

Våren 2010

Sektionen för lärarutbildning
Psykologi III

Sociala Kategorier och
Stereotypa
Uppfattningars
Inverkan vid Bedömning
av Gymnasieelever

Författare
Karin Riddersporre

Handledare
Tobias Johansson

Sociala Kategorier och Stereotypa Uppfattningars Inverkan vid Bedömning av Gymnasieelever

Karin Riddersporre
Högskolan Kristianstad

Stereotypa uppfattningar om människor grundar sig på egenskaper som associeras med sociala kategorier. Denna tillhörighet ger olika referenspunkter, så kallade standards, vid bedömning av andra. Utifrån teorin om shifting standards och antagandet att gymnasieprogram kan ses som sociala kategorier, undersöktes om och hur bedömningar av elever från olika typer av gymnasieprogram varierade och påverkades. Åttiofem personers enkätsvar som analyserades. Antingen bedömdes en typisk elev samt en specifik elev, eller enbart en specifik elev, från ett studie- eller yrkesförberedande program. Hypotesen var att en aktiverad effekt av en generell standard för en programtyp skulle skilja sig åt mellan programtyperna. Det fanns en signifikant skillnad i uppfattning om den typiska eleven från respektive programtyp. Dock fanns inga signifikanta skillnader kring den specifika eleven mellan programtyperna och således aktiverades ingen effekt av standarden.

Att människor har åsikter och uppfattningar om andra personer uppfattas av många som något självklart. Hur vi bedömer människor i vår omgivning beror till stor del på deras grupptillhörighet och de personligheter, attityder och beteenden som vi antar att personer inom denna grupp har. Flera faktorer avgör denna kategorisering, exempelvis kön, ålder, etnicitet och yrke (Kunda & Thagard, 1996). I vissa sammanhang är indelningen av människor i olika grupper eller kategorier extra tydlig. Ett sådant sammanhang är den svenska gymnasieskolan. Den här studien fokuserar på den typ av kategorisering som sker av elever i gymnasiet; hur människors bedömningar och uppfattningar av elever varierar med vilken typ av gymnasieprogram eleven går. Syftet är att se om och hur bedömningar och uppfattningar av elever från olika typer av gymnasieprogram varierar och påverkas utifrån shifting standards, en teori som kommer presenteras senare i studien.

På gymnasiet sker en tydlig kategorisering utifrån programtillhörighet då elever väljer olika program. Detta är en speciell omständighet som inte existerar inom grundskolan. Andersson (1999) menar att skolan sorterar människor genom att vissa elever kommer in på teoretiska utbildningar som leder till högstatusyrken i samhället, medan andra kommer in på utbildningar som leder till yrken med lägre status. Samtidigt måste påpekas att valet av gymnasieprogram är ett aktivt val av eleverna, där ändå en rad faktorer kan påverka hur elever väljer. Exempel på detta är betyg från grundskolan (Svensson, 2001) och traditionella könsmonster (Skolverket, 2009b). Hösten 2011 kommer det att ske en reform av Gymnasieskolan (Gy2011). Skolverket arbetar på regeringens uppdrag med att ge exempel på bland annat examensmål och programstrukturer. Den nya gymnasieskolan kommer att bestå av 18 nationella program, varav tolv är yrkesförberedande program och sex är högskoleförberedande program. De största förändringarna är att programstrukturen förändras samt att det kommer att bli en större skillnad mellan yrkes- och högskoleförberedande program, vilket är av stort intresse för denna uppsats. Elever på yrkesförberedande program ska vara bättre förberedda för yrkeslivet och direkt kunna påverka en yrkesbana efter utbildningen. De högskoleförberedande programmen ska bättre förbereda eleverna för

högskolestudier. Detta innebär i praktiken att eleverna i den nya gymnasieskolan kan få antingen en yrkes- eller en högskoleförberedande examen vilket ger en tydligare uppdelning programtyperna emellan (Skolverket, 2009a).

Social Kategorisering

Termen social kategorisering beskriver den process som sker när individer upplever andra som sociala kategorier. Den information som vi grundar våra bedömningar om andra på, kan bland annat delas in i stereotyper och individualiserad information. Stereotyper refererar till sociala kategorier som kön, ålder, yrke, ras eller andra kategorier som vi tror är associerade med specifika egenskaper eller beteenden (Corneille & Judd, 1999; Kunda & Thagard, 1996). Till skillnad från stereotyper som grundar sig på kategorisk information utifrån gruppstillhörighet, är individualiserad information en persons unika beteende och egenskaper (Abele & Petzold, 1998). Individualiserad information kan beskrivas som allt annat vi vet om en individ utöver dennes gruppstillhörighet, t.ex. beteende, personlighet och familjesituation. Gränsen mellan stereotyper och individualiserad information är inte alltid tydlig och det är ibland svårt att avgöra på vilken typ av information vi grundar våra bedömningar. Människor kan använda båda typerna av information när de gör sina värderingar. Dock verkar det som att den kategoriska informationen har stort inflytande när det inte finns tillräckligt med information om individen (Kunda & Thagard, 1996). Människor tenderar också att göra det enkelt för sig när de ska bedöma andra. Detta görs genom att kategorisera andra som medlemmar av grupper eftersom det kräver större ansträngning att individualisera dem (Neuberg, & Fiske, 1987). Kategorisk information kan som sagt vara information om att en specifik person tillhör en grupp av människor, t.ex. en yrkeskategori. Denna information väcker en representation av denna kategori i minnet. Minnesrepresentationen innehåller såväl typiska exempel och fall utifrån kategorin som gränsdragningar för kategorin (Scherman et al., 2009).

När individer, som är jämförbara på alla plan förutom att de tillhör olika grupper, bedöms i en riktning som överensstämmer med de förväntningar eller stereotypa bilder som finns av respektive gruppstillhörighet, uppkommer en stereotyping effect. Som exempel kan nämnas fall där män bedöms som bättre ledare än kvinnor eftersom det är den förväntningen och stereotypen som finns mot de specifika grupperna (Biernat, 2003). Biernat menar att en sådan bedömning kan göras trots att det inte finns några större individuella skillnader mellan männen och kvinnorna i fråga. Detta kan ses som ett stöd för att stereotyper används för att bedöma individer (Biernat, 2003). Kategorisk information bidrar till att skapa en överdriven uppfattning om såväl skillnader mellan kategorier som likheter inom kategorier. Denna accentuation effect av mellangruppsliga skillnader och inomgruppsliga likheter bidrar till att gruppstereotyper utvecklas och vidmakthålls (Corneille & Judd, 1999). Ett grundantagande för denna effekt är att det föreligger verkliga skillnader mellan kategorierna, och att processen i fråga bidrar till att betona och framhäva dessa skillnader (Sherman et al., 2009).

Stereotyper formas av attribut där mellangruppsliga skillnader är stora och inomgruppsliga skillnader är små (Sherman et al., 2009). När stereotyper används för att bedöma individer blir resultatet att individen assimileras till stereotypen (Biernat, 2003). Dessutom verkar det som att vi lättare lägger märke till gruppmedlemmar som förstärker mellankategoriska skillnader och inomkategoriska likheter. Dessa individer får också större inflytande över våra bedömningar och är lättare för oss att komma ihåg (Sherman et al., 2009).

En contrast effect å andra sidan, uppstår när människor bedöms vara mindre lika medelvärdet för den kategori de tillhör. Detta kan ske när människor mer bedöms gentemot gruppstillhörighet än om de bedöms enbart utifrån individualiserande information (Abele & Petzold, 1998). Detta innebär att stereotyper inte enbart betonar och spär på likheter inom kategorier, eller skillnader mellan kategorier. Stereotyper är verksamma även om en individ

från en kategori som vanligtvis värderas lägre bedöms mer fördelaktigt eller mer extremt än en annan, på grund av en inomgruppsstandard, där båda personerna är jämförbara förutom på just den stereotypa kategoriindelningen (Skowronski & Carlston, 1989). Detta, menar Biernat (2003), beror på att stereotyper fungerar som en standard, utifrån vilken individer blir bedömda.

Shifting Standards

När individer bedöms på specifika dimensioner utifrån stereotypa uppfattningar, görs detta utifrån en så kallad inomgruppsstandard (Biernat, 2003). Människor använder omedvetet kognitiva scheman som referenspunkter när de gör subjektiva bedömningar av medlemmar från stereotypiserade sociala grupper. Teorin bakom modellen om skifte av standards är att människor använder olika referenspunkter när de gör subjektiva bedömningar av andra, beroende på den sociala identiteten hos den person som är målet för bedömningen (Biernat, Manis & Nelson, 1991; Gushue, 2004). Biernat och Manis (1994) menar att individer använder dessa olika referenspunkter helt rutinmässigt för människor från olika sociala kategorier. Beskrivningar och bedömningar av andra utgår ofta från olika adjektiv och dessa bedömningar är många gånger subjektiva. Collins, Biernat och Eidelman (2009) menar att dessa typer av subjektiva beskrivningar ofta behöver någon slags referens eller standard för att de ska förstås. Vi utgår då från den genomsnittliga människan, vilket innebär att antingen har vi mer eller så har vi mindre av en viss egenskap än vad vi uppskattar att genomsnittet för denna grupp har. Collins et al. nämner som exempel uppfattningen att en person är smart. Med detta menar vi förmodligen att personen är smartare än genomsnittet. Men Collins et al. betonar att den referensram vi använder för att bedöma någon som smart, kan se olika ut beroende på den sociala tillhörigheten för en individ. Hur människor bedömer och tolkar ordet det subjektiva ordet smart, beror till stor del på individens sociala kategoritillhörighet och de stereotyper som finns riktade mot denna kategori. Collins et al. (2009) menar att detta är en av premisserna för modellen om skifte av standards; bedömningar av andra görs med avseende på gruppsspecifika, stereotypbaserade standards.

Studier kring skifte av standards har visat att medlemmar från en social kategori bedöms med hänsyn till andra medlemmar av den sociala kategorin, när bedömningarna görs subjektivt, exempelvis med en Likertskala (Biernat, Manis & Kobryniewicz, 1997). Biernat (2003) menar att utifrån den stereotypa uppfattningen att män är bättre ledare än kvinnor så kan inte bedömningen av ledarskapsförmåga för två individer ur dessa grupper ses som direkt jämförbar på en subjektiv skala. Bedömningen *bra* för kvinnor behöver inte betyda samma sak som *bra* för män, eftersom bedömningen görs med hänsyn till den inomgruppsliga standarden. Om mätningen görs med en subjektiv bedömningsskala tenderar människor att bedöma ledarskapsförmågan hos en kvinna utifrån den (lägre) standard som finns för kvinnor, och ledarskapsförmågan hos en man utifrån den (högre) standard som finns för män. Detta sker på grund av de stereotyper som finns mot dessa kategorier kring denna faktor enligt Biernat (2003). Biernat (2003) menar också att denna typ av standardskiften är möjliga på grund av språkets och bedömningars subjektivitet. Om samma bedömning istället görs med en objektiv mätning, kommer bedömningen att göras relativt till den andra kategorin, och mer utifrån det ”verkliga” livet (Garcia & Ybarra, 2007).

Två grupper kan skiljas åt utifrån en viss stereotypisk faktor. Detta innebär att grupperna skiljer sig åt dels vad gäller det ”typiska” värdet för respektive grupp på en viss faktor, dels på det intervall av värden som kan förväntas för den specifika gruppen (Biernat och Manis 1994). I en studie av Kobryniewicz och Biernat (1997) visades att bedömningar av arbetssökande skiljde sig åt beroende på hur bedömningen utfördes. Försökspersonerna fick läsa jobbansökningar från antingen en kvinnlig eller en manlig arbetssökande, till antingen typiskt manliga eller typiskt kvinnliga jobb. Utifrån subjektiva bedömningar kunde en slutsats dras att båda könen ansökningar ansågs som *ganska bra* till de båda typerna av arbeten, d.v.s.

inga könsstereotyper verkade finnas. Dock visade det sig att mer objektiva bedömningar, t.ex. uppskattningar av sökandes resultat på standardiserade test, gav ett omkastat mönster; dessa bedömningar avslöjade en preferens för det könet som passade till rätt jobb. Kvinnor bedömdes alltså passa bättre för det typiskt kvinnliga yrket och vice versa i enlighet med de stereotyper som finns. Detta innebär att trots att den första bedömningen på ytan ser ut att vara helt utan inblandning av stereotyper, så har den bedömningen uppstått på grund av ett skifte av standards och inomkategoriska jämförelser.

Biernat och Manis (1994) fann en stereotyp föreställning att män tjänar mer pengar än kvinnor när de använde sig av en objektiv skala och lät deltagare bedöma hur mycket pengar specifika personer tjänade per år. I studien fick deltagare göra bedömningar utifrån bilder av personer, där hälften av deltagarna gjorde subjektiva bedömningar och hälften gjorde objektiva bedömningar. Deltagarna hade tydliga genusstereotyper om finansiell status utifrån objektiva bedömningar om årsinkomst. Dock visade det sig att kvinnor bedömdes högre med en subjektiv skala 1-7 där deltagarna fick bedöma en person på en skala från *financially very unsuccessful* till *financially very successful*. Subjektiva bedömningar ledde till att deltagarna gjorde implicita jämförelser inom grupperna, vilket i sin tur ledde till högre bedömning av kvinnorna; en contrast effect. Biernat och Manis (1994) menar att detta beror på att människor helt rutinmässigt och omedvetet skiftar sina standards i enlighet med de förväntningar som finns för medlemmar av olika sociala grupper.

Biernat (2003) menar att ett grundantagande för modellen om skifte av standard, är att standards är lägre för den grupp som har blivit stereotypiserad som bristfällig eller otillräcklig på en faktor. I enlighet med stereotypa föreställningar om att en grupp är mer bristfälliga eller otillräckliga på en viss faktor blir standards lägre för gruppen på samma faktor. Således är det dessa olika standards som skapar de typer av bedömningseffekter som redovisats tidigare.

Syfte och Problemformulering

Syftet är att se om och hur bedömningar och uppfattningar av elever från olika typer av gymnasieprogram varierar och påverkas utifrån teorin om shifting standards. I inledningen har det enbart redovisats generell forskning kring sociala kategorier och shifting standards utifrån de vanligaste förekommande variablerna såsom kön eller yrke. Men baserat på teorierna om sociala kategorier och shifting standards, borde även gymnasieprogram kunna ses som sociala kategorier som det finns olika stereotypa föreställningar om. Trots allt verkar praktisk och teoretisk utbildning värderas olika, samtidigt som skillnaderna programtyperna emellan kommer att betonas ytterligare i och med den nya gymnasieskolan. Eftersom standards är lägre för den grupp som har blivit stereotypiserad som bristfällig eller otillräcklig på vissa specifika faktorer, bör en subjektiv mätning inte röja några större skillnader i bedömning mellan grupperna alternativt en contrast effect. En mer objektiv mätmetod bör visa en större skillnad i bedömning mellan grupperna då standards kan bidra till effekter vid bedömningen av andra (Biernat, 2003).

En design utformades där deltagare från olika populationer, både yrkesverksamma lärare, lärarstudenter, samt andra yrkesverksamma utan koppling till läraryrket, antingen fick bedöma en elev från ett studieförberedande program eller ett yrkesförberedande program. Dessutom fick deltagarna antingen göra en generell bedömning av en typisk elev från en viss programtyp eller så gjordes ingen generell bedömning. Med generell menas att bedömningen inte gjordes av någon specifik elev. Deltagarna fick endast information om programtillhörighet och skulle utifrån detta göra en bedömning av hur en typisk sådan elev kan tänkas vara. Den generella bedömningen var tänkt att inducera en standard för en viss programtyp, och de olika programtyperna skulle visa om standarden varierade och gav någon effekt. Detta utgår från att olika grupper har olika standards och att en generell bedömning av en typisk elev från en viss typ av program aktiverar en standard i större utsträckning eftersom eleven då delvis bedöms utifrån den inducerade standarden. Följande frågeställningar testas:

1. Har de olika programtyperna olika standards?
2. Aktiverar en generell bedömning en standard i större utsträckning än om ingen generell bedömning görs, för respektive programtyp?
3. Skiljer sig den eventuella effekt som aktiveras av den generella standarden åt mellan programtyperna?

Hypotesen är att en aktiverad effekt av en generell standard för gruppen yrkesförberedande program kommer att skilja sig från en aktiverad effekt av en generell standard för gruppen studieförberedande program.

Metod

Deltagare

Deltagarna bestod av lärarstuderande, yrkesverksamma lärare på gymnasiet samt andra yrkesverksamma personer utan någon uttalad koppling till läraryrket. På grund av tre outliers på en fråga fick tre deltagare uteslutas ur studien, vilket ledde till att resultatet grundar sig på 83 personers enkätsvar. Dessa deltagare randomiserades till fyra olika grupper med 18, 20, 22, och 23 personer i respektive grupp. Alla deltagare var antingen studerande eller yrkesverksamma. Åldersfördelningen bör alltså vara mellan 18-65 år. Ingen information angående deltagarna insamlades då detta inte är relevant för studiens syfte, dock är detta en svaghet för studien då en replikering kompliceras.

Material

För att kunna studera en skillnad i uppfattning mellan olika programtyper vad gäller standards, delades deltagarna in i fyra olika grupper och fyra olika enkäter utformades och användes (se Appendix). Enkäterna utformades för denna studie, utifrån frågeställningarna. Två av grupperna fick göra en generell bedömning av en typisk elev, medan de andra två grupperna inte gjorde någon sådan bedömning. Den generella bedömningen testade för deltagarnas uppfattning och stereotypa bild av den typiska eleven från en viss programtyp (yrkesförberedande/studieförberedande). Deltagarna gavs i enkäten information att de skulle föreställa sig en ”typisk” elev från ett studieförberedande/yrkesförberedande program och dennes beteende och förutsättningar inom kärnämneskurserna på gymnasiet. Denna formulering valdes eftersom alla elever läser dessa kurser oavsett program, samt för att ge deltagarna något att utgå ifrån i sina bedömningar. Deltagarna fick sedan bedöma den typiska eleven utifrån följande faktorer: 1) ansträngning, 2) förutsättning, och 3) prestation. Dessa tre faktorer valdes då de är faktorer hos eleven vilka lärares förväntningar ofta riktas mot (Madon, Jussim & Eccles, 1997).¹ För att mäta detta ställdes följande frågor: 1) Hur mycket anstränger sig eleven för att klara skolarbetet?, 2) Hur stora är elevens förutsättningar för att klara skolarbetet?, 3) Hur väl presterar eleven i skolan? På samtliga items fick deltagarna ange ett alternativ mellan 1 (*väldigt mycket/stora/väl*) och 7 (*väldigt lite/små/dåligt*). På ett fjärde item fick deltagarna dessutom själva ange det typiska betygsmedelvärdet för en elev för den specifika programtypen med en siffra mellan 0-20, vilket är de värden som elevers slutbetyg kan variera mellan.

Bedömningen angående en specifik elevs prestation var likadant utformad för alla grupperna. Den enda skillnaden var programtypen. De deltagare som inte gjorde någon generell bedömning, gjorde endast denna del. De som hade gjort en generell bedömning gjorde denna del också. Den specifika delen gick ut på att deltagarna fick läsa en del ur en elevtext, som var hämtad och smått omarbetad samt förkortad, från elevforumet Mimers

¹ Madon, Jussim och Eccles (1997) anger ordet *talent* istället för förutsättning. Även om förutsättning inte är en direkt synonym till talang, föll ändå valet på ordet förutsättning då det ansågs mer lämpligt och passande med avseende på sammanhanget.

Brunn på Internet. Alla grupper hade exakt samma text. För de två grupperna som tidigare hade gjort en generell bedömning gavs informationen att texten var skriven av en elev från samma typ av program som tidigare. För de grupper som inte hade gjort någon generell bedömning gavs enbart information om att texten var skriven av en elev från en viss typ av program (yrkesförberedande/studieförberedande). Deltagarna fick sedan bedöma sitt helhetsintryck av texten på en subjektiv skala mellan 1 (*väldigt bra*) och 7 (*väldigt dålig*). Slutligen fick deltagarna göra ytterligare en bedömning av texten genom att ange ett värde i kronor som de tyckte att texten skulle vara värd, om den hade valts ut som ett av flera bidrag till en novellsamling av elever på gymnasiet. Tanken med denna fråga var att ge deltagarna en chans till en mer konkret bedömning som förhoppningsvis skulle leda till större effekter. Deltagarna fick utgå från en elevtext eftersom det skulle möjliggöra bedömningen av en specifik person från en viss kategori. Allt detta ledde till att enkäten hade totalt antingen sex eller två mätbara items.

Procedur

Deltagarna rekryterades genom kontakt på e-post eller direkt tillfrågan. Vissa av deltagarna gjorde enkäten i samband med att de tillfrågades, andra genomförde enkäten vid ett senare tillfälle och lämnade in den när de var klara. Alla deltagare gavs en kort muntlig instruktion samt att det var helt frivilligt att delta. Ingen ersättning gavs. Det fanns inga frågor på enkäten som skulle kunna avslöja deltagarna identitet.

Antingen fick försöksdeltagarna göra en generell bedömning av en typisk elev från en viss programtyp, eller så gjordes ingen generell bedömning. Dessutom bedömde alla deltagarna antingen en specifik elev från ett yrkesförberedande program, eller en specifik elev från ett studieförberedande program. Grupperna uppdelades alltså på följande sätt: 1) Generell bedömning yrkesförberedande program, specifik bedömning yrkesförberedande program, 2) Ingen generell bedömning, specifik bedömning yrkesförberedande program, 3) Generell bedömning studieförberedande program, specifik bedömning studieförberedande program, 4) Ingen generell bedömning, specifik bedömning studieförberedande program. Tanken med de generella bedömningarna av en typisk elev från en viss typ av program, var att dessa i större grad skulle inducera en standard och en stereotyp bild av den specifika eleven från respektive program.

Då studien består av fyra olika testgrupper, randomiserades deltagarna till de olika grupperna genom en turordning där en enkät från varje grupp besvarades innan nästa omgång började (A1, B1, C1, D1, A2, B2... o.s.v.) Vid tillfällena där en grupp av människor rekryteras på samma gång, användes endast två olika enkäter för samma typ av program (A1, B1, A2, B2, A3, B3... o.s.v.). Detta gjordes för att minska risken för att deltagarna skulle upptäcka att samma text användes för en bedömning av olika elever. Deltagare från olika grupper bör således ha spridits jämt över grupperna. Då det inte finns några sätt att identifiera enkäterna kan dock detta inte garanteras.

Resultat

Generell Bedömning

Den första frågeställningen var om de olika programtyperna har olika standards. För att testa detta samt inducera en tänkt standard för olika programtyper fick några av deltagarna göra en generell bedömning av en typisk elev från en viss programtyp. Det fanns signifikanta skillnader mellan studieförberedande program ($n=18$) och yrkesförberedande program ($n=20$) på samtliga faktorer; bedömningen för ansträngning, $t(36) = -2.96$, $p < .005$, $\eta^2 = .20$, förutsättning $t(36) = -3.15$, $p < .003$, $\eta^2 = .22$, och prestation $t(36) = -3.50$, $p < .001$, $\eta^2 = .25$. Medelvärden för respektive faktor kan ses i Tabell 1. Den typiska eleven från ett

yrkesförberedande program bedömdes anstränga sig mindre, ha sämre förutsättningar samt prestera sämre än en elev från ett studieförberedande program. Dessutom ansågs den typiska eleven från ett yrkesförberedande program ha lägre betygsmedelvärde än den typiska eleven från ett studieförberedande program, $t(29.97) = 4.49$, $p < .001$, $\eta^2 = .35$.

Tabell 1

Medelvärden och standardavvikelser för den generella bedömningen av en typisk elev

	Studieförberedande (n=18)		Yrkesförberedande (n=20)	
	M	SD	M	SD
Ansträngning	2.78*	0.88	3.75*	1.12
Förutsättning	1.89*	0.76	2.75*	0.91
Prestation	2.72*	0.89	3.75*	0.91
Betygsmedelvärde	15.64	1.35	12.78	2.47

*Not. 1=väldigt mycket/stora/väl – 7=väldigt lite/små/dåligt.

Specifik Bedömning

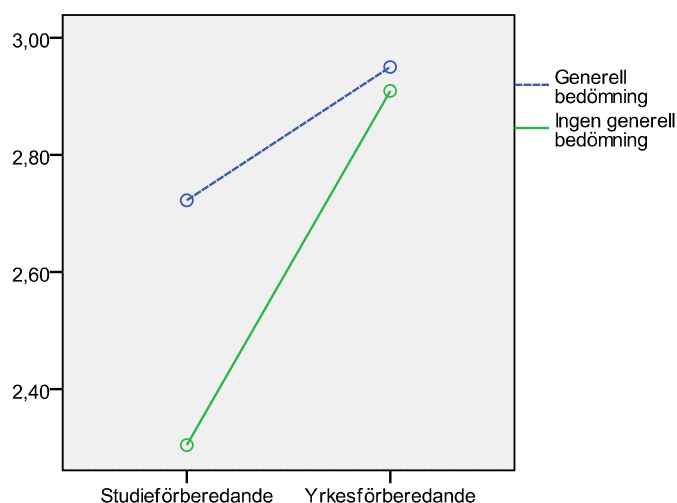
Den andra frågeställningen var om en generell bedömning aktiverar en standard i större utsträckning än om ingen generell bedömning görs, för respektive programtyp. För att testa detta gjordes t-test mellan programtyperna för att se om det fanns signifikanta skillnader i medelvärden för bedömningen av den specifika eleven. För studieförberedande program, bedömde gruppen som först hade gjort en generell bedömning den specifika eleven som ickesignifikant sämre (n=18, $M=2.72$, $SD = 1.36$) än den grupp som enbart bedömde eleven (n=23, $M=2.31$, $SD=1.11$). Den effekt som var tänkt att aktiveras av den generella standarden, var inte signifikant på skalan som testade för uppfattning om eleven utifrån helhetsintryck av texten för studieförberedande program, $t(39) = 1.08$, $p > .29$. Det fanns heller ingen effekt för yrkesförberedande program på samma skala, $t(40) = .11$, $p > .92$. Gruppen som först hade gjort en generell bedömning bedömde den specifika eleven som ickesignifikant sämre (n=20, $M=2.95$, $SD=1.28$) än den grupp som enbart bedömde eleven (n=22, $M=2.91$, $SD=1.23$).

För att kontrollera om en generell bedömning aktiverar en standard i större utsträckning än om ingen generell bedömning görs för respektive programtyp gjordes t-test även för den mindre subjektiva skalan. Här mättes deltagarnas specifika bedömning av en elev utifrån textens värde i kronor. På denna item fanns tre outliers. Samtliga outliers fanns i olika grupper. En analys utfördes både med och utan dessa. Trots att borttagandet av outliers inte gav något signifikant resultat togs ändå beslutet att utesluta dem eftersom de hade stor inverkan på medelvärdenas inbördes ordning.² Högst värderad i kronor blev texten av den grupp som först hade gjort en generell bedömning av en typisk elev från ett yrkesförberedande program ($M = 485.13$, $SD = 431.86$), följt av ingen generell bedömning studieförberedande program ($M = 392.61$, $SD = 270.79$), ingen generell bedömning yrkesförberedande program ($M = 375.23$, $SD = 290.32$) och slutligen generell bedömning studieförberedande program ($M = 306.67$, $SD = 159.52$). Den generella bedömningen för

² De tre deltagarnas resultat uteslöts ur alla analyser p.g.a. outliers på denna fråga. Det var dock bara i denna analys som den inbördes ordningen ändrades vid borttagandet av outliers. På samtliga andra bibehölls samma inbördes ordning och i princip samma signifikansnivåer. Följande medelvärden och standardavvikelser framkom utan borttagandet av outliers för detta item: Ingen generell bedömning studieförberedande program ($M=584.58$, $SD=977.06$), ingen generell bedömning yrkesförberedande program ($M=576.30$, $SD=1005.18$), generell bedömning studieförberedande program ($M=553.68$, $SD=1087.83$) och slutligen generell bedömning yrkesförberedande program ($M=485.13$, $SD=431.86$). Inga signifikanta skillnader fanns mellan de två olika studieförberedande grupperna $t(41) = -.99$, $p > .92$, eller de två yrkesförberedande grupperna $t(41) = -.38$, $p > .71$.

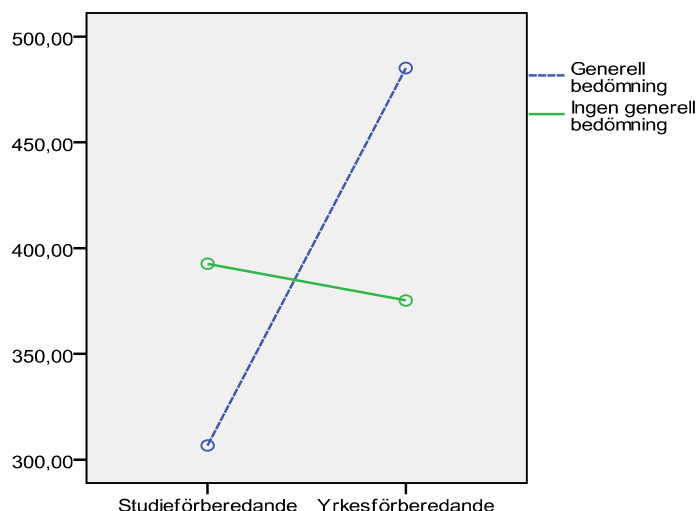
respektive program gav dock ingen signifikant effekt, varken för de två studieförberedande grupperna $t(39) = -1.19, p > .24$ eller för de två yrkesförberedande grupperna $t(40) = .98, p > .34$.

Den sista frågeställningen var om den eventuella effekt som aktiveras av den generella standarden skiljer sig den mellan programtyperna. Detta testades både för den subjektiva och för den mindre subjektiva skalan. För att se om det fanns några skillnader i effekt mellan programtyper och typ av bedömning på Likertskalet gjordes en 2 X 2 (programtyp X typ av bedömning) ANOVA. Det fanns inga signifikanta huvudeffekter för vare sig programtyp $F(1, 79) = 2.32, p > .13$, eller om en generell bedömning hade gjorts eller ej $F(1, 79) = .71, p > .40$. Ingen interaktionseffekt fanns heller mellan dessa på den mer subjektiva bedömningen, $F(1, 79) = .48, p > .49$. Den tänkta induceringen av en standard utifrån en generell bedömning gav således inte någon effekt i jämförelse med den grupp som inte gjorde någon generell bedömning på den mer subjektiva skalan, även om de båda yrkesförberedande programmen bedömdes som sämre än de två studieförberedande (se Figur 1).



Figur 1. Uppskattade medelvärden för bedömning av den specifika eleven utifrån textens värde på Likert-skalan, 1= väldigt bra – 7=väldigt dålig).

En 2 X 2 (programtyp X typ av bedömning) ANOVA gjordes också för den specifika bedömningen på den mindre subjektiva skalan där bedömningen gjordes i kronor. Inte här fanns några signifikanta huvudeffekter, varken för programtyp $F(1, 79) = 1.43, p > .24$ eller för om en generell bedömning hade gjorts eller ej $F(1, 79) = .03, p > .86$. Det fanns heller ingen signifikant interaktionseffekt $F(1, 79) = 2.12, p > .15$, även om det syntes en skillnad mellan programtyperna (Se Figur 2). Den eventuella effekt som skulle aktiveras av den generella standarden skiljde sig således inte åt mellan programtyperna.



Figur 2. Uppskattade medelvärden för bedömning av den specifika eleven utifrån textens värde i kronor.

Diskussion

Denna studie utgår från teorierna om sociala kategorier och shifting standards. Utifrån antagandet att gymnasieprogram kan ses som sociala kategorier var hypotesen att effekten som aktiveras av en generell standard för gruppen yrkesförberedande program skulle skilja sig från den effekt som aktiveras av en generell standard för gruppen studieförberedande program. Dock framkom ingen sådan effekt för någon av programtyperna. Inte heller fanns någon effekt som skiljde sig åt mellan programtyperna, även om det fanns en antydning om det. Hypotesen måste därför förkastas. Det fanns dock en skillnad mellan uppfattningen om programtyperna, där elever på yrkesförberedande bedömdes lägre än elever på studieförberedande program på samtliga faktorer.

Den första frågeställningen var om de olika programtyperna har olika standards. I denna studie framkom att olika typer av program bedömdes och uppfattades på olika sätt av deltagarna, med avseende på de faktorer som tidigare forskning har visat att lärare har olika förväntningar om (Madon, Jussim & Eccles, 1997). Elever från yrkesförberedande program bedömdes anstränga sig mindre, ha sämre förutsättningar, samt prestera sämre än eleverna från studieförberedande program. Dessutom bedömdes elever från yrkesförberedande program ha lägre betygsmedelvärde. Enligt Corneille och Judd (1999) kan gruppstereotyper vidmakthållas eftersom kategorisk information bidrar till att skapa en överdriven uppfattning om skillnader mellan kategorier men också likheter inom kategorier. I denna studie framkom såväl tydliga skillnader mellan de olika grupperna som en enighet om likheter inom grupperna, eftersom det framkom signifikanta skillnader mellan studie- och yrkesförberedande program. Sherman et al., (2009) menar att för att mellangrupsliga skillnader samt inomgruppsliga likheter ska kunna vidmakthållas bör det föreligga verkliga skillnader mellan grupperna. Frågan är för denna studie om det verkligen föreligger verkliga skillnader, eller om det handlar om skillnader i stereotypa uppfattningar? När det inte finns tillräckligt med information om individen, tenderar människor att utgå från den kategoriska informationen (Kunda & Thagard, 1996). I studien fanns det ingen individualiserad information att utgå från eftersom frågorna handlade om en typisk elev. Bedömningarna måste ha gjorts helt utifrån en föreställning om hur elever från olika program är, från någon form av minnesrepresentation för respektive kategori (Sherman et al., 2009). Oavsett vilket

kan nämnas att vid Skolverkets senaste sammanställning av gymnasieelevers slutbetyg för läsåret 2008/2009, fanns det en skillnad mellan programtyperna i betygsmedelvärde. Av de yrkesförberedande programmen var det högsta betygsmedelvärdet 14.4 poäng, medan det högsta var 16.2 för de studieförberedande. Det lägsta var 11.8 för yrkesförberedande medan det var 14.1 för studieförberedande program ("Betyg och studieresultat", 2010). Vad gäller betygsmedelvärde finns det således en verklig skillnad mellan programmen.

Som nämndes tidigare är ett grundantagande för modellen om skifte av standard, att standards är lägre för den grupp som är stereotypiserad som bristfällig eller otillräcklig på en faktor. Det är i sin tur dessa standards som skapar olika typer av bedömningseffekter (Biernat, 2003). I studien blev de yrkesförberedande programmen värderade annorlunda gentemot de studieförberedande, vilket enligt hypotesen borde ha lett till skillnader i bedömning av den specifika eleven. Dock framkom inga signifikanta skillnader mellan grupperna på någon av bedömningsskalorna. Den andra frågeställningen var om en generell bedömning aktiverar en standard i större utsträckning än om ingen generell bedömning görs, för respektive programtyp. Den generella bedömningen aktiverade inte någon standard i större utsträckning än om ingen generell bedömning gjordes för någon programtyp. Även om grupperna bedömdes olika på den generella bedömningen, ledde denna olikhet inte till något resultat som kan stödja att standards för gymnasieprogram är aktiva vid bedömning av en specifik elev. Förvisso blev det inte någon skillnad i bedömning mellan grupperna på Likert-skalan, men det blev heller ingen skillnad på den mindre subjektiva bedömningen. Om en stereotyp bild av den typiska eleven från ett visst program hade spelat in, borde detta ha avslöjats när bedömningen var mindre subjektiv, vilket var fallet i de studier som redovisades tidigare. Detta eftersom en subjektiv mätmetod ofta leder till att deltagare gör implicita jämförelser inom grupperna, medan en mer objektiv mätmetod kan röja skillnader som kommer sig av jämförelser mellan grupperna (Biernat & Manis, 1994). Tanken med de generella bedömningarna var att de skulle inducera en standard för ett program och att bedömningen skulle göras utifrån bilden av eleven från ett visst program i färskt minne, inte att bedömningen skulle präglas av standarden i sig. En generell standard för ett studieförberedande program borde enligt hypotesen ha inducerat en annan bild av en elev än standarden för ett yrkesförberedande program. I enlighet med den inomgruppsliga jämförelsen borde den specifika eleven ha bedömts lägre utifrån den generella standarden för studieförberedande program på den subjektiva skalan, men högre utifrån den generella standarden för yrkesförberedande program. Denna slutsats kunde dock inte dras utifrån resultatet. Intressant är dock att den specifika eleven från ett yrkesförberedande program bedömdes lägst på den subjektiva skalan där en generell bedömning hade gjorts, men att texten för samma grupp sedan bedömdes vara värd mest i pengar.

Den tredje frågeställningen var om den eventuella effekt som aktiveras av den generella standarden skiljer sig åt mellan programtyperna. Här fanns inget signifikant resultat. Om en effekt hade aktiverats av den generella standarden så skulle yrkesförberedande program ha bedömts som bättre, eller lika bra som studieförberedande program på den mer subjektiva skalan. Som Biernat (2003) poängterar är grundtanken med shifting standards att en grupp bedöms i jämförelse med andra inom samma grupp vid en mer subjektiv bedömning. Därför borde alltså yrkesförberedande program ha bedömts som bättre eller lika bra som studieförberedande program på den mer subjektiva skalan, om en stereotypisk effekt hade uppkommit. Tvärt emot borde den mindre subjektiva skalan ha röjt stereotyperna om de olika programtyperna, och studieförberedande program borde således ha bedömts som bättre än yrkesförberedande program. Här måste dock en viktig och bristfällig faktor poängteras: Eftersom deltagarna fortfarande bedömde en specifik elev utifrån samma text som innan, kan detta ha minimerat möjligheterna till en bedömning i den riktning som var avsikten med studien. Möjligheten finns att deltagarna fortfarande använde en inomgruppslig standard; ”med

tanke på elevens programtillhörighet tycker jag att texten är värd såhär mycket". För att ha fått en tydlig effekt borde det ha funnits något annat sätt att mäta uppfattningen om eleven på den mindre subjektiva skalan. Det bästa för studien hade troligen varit att göra den mätningen utifrån en annan variabel. Tidigare studier har bland annat använt sig av uppskattning av poäng på standardiserade och allmänt kända test (t.ex. Biernat, 2003), men detta är svårt att använda sig av då det inte finns någon direkt motsvarighet i Sverige. Det är tänkbart, och högst troligt, att ett förflyttande från fokus på texten vid den specifika bedömningen, till en helt annan variabel hade gett ett annat och mer rättvist resultat. Till exempel kunde pengar ha använts som mätvärde, men för en annan typ av bedömning. Den bedömning som gjordes av den specifika eleven utifrån värde i kronor visade ingen signifikant interaktionseffekt, men hade ändå ett p-värde strax över 0.1 vilket ger en indikation på att någon typ av interaktion kan finnas. Trots att ingen slutsats kan dras här, så finns det en antydning om skillnader i effekter.

Samtidigt måste frågan ställas om vad som egentligen mäts i de olika grupperna. Det finns fler faktorer som var bristfälliga i studien. Möjligheten finns att den generella bedömningen inte alls aktiverade någon standard, eller att bedömningen av en text inte var rätt metod för att avslöja eventuella stereotypier. Det är troligt att för mycket fokus lades på texten som sådan istället för att fungera som något att bedöma den specifika eleven på. Därtill finns det anledning att fundera över om deltagarna på något sätt blev styrda av någon typ av social önskvärdhetsbias. Den generella bedömningen avslöjade inget sådant, men möjligheten finns att deltagarna bedömde texten som annorlunda än vad de egentligen tyckte, i och med den tidigare standarden. Dessutom kan den mer konkreta bedömningen ha uppfattats som svår då det inte fanns någon referensram för ett pengavärde att utgå ifrån. Vad som är rimligt för en individ kan ju vara något helt annat för en annan.

Även om de olika programtyperna kan ses som olika sociala kategorier med olika stereotyper riktade mot sig, finns det i denna studie inget stöd för några standardskillnader eller effekter, vare sig inom eller mellan programtyperna. Med tanke på de olikheter som ändå fanns mellan programtyperna på den generella bedömningen, borde det finnas möjligheter att finna någon typ av effekt med en annan mätmetod, även om ingen sådan fanns i denna studie. Eftersom elever ständigt är föremål för bedömning är det av största intresse att se om och hur denna bedömning kan påverkas av inom- och mellangrupsstandarder – speciellt med tanke på de uttalade skillnader som kommer att bli programtyperna emellan.

Referenser

- Abele, A. E., & Petzold, P. (1998). Pragmatic use of categorical information in impression formation. *Journal of Personality and Social Psychology*, *75*, 347-358.
- Andersson, B-E. (1999). *Spräng skolan!* Jönköping: Brain Books.
- Betyg och studieresultat i gymnasieskolan läsåret 2008/09. (2010). Retrieved May 8, 2010, from <http://www.skolverket.se/sb/d/1721#paragraphAnchor0> (Tabell 2A)
- Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist*, *58*, 1019-1027.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*, 5-20.
- Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, *60*, 485-499.
- Biernat, M., Manis, M., & Kobrynowicz, D. (1997). Simultaneous assimilation and contrast effects in judgments of self and others. *Journal of Personality and Social Psychology*, *73*, 254-269.
- Collins, E. C., Biernat, M., & Eidelman S. (2009). Stereotypes in the communication and translation of person impression. *Journal of Experimental Social Psychology*, *45*, 368-374.
- Corneille, O., & Judd, C. M. (1999). Accentuation and sensitization effects in the categorization of multifaceted stimuli. *Journal of Personality and Social Psychology*, *77*, 927-941.
- Garcia, S. M., & Ybarra, O. (2007). People accounting: Social category-based choice. *Journal of Experimental Social Psychology*, *43*, 802-809.
- Gushue, G. V. (2004). Race, color-blind racial attitudes, and judgment about mental health: A shifting standards perspective. *Journal of Counseling Psychology*, *51*, 398-407.
- Kobrynowicz, D., & Biernat, M. (1997). Decoding subjective evaluations: How stereotypes provide shifting standards. *Journal of Experimental Social Psychology*, *33*, 579-601.
- Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, *103*, 284-308.
- Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology*, *4*, 791-809.
- Neuberg, S. L., & Fiske, S., T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, *53*, 431-444

- Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. V., & Conrey, F. R. (2009). Attentional processes in stereotype formation: A common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology, 96*, 305-323.
- Skolverket (2009a). *En ny gymnasieskola. Gy2011*. Trycksak om gymnasiereformen. Stockholm: Fritzes.
- Skolverket (2009b). Skolverkets lägesbedömning 2009. Rapport 337. Stockholm: Fritzes.
- Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131-142.
- Svensson, A. (2001). Består den sociala snedrekryteringen? Elevernas val av gymnasieprogram hösten 2008. *Pedagogisk Forskning i Sverige, 3*, 161-172.

Appendix

Denna studie handlar om elevuppfattningar och bedömningar. Medverkan är helt frivillig och du kan när som helst välja att avbryta. Tänk på att det inte finns något rätt och fel. Det är din egen uppfattning som är det viktiga.

Föreställ dig en ”typisk” elev från ett STUDIEFÖRBEREDANDE program och dennes beteende och förutsättningar inom kärnämneskurserna på gymnasiet. Var vänlig och ringa in det alternativ som du bäst tycker överensstämmer med din uppfattning.

1. Hur mycket *anstränger* sig eleven för att klara skolarbetet?

1	2	3	4	5	6	7
Väldigt mycket						Väldigt lite

2. Hur stora är elevens *förutsättningar* för att klara skolarbetet?

1	2	3	4	5	6	7
Väldigt stora						Väldigt små

3. Hur väl *presterar* eleven i skolan?

1	2	3	4	5	6	7
Väldigt väl						Väldigt dåligt

4. Gymnasieelever kan ha ett betygsmedelvärde mellan 0-20 poäng. Med avseende på den ”typiska” eleven från ett studieförberedande program, hur skulle du bedöma elevens totala betygsmedelvärde?

Var vänlig och ange ett heltal mellan 0-20: _____

VÄND

Nedan är ett utdrag ur en text som är skriven av en elev från samma typ av program som nämndes ovan, alltså ett studieförberedande program. Var vänlig och läs igenom texten. När du har gjort det kommer du att få bedöma texten.

Novell

- Va, fan gör du din jävla sopa?! Kan du inte ens träffa bollen? Glåpord och negativa kommentarer följer efter det missade skottet. Simon, pojken som är målet för glåporden försöker att inte påverkas, det är ju bara ord, försöker han intala sig själv. Men det är omöjligt att inte bli berörd när man gång på gång får veta vilken värdelös människa man är. Ändå väljer pojken att spela vidare i laget, varför? Ibland är det mer som spelar in än ens egen vilja. Det regnar ute. Arvid, En av pojkarna från laget har duschat färdigt. utan en tanke på att han bara någon halvtimma tidigare skurit ännu ett djupt sår i en ensam människas inre, börjar gå hemåt, obekymrad av regnet. Kanske har han gjort det för många gånger för att lägga en tanke på den hemska sak han nyss gjort. Han har inte heller någon tanke på att sluta tortera den stackars människan, alla är ju med honom. Och dessutom borde pojken ha vant sig nu, det har ju faktiskt fortgått ett bra tag nu.

Simon är redan hemma genomblöt av regnet. Det var länge sedan han slutade duscha med de andra pojkarna. Han vill bara ifrån träningslokalen och alla människor där inne så snabbt som möjligt, han hatar dem alla. Men frågan är om han inte hatar sig själv minst lika mycket. Han hoppas att alla hans bekymmer ska ha sköljts bort av regnet när han torkar, men vet innerst inne att det egentligen bara finns ett sätt att få slut på lidandet. Han suckar och går med tunga steg upp för trappen i lägenheten. När han kommer innanför dörren är hans far mycket snabb att fråga hur det gick på träningen. Visst gjorde du några mål? Visst satt alla passningar som de skulle? Du lät väl ingen ta bollen av dig? Simon svarar likadant som alltid.

Nedan ombeds du bedöma texten. Var vänlig och ringa in det alternativ som du bäst tycker överensstämmer med texten ovan.

5. Vilket är ditt helhetsintryck av texten?

1	2	3	4	5	6	7
Väldigt bra						Väldigt dålig

6. Tänk dig att ovanstående text har valts ut som ett av flera bidrag till en novellsamling av elever på gymnasiet. Eleverna ges en ersättning för sitt bidrag utifrån textens kvalitet. Hur stor summa pengar tycker du att texten är värd?

Var vänlig och ange en summa: _____

TACK FÖR DIN MEDVERKAN!