utfördes där ett mansansikte manipulerades med hjälp av TSI (Transformed Social Interaction). Forskarna använde digital teknik för att förändra ansiktet så att fotot liknade väljarens (afrikanskt, latinskt etc.).

Deltagarna i experimentet fick sedan se fotot av kandidaten, antingen originalet eller det manipulerade i 20 sek. De fick sedan svara på frågor. Vid experimentet med den manipulerade fotot steg rösterna på kandidaten med 20 % för männen men sjönk med 20 % för kvinnorna i frågan om skulle du rösta på kandidaten? Känslotermometern steg 12 % för män och sjönk 10 % för kvinnor. Männen var alltså mer benägna att rösta på kandidater som liknade dem själva och tillhörde samma grupp. Kvinnorna däremot verkade tycka att kandidaten som liknade dem var mer fränstötande.

Tre möjliga förklaringar till resultatet diskuterades: kvalitén på de manipulerade fotona, det var ju ett manligt ansikte som fick kvinnliga drag. Intressant vore om kandidaten istället hade varit kvinna, vilket inte lät sig göras i detta test. Undermedvetet kan kvinnorna också ha "straffat" kandidaten för att han såg kvinnlig ut. Det kan också finnas en självkategoriserande teori, människor befinner sig i ett ramverk där man jämför sig med människor i samma grupp och ser likheter medan man ser olikheter och kontraster på dem som står utanför gruppen, (boomerangeffekten). Det var inte möjligt att säga vilken av de tre varianterna som var mest trovärdig, men det fanns ett klart samband mellan någon typ av likhet eller "närvaro av en själv". Undersökningen visade att kandidater kan få så mycket som 20 % av rösterna enbart på visuell information.¹⁶

I vår undersökning kan det innebära att lärare kan väga in likhetsfaktorer och kanske bli påverkade till att sätta ett annat betyg när eleven är känd.

¹⁶ Bailenson, J N, Garland P, Iyengar S, Yee N. (2006). *Transformed Facial Similarity as a Political Cue: A Preliminary Investigation*

Forskning i början av förra seklet

Redan för 100 år sedan upptäcktes problem med betygsättning. Forskaren Max Meyer uppmärksammade att inget enhetligt system fanns för betygsättning. En lärare upptäckte stora svårigheter att betygsätta en klass då den bestod av yngre oerfarna studenter och äldre studievana studenter. Alla i klassen fick samma prov, eleverna jämfördes med varandra. Istället för att jämföra de yngre med varandra så jämfördes de alltså med de äldre mer erfarna. Max Meyer gjorde då ett experiment där differentiering genomfördes, detta innebar att klassen delades in i två grupper där eleverna fick olika prov och varsin betygsskala. På ett universitet i Missouri användes betygsskalan där 25 % av eleverna skulle vara (superior) överlägset bäst, 50 % skulle vara medel och 25 % under medel (inferior). Elevernas bakgrund avgjorde vilken klass de kom att tillhöra. Det visade sig att två tredjedelar av det man då ansåg vara problem med betygsättning försvann. Övriga tredjedelen var svårare att få bukt med på grund av skeptiska rektorer och motsträviga lärare som inte ville använda betygsskalan och dela in eleverna i grupper. Max Meyer upptäckte också i sina undersökningar att kvinnor ofta var duktigare i skolan men, de hade tre saker som hindrade dem i sina yrkesmässiga framgångar: diskriminering, tradition och barnafödande.¹⁷

Sammanfattning av tidigare forskning

I det något begränsande underlag vi fått fram i tidigare forskning hävdar man bestämt att lärare tar personliga hänsyn vid bedömning. Saker som elevers attityder, påtryckningar och ursprung påverkar läraren i betygsättningen.

Mot bakgrund av tidigare forskning framgår det att:

- Högutbildade lärare söker sig till skolor med högpresterande elever
- Lärare tenderar att påverkas i sin betygssättning av elever på grund av vad betyget får för konsekvenser
- Lärare påverkas i betygsättningen av elevens påtryckningar, positivt eller negativt
- Lärare tenderar att mildra negativa sociala konsekvenser genom att sätta högre betyg, så kallad omvänd diskriminering
- Människor över lag tenderar att gynna dem som mest liknar en själv och alltså tillhör samma grupp

Det finns en medvetenhet hos lärare om yttre påverkans faktorer vid betygsättning. De lärare vi samtalat med hade också en medvetenhet om problematiken med betygsättningen.

¹⁷ Meyer, M. (1914). *The limit of uniformity in the grading of college students by different teachers*

Betygens historia

Det har skett stora förändringar i betygssystemet de senaste sextio åren. Det har funnits tre olika betygssystem sedan 1950-talet, det absoluta, det relativa och det mål- och kunskapsrelaterade betygssystem.

Absoluta betygssystemet

Det absoluta betygssystemet bestod av en sjugradig skala, A – C.

- A - Berömlig
- a - Med utmärkt beröm godkänd
- AB - Med beröm godkänd
- Ba - Icke utan beröm godkänd
- B - Godkänd
- BC - Icke fullt godkänd
- C - Otillräcklig

Eleverna fick även betyg i uppförande och ordning i en tregradig betygsskala.

*”Varje betyg angav ett visst mått av kunskap... Betygssystemet kritiserades för att det inte fanns några anvisningar som angav vilka kunskapskrav som de olika betygsstegen motsvarade”.*¹⁸

Relativa betygssystemet

På 1960-talet infördes det relativa betygssystemet. Detta system byggde på den matematiska teorin om normalfördelning. Eleverna rangordnades sinsemellan. Betyget fördelades procentuellt enligt en femgradig skala från 1 – 5.

Betyg	1	2	3	4	5
Procent	7%	24%	38%	24%	7%

Även detta betygssystem fick kritik och ifrågasattes på flera punkter. Kritiken gällde framför allt om att betyget inte sa så mycket om elevens kunskap i ämnet, utan i förhållande till andra elevers kunskap. Samma kunskaper i ämnet kunde alltså betyg sättas olika i olika klasser.

Mål- och kunskapsrelaterade betygssystem

Detta betygssystem ger bättre information av elevens kunskapsutveckling och kunskapsprofil. Elevernas kunskaper bedöms i relation till förbestämda kunskapskvalitéer som är specificerade i mål och kriterier. Betygskriterierna beskriver de kvalitéer som gäller för respektive betyg som är indelade i G - MVG.

- G - Godkänd
- VG - Väl godkänd
- MVG - Mycket väl godkänd

¹⁸ Skolverket

Orättvisa betyg

Om eleven anser att han fått ett orättvist betyg så finns det inget sätt att överklaga betyget. ”När betyget väl är satt kan man inte ändra det även om det är helt uppenbart orättvist//...//Det är därför väldigt viktigt att man får prata med sin lärare om vad läraren tänker sätta för betyg innan de sätts”.¹⁹ Betyget skall aldrig komma som en överraskning för eleven. De gymnasieelever som fått IG på någon kurs kan gå upp till provning under skoltiden. Om eleven däremot fått G eller VG och vill provas för högre betyg får vänta med det till efter gymnasiet.

Prov

Det finns inga centrala bestämmelser för prov. Det finns inget som säger att man måste ha dem över huvudtaget det passar inte alla elever. Däremot är skolan skyldig att anordna nationella prov i svenska, engelska och matte på gymnasiet, dessa är desamma för alla elever i Sverige och finns till för att man ska få ett likvärdigt betyg oavsett i vilken skola eller del av landet man bor. Prov är allmänna handlingar och innebär att du som elev kan få tillbaka det utan att förklara varför, såvida det inte omfattas av sekretess.²⁰ Prov får endast utgöra en del av bedömningen av den enskilde eleven, de har konstruerats för att bestämma om den studerande uppnått vissa kunskapsnivåer som definierats via kursplaner och betygskriterier och inte för att rangordna individer. Det finns inget stöd för att betygsätta personliga egenskaper så som, kön, etnisk eller social bakgrund, i vare sig skollagen, läroplanerna, kursplanerna eller betygskriterierna. Men från tidigare forskning om hur lärare betygsätter vet man att personlighetsfaktorer ändå vägs in.²¹

Svårigheter vid bedömning och betygsättning

Eftersom varje elev är unik måste all bedömning vara individrelaterad och inte generaliserande. Målrelaterad bedömning ger information om hur väl eleven nått uppställda mål. Syftet med individrelaterad bedömning är att bekräfta individen i dennes personliga existens. Bedömning av eleverna och deras arbete bör vara allsidig. Vid förhör och redovisningar bör förutom elevernas detaljkunskaper även hänsyn tas till deras förståelse för orsakssammanhangen. Betygen kan vara livsavgörande för den enskilde eleven genom att de används som urvalsmetod vid ansökan om tillträde till högre studier.

I det nuvarande bedömningssystemet måste läraren på ett annat sätt än tidigare kunna bedöma elevers kunskaper och verbalisera sin bedömning. Detta arbete kräver både kompetens och mycket tid för att kunna göras omsorgsfullt. Läraren måste vara övertygad om att varje elev har tillräckligt med kunskaper inom varje kunskapsområde. Därför bör lärarna få ökad tid för bedömning och analys vilket också den elev som skall bli bedömd bör få.²²

Svårigheter med betygsättningen kan vara:

- Att inte jämföra med andra elever
- Att bortse från värderingar och föreställningar som lärare har om eleverna
- Att tolka och utläsa kunskapsmålen i kursplanerna
- Att inte påverkas av ledningens krav på att hävda sig i konkurrensen skolor emellan

¹⁹ Åborg, Y. (2000). *Elevrätt – en bok om elevers rättigheter*

²⁰ Åborg, Y. (2000). *Elevrätt – en bok om elevers rättigheter*

²¹ Andersson, H. (1999). *Varför betyg?*

²² Lindström, L. & Lindberg, V. (2005). *Pedagogisk bedömning. Om att dokumentera, bedöma och utveckla*

Konsekvenser

Elever kan ge oväntade svar på frågor i ett prov. Svaren kan vara bättre än vad läraren förväntar sig vilket kan ge svårigheter vid rättningen och göra den mer godtycklig beroende på hur kunnig eller skicklig läraren är.²³ För att inga ovidkommande faktorer ska påverka bedömningen bör eleverna vara anonyma för bedömaren, d v s läraren skall inte kunna associera en viss prestation till särskild elev. Men även om läraren inte kan sammankoppla proven med personer, så finns det en möjlighet att handstil, språkfel eller språklig stil, gör att läraren associerar svaret med ett visst kön eller en samhällsgrupp. Lärare kan omedvetet medvetet låta sig påverkas av sina egna värderingar vid bedömning.²⁴

Enligt en tidigare undersökning fann man att flickor i allmänhet förväntades prestera sämre i matte och tekniska ämnen av sina lärare som gav dem lägre betyg. I de könsrollsmönster som finns anses matte och teknik vara okvinnligt och kan skapa svårigheter i en mansdominerad värld. Detta kan ge flickan dåligt självförtroende och få henne att prestera dåligt i matte och tekniska ämnen. I en undersökning som gjorts där tre miljoner elever deltog noterades att flickor och pojkar oftast var lika duktiga i dessa ämnen fram till högstadiet, där pojkar var något bättre på standardproven. Många flickor hoppade av studier med inriktning på matte och teknik.²⁵

Bedömningar som används rätt kan stimulera lärandet och stärka elevens tillit till sin egen förmåga. Men all bedömning stimulerar inte till lärande och tillit till förmågan. En risk i bedömningen är om den ensidigt används summativt. Då kan eleverna uppleva bedömningen som stämplande eller dömande.²⁶

Metod och praktiskt genomförande

Vi vet att betygsättning kan gå till på lite olika sätt men det saknas forskning i bedömning av enskilda prov. Vi ville ta reda på hur det gick till ute i skolorna vid bedömningssituationer och vi tänkte göra en experimentliknande undersökning där lärare skulle få rätta ett prov. Vi kallar vår undersökning experimentliknande, då undersökningen ”inte är experiment i strikt mening eftersom forskaren inte har fullständig kontroll över experimentsituationen”.²⁷ Urvalet skedde inte heller med hjälp av slumpen och endast två av lärarna fick agera sin egen kontrollgrupp.

Vår idé var att utforma ett prov i Hållbar utveckling som inte är ett ämne i sig men skall ingå i varje kurs. Vi gjorde på intet sätt anspråk på att vara experter inom området utan ville ta hjälp av personer med stor kännedom i området. Vi försökte först ta kontakt med Helene Wåhlander som arbetar med Hållbar utveckling på Pedagogiska institutionen, tyvärr var hon inte anträffbar förrän i december. Nästa steg var att vi tog kontakt med Universeum, Renova, Världsnaturfonden, Greenpeace, Ekocentrum med flera, de hade inget material vi kunde använda oss av. Till slut fick vi kontakt med Karin Theorin-Mägi som är skolinformatör på Kretsloppskontoret. Av henne fick vi hjälp att utforma frågor till ett prov som vände sig till

²³ Egidius, H. (1992). *Betygsättning och utvärdering i skola och utbildning*

²⁴ Korp, H. (2003). *Kunskapsbedömning – hur, var och varför*

²⁵ Steele, C.M. (1997). *A Threat in the Air, How stereotypes shape intellectual identity and performance*

²⁶ Lindström, L & Lindberg, V. (2005). *Pedagogisk bedömning. Om att dokumentera, bedöma och utveckla*

²⁷ Esaiasson, P m. fl. (2004). *Metodpraktikan – Konsten att studera samhälle, individ och marknad*

elever på gymnasienivå. På grund av att vi valde Hållbar utveckling som ämne till provet fick vi ett bredare urval av lärare som kunde delta i undersökningen. Vi behövde inte begränsa oss till kärnämnes- eller karaktärsämneslärare på något speciellt program. Vi ville även arbeta med ett tema som vi kommer att använda oss av i fortsättningen med tanke på att vi är tre yrkeslärare med olika bakgrund.

Nästa steg i processen var att vi tog kontakt med ett flertal rektorer på gymnasieskolor i Göteborgsregionen. Det var svårt att rekrytera lärare till vårt experiment för att rektorerna på skolorna var oroliga för hur deras personal skulle reagera över att utsättas för undersökningen. Arbetsbelastningen är mycket hög under höstterminen med nationella prov som ska genomföras och rättas. Under perioden som vår undersökning skulle genomföras var det dessutom gymnasiedagar i Göteborg. Detta medförde att det blev svårare att hitta personer som ansåg sig ha tid att ställa upp. Till slut fick vi namn på en lärare som kunde tänkas delta. Instruktioner inför undersökningen fick lärarna i både skriftlig och muntlig form. Eftersom det började bli ont om tid valde vi att den första läraren rätta samtliga proven då vi var osäkra på om vi skulle få fler lärare till vår undersökning. Detta innebar att vi inte han korrigerade texten i den redan utformade lärarinstruktionen (se bilaga 1:1). Så småningom fick vi ytterligare fyra lärare till undersökningen vilket gjorde att vi valde att använda samma tillvägagångssätt.

Rektorerna blev informerade om det egentliga syftet med studien enligt de forskningsetiska reglerna. Däremot var vi tvungna att undanhålla delar av informationen för de deltagande lärarna, för att vi skulle få ett så tillförlitligt resultat som möjligt.

Vi gjorde en undersökning på tre gymnasieskolor och fyra olika program i Göteborgsregionen. Eleverna skulle få göra ett prov med 8 enkla frågor som deras lärare sedan skulle få rätta och betygsätta i två till tre omgångar. Under tiden eleverna skrev provet hade läraren möjlighet att fylla i en enkät med enkla frågor med uppgifter om kön, ålder, yrkesverksamma år, undervisningsämnen och den tid de arbetat på skolan, (se bilaga 2:2). I undersökningen deltog 5 lärare i åldrarna 30-60 år och 91 elever i åldrarna 16-19 år. Vid det första rättningstillfället var provet från sina egna elever i originalform. Alltså handskrivet och med elevernas namn och klass på ett försättsblad, eleven skulle vara känd för läraren. Vid det andra rättningstillfället hade vi skrivit om svaren på alla 91 proven till datautskrift. Samtliga lärare fick återigen rätta proven från sina egna elever men vid det här tillfället var namn och klass borttaget från proven. För att vi skulle veta vem som skrivit vilket prov kodade vi dem, sammanlagt blev det 182 rättade prov. Två lärare fick rätta provet en tredje gång och då var det handskrivet och i originalform, totalt rättades 68 prov vid det tillfället. Det totala antalet rättningar blev således, 250 stycken. Vi ville köra världen i repris och undersöka om rättningen blev lika vid alla rättningstillfällena. På så sätt fungerade första rättningen som kontrollgrupp till tredje rättningen.

Datamängden analyserades i statistikprogrammet SPSS.²⁸ Vi valde att arbeta med SPSS för att få så stor tillförlitlighet i resultatet som möjligt. Detta medförde dock att vi inte kunde arbeta hemifrån utan att vi var hänvisade till universitetet som tillhandahåller programmet. Resultaten kommer att redovisas i tabeller och visar en jämförelser mellan olika variabler.

Den information som lärarna fick var att vi ville undersöka dels elevernas kunskaper i området Hållbar utveckling och dels huruvida det var lättare och mer tidsbesparande att rätta

²⁸ SPSS Scandinavia AB

ett datautskrivet prov. Eleverna fick veta att de skulle göra ett oförberett prov och att de skulle svara så utförligt på frågorna som möjligt. Vi kunde inte meddela dem innan provet att resultatet inte skulle vara betyggrundande då risken skulle ha varit stor att de undvikit att delta över huvud taget. Vi berättade därför först efter de svarat på frågorna att resultatet inte skulle påverka deras betyg. Men vi informerade givetvis även om att Hållbar utveckling skall ingå i deras kurser, för att de inte skulle känna att de gjort något förgäves.

Som tack för att lärare och elever ställde upp i undersökningen bjöd vi på choklad och pepparkakor.

Forskningsetik - Etiska övervägande

För att kunna genomföra denna undersökning har vi tagit hänsyn till de forskningsetiska regler som gäller. Innan vi utförde undersökningen tog vi kontakt med skolledare på olika gymnasieskolor och talade om vårt egentliga syfte med undersökningen och hur den skulle genomföras. Vi erbjöd oss att lämna ut en skriftlig information till eleverna men detta ansågs inte vara nödvändigt av rektorerna. Eftersom undersökningen inte innefattar privata frågor eller är av etisk känslig karaktär räckte det med att få samtycke på detta sätt. De deltagande lärarna fick en mindre detaljerad information eftersom det fullständiga syftet skulle äventyra undersökningens resultat.

Vi talade om för lärarna att deltagandet var frivilligt och att de när som helst fick avbryta sin medverkan. De i sin tur delgav eleverna samma information. De uppgifter som samlats in ska inte användas till annat än denna studie, eftersom forskningsmaterialet för ändamålet inte får användas eller utlånas för kommersiella- eller andra ickevetenskapliga syften. Dessutom var alla som deltog anonyma så att ingen skulle kunna utläsa vem som var med i undersökningen.

Efter att experimentet genomfördes bestämde vi ett möte med varje deltagande lärare där vi gav dem den fullständiga informationen som var att undersöka om lärare påverkades av att de visste vem eleven var vid rättning av ett prov. Vi erbjöd dem att ta del av rapporten när den var färdig.²⁹

²⁹ Esaiasson, P m. fl. (2004). *Metodpraktikan – Konsten att studera samhälle, individ och marknad*

Resultat

Ingen har tidigare gjort en undersökning som går ut på att mäta resultatet av rättning av ett enskilt, enkelt utformat prov. Vår nyfikenhet väcktes efter rundfrågning bland vänner och kollegor om deras tankar om betyg och bedömning. De allra flesta av dem vi frågade hade en åsikt om bedömning och ansåg att den inte alltid var likvärdig. Om nu detta är en utbredd åsikt bland människor i allmänhet, varför finns det då så lite forskning i ämnet? Detta ledde till att vi ville genomföra denna undersökning. Samtidigt som det dök upp fler frågor kring området. Är det rent utav så att man har sådan respekt för skolväsendet och inte vill trampa lärarprofessionen på tårna? Eller är forskning i skolvärlden inte lika prestigefull?

Provet i Hållbar utveckling var synnerligen enkelt utformat och borde vara enkelt att bedöma. Sannolikheten skulle vara stor att bedömningen enbart utgick ifrån elevens prestation. Blir det då stora skillnader i resultatet så finns det verkligen anledning att tro att lärare väger in personligheten hos eleven vid bedömning. Med tanke på tidigare forskning, då särskilt Cross & Frary (1996), som redovisar hur lärare och elever anser att bedömning går till, ansåg vi ändå att det vore troligt att vi skulle se skillnader även i ett enkelt utformat prov.

Undersökningen har skett i flera steg. Då vi redan från början märkte skillnader i resultatet i den första insamlade datamängden. Vi ville delvis se om det fanns några svagheter i vår studie och även pröva vårt antagande på flera sätt.

Provet vi utformade innehöll följande åtta frågor:

- Vad kan man kompostera? Räkna upp 5 saker
- Varför ska du kompostera ditt matavfall istället för att slänga det i soppsåsen?
- Vad kan du som konsument tänka på att göra för miljön när du handlar?
- Hur kan du hjälpa miljön med att minska utsläpp av koldioxid och tungmetaller?
- Avloppsreningsverket i Göteborg renar avloppsvatten från slam (organiska rester). Vad gör man med slammet?
- Motivera varför det är bättre ur miljösynpunkt att dricka kranvatten jämfört med att köpa vatten på flaska.
- Avloppsreningsverken har problem med att vi spolar ner fel saker i avloppet. Räkna upp 5 exempel på saker som inte ska slängas i avloppet.
- Vad gör man med gamla läkemedel eller mediciner som har blivit över?

Lärarna skulle rätta samma prov vid två tillfällen första gången var det handskrivet och i original och vid det andra tillfället var det skrivet på dator och eleverna var anonyma. Lärarna fick fria händer att själva utforma en rättningssmall. Från början hade vi sex lärare som ville delta i vår undersökning. En av lärarna kunde dock inte genomföra sin uppgift inom tidsramen på grund av yttre omständigheter. Vi fick därför avböja dennes medverkan vilket resulterade i att vi fick totalt fem deltagande lärare.

Eftersom lärarna utformade sin egen rättningssmall medförde detta att vi fick vitt skilda system att bedöma ett prov. En fråga kunde hos en lärare vara värd en poäng medan det hos en annan blev fem poäng. En av de deltagande lärarna använde ett eget bedömningssystem där de rätta svaren fick olika antal plustecken beroende på frågornas karaktär. Detta gör att den totala poängsumma varierar på proven. När vi samlat in alla data jämförde vi alla 91 handskrivna svaren med de 91 datautskrivna, i SPSS.

Tabell 1:1

Tabellen visar alla frågor på samtliga prov i undersökningen med två rättningar.

	Total poäng	Medelpoäng	Antal prov
Handskrivna prov	979,5	10,76	91
Datautskrivna prov	1002	11,01	91

Vi fann inga nämnvärda skillnader. Vilket i sig är intressant då betygsättningen verkar fungera på ett tillfredsställande sätt och lärarna bedömer prestationer och inte personer.

Det vi såg var att totalpoängssumman var högre på proven som var datautskrivna än på dem som var handskrivna. Skillnaden var att de datautskrivna proven blev 0.25 poäng högre per prov. Ur forskningssynpunkt kan det också vara intressant att skillnaden inte är anmärkningsvärt, då tidigare undersökningar vi tagit del av pekar på motsatsen. Men eftersom det ändå fanns en viss skillnad bestämde vi oss för att analysera svaren fråga för fråga.

Analys av anonymitetseffekt

Vi delade in svaren på frågorna i två kategorier, frågor som har litet tolkningsutrymme och frågor som har stort tolkningsutrymme där svårigheten i att bedöma svaren ökar. Vi har också delat upp resultatet i antalet poäng på rättningen och betyg i en egen tabell.

Tabell 2:1, visar resultatet på frågor med litet tolkningsutrymme. Frågorna är följande,

- Vad kan man kompostera? Räkna upp 5 saker
- Avloppsreningsverken har problem med att vi spolar ner fel saker i avloppet. Räkna upp 5 exempel på saker som inte ska slängas i avloppet.
- Vad gör man med gamla läkemedel eller mediciner som har blivit över?

Tabell 2:2, visar resultatet på frågor med stort tolkningsutrymme. Frågorna är följande,

- Varför ska du kompostera ditt matavfall istället för att slänga det i soppåsen?
- Vad kan du som konsument tänka på att göra för miljön när du handlar?
- Hur kan du hjälpa miljön med att minska utsläpp av koldioxid och tungmetaller?
- Avloppsreningsverket i Göteborg renar avloppsvatten från slam (organiska rester). Vad gör man med slammet?
- Motivera varför det är bättre ur miljösynpunkt att dricka kranvatten jämfört med att köpa vatten på flaska.

Tabell 2:3, visar provets:

- totala poäng
- betygsättning.

Tabell 2:1**En jämförelse av resultaten mellan handskrivet- respektive datautskrivet prov****Frågor med litet tolkningsutrymme**

	Genomsnittlig poäng per fråga	Medelvärdes-skillnad	Signifikans	Antal poängsatta frågor
Fråga 1 Handskrivet	3,09	0	1,0	84
	Datautskrift			
Fråga 7 Handskrivet	3,20	0,03	0,71	76
	Datautskrift			
Fråga 8 Handskrivet	1,25	-0,03	0,18	44
	Datautskrift			

Kommentar: Dessa frågor var av enklare karaktär med lättolkade svarsalternativ. Det som vi kan utläsa av värdena i tabellen är att det finns en märkbar skillnad i fråga 8. Knappt hälften har fått poäng på svaren på frågan, endast 44 av 91.

Tabell 2:2**En jämförelse av resultaten mellan handskrivet- respektive datautskrivet prov****Frågor med stort tolkningsutrymme**

	Genomsnittlig poäng per fråga	Medelvärdes-skillnad	Signifikans	Antal poängsatta frågor
Fråga 2 Handskrivet	1,30	0,72	0,25	69
	Datautskrift			
Fråga 3 Handskrivet	1,61	0,11	0,03	79
	Datautskrift			
Fråga 4 Handskrivet	1,58	0,01	0,92	61
	Datautskrift			
Fråga 5 Handskrivet	1,28	0,02	0,77	27
	Datautskrift			
Fråga 6 Handskrivet	1,26	-0,12	0,08	48
	Datautskrift			

Kommentar: Frågorna som redovisas i tabellen var av mer analytisk karaktär. I fråga 3 finns det en signifikant skillnad vid rättningarna, fråga 6 har också en anmärkningsvärd skillnad. Båda värdena är markerade med fet stil.

Tabell 2:3**En jämförelse av resultaten mellan handskrivet- respektive datautskrivet prov****Tabell med total poäng och betyg på proven**

	Genomsnittlig poäng per elev	Medelvärdes-skillnad	Signifikans	Antal poängsatta frågor
Total P. Handskrivet	10,76	-0,25	0,15	91
	Datautskrift			
Betyg Handskrivet	2,79	-0,04	0,56	91
	Datautskrift			

Kommentar: Denna tabell visar en jämförelse av skillnader mellan två rättningstillfällen, handskrivet prov respektive datautskrivet. Tabellen visar den totala poängsumman som eleverna fick på provet samt det betyg som sattes.

Kommentarer till tabellerna 2:1 – 2:3

Det finns ingen alls eller ingen nämnvärd skillnad i resultatet på fråga 1 och 7. Däremot i fråga 8 finns det en märkbar skillnad. Det finns dock en möjlighet att slumpen orsakat skillnaden. När vi går vidare i analysen av nästa tabell ser vi att det är två frågor där det finns en anmärkningsvärd skillnad. Frågorna var 3 och 6. Här kan det vara anonymitetseffekten som utgör skillnaden. Om man som lärare vet vem eleven är kan det medföra att man höjer eller sänker poängen. Enligt Cross & Frary (1996) så kan det handla om omvänd diskriminering där läraren vill kompensera de negativa sociala konsekvenserna. Andra tänkbara faktorer som kan finnas i skillnaden kan vara att frågorna är av mer analyserande karaktär och kan ge större tolkningsutrymme för läraren vid rättning.

Om vi utgår från att det blir skillnad vid bedömningen när eleven är känd för läraren så innebär resultatet av fråga tre att läraren tenderar att sätta högre poäng för denne. På fråga sex är det fler elever som fått högre poäng i det datautskrivna provet. Detta kan dock bero på att det är färre elever som fått poäng på frågan.

Resultatet av poängen och betygen, (se bilaga 4:1 för betyg), som resovisas i tabellen visar att det inte finns någon generell tendens bland lärarna att sätta olika betyg vid de olika rättningstillfällena. En tänkbar förklaring kan vara att de vida betygsspann som lärarna gått efter var så stora att betygen inte påverkades. Det vi däremot kan se är att poängskillnaden varierade mellan rättningarna. Så här långt visar vårt analysmaterial en tendens till att vår hypotes, att läraren blir påverkad av att veta vem eleven är vid bedömning av ett prov stämmer. Vi vill dock utsätta materialet för en djupare analys och för att pröva andra tänkbara faktorer som kan påverka resultatet.

Analys av andra tänkbara faktorer

Eftersom vi upptäckte en skillnad i rättningen mellan det handskrivna provet och det datorutskrivna provet men inte kunde utläsa om den orsakades av anonymitetseffekten så fick två lärare rätta det handskrivna provet en gång till. De fick använda samma rättningssmall men fick inte se hur de rättat proven vid det första tillfället. Vi ville på detta sätt se om vårt antagande om anonymitetseffekten höll eller om det var någon annan faktor som påverkade lärarna.

Följande datamängd består av 68 prov som rättades en andra gång. Eleverna var alltså kända även vid det här rättningstillfället.

Tabell 3:1

Här visas resultaten av de handskrivna proven som blev rättade vid två tillfällen

Frågor med litet tolkningsutrymme

	Genomsnittlig poäng per fråga	Medelvärdes-skillnad	Signifikans	Antal poängsatta frågor
Fråga 1 Handskrivet1	3,39	0,02	0,77	41
	3,37			
Fråga 7 Handskrivet1	3,32	-0,05	0,53	38
	3,37			
Fråga 8 Handskrivet1	1,33	-0,11	0,10	18
	1,44			

Kommentar: Det finns ingen signifikant, men ändå tydlig skillnad i bedömningen, den har vi markerat i fet stil.

Tabell 3:2

Här visas resultaten av de handskrivna proven som blev rättade vid två tillfällen

Frågor med stort tolkningsutrymme

	Genomsnittlig poäng per fråga	Medelvärdes-skillnad	Signifikans	Antal poängsatta frågor
Fråga 2 Handskrivet1	1,08	-0,11	0,23	33
	1,18			
Fråga 3 Handskrivet1	1,41	-0,03	0,70	39
	1,44			
Fråga 4 Handskrivet1	1,69	0,14	0,10	29
	1,55			
Fråga 5 Handskrivet1	1,09	-0,09	0,17	11
	1,18			
Fråga 6 Handskrivet1	1,17	0,13	0,19	24
	1,04			

Kommentar: Det finns skillnad i bedömningen, som vi har markerat i fet stil.

Tabell 3:3

Här visas resultaten av de handskrivna proven som blev rättade vid två tillfällen

Tabell med total poäng och betyg på proven

	Genomsnittlig poäng per elev	Medelvärdes-skillnad	Signifikans	Antal poängsatta frågor
Total P. Handskrivet1	10,09	-0,22	0,26	48
	10,31			
Betyg Handskrivet1	2,15	-0,10	0,32	48
	2,25			

Kommentar: Tabellen visar den totala poängsumman som eleverna fick på provet samt det betyg som sattes.

Kommentarer till tabell 3:1 – 3:3

I den här delen av undersökningen finns det ingen jämförelse av anonymitetseffekten. Skillnaden på fråga 8 är här större än i tabell 2:1 trots att det är färre deltagare och att dessa är kända för läraren.

Det vi utläser av nästa tabell är att fråga 4 inte visar någon signifikant men anmärkningsvärd skillnad. Värdena i fråga 5 och 6, är inte heller signifikanta men skiljer sig ändå till viss del. Trots att eleven är känd vid båda tillfällena så skiljer sig bedömningen anmärkningsvärt. Detta innebär att anonymitetseffekten inte kan uteslutas men, hur kommer det sig att man bedömer samma svar på en fråga så olika?

Det man klart kan utläsa är att läraren inte är konsekvent i sin bedömning utan att ett rätt kan bli fel och tvärt om. Det skiljer sig märkbart mellan de två proven. Läraren läser svaren för en tredje gång och kan därmed ha tänkt över och omvärderat vad som är rätt eller fel svar. Ytterligare en aspekt är att lärarna själva inte utformat provet, och kan då ha svårigheter i att bedöma vad som är en bra prestation eller inte. Felen kan också bero på att frågan är formulerad på ett sådant sätt att svaren blir svårtolkade.

Resultatet av denna utökade del gör att vår nyfikenhet väcks för att se om det är andra faktorer som spelat in vid bedömningen och att vi ytterligare vill klargöra resultaten i vår undersökning.

Fördjupning av undersökningen

Som vi tidigare nämnt utförde vi undersökningen i flera steg för att göra en djupare analys av datamängden. Vi ville se om anonymitetseffekten yttrar sig hos vissa kategorier av lärare med olika bakgrund och erfarenheter. Eller om bedömningen blir annorlunda i vissa situationer och under andra omständigheter. Vi undersökte om följande betingelser påverkar resultatet vid rättning av provet.

- samma/olika kön på lärare och elev
- lärarens yrkesverksamma år
- lärarens kännedom om Hållbar utveckling
- rättningstid

Lika barn leka bäst?

Den centrala frågan är fortfarande om det finns en anonymitetseffekt vid betygsättning av ett enkelt prov. Här vill vi se om betygen blir annorlunda om man är lik eller olik eleven. I denna undersökning innebär det, samma kön eller olika kön.

I samhällsvetenskapen har forskning bedrivits på hur likheter påverkar människor, Bailensson, m.fl. (2006). Människor tenderar att gynna dem som mest liknar en själv och alltså tillhör samma grupp. Med tanke på omfattningen av vårt material så hade vi möjlighet att analysera ytterligare faktorer. Vilket vi också gjorde, resultatet redovisar vi i följande tabeller.

Tabell 4:1

I tabellen visas en jämförelse mellan kön på lärare och bedömning av elever

Frågor med litet tolkningsutrymme

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 1 Hand & Data Samma kön	0,28	0,13	0,39	25
	Olika kön			0,15
Fråga 7 Hand & Data Samma kön	0,35	-0,05	0,75	24
	Olika kön			0,40
Fråga 8 Hand & Data Samma kön	0	-0,05	0,18	15
	Olika kön			0,05

Kommentar: Det finns inga signifikanta skillnader i bedömningen av dessa frågor.

Tabell 4:2

I tabellen visas en jämförelse mellan kön på lärare och bedömning av elever

Frågor med stort tolkningsutrymme

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 2 Hand & Data Samma kön	0,41	0,24	0,07	24
	Olika kön			0,18
Fråga 3 Hand & Data Samma kön	0,20	-0,01	0,94	27
	Olika kön			0,21
Fråga 4 Hand & Data Samma kön	0,61	0,41	0,02	19
	Olika kön			0,19
Fråga 5 Hand & Data Samma kön	0,21	0,14	0,24	12
	Olika kön			0,67
Fråga 6 Hand & Data Samma kön	0,11	-0,14	0,20	18
	Olika kön			0,25

Kommentar: Det finns en signifikant skillnad i bedömningen av fråga 4, som vi har markerat i fet stil. Vi har även markerat fråga 2, värdet är dock inte signifikant.

Tabell 4:3

I tabellen visas en jämförelse mellan kön på lärare och bedömning av elever

Tabell med total poäng och betyg på proven

	Genomsnittlig skillnad per elev	Medelvärdes-skillnad	Signifikans	Antal
Total P. Hand & Data Samma kön	1,59	-0,51	0,06	28
	Olika kön			1,08
Betyg Hand & Data Samma kön	0,39	-0,19	0,25	28
	Olika kön			0,21

Kommentar: Skillnaden i bedömningen på totalpoängen är nästintill signifikant, vilket vi har markerat i fet stil.

Kommentar till tabell 4:1 – 4:3

Liksom tidigare är det fråga 8 som visar störst skillnader och kan fortfarande bero på det låga antalet prov som fått poäng. Men så plötsligt så ser vi en stor skillnad i resultaten på fråga 2 och 4, som är av analytisk karaktär. Detta medför att även den slutgiltiga totalpoängen ändras och därför visas en nära nog, signifikant skillnad även där. Se tabell 4:3.

Sannolikheten är alltså stor att lärare gynnar elever av samma kön genom att ge dem högre poäng på proven. Enligt den samhällsvetenskapliga forskningen, Bailenson, m.fl. (2006), så befinner sig människor i ett ramverk där man jämför sig med människor i samma grupp och ser likheter, medan man ser olikheter och kontraster på dem som står utanför gruppen.

Nu när vi kommit så här långt så kan vi konstatera att visst påverkas lärare av sina elever. Vi vill ta reda på om det trots det vi redan kommit fram till finns ytterligare faktorer som stärker vår teori. Vi kontrollerar också om anonymitetseffekten är större hos vissa kategorier av lärare och om man påverkas under andra omständigheter.

Trygg och säker i yrkesrollen?

Här vill vi undersöka om det finns skillnader i hur man som lärare påverkas av yttre faktorer om man har lång erfarenhet i yrket. Vi har jämfört lärarnas år som yrkesverksamma i två kategorier, under eller över 10 års yrkeserfarenhet.

Tabell 5:1

Tabellen visar jämförelsen yrkesverksamma år och bedömning av elever

Frågor med litet tolkningsutrymme

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 1 Hand & Data <10år	0,17	-0,04	0,73	41
	>10år			0,21
Fråga 7 Hand & Data <10år	0,39	0,01	0,93	38
	>10år			0,38
Fråga 8 Hand & Data <10år	0,08	0,08	0,19	19
	>10år			0

Kommentar: Det finns en skillnad i fråga 8 som urskiljer sig marginellt. Men som vi tidigare nämnt så är det få elever till antal som fått poäng på frågan.

Tabell 5:2

Tabellen visar jämförelsen yrkesverksamma år och bedömning av elever

Frågor med stort tolkningsutrymme

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 2 Hand & Data <10år	0,25	-0,02	0,85	34
	>10år			0,27
Fråga 3 Hand & Data <10år	0,13	-0,16	0,09	39
	>10år			0,29
Fråga 4 Hand & Data <10år	0,24	-0,17	0,26	33
	>10år			0,41
Fråga 5 Hand & Data <10år	0,13	-0,01	0,94	12
	>10år			0,13
Fråga 6 Hand & Data <10år	0,17	-0,06	0,64	21
	>10år			0,22

Kommentar: Det finns skillnad i bedömningen, som vi har markerat i fet stil.

Tabell 5:3

Tabellen visar jämförelsen yrkesverksamma år och bedömning av elever

Tabell med total poäng och betyg på proven

	Genomsnittlig skillnad per elev	Medelvärdes-skillnad	Signifikans	Antal
Total P. Hand & Data <10år	1,26	0,05	0,82	48
	>10år			1,20
Betyg Hand & Data <10år	0,19	-0,16	0,26	48
	>10år			0,35

Kommentar: Ingen nämnvärd skillnad i resultatet.

Kommentarer till tabell 5:1 – 5:3

Resultatet från den här analysen visar att antalet yrkesverksamma år inte hänger ihop med anonymitetseffekten. Det påverkar inte heller betygsättningen, se tabell 5:3. Medelvärdeskillnaden är liten mellan de två gruppernas betygsättning. Det pekar i och för sig på att man inte blir säkrare på att rätta prov för att man har längre yrkeserfarenhet.

Vi går vidare i vår analys och vill veta hur kännedomen i Hållbar utveckling utgjorde någon skillnad.

Hållbar betygsättning?

Med tanke på att vi valde att utforma ett prov i Hållbar utveckling så ville vi undersöka huruvida lärarnas kunskaper i ämnet påverkar bedömningen. Lärarnas förkunskaper varierade från att ha begränsade kunskaper till att ha fått kompetensutveckling i form av kurser i ämnet. Men det finns ju inget enhetligt material för skolorna att tillgå i området. Det upptäckte vi då vi behövde hjälp att utforma vårt prov. Det kan också bero på skolornas policy och deras engagemang kring ekologi, ekonomi och sociala frågor som berörs i Hållbar utveckling.

Tabell 6:1**Påverkar lärares kännedom om Hållbar utveckling bedömningen av proven?****Frågor med litet tolkningsutrymme**

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 1 Hand & Data Inte god	0	-0,33	0,01	36
	God 0,33			48
Fråga 7 Hand & Data Inte god	0,22	-0,27	0,07	29
	God kön 0,49			47
Fråga 8 Hand & Data Inte god	0	-0,05	0,18	13
	God 0,05			31

Kommentar: Det finns en signifikant skillnad i bedömningen i fråga 1 samt en nästintill signifikant skillnad i fråga 7, värdena har vi markerat i fet stil.

Tabell 6:2**Påverkar lärares kännedom om Hållbar utveckling bedömningen av proven?****Frågor med stort tolkningsutrymme**

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 2 Hand & Data Inte god	0,12	-0,22	0,02	25
	God 0,34			44
Fråga 3 Hand & Data Inte god	0,08	-0,20	0,01	26
	God 0,27			53
Fråga 4 Hand & Data Inte god	0,20	-0,21	0,14	27
	God 0,41			34
Fråga 5 Hand & Data Inte god	0	-0,16	0,03	5
	God 0,16			22
Fråga 6 Hand & Data Inte god	0,09	-0,16	0,15	16
	God 0,25			32

Kommentar: Det finns signifikanta skillnader i bedömningen på tre av frågorna, som vi har markerat i fet stil.

Tabell 6:3**Påverkar lärares kännedom om Hållbar utveckling bedömningen av proven?****Tabell med total poäng och betyg på proven**

	Genomsnittlig skillnad per elev	Medelvärdes-skillnad	Signifikans	Antal
Total P. Hand & Data Inte god	0,68	-0,94	0	37
	God 1,62			54
Betyg Hand & Data Inte god	0,24	-0,03	0,84	37
	God 0,28			54

Kommentar: Det finns en signifikant skillnad i bedömningen, den har vi markerat i fet stil.

Kommentarer till tabeller 6:1 – 6:3

I samband med undersökningen delade vi ut en enkät, (se bilaga 2:1 och 2:2), där vi bad lärarna svara på frågor bland annat kring deras kännedom om Hållbar utveckling och om de berört ämnet i sin undervisning.

Vi ser en signifikant skillnad i den mycket enkelt formulerade fråga 1. Vilket kan tolkas som att lärare med god kännedom i ämnet har svårare för att bedöma svaret på frågan. De två andra frågorna ger också utslag men dessa värden är dock inte signifikanta.

I tabellen 6:2 utläser vi också skillnader mellan fråga 2, 3 och 5 som alla är signifikanta, vilket innebär att lärare som har god kännedom om Hållbar utveckling har störst variationer i sin bedömning av frågornas svar. De har alltså svårt för att vara konsekventa i sin betygsättning när de rättar.

Tabellen 6:3 visar att det finns en signifikant skillnad som innebär stora variationer av resultatet av rättningen på vissa frågor trots att lärarens kännedom om Hållbar utveckling är god. Det skiljer i genomsnitt 1 poäng per rättat prov.

De lärare som anser sig ha god kännedom om Hållbar utveckling har också visat sig påverkas mer av att känna eleverna vid bedömningen. Vilket ger extra tyngd i vårt antagande om anonymitetseffekten.

I och med att vi inte kunde berätta det egentliga syftet med vår undersökning i vår första kontakt med lärarna. Så bad vi dem att mäta tiden det tog att rätta proven. För att inte lärarna skulle gjort detta förgäves så valde vi att använda oss av den informationen i vår analys.

Hare eller snigel?

Eftersom anonymitetseffekten fortfarande inte kan uteslutas valde vi att jämföra om tiden som lades ner på rättningen påverkar sättet att bedöma. Som rubriken visar så delade vi in lärarna i två grupper beroende på hur lång tid de rättat proven. Den snabba gruppen, hararna, rättade under 1,5 minuter och den långsamma gruppen, sniglarna, rättade proven på mer än 1,5 minuter. Se tabell 8:1 och 8:2.

Tabell 7:1**Jämförelse av rättningstid på proven och om det påverkar bedömningen av proven****Frågor med litet tolkningsutrymme**

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 1 Hand & Data Snabb	0,10	-0,21	0,08	49
	Långsam			0,31
Fråga 7 Hand & Data Snabb	0,43	0,09	0,58	37
	Långsam			0,35
Fråga 8 Hand & Data Snabb	0	-0,07	0,19	23
	Långsam			0,07

Kommentar: Det finns en stor skillnad i bedömningen av fråga 1, som vi har markerat i fet stil.

Tabell 7:2**Jämförelse av rättningstid på proven och om det påverkar bedömningen av proven****Frågor med stort tolkningsutrymme**

	Genomsnittlig skillnad per fråga	Medelvärdes-skillnad	Signifikans	Antal
Fråga 2 Hand & Data Snabb	0,18	-0,18	0,10	37
	Långsam			0,36
Fråga 3 Hand & Data Snabb	0,24	0,06	0,54	40
	Långsam			0,18
Fråga 4 Hand & Data Snabb	0,67	-0,30	0,04	30
	Långsam			0,47
Fråga 5 Hand & Data Snabb	0,07	-0,14	0,24	15
	Långsam			0,21
Fråga 6 Hand & Data Snabb	0,15	-0,09	0,45	23
	Långsam			0,24

Kommentar: Det finns stora skillnader i bedömningen, som vi har markerat i fet stil. Det är dock endast fråga 4 som är signifikant.

Tabell 7:3**Jämförelse av rättningstid på proven och om det påverkar bedömningen av proven****Tabell med total poäng och betyg på proven**

	Genomsnittlig skillnad per elev	Medelvärdes-skillnad	Signifikans	Antal
Total P. Hand & Data Snabb	0,88	-0,80	0,001	50
	Långsam			1,67
Betyg Hand & Data Snabb	0,26	0	0,95	50
	Långsam			0,27

Kommentar: Det finns en signifikant skillnad i bedömningen, den har vi markerat i fet stil.

Kommentarer till tabeller 7:1 – 7:3






Finns det då något samband mellan rättningstid och olikheterna i poängsättningen? Ja, man skulle kunna tro att de lärare som tog längre tid på sig att rätta skulle också vara mer noggranna i sin bedömning. Men tvärt emot våra förutfattade meningar så skiljde sig deras bedömningar mest. Resultatet visar att på fråga 1 och 8 så är det den snabba gruppen av lärare, hararna, som rättar med större säkerhet.

Det finns en signifikant skillnad i den totala poängsättningen av proven. Ju längre tid läraren tar på sig att rätta desto större skillnad blir det i resultatet. Det kanske är så att de snabba lärarna är lite mer effektiva och har mer rutin på att rätta prov? Den genomsnittliga skillnaden på elevernas totalpoäng på provet var 1,67 poäng.

Illustration av rättningstiden

Tabell 8:1






De deltagande lärarnas rättningstid på det handskrivna provet

	Lärare 1	Lärare 2	Lärare 3	Lärare 4	Lärare 5
Under 1,5 minut					
Över 1,5 minut					

Kommentar: Tabellen visar genomsnittlig tid per rättat handskrivet prov.

Tabell 8:2

De deltagande lärarnas rättningstid på det datautskrivna provet

	Lärare 1	Lärare 2	Lärare 3	Lärare 4	Lärare 5
Under 1,5 minuter					
Över 1,5 minuter					

Kommentar: Tabellen visar genomsnittlig tid per rättat dataskrivet prov.

För att illustrera tiden av rättningen och förtydliga vår rubrik om harar och sniglar bestämde vi oss för att använda oss av symboler istället för kryss i rutorna. Detta innebär inte att vi tar ställning för eller emot och anser att något är bättre än det andra. Vi menar på intet sätt att ironisera över lärarnas sätt att arbeta.

Lärarnas egna kommentarer

I samband med att vi hämtade den samlade datamängden och att vi delgav lärarna det egentliga syftet med undersökningen fick vi ta del av spontana åsikter och erfarenheter lärarna hade av rättning och betygsättning. Under det informella samtalet passade vi på att ställa frågor liknande, kände du igen vilken elev som svarat vad? Tror du att du blir påverkad av att veta vem eleven är vid rättningen? Dokumentationen av samtalen skedde i form av minnesanteckningar. Samtliga lärare var mycket intresserade av vår undersökning och var väldigt tillmötesgående i sitt agerande. En lärare hade även skrivit ett följebrev som vi delvis citerat. Vi tycker att det är viktigt att förmedla lärarnas åsikter som ett komplement för att ge ett djup i rapporten. Det är viktigt att lyfta fram lärarnas egna reflektioner kring bedömning. Vi är mycket tacksamma för lärarnas hjälp att ta sig tid när deras arbetsbelastning redan var hög. Lärarnas egna kommentarer till undersökningen lyder som följer.

Samtal med lärare 1

Under tredje rättningen så fanns mycket tankar på betygsättningen. Läraren berättade spontant om en elev som tidigare gjort ett arbete på dator. Eleven hade läs- och skrivsvårigheter men använde sig ändå av ord som huruvida i sin uppsats. Detta gjorde läraren misstänksam, efter en koll på nätet upptäcktes att eleven kopierat, klippt och klistrat sitt arbete. Eleven får göra om arbetet i vår då inget betyg kunde sättas. Läraren ville inte sätta IG då elevens praktikplats skulle äventyras. Hade eleven "varit okänd" så kanske inte denna upptäckt gjorts, eftersom elevens arbeten inte alltid kontrolleras.

Läraren beskrev också att vissa elever skriver jättebra essäer i svenskan som man blir jätteglad över, de är fantastiskt bra... men de har inte hållit sig till ämnet när de svarat på frågan, de har skrivit om andra saker de gillat bättre. Då blir det IG ändå. Slutligen fick vi ett konstaterande: Såklart man väger in vilken elev det är när man rättar. Jag tänker på det hela tiden, och jag har lärt mig en hel del av er undersökning.

"Jag tycker alltid att det är svårt att rätta ett prov som jag inte själv har konstruerat. När jag sätter ihop ett prov själv vet jag på förhand ungefär vilka svar jag har tänkt mig. Dessutom jobbar jag mer med öppna frågor, av typen *'berätta det du vet om kompostering!'* eller *'Hur tror du att miljöarbetet i Sverige kommer att se ut om femtio år?'*. Då kan alla elever svara något. Ofta ger jag ett samlat betyg på uppgiften än rena poäng. Detta gör jag för att kunna ge mer rättvisande betyg. Det finns alltid elever som kan plugga in en massa fakta och sedan skriver 'alla rätt' på provet. Det betyder inte alltid att de har förstått uppgiften. De elever som kan diskutera och tänka själva har en djupare förståelse än de som endast kan rapa upp fakta. Tycker jag... (Givetvis finns det elever som både kan lära sig fakta **och** analysera, diskutera)".³⁰

Samtal med lärare 2

Vid samtal med lärare 2 efter det att vi informerat om det egentliga syftet med vår undersökning kom det fram mycket spontana tankar. Bland annat om hur man ska bedöma elevens inlämnade skriftliga arbeten. Tidigare hade läraren kontrollerat sina elevs inlämningar för att upptäcka fusk. Men efter att ha påträffat allt för många elever som hämtat arbeten från Internet så är lärarens kommentar som följer, *"Jag har slutat att söka på nätet för att kolla upp mina elevs inlämnade uppgifter, jag blir så besviken..."*.

Läraren hade förövrigt vid andra rättningstillfället reflekterat över hur lätt det var att känna igen vissa elever på hur de svarat. (Detta gällde fråga 7, om vad som inte får spolas ner i

³⁰ Lärare 1 delger sina egna kommentarer till undersökningen.

toaletten, en fråga med litet tolkningsutrymme). Det var också mycket tänkvärt att rätta proven fler gånger än en för att tänka igenom svaren.

Samtal med lärare 3

Läraren som för övrigt är mycket intresserad av ämnet Hållbar utveckling, tyckte det var mycket roligt att få ett avbrott i den ordinarie lektionen för att eleverna skulle få göra provet. Redan idag använder sig läraren av ett miljömedvetet ”tänk” i undervisningen. Läraren hade inte upplevt det som konstigt eller jobbigt att rätta det andra tillfället då elevens namn var okänt. ”*Jag brukar inte tänka på vem som skrivit vad...*” Vid det första rättningstillfället hade läraren vikt bort försättsbladet till provet, där eleven skrivit sitt namn, och rättat alla på samma sätt. (Provets utformning kan ses i bilagorna).

Samtal med lärare 4

Läraren valde att bedöma varje fråga lika, tolkade frågorna som lika mycket värda och att de krävde samma mängd av tankekraft för eleverna. Läraren uttryckte själv att denne sällan skriver ut poäng på prov av sådan här enkel karaktär. Utan försöker istället bara se om eleverna svarat tillfredsställande på frågorna. Prov och läxförhör används endast som en mycket liten del vid den totala bedömningen innan betygsättning.

Samtal med lärare 5

Lärare 5 tänkte inte annorlunda vid rättningen andra gången då provet var datautskrivet utan bedömde som tidigare varje svars innehåll med poäng som motsvarade lärarens krav. Läraren kände heller inte igen de datautskrivna svaren.

Läraren satte plustecken i kanten för varje fråga i stället för siffror. Denne tyckte att det var bättre eftersom värdet av frågorna varierade från enkla uppräkningsfrågor av G-karaktär till VG-frågor där det ingick att analysera mera. Detta bidrog också till att det var svårare att poängsätta detta prov.

Samtal med utebliven lärare

En lärare kunde inte delta i undersökningen på grund av andra omständigheter. Vi träffade ändå läraren när vi samlat in vårt material, och talade då om det egentliga syftet med undersökningen. Läraren berättade om sitt sätt att vanligtvis rätta fråga för fråga över hela klassen, och utläsa svaren efter en mall med kriterier för innehåll per fråga sedan bedöma svaren. Läraren trodde sig inte bli påverkad i så stor utsträckning av att veta vem eleven var vid bedömning av prov, utan att det endast var prestationerna som avgjorde. Det är det stora hela som avgör det slutliga betyget. För att veta om man har bedömt inom ramen av de mål som eleverna skall uppnå inom ett ämne läser lärarna i arbetslaget varandras prov. Vi fick också höra av läraren att det fanns en lärare på skolan som kodade sina elevers prov för att vara så rättvis i sin bedömning som möjligt. Enligt läraren är detta tidskrävande och något som få skulle vilja göra. Dessutom tillade läraren att det var bra att arbeta i både ”låg-” och ”högpresterande” klasser på olika program, för att veta var man skall lägga sig bedömningsmässigt. Det kunde annars finnas risk att man som lärare höjde alternativt sänkte kraven på eleverna i de olika programmen.

Diskussion

Lärare har ett stort ansvar att sätta sig in i lagstiftning och betygskriterier. Lärare måste också ha en medvetenhet om de konsekvenser betygsättning kan få. Det är viktigt med bra ledning och arbetslag där man som lärare kan stötta varandra och jämföra material och betygsättning. Det känns lite skrämmande att tänka på att betygsättningen kan ge elever konsekvenser som påverkar deras framtid. Syftet med vår undersökning var att undersöka om man som lärare blir påverkad av att man vet vem eleven är vid rättning av ett prov. Vi ville också ta reda på om bedömningen sedan blir den samma om eleven är anonym eller inte. Som vi tidigare nämnt så finns det väldigt begränsat med forskning inom området. Enligt den Norska studien, Bonesrønning (1999) så kan betyg vara både hämmande och uppmuntrande. Även Cross & Frary (1996) har noterat att lärare ofta värderar elever efter attityd, ansträngning och prestation. Detta framkom också i vårt samtal med Lärare 1, som valt att skjuta upp betygsättningen i en kurs för att inte elevens praktikplats skulle äventyras.

Vi lägger ingen värdering i huruvida det är bättre eller sämre för eleven att vara känd eller okänd för läraren. Däremot menar vi att det blir en mer likvärdig bedömning om alla eleverna vid ett rättningstillfälle är okänd för läraren.

Vilka orsaker spelade då in i bedömningen i vår undersökning? Resultaten varierade mycket vid de olika rättningstillfällena, det kan inte enbart vara slumpen som gav skillnaderna. Det fanns många saker som kunde påverka att resultatet av rättningarna blev olika. En del lärare hade många prov, vilket kanske stressade dem och de slarvade med rättningen. Det kan ha varit så att de blev trötta efter en stunds rättande samt att tiden på dagen kunde ha inverkan på rättningen och ifall det var före eller efter lunch. Detta är yttre omständigheter som man inte kan påverka eller mäta. Givet det man kommit fram till i tidigare forskning så är sannolikheten stor att vårt resultat av undersökningen överensstämmer med tidigare antaganden.

Är det så att lärare tänker och känner annorlunda för innehållet när de rättar andra gången? Läraren kan ha funderat över frågan mellan rättningarna och omvärderat vad som är rätt eller fel svar. Men det kan också vara anonymitetseffekten som gör att läraren sätter högre poäng när eleven är okänd. Elevens bakgrund kan också ha haft betydelse för hur lärarna bedömde, då de vid rättningen av de datautskrivna i flera fall satte högre poäng. Trots att andra rättningen var datautskriven så kan det finnas en chans att språklig stil, Korp (2003), gör att läraren känner igen svaren hos eleven och påverkas av sina värderingar. Lärare 3 hävdar dock sin objektivitet och tittar inte på vem som skrivit vad.

Det kan också vara likhetseffekten som gjort att läraren satt olika poäng beroende på om det var handskrivet eller datautskrivet prov som bedömdes. Enligt den amerikanska undersökningen av Bailenson, m.fl. (2006) om likhetseffekten kom forskarna fram till att man tenderar att gynna dem som liknar en själv. I vår undersökning kunde vi utläsa signifikanta skillnader när läraren och eleven hade samma kön. Även här finns det en möjlighet att anonymitetseffekten kan spela roll. Vet man inte vem eleven är vid bedömningstillfället finns det inte heller någon möjlighet att gynna den som är av samma kön eller samhällsgrupp.

Det finns inga systematiska skillnader när det gäller lärares erfarenhet. Detta innebär att de lärare i vår undersökning som hade lång yrkeserfarenhet på intet sätt var säkrare på sitt sätt att rätta proven.

I resultatet av vår undersökning framkom en signifikant skillnad i poängsättningen hos de lärare som ansåg sig ha god kännedom om Hållbar utveckling. Elever kan ge oväntade svar på frågor i ett prov. Svaren kan vara bättre än vad läraren förväntar sig vilket kan ge svårigheter vid rättningen och göra den mer godtycklig beroende på hur kunnig eller skicklig läraren är. Egidius (1992). Detta kan jämföras med den norska studien, Bonesrønning (1999), där resultatet visade på motsatsen. Där fanns det stora skillnader mellan betygsättningen i matte, elever fick högre betyg av de lärare som hade högre kompetens inom området.

Man skulle kunna tänka sig att lärare som tar lång tid på sig att rätta ett prov också är omsorgsfullare men det visade sig vara tvärtom i vår undersökningsgrupp. Den genomsnittliga skillnaden i poäng per elev var 1,67 på de långsamma lärarna, jämfört med 0,88 för de snabba. Vad vi uppmärksammat är vikten i att vara noggrann när man rättar. Att ta god tid på sig är inte avgörande för att göra en likvärdig bedömning. Att vara snabb betyder inte att man är mindre omsorgsfull eller att man slarvar.

Rekryteringen till vår undersökning är inte att betrakta som slumpmässig och därför är inte resultatet generaliserbart till alla lärares betygsättning. De lärare som tillslut deltog i undersökningen gjorde det frivilligt. Man kan ställa sig frågan om det blir skillnad med lärare som är tvingade att delta i en liknande undersökning. Det finns dock en rimlighet i resonemanget att lärare har svårt att bortse ifrån elevens egenskaper när de sätter betyg. Alltså kan vi inte utesluta anonymitetseffekten.

En rektor på en friskola i Göteborgsområdet var mycket intresserad av vår undersökning och tyckte det var synd att vi inte gjorde den i våras. Skolan hade då arbetat mycket med just betygsättning, och lärarna hade fått fortbildning i ämnet. Rektorn menade att han kunde ha använt oss som resurs och bollplank. Tyvärr var denna friskola just nu inne i ett annat projekt och rektorn ansåg att tiden var illa vald för att lägga mer arbetsbörda på sina lärare. Det var dessutom de årliga gymnasiedagarna som upptog tiden för lärarna. Om vi hade fått möjligheten att utföra vår undersökning på skolan så hade det varit intressant att se om resultaten blev lika vid de båda rättningstillfällena.

Provresultat används som en del vid betygsättning. Man vill under kursens gång mäta elevernas kunskapsnivå i ämnet. Beroende på hur proven är utformade finns det olika saker att väga in vid bedömningen. Skolverkets allmänna råd visar övergripande vad som ska bedömas vid olika tillfällen. Vid ett skriftligt prov är det endast det som står på pappret som skall betygsättas.

Vår slutsats blir nu att det är särskilt viktigt att lärare har en gemensam syn på betygssättning och vad som bör krävas av eleverna. Lärarna i undersökningen använde sig av en egenhändigt utformad rättningsmall. Trots att de hade samma mall vid alla rättningstillfällen så rättade de olika vissa gånger. Ett steg i rätt riktning kan vara kodade prov eller att eleverna skriver under pseudonym. Man kan även som lärarlag eller ämnesvis byta elevers prov med varandra och rätta tillsammans för att därigenom få en samsyn på om man bedömt likvärdigt. Återigen måste vi poängtera vikten av att vara noggrann vid rättningstillfällen och vid bedömning.

Slutsatsen blir också att vår frågeställning inte höll när det gäller lärarnas betygsättning på provet. Däremot så blev det anmärkningsvärda skillnader i poängsättningen. Vilket visar att vårt syfte med undersökningen har uppnåtts och att det tillfört något nytt till forskningen.

Svagheter

Den begränsade tid vi hade för undersökningen innebar svårigheter i att rekrytera deltagare, under perioden pågick gymnasiedagar där gymnasieskolorna i regionen var engagerade. Detta innebar att vi endast fick fem deltagande lärare. Även nationella prov inföll under perioden samt att jullovet var i antågande, det medförde att det var svårt att komma i kontakt med rektorer och lärare på skolorna. Det dröjde att få in datamängden från lärarna vilket medförde att det drog ut på tiden för att analysera materialet. En akilleshäls med den experimentella metoden är att den som deltar är medveten om att den deltar i en undersökning.

Konsekvenserna av att vi begränsade undersökningen är flera, vi kan inte få samma säkerhet i vårt resultat.

- Undersökningen saknar balans i deltagare, det är övervägande män bland både lärare och elever
- Frågorna i testet skulle kunna vara formulerade på ett annat sätt
- Svarsfrekvensen var låg på vissa frågor. Detta kunde ha undvikits om vi hade gjort en pilotstudie för att testa svårighetsgraden i frågorna
- Eleverna visste att resultatet på provet inte skulle påverka deras kursbetyg och kan ha valt att slarva igenom frågorna
- Det var inte lärarna själva som utformat provet vilket kan ha inneburit svårigheter för dem att tolka svaren
- Tidsaspekten kan ha stressat lärarna
- Det är svårt att få exakt samma förutsättningar vid alla rättningstillfällen

Hur vi kunde ha gått tillväga:

Frågorna skulle ha provats på en grupp elever före själva undersökningen för att kontrollera formuleringarna på frågorna. Då hade vi haft en möjlighet att justera frågorna och kunnat ha utforma ett mer genomtänkt prov med mer svarsutrymme som hade gjort tolkningsutrymmet större för lärarna i studien.

Skulle vi kunna genomföra denna studie på något annat sätt?

- Ja, vi kunde ha gjort så att läraren rättade halva klassens prov som var handskrivet och den andra halvans prov som var datautskrivet. Därefter skulle läraren göra tvärt om för att se om det blev en skillnad
- Vi kunde också ha låtit en annan lärare som inte känner eleverna rätta halva klassens prov och den lärare som har klassen i vanliga fall rättat den andra halvan.

Förslag på framtida forskning

Studien skulle kunna genomföras under en hel gymnasieperiod det vill säga under tre års tid. Förslagsvis kunde cirka 200 lärare ifrån olika gymnasieskolor i landet delta. Låt säga att vi väljer ett kärnämne där lärarna själva får utforma proven. Efter rättningen av provet byter man med en kollega och rättar dennes prov och omvänt vid varje provtillfälle.

Det skulle också vara intressant att göra en enkätundersökning liknande den amerikanska studien, Cross & Frary (1996) som vi omnämnt i den tidigare forskningen. Lärare kan med egna ord beskriva vad man väger in vid betygsättning av ett prov. Frågorna skulle också kunna vara utformade så att lärarna svarade på olika påståenden om betygsättning. Detta enligt en skala på en känslotermometer, där man instämmer eller inte instämmer i påståendena.

Referenslista

Litteratur

- Aspelin, J. (1999). *Banden mellan oss – Ett socialpsykologiskt perspektiv på lärare – elev - relation*. Stockholm: Brutus Östlings Bokförlag Symposion
- Bergem, T. (2000). *Lärare i etikens motljus*. Lund: Studentlitteratur.
- Egidius, H. (1992). *Betygsättning och utvärdering i skola och utbildning*. Malmö: Gleerups
- Esaiasson, P m. fl. (2004). *Metodpraktikan – Konsten att studera samhälle, individ och marknad*. Stockholm: Nordstedts juridik
- Korp, H. (2003). *Kunskapsbedömning – hur, var och varför*. Stockholm: Myndigheten för skolutveckling.
- Kveli, A-M. (1994). *Att vara lärare*. Lund: Studentlitteratur.
- Lalander, P. & Johansson, T. (2002). *Ungdomsgrupper i teori och praktik*. Lund: Studentlitteratur.
- Lindström, L & Lindberg, V. (2005). *Pedagogisk bedömning. Om att dokumentera, bedöma och utveckla kunskap*. Stockholm: HLS Förlag.
- Lytsy, A. (2006). *Hur betygsätta en verkstad?* Växjö: Grafiska Punkten.
- Läraryrket. (2004). *Lärarens handbok*. Stockholm: Tryckindustri Information.
- Romhed, R. (2006). *A race to the bottom*. Ord och bild. Göteborg: SG Zetterqvist AB
- Sjödén, S. (2001). *Psykologi för gymnasieskolan*. Borås: Natur och kultur
- Skolverket. (2001). *Bedömning och betygsättning. Kommentarer med frågor och svar*. Stockholm: Liber distribution.
- Skolverket. (2002). *Att bedöma eller döma. Tio artiklar om bedömning och betygssättning*. Stockholm: Liber distribution.
- Skolverket. (2004). *Likvärdig bedömning och betygsättning*. Stockholm: Ljungbergs tryckeri
- Stukát, S. (2005). *Att skriva examensarbete inom utbildningsvetenskap*. Lund: Studentlitteratur
- Åborg, Y. (2000). *Elevrätt – en bok om elevers rättigheter*. Stockholm: Nordstedts Juridik AB

Artiklar

Bailenson, J N, Garland, P, Iyengar, S, Yee, N. (2006). *Transformed Facial Similarity as a Political Cue: A Preliminary Investigation*. Political Psychology.

Bonesrönning, H. (1999). *The variation in teachers grading practises: causes and consequences*. Economics of Education Review

Cross, L H. and Frary, R B. (1996). *Hodgepodge Grading: Endorsed by Students and Teachers Alike*. National Council on Measurement in education. New York

Meyer, M. (1914). *The limit of uniformity in the grading of college students by different teachers*. AAAS (American Association for the Advancement of Science. Washington, DC)

Steele, C. M. (1997). *A Threat in the Air, How stereotypes shape intellectual identity and performance*. American Psychologist.

Internetreferenser

Från Skolverkets hemsida, <http://www.skolverket.se/sb/d/204/a/4763>

Från Skolverkets hemsida, <http://www.skolverket.se/sb/d/208/a/6338>

Från SPSS Scandinavia AB, <http://www.spss.com>

Lärarinstruktioner

Till lektionen vill vi att Ni har med Er en klasslista, gärna med plats för anteckningar.

Under tiden testet genomförs, vill vi att Ni fyller i en mycket kortfattad enkät om ålder, kön, yrkesverksamma år och liknande frågor. All information kommer att behandlas enligt forskningsetiska regler och Era uppgifter kommer att vara anonyma.

Efter det att testet har genomförts kommer vi att ta hälften av dem och överföra dem till datautskrift. Under tiden kommer Ni ha möjlighet att rätta den första hälften av testen och betygsätta dem. Glöm inte att mäta hur lång tid det tar att rätta proven. Samt att efter Ni har rättat samtliga test, skriver en mycket kortfattad sammanfattning om Ni upplevde att det var lättare att rätta något av testen.

Svaren på frågorna finner Ni på nästa sida. Vi lämnar Er fria händer att tolka elevernas svar utifrån vår rättningsmall Ni får sätta en egen poängskala på frågorna och sedan betygsätta Era elevers testresultat, från IG - MVG. Ni har även möjlighet att sätta halva poäng om Ni finner att ett svar är ”nästan” rätt.

Vi kommer till Er och genomför testet i Er klass. Vi håller även en kort information till eleverna om Hållbar utveckling. Detta sker under vecka 48.

Resultatet på elevernas test kommer inte att vara betyggrundande. Detta kommer att meddelas först efter genomfört test i och med att vi vill att eleverna ska prestera sitt bästa.

Vi hoppas att med denna undersökning och detta test kan inspirera eleverna till att handla utifrån ett mer hållbart förhållningssätt och främja skolans fortsatta arbete med Hållbar utveckling.

Om Ni har några frågor kring undersökningen kontakta någon av oss,

Karin Svensson
Lärarstuderande
Göteborgs Universitet
Tel: xxxx – xx xx xx

Gill Brink
Lärarstuderande
Göteborgs Universitet
Tel: xxxx – xx xx xx

Git Öjergren
Lärarstuderande
Göteborgs Universitet
Tel: xxxx – xx xx xx

1. Vad kan man kompostera? Räkna upp 5 saker.

(T ex, kaffesump, äggskal, skal och skrottar av frukt och grönt och övrigt organiskt material)

2. Varför ska du kompostera ditt matavfall istället för att slänga det i soppåsen?

(För att minska sopmängden och för att få jord och naturligt gödsel)

3. Vad kan du som konsument tänka på att göra för miljön när du handlar?

(KRAV, lokalproducerat, rättvisemärkt, Svanen märkning, Bra miljöval)

4. Hur kan du hjälpa miljön med att minska utsläpp av koldioxid och tungmetaller?

(Åka kollektivt, gå och cykla mer)

5. Avloppsreningsverket i Göteborg renar avloppsvattnet från slam (organiska rester).
Vad gör man med slammet?

(Man rötter det och utvinner biogas för att driva bussar av resterna gör man sedan mull/biogödsel).

6. Motivera varför det är bättre ur miljösynpunkt att dricka kranvatten jämfört med att köpa vatten på flaska.

(Inga transporter, mindre förpackningar)

7. Avloppsreningsverken har problem med att vi spolar ner fel saker i avloppet.
Räkna upp 5 exempel på saker som inte ska slängas i avloppet.

(T ex mediciner, tops, lösningsmedel, blöjor, tamponger, kondomer, hushållspapper, inget annat än det som kommer från kroppen och toalettpapper)

8. Vad gör man med gamla läkemedel eller mediciner som har blivit över?

(Lämnar in till Apoteket)



GÖTEBORGS UNIVERSITET
Institutionen för pedagogik och didaktik

Bilaga 2:1

Hej!

Enkäten Du har fått kommer att användas i vårt examensarbete för att kunna analysera studiens resultat.

Enkäten har vi medvetet valt att formulera kortfattat för att den inte ska ta för mycket av Er tid i anspråk. All information som framkommer från studien och de svar som Du lämnar på enkäten kommer att vara anonyma.

Tack för Er medverkan,

Karin, Gill & Git

Lärarenkät

Är Du man eller kvinna?

Man Kvinna

Jag är _____ år gammal

Jag har arbetat i _____ år som lärare

Jag har arbetat i _____ år på den här skolan

Jag undervisar i följande ämnen _____

Vad har Du för kännedom om Hållbar utveckling?

Har Du berört ämnet Hållbar utveckling i Din undervisning?



GÖTEBORGS UNIVERSITET
Institutionen för pedagogik och didaktik

Bilaga 3:1

Test i Hållbar Utveckling

HT 2006

Namn: _____ Klass: _____

1. Vad kan man kompostera? Räkna upp 5 saker.

2. Varför ska du kompostera ditt matavfall istället för att slänga det i sopåsen?

3. Vad kan du som konsument tänka på att göra för miljön när du handlar?

4. Hur kan du hjälpa miljön med att minska utsläpp av koldioxid och tungmetaller?

5. Avloppsreningsverket i Göteborg renar avloppsvattnet från slam (organiska rester).
Vad gör man med slammet?

6. Motivera varför det är bättre ur miljösynpunkt att dricka kranvatten jämfört med att köpa vatten på flaska.

7. Avloppsreningsverken har problem med att vi spolar ner fel saker i avloppet.
Räkna upp 5 exempel på saker som inte ska slängas i avloppet.

8. Vad gör man med gamla läkemedel eller mediciner som har blivit över?

Kodbok

Betyg:

IG = 1

G- = 2

G = 3

G+ = 4

G ++ = 5

VG- = 6

VG = 7

VG+ = 8

Bedömning:

(+) = 1

+ = 2

+ (+) = 3

+ + = 4

Rättningsid:

Snabb Under 1,5 minut/elev = 1

Långsam Över 1,5 minut/elev = 2

Yrkesverksamma år:

Kort tid Under 10 år = 1

Lång tid Över 10 år = 2

Kännedom om Hållbar utveckling:

Inte så god kännedom = 1

God kännedom = 2

Könsfördelning elever – lärare

Samma kön på lärare och elev = 1

Olika kön på lärare och elev = 2

Elever och lärares kön

Kvinna = 1

Man = 2