



**GÖTEBORGS UNIVERSITET**

## Det fria rummets baksida

- En experimentell studie av hur elevens programtillhörighet påverkar lärarens bedömning.

Robert Andersson  
Jennie Gunnarsson

LAU690

Handledare: Anders Fredriksson

Examinator:

Rapportnummer:

Vi vill rikta ett stort

**TACK**

till alla som gjort detta examensarbete möjligt.

Tack Anders och Klas som lyckades med den svåra handledarkonsten  
att med omväxlande uppmuntran och skarp kritik puffa oss i rätt riktning.  
Vi är också mycket tacksamma till alla lärare som ställt upp med tid och intresse!  
Utan er hade det inte blivit någon undersökning, och vi hoppas att ni också fått ut  
något positivt av ert deltagande. Att uppmärksamma ett problem är  
alltid en bra början till lösning.

## Abstract

### **Examensarbete inom lärarutbildningen**

**Titel:** Det fria rummets baksida – En experimentell studie av hur elevens programtillhörighet påverkar lärarens bedömning.

**Författare:** Robert Andersson & Jennie Gunnarsson

**Termin och år:** HT 2009

**Kursansvarig institution:** Sociologiska institutionen

**Handledare:** Anders Fredriksson

**Examinator:**

**Rapportnummer:**

**Nyckelord:** likvärdighet, betyg och bedömning, handlingsutrymme, experiment, programtillhörighet, gymnasieskolan.

Med ett intresse som grundar sig i praktiska erfarenheter av att förutsättningarna för undervisning och betygsättning är mycket olika på yrkesinriktade respektive studieinriktade program i den svenska gymnasieskolan undersöker vi i detta examensarbete om lärarens vetskap om elevens identitet påverkar bedömningen av elevens kunskaper.

Flera undersökningar visar att bedömningen i skolan inte är likvärdig, trots att likvärdigheten som hörnsten för en rättvis skola betonas av styrdokument och trots att dess vikt blir särskilt tydlig i bedömningssituationen där rättvisa betyg är ett krav. Problemen med likvärdighet kan spåras till lärarens betydande handlingsutrymme, vilket är en förutsättning för yrkesutövande i den komplexa situation läraren dagligen befinner sig i; i det fria rummet har läraren i samspråk med eleverna möjlighet att skraddarsy undervisningen efter eleverna.

För att undersöka huruvida lärarens vetskap om elevens programtillhörighet påverkar läraren i bedömningssituationen konstruerade vi ett experiment där 65 gymnasielärare i naturkunskap från hela Sverige deltog. Genom att samtliga lärare bedömde exakt samma elevsvar men med olika information om elevens programtillhörighet såg vi en tydlig skillnad där en elev som gick ett yrkesinriktat program bedömdes med högre betyg än en elev på ett studieinriktat program, för exakt samma prestation. Vidare kunde en stor spridning i bedömning konstateras för samma elevsvar, där alla betygssteg från IG till MVG fanns representerade.

Vår undersökning bekräftar tidigare forskning i slutsatsen att det finns ett problem för likvärdigheten i den svenska skolan. Vi har genom experiment kunna lyfta ut en faktor och tydligt visa att elevens programtillhörighet har betydelse för elevens bedömning, och efterlyser fler experimentella studier inom området för att vidare utforska kausala samband vid betygsättning. Brister i likvärdighet vid betygsättning kan och bör förbättras genom diskussioner mellan lärare på både skolnivå och nationell nivå.

## Innehållsförteckning

|   |    |
|---|----|
| Tabell- och figurförteckning .....        | 5  |
| 1. Inledning.....                         | 6  |
| 2. Tidigare forskning .....               | 8  |
| 2.1 Översikt .....                        | 8  |
| 2.2 Olika mått .....                      | 9  |
| 2.2 Skolans betydelse .....               | 9  |
| 2.3 Elevens betydelse .....               | 10 |
| 2.4 Lärarens betydelse.....               | 11 |
| 2.5 Programtillhörighetens betydelse..... | 12 |
| 2.6 Tidigare experimentella studier.....  | 13 |
| 2.7 Sammanfattning .....                  | 14 |
| 3. Teoretiskt ramverk .....               | 14 |
| 3.1 Likvärdighet .....                    | 15 |
| 3.2 Handlingsutrymme .....                | 17 |
| 3.3 Sammanfattning .....                  | 19 |
| 4. Design.....                            | 19 |
| 4.1 Experimentet .....                    | 20 |
| 4.2 Uppgiftens utformning .....           | 23 |
| 4.3 Analysförberedelser .....             | 24 |
| 5. Resultat.....                          | 26 |
| 6. Avslutande diskussion.....             | 30 |
| 7. Referenser.....                        | 33 |
| 7.1 Elektroniska källor .....             | 35 |
| 8. Bilaga: uppgift till lärarna .....     | 36 |

## Tabell- och figurförteckning

|  |      |
|--|------|
| Tabell 1: Beskrivning av experimentgruppens sammansättning                     | s.23 |
| Tabell 2: Randomiseringskontroll   | s.25 |
| Tabell 3: Skillnad i bedömning mellan HR- och NV-gruppen?                      | s.26 |
| Tabell 4: Skillnad i betyg mellan HR- och NV-gruppen inom gruppen ”teoretisk”? | s.27 |
| Figur 1: Skillnad i bedömning mellan HR- och NV-gruppen?                       | s.26 |
| Figur 2: Skillnad i betyg mellan HR- och NV-gruppen inom gruppen ”teoretisk”?  | s.27 |
| Figur 3: Antal satta betyg per betygssteg                                      | s.29 |
| Figur 4: Effekt av kön på bedömning  | s.29 |

# 1. Inledning

Betygen har alltid varit en kontroversiell fråga i svensk utbildningspolitik.<sup>1</sup> Aspekter som debatterats har varit om betygen gynnar förutsättningarna för det pedagogiska uppdraget eller inte<sup>2</sup>, vid vilken ålder man bör introducera betyg, i vilken mån betyg egentligen kan sägas motsvara en viss kunskap hos eleven<sup>3</sup>, hur fingerad skalans skall vara samt huruvida betygen kan anses vara rättvisa eller inte.<sup>4</sup> Inställningen till om man skall se på betygen som ett nödvändigt ont eller som en naturlig och viktig del av den pedagogiska verksamheten i skolan varierar dessutom i detta sammanhang mellan lärare och kan i förlängningen tänkas påverka bedömningspraktiken. I en skola där elever med olika bakgrund och förutsättningar skall ges möjlighet att tillägna sig de ämnen som är gemensamma för alla gymnasieprogram utmanas läraren i sin yrkesutövning när det gäller läroplanens riktlinjer för likvärdighet. Läroplanen betonar att likvärdig utbildning inte innebär ”att undervisningen skall utformas på samma sätt överallt eller att skolans resurser ska delas lika. Hänsyn ska tas till elevernas olika förutsättningar, behov och kunskapsnivå.”<sup>5</sup> Samtidigt som denna hänsyn till olika förutsättningar kräver en hög grad av flexibilitet på skol- och lärarnivå kvarstår kravet på att betygen skall vara jämförbara.

För att möjliggöra anpassningen av undervisningen och hanterandet av den mängd små beslut rörande eleverna som det dagliga arbetet innebär måste läraren i sitt yrkesutövande ges en viss grad av handlingsutrymme.<sup>6</sup> Relaterat till betygsättning som kan betraktas som den mest tydliga myndighetsutövande delen av läraryrket blir handlingsutrymmet ett problem när samma prestationer och kunskaper hos eleven värderas olika av olika lärare. Då betygen är viktiga för elevens möjligheter att läsa vidare kan betyg som inte uppfattas som rättvisa leda till en betydande legitimitetsförlust för hela systemet.<sup>7</sup> Betygens dubbla funktion av att vara en utvärdering av elevens kunskaper och samtidigt ett urvalsinstrument kan här hamna i konflikt med varandra.

På senare tid har frågan om betygsinflation och skillnader i betygsättningspraxis mellan lärare och skolor hamnat i centrum av den offentliga debatten. Dagens Nyheter skriver den 30:e november 2009 att gymnasieelevers betyg inte kan anses vara rättvisa och att betygsättningen för den enskilde eleven har blivit något av ett lotteri.<sup>8</sup> Artikeln baseras på en rapport från Skolverket som genom att jämföra resultat på det nationella provet med slutbetyget visar på stora brister när det gäller likvärdig betygsättning.<sup>9</sup> Det är att betrakta som särskilt problematiskt om dessa skillnader beror på systematiska skillnader som går att relatera till vilket program eleven går eller vilken social bakgrund eller kön eleven har.<sup>10</sup>

---

<sup>1</sup> Wedman, Ingemar (1983) *Den eviga betygsfrågan: Historiskt och aktuellt om betygsättning i skolan*.

Stockholm: Liber Utbildningsförlaget. Andersson, Håkan (1999) *Varför betyg? Historiskt och aktuellt om betyg*. Lund: Studentlitteratur e-bok

<sup>2</sup> Lindberg, Viveka & Lars, Lindström (Red.) (2008) *Pedagogisk bedömning – Om att dokumentera, bedöma och utveckla kunskap*.

<sup>3</sup> Linde, Göran (2003) *Kunskap och betyg*. Lund: Studentlitteratur

<sup>4</sup> Korp, Helena (2006) *Lika chanser i gymnasiet? En studie om betyg, nationella prov och social reproduktion*.

<sup>5</sup> Läroplan för de frivilliga skolformerna Lpf 94 1:2

<sup>6</sup> Lipsky, Michael (1980) *Street-Level Bureaucracy – Dilemmas of the individual in public services*. New York: Russell Sage Foundation

<sup>7</sup> Korp (2006) s. 9

<sup>8</sup> DN 091130 ”Gymnasieelevers betyg inte rättvisa”

<sup>9</sup> Skolverket (2009) *Likvärdig betygsättning i gymnasieskolan – En analys av sambandet mellan nationella prov och kursbetyg*. Rapport 338

<sup>10</sup> Ibid, s. 31

Om betygsättningen inte är likvärdig finns det ett stort problem i den svenska skolan. De olika rapporter som har producerats och den tidigare forskning som bedrivits pekar ut olika anledningar till att brister kan uppstå. Olika perspektiv och tolkningar av begreppet likvärdighet framträder i forskning och styrdokument. Dessutom är situationen den betygsättande läraren befinner sig i mångtydig och komplex. Vi vill i vår uppsats använda oss av experiment som ett sätt att undersöka och beskriva denna situation. Genom att renodla en aspekt som vi på förhand kan anta skulle kunna ha betydelse kan vi om experimentet är väl utfört konstanthålla för andra faktorer.<sup>11</sup> Utgångspunkten för vårt intresse för problemområdet och vad som kan tänkas påverka betygsättningen var våra erfarenheter från praktikperioden under lärarutbildningen. Förutsättningarna för undervisning och betygsättning var mycket olika på de yrkesinriktade respektive studieinriktade program vi praktiserade inom. Vi tänker oss att lärarens förståelse av eleven har stor betydelse när bedömning av en uppgift skall utföras. I denna uppsats vill vi undersöka om lärarens vetskap om elevens programtillhörighet påverkar bedömningen i den ena eller andra riktningen. Om vi upptäcker en skillnad i bedömning av samma uppgift mellan lärare som enbart kan anses vara grundad i denna vetskap vill vi hävda att vi har visat på ett grundläggande problem som i sin tur har stor betydelse för i vilken mån man kan hävda att betygsättningen är likvärdig i systemet som helhet.

Syftet med detta examensarbete är alltså att undersöka hur lärarens vetskap om elevens identitet påverkar bedömningen av elevens kunskaper. Genom att avgränsa bedömningssituationen så att endast en faktor kan påverka vill vi mäta om lärarens syn på elever med olika *programtillhörighet* påverkar bedömningen och därmed i förlängningen betygens likvärdighet.

Först skall vi redogöra för vad den tidigare forskningen har kommit fram till när det gäller olika övergripande faktorer som kan påverka betygsättningen. I detta avsnitt vill vi också visa på intressanta jämförelsepunkter mellan vår undersökning och den forskning som går att relatera till lärarens vetskap om elevens programtillhörighet. Därefter kommer vi att närmare utveckla begreppen *likvärdighet* och *handlingsutrymme*. Resonemanget avser att beskriva de förutsättningar av systemkaraktär som omger läraren i bedömningssituationen. Experimentets konstruktion och de val vi har ställts inför i forskningsprocessen kommer att framställas under rubriken *Design*. I resultatredovisningen redovisas de viktiga slutsatser vi kan härleda från analysen av materialet vi fått in under experimentet. Vi kommer att visa att elever som går hotell- och restaurangprogrammet får genomsnittligt högre betyg än elever som går naturvetarprogrammet för exakt samma prestation, att bedömningen av exakt samma elevsvar resulterar i allt från IG till MVG samt att manliga lärare under vissa förutsättningar sätter högre betyg än kvinnliga lärare på samma elevsvar. Avslutningsvis kommer vi att återkoppla till den tidigare forskningen och den övergripande diskussionen om likvärdighet.

---

<sup>11</sup> Både "legitima" (som har att göra med bredare definitioner av likvärdighet och hänsynstagande till andra faktorer än resultatet på testet/provet) och "icke-legitima" (strukturer och förståelse/fördomar).

## 2. Tidigare forskning

### 2.1 Översikt

Forskningen på området har använt olika mått för att beräkna en avvikelse från vad som kan antas vara den rättvisa bedömningen. Dessutom har olika faktorer lyfts fram som förklaringar till denna avvikelse. Förklaringarna kan delas in i tre övergripande områden. För det första kan olikheter i betygsutfall handla om elevens prestation i sig. Denna prestation kan i sin tur förklaras av elevens bakgrund, ofta beskriven i termer av klass, kön eller etnicitet som i sin tur kan antas påverka viktiga faktorer för inläring som intresse och motivation. För det andra kan olikheterna bero på skolans organisations- eller driftsform, pedagogiska inriktning eller storlek. För det tredje kan olikheterna bero på egenskaper hos den betygsättande läraren.

Enligt den forskningsöversikt som Helena Korp genomförde på uppdrag av Skolverket 2003 är den internationella forskningen kring kunskapsbedömning mycket omfattande och tilldrar sig allt mer intresse relativt andra utbildningsvetenskapliga områden. Däremot identifierar hon stora brister inom den svenska och nordiska forskningen när det gäller kunskapsbedömning och efterlyser arbeten som sätter in prov och bedömning i ett teoretiskt, historiskt och samhällsligt sammanhang samt ”empiriska studier av svensk bedömningspraktik”.<sup>12</sup> Även Viveka Lindholm som har publicerat en översikt när det gäller svensk forskning om bedömning och betyg under perioden 1990-2005 menar att det saknas ”beskrivningar av den konkreta bedömningspraktiken i skolorna”.<sup>13</sup> Hon vill också lyfta fram att forskningen domineras av en psykometrisk inriktning samt att de flesta studier behandlar grundskolan snarare än gymnasieskolan.

Ett antal svenska avhandlingar har under de senaste fem åren ägnats åt att undersöka olika aspekter av problematiken kring likvärdig utbildning och rättvis/jämförbar betygsättning i ett målrelaterat betygssystem. Förutom dessa vill vi också förhålla oss till den tidigare nämnda rapporten från Skolverket samt ett antal tidigare examensarbeten som med liknande problemområde som vårt använt sig av den experimentella metoden. Vår genomgång av detta material bekräftar Lindholms påstående om att de flesta tidigare arbeten är inriktade på grundskolan.<sup>14</sup> Detta faktum kombinerat med vårt eget yrkesmässiga intresse har bidragit till vårt val att förlägga undersökningen till gymnasieskolan.

I det följande skall vi titta lite närmare på vad tidigare forskning säger om olika faktorer som har med *skolan*, *elev* eller *läraren* att göra för att därefter rikta ljuset mot vår påverkansvariabel *programtillhörighet*. Tänkbara samband och förklaringar till varför denna variabel skulle kunna tänkas ha betydelse i bedömningssituationen kommer att behandlas.

---

<sup>12</sup> Korp, Helena (2003) *Kunskapsbedömning – hur, vad och varför* Stockholm: Myndigheten för skolutveckling. s. 18

<sup>13</sup> Lindholm, Viveka (2005) *Svensk forskning om bedömning och betyg 1990-2005* **Studies in Educational Policy and Educational Philosophy E-tidskrift** 2005:1 s. 20

<sup>14</sup> En huvudanledning till detta förhållande är förmodligen att det inte var möjligt förrän 2009 att koppla individens personnummer till provresultat på gymnasiet. Den intressanta jämförelsen av hur resultat från ett lägre stadium korrelerar med resultat från ett högre möjliggörs till exempel. Skolverket (2009) s. 33



## 2.2 Olika mått

De forskare som velat undersöka olika faktorerers betydelse i förhållande till en avvikelse från det rättvisa betyget har använt två olika typer av mått. Det första måttet bygger på en jämförelse mellan resultat på det nationella provet och slutbetyg. Detta gäller både för Alli Klapp Lekholms avhandling *Grades and grade assignment - effects of student and school characteristics*<sup>15</sup> och för rapporten från Skolverket från i år (2009) *Likvärdig betygsättning i gymnasieskolan – En analys av sambandet mellan nationella prov och kursbetyg*.<sup>16</sup> Den andra typen av mått är en jämförelse mellan slutbetyg och resultat på högskoleprovet. I avhandlingen *Criterion-Referenced Measurement for Educational Evaluation and Selection* har Christina Wikström använt måttet för att testa den kriterierelaterade betygsättningens värde som urvalsinstrument. Med utgångspunkt i måttet undersöker hon även ett antal faktorerers samvariation med den uppmätta avvikelsen.<sup>17</sup>

Gemensamt för dessa mått är att de använder statistiska uppgifter som är hämtade från verkliga resultat och verklig betygsättning. Om det går att fastställa skillnader i avvikelse som samvarierar med olika faktorer som inte borde ha betydelse kan man hävda att resultaten tyder på en orättvis/ojämförbar betygsättning. Vårt experiment kan sägas utgöra ett liknande mått med den viktiga skillnaden att deltagarna vid bedömning av den uppgift vi konstruerat befinner sig i en artificiell situation. Vi kan därför inte hävda att resultaten i experimentet direkt kan överföras till en verklig situation. Detta leder till att de olika måtten inte är fullt ut jämförbara.

## 2.2 Skolans betydelse

Efter friskolereformen under 1990-talet har ökad uppmärksamhet riktats mot vilken betydelse skolans driftsform och konkurrenssituation har för olika aspekter av verksamheten.

Skolverkets rapport visar att det inte går att hitta någon större skillnad i nettoavvikelse mellan friskolor och kommunala skolor. Dessutom pekar en jämförelse mellan siffrorna för år 1999 och år 2008 att ingen förändring skett under den period då antalet fristående skolor ökat som mest. Slutsatsen blir enligt Skolverket att det inte finns skäl att tro att fristående skolor sätter betyg på ett annat sätt än de kommunala skolorna.<sup>18</sup> Lekholm som använder ett liknande mått testar i sin undersökning olika faktorer mot varandra. Hon finner ett samband mellan fristående skolor och hög avvikelse men detta samband försvagas kraftigt när faktorn kontrolleras mot uppgifter om föräldrarnas utbildningsnivå.<sup>19</sup> På skolnivå betyder detta att skolor som har många elever med föräldrar som har en låg utbildningsnivå tenderar att ta större hänsyn till andra faktorer än de rent kognitiva.<sup>20</sup> När Wikström jämför kommunala och fristående skolor med olika grad av skoltäthet som ett mått på hur stor konkurrens de kan antas vara utsatta för finner hon att detta har betydelse. För kommunala skolor blir det en relativt liten effekt medan det får stor effekt för de fristående skolorna. Hennes slutsats blir att

---

<sup>15</sup> Klapp Lekholm, Alli (2008) *Grades and grade assignment - Effects of student and school characteristics*. Göteborg studies in educational science, Vol. 269. Göteborg: Acta Universitatis Gothoburgensis

<sup>16</sup> Skolverket kallar detta mått för *nettoavvikelse*. Skolverket (2009) s. 21ff

<sup>17</sup> Wikström, Christina (2005) *Criterion-Referenced Measurement for Educational Evaluation and Selection*. Avhandling. Department of Educational Measurement Umeå University No 1

<sup>18</sup> Skolverket (2009) s. 58ff

<sup>19</sup> Klapp Lekholm (2008) s. 76

<sup>20</sup> Ibid, s. 86f

användandet av betyg som ett kvitto på att skolan har hög kvalitet i en konkurrenssituation kan leda till betygsinflation.<sup>21</sup>

Meningarna är alltså delade inom forskningen om vilken betydelse friskolor och konkurrens har för likvärdigheten i det svenska skolsystemet. Ett resultat som är relevant att diskutera i förhållande till vår undersökning är Lekholms jämförelse mellan effekt på nettoavvikelsen av att eleven går på en fristående skola med effekten av elevens föräldrars utbildningsbakgrund. Det tydliga mönstret av att elever som går yrkesprogram har föräldrar med lägre utbildningsbakgrund än elever på studieförberedande<sup>22</sup> kan tillsammans med Lekholms resultat sägas stärka relevansen av vårt val av faktor. Att föräldrarnas utbildningsbakgrund har en större betydelse än vilken skola eleven går på tyder på att faktorn *programtillhörighet* är viktig. I nästa steg är det möjligt att lärarens förförståelse om elevens bakgrund påverkar betygsättningen. Läraren behöver kategorisera eleverna för att skapa ett underlag för undervisningens utformning. Även om många lärare i sin professionalitet inkluderar en strävan efter att undvika stereotyper och fördomar kan det i skolmiljön vara svårt att undvika helt.<sup>23</sup>

### 2.3 Elevens betydelse

De olika gruppegenskaper i form av klasstillhörighet, kön och etnicitet som eleverna har med sig in i skolan skall egentligen inte kunna påverka en tänkt objektiv bedömningsituation.<sup>24</sup> Den tidigare forskningen ger motstridiga besked om vilken betydelse de olika faktorerna har i sammanhanget. I sin rapport finner Erica Lindahl ett tydligt positivt samband med ett högre slutbetyg i förhållande till resultat på det nationella provet för både egenskaperna kön och etnicitet<sup>25</sup> men samtidigt menar Skolverket att denna påverkan är svag. Skolverket hävdar att sambandet försvagas eftersom effekterna av kön inte kan återfinnas på skol eller klass nivå.<sup>26</sup> För etnicitet har man använt sig av definitionen: *annat modersmål än svenska* och finner inte heller här några skillnader.<sup>27</sup> Eftersom Lindahl använder sig av en annan definition: *födda utanför Norden* blir resultaten dock inte helt jämförbara. Inte heller Lekholm menar att det går att belägga ett samband mellan vilket kön eleven tillhör och avvikelser mellan slutbetyg och resultat på högskoleprovet. Hon inför en variabel som beskriver vilket generellt intresse för lärande och motivation för skolarbete eleven har. När de bägge faktorerna ställs mot varandra förklarar variabeln intresse och motivation nästan hela den tidigare uppmätta könsskillnaden.<sup>28</sup>

---

<sup>21</sup> Wikström (2005) s. 33

<sup>22</sup> Broady, Donald & Mikael Börjesson (2006) "En social karta över gymnasieskolan" ur *Ord & Bild 3 – 4 2006 Tema - Skolan*

<sup>23</sup> Michael Lipsky har till exempel i sin teori om gatunivåbyråkrater hävdad att institutioner som har till uppgift att behandla människor av nödvändighet skapar sin klient utifrån vissa givna kategorier. För lärare kan denna effekt leda till att förväntningarna blir olika för elever med olika bakgrund. Detta kan i sin tur påverka förutsättningarna för lärande. Lipsky (1980) s. 59-67

<sup>24</sup> En stor mängd forskning har som utgångspunkt att det är svårt att föreställa sig en helt objektiv bedömningsituation och att kunskapens sociala karaktär medför att strukturella faktorer hela tiden påverkar förhållandet mellan elev och skolmiljö (inklusive lärare).

<sup>25</sup> Lindahl, Erica (2007) *Systematiska avvikelser mellan slutbetyg och provresultat – Spelar elevens kön och etniska bakgrund roll?* Institutet För Arbetsmarknadspolitisk Utvärdering Rapport 2007:22

<sup>26</sup> Skolverket har undersökt skillnader på flera nivåer. Det handlar alltså om skillnader mellan olika klasser och mellan olika skolor.

<sup>27</sup> Skolverket (2009) s. 8

<sup>28</sup> Klapp Lekholm (2008) s. 73

Av dessa tre övergripande kategorier är klasstillhörigheten underrepresenterad i den senare forskning som har bedrivits om betyg och betygsättning. Det är framförallt till denna kategori som vår undersökning kan relateras. Eftersom eleverna på yrkesinriktade program till största delen har arbetarklassbakgrund och eleverna på studieinriktade i högre grad har medelklassbakgrund skulle perspektiv som tar hänsyn till detta faktum kunna föreslå förklaringar till eventuella skillnader mellan grupperna när det gäller lärarens betygsättningspraxis. På en strukturell nivå har teorier föreslagits genom två dominerande perspektiv. Inom ett *konsensusinriktat* perspektiv tenderar forskare att uppfatta skolan som funktionell i förhållande till samhällets behov av arbetskraft. Graden av motsättningar mellan samhälle och skola ses som låg. En annan tradition i form av ett *konfliktinriktat* perspektiv menar tvärtom att skolan i många fall bidrar till att reproducera de ojämlika strukturer som man uppfattar finns i samhället.<sup>29</sup> I detta sammanhang kan våra resultat tolkas ur bägge dessa perspektiv. Bristen på rättvisa betyg blir ett problem inom konsensusperspektivet i första hand för att villkoren för lika möjligheter i konkurrensen om platser på högre utbildning blir olika. De lösningar som föreslås handlar om att utforma reliabla och valida bedömningsinstrument samt att skapa standarder för jämförbarhet mellan olika lärares bedömningar. Inom det konfliktinriktade perspektivet hävdar man att det inte går att separera undervisning från betyg och bedömning. Läraren är här en del av det system som reproducerar de skillnader på gruppnivå som finns i samhället.<sup>30</sup> En viktig aspekt av detta perspektiv som vi vill lyfta fram i det här arbetet är lärarens förväntningar. Denna faktor har visat sig ha stor betydelse för skolframgång i allmän mening.<sup>31</sup> När det gäller betyg och bedömning kan lärarens positiva eller negativa förväntningar grundade i vetskap om elevens klassbakgrund tänkas påverka utfallet.

## 2.4 Lärarens betydelse

I det svenska skolsystemet ges läraren särskilt stort ansvar när det gäller betygsättning genom avsaknaden av oberoende examinationsformer.<sup>32</sup> Trots detta är det svårt att hitta undersökningar av betygens likvärdighet som fokuserar på läraren. De avhandlingar som finns använder sig i huvudsak av intervjumetoder för att beskriva hur lärare tänker kring betygsättning. Helena Korp har undersökt hur förutsättningarna för undervisning och betygsättning skiljer sig åt mellan de studieinriktade och de yrkesinriktade programmen. Hon menar att risken är stor att det skapas ”negativa spiraler” som framförallt missgynnar yrkeseleverna. Hennes undersökning visar att det ofta saknas ett samarbete mellan kärnämnes och karaktärsämneslärare. Detta gör att det blir svårt att ”infärga”<sup>33</sup> de ämnen som inte har direkt med yrkesutbildningen att göra för att försöka bryta den negativa inställning som många elever har till kärnämnen. Risken är stor att eleverna minimerar sin arbetsinsats till den lägsta kravnivån för ett godkänt betyg samtidigt som lärarna sänker kraven.<sup>34</sup> Korp hävdar vidare att det inte är ovanligt att lärarna ”begränsar kursen till ett visst betygsspann beroende

---

<sup>29</sup> Korp (2006) s. 10-19

<sup>30</sup> Ibid, s. 29ff

<sup>31</sup> Gipps, Caroline & Patricia Murphy (1994) *A fair test? – Assessment, achievement and equity*. Philadelphia: Open University Press s. 4

<sup>32</sup> Det har i debatten under senare tid föreslagits att återinföra en studentexamen för gymnasieskolan. De internationella IB-programmen har redan en variant av examen där utomstående examinatore sätter betyg på ett genomfört slutprov. DN 090817 ”Konkurrens om elever ger orättvisa betyg”

<sup>33</sup> Med infärgning avses att arbeta med att låta programmets huvudinriktning prägla även de ämnen som är mer allmänna. Exempelvis kan man arbeta med tekniska begrepp i Engelska på fordonsprogrammet eller med livsmedelskemi på hotell och restaurangprogrammet.

<sup>34</sup> Korp (2006) s. 142f

på vilken ambitionsnivå och vilka kunskapsmässiga förutsättningar eleverna *antas* ha”.<sup>35</sup> I intervjuerna med de lärare som undervisar på både yrkesinriktade och studieinriktade program kunde hon se att dessa lärare använde sig av mycket skilda typer av kategoriseringar för att beskriva elever från de olika programmen. Där elever från det naturvetenskapliga programmet beskrivs som ambitiösa och intresserade beskrivs elever från Barn och Fritidsprogrammet samt Byggprogrammet som svagpresterande och omogna.<sup>36</sup>

Ytterligare en studie som fokuserar på läraren och betygsättning är Bengt Selgheds avhandling *Ännu icke godkänt – Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen* från 2004. Genom intervjuer med 30 lärare som har erfarenheter av att sätta betyg i årskurs 8 och 9 beskriver han olika sätt att uppfatta det mål- och kunskapsrelaterade betygssystem som infördes i grundskolan 1995. Resultaten visar att lärarna har haft stora problem med att översätta sin äldre betygsättningspraxis till det nya systemets intentioner. Han menar vidare att lärarnas olika sätt att uppfatta betygssystemet pekar på en institutionell tröghet i skolan och att systemet möjligen är alltför komplicerat i sin utformning.<sup>37</sup>

När Skolverket i en statistisk flernivåanalys jämförde nettoavvikelsen på skolnivå med samma avvikelse på klass nivå konstaterades att det hade större betydelse vilken klass eleven gick i än vilken skola. Den slutsats som Skolverket drar av detta samband är att det för eleven i mycket högre grad spelar roll vilken lärare man har än vilken skola man går på.<sup>38</sup>

Det intryck vi hade med oss från praktikperioden under lärarutbildningen av de problem och utmaningar läraren ställs inför när det gäller att undervisa och betygsätta i samma ämne på både yrkes- och studieförberedande program bekräftas av Korps studie. Skolverket pekar i sin rapport ut läraren som central när det gäller uppmätta avvikelser mellan slutbetyg och resultat på det nationella provet. Om Selgheds beskrivning av att olika lärare uppfattar det nya betygssystemet på olika sätt gäller även för gymnasieskolan i allmänhet kan detta vara en delförklaring till avvikelsen.

## **2.5 Programtillhörighetens betydelse**

Som tidigare har beskrivits skiljer sig förutsättningarna för undervisning och betygsättning mellan de yrkesinriktade och de studieinriktade programmen. Bland de forskare som studerat betydelsen av elevens *programtillhörighet* i förhållande till rättvis betygsättning återfinns vi en rad olika samband och slutsatser som vi bedömer som relevanta. Dessa undersökningar grundar sig i uppgifter om eleven och undersöker inte lärarens roll i sammanhanget. Wikström undersöker hur programtillhörigheten och kurssammansättning samvarierar med hennes avvikelsemått.<sup>39</sup> Hon finner att elever på yrkesinriktade program tenderar att få ett högre betyg jämfört med måttet och hävdar att elever på naturvetenskaplig linje blir särskilt missgynnade när det gäller urval till högre studier.

Den nya rapporten från Skolverket tar sin utgångspunkt i studier genomförda 2001 och 2002 där resultaten visade att elever på yrkesinriktade program fick högre betyg i förhållande till de

---

<sup>35</sup> Korp (2006), s. 232

<sup>36</sup> Ibid

<sup>37</sup> Selghed, Bengt (2004) *Ännu icke godkänt – Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Malmö Studies In Educational Sciences No. 15 s, 188ff

<sup>38</sup> Skolverket (2009) s. 43

<sup>39</sup> Wikström (2005) s. 35

nationella proven i matematik.<sup>40</sup> Detta generella samband återfinns inte i den nyare rapporten. Två typer av resultat lyfts fram av Skolverket för att belägga detta påstående. För det första valdes en skola som motsvarade riksgenomsnittet ut och skillnaderna mellan klasserna listades. Här kunde man se att det fanns en betydande variation inom grupperna yrkesinriktade och studieinriktade program.<sup>41</sup> För det andra undersöktes statistisk samvariation mellan prestationsnivå och nettoavvikelse. Olika resultat erhöles för de olika ämnen som ingick i studien. Gemensamt för samtliga mätningar var dock att spridningen för nettoavvikelse i gruppen yrkesinriktade program var stor.<sup>42</sup> Ett program sticker dock ut i Skolverkets undersökning. För det naturvetenskapliga programmet, som har högst betygsmedelvärde av samtliga program, uppmättes den i särklass största nettoavvikelsen (i positiv riktning).<sup>43</sup>

Vår undersökning passar väl in i den bild forskningen ger på detta område. Konkreta påståenden om att elever på yrkesinriktade program gynnas (Wikström) eller att elever på naturvetenskapliga programmet får högre betyg än prestationen motiverar (Skolverket) kan direkt relateras till vårt experiment. Förklaringar till dessa samband kan sökas i olika riktningar. Det är dock sannolikt att läraren och lärarens uppfattningar om eleven har betydelse. Vi vill därför undersöka om lärarens vetskap om elevens programtillhörighet påverkar bedömningen i den artificiella bedömningsituationen. Även om det är viktigt att påpeka att vi inte testat de tidigare föreslagna sambanden kan vi i våra slutsatser resonera om i vilken mån vårt resultat stärker eller försvagar dessa samband.

## **2.6 Tidigare experimentella studier**

Under vår forskningsgenomgång hittade vi tre studentuppsatser som ligger nära vårt intresseområde och metodval. Två av dessa studier har genom experiment undersökt om det går att påvisa inverkan av elevens kön respektive etnicitet på bedömning av konstruerade provsvar inom kursen samhällskunskap A. Resultaten visar inga statistiskt signifikanta skillnader mellan grupperna, men väl andra intressanta indikationer. I studien om kön av Anna-Maria Fjellman kunde en stor spridning observeras i bedömningen.<sup>44</sup> Denna spridning var oberoende av kön, men intressant då den visar hur utlämnade elever är åt vilken lärare de blir bedömda av. Henrik Gullberg genomförde ett experiment för att visa på eventuella skillnader i bedömning beroende på elevens etnicitet, men hittade ingen statistiskt signifikant skillnad mellan experimentgrupperna.<sup>45</sup> De olika grupperna fick till uppgift att bedöma en kille med ”svenskt utseende” och en kille med ”utländskt utseende”. När experimentgrupperna sedan delades upp efter ålder såg han dock en signifikant skillnad där äldre lärare gav killen med utländskt utseende lägre betyg än yngre lärare gjorde.

Den studie som ligger närmast vår är ett experiment av David Johansson som tittar på hur lärarens vetskap om elevens programtillhörighet påverkar bedömningen av en uppgift i

---

<sup>40</sup> Skolverket (2009) s. 53

<sup>41</sup> Ibid, s. 43

<sup>42</sup> Ibid, s. 55

<sup>43</sup> Ibid, s. 57

<sup>44</sup> Fjellman, Anna-Maria (2008) *Duktiga flickor och tuffa pojkar – en experimentell studie om könskillnader i betygsättning*. Examensarbete inom lärutbildningen i Göteborg. Göteborgs Universitet

<sup>45</sup> Gullberg, Henrik (2008) *Olika etnicitet, lika rätt? – En experimentell studie om etnicitetens påverkan på betyg och bedömning*. Examensarbete i Samhällskunskap Göteborgs Universitet

samhällskunskap.<sup>46</sup> Hans resultat visar ingen signifikant skillnad mellan experimentgrupperna. Ett experiment säger sällan särskilt mycket på egen hand dock, utan upprepningar och liknande experiment är ett måste för att bekräfta resultaten. Vi har därför valt att genomföra ett experiment som liknar Johanssons, med det undantaget att vi använder oss av kärnämnet naturkunskap A istället för samhällskunskap. Man kan tänka sig att dessa båda ämnen skiljer sig på flera punkter i bedömningsituationen vilket kan orsaka skillnader i utfallet mellan våra två experiment. Naturkunskap ses i allmänhet som ett mer faktabetonat ämne än samhällskunskap vilket kan ge skillnader i bedömning. Bedömningsformerna skiljer sig också då lärare i samhällskunskap vanligtvis använder sig av uppsatsskivning som betygsunderlag i högre grad än lärare i naturkunskap.

## 2.7 Sammanfattning

I detta kapitel har vi visat att forskningen pekar ut vissa faktorer som tänkbara påverkansfaktorer när det gäller betygsättning. Bland dessa identifierar vi lärarens vetskap om elevens programtillhörighet som särskilt intressant att undersöka. Bristen på forskning under senare tid som ägnats åt klasstillhörighetens betydelse bidrar i detta sammanhang till att vi anser att denna inriktning är vetenskapligt relevant. Även de resultat som kopplar samman elevens föräldrars utbildningsbakgrund som stark påverkansfaktor för en nettoavvikelse för betygsättningen kan indirekt förbindas med programtillhörigheten. Att förutsättningarna för undervisning och betygsättning skiljer sig åt mellan yrkesinriktade och studieinriktade program bekräftas av framförallt Helena Korps avhandling. Även om vi inte i denna uppsats avser att *förklara* ett eventuellt resultat har vi föreslagit att lärarens förväntningar, förförståelse och kategorisering av eleven skulle kunna ingå i vidare teoribildning på detta fält. Vidare vill vi hävda att vårt resonemang om att använda experimentet som ytterligare ett mått på rättvist/jämförbart betyg kan bidra till metodutvecklingen inom detta område. De tidigare experimentella studier som vi har tagit del av har inte lyckats påvisa någon effekt av faktorerna kön, etnicitet eller programtillhörighet. Statistiska studier som mer specifikt använder programtillhörigheten som påverkansfaktor ger tvetydiga resultat och detta ger oss ytterligare anledning att genomföra vår studie.

## 3. Teoretiskt ramverk

I detta kapitel vill vi med hjälp av de centrala begreppen *likvärdighet* och *handlingsutrymme* diskutera viktiga aspekter av betygssystemet och styrsystemet som den betygsättande läraren har att ta ställning till i sitt yrkesutövande. Vi vill på detta sätt placera in *bedömningsituationen* och därmed *experimentet* i ett sammanhang.

---

<sup>46</sup> Johansson, David (2009) *Påverkas lärare av elevers programtillhörighet i betygsprocessen? – En experimentell studie om huruvida lärares vetskap om elevers programtillhörighet påverkar betygsättningen*. Examensarbete i samhällskunskap Göteborgs Universitet

### 3.1 Likvärdighet

För att läraren skall kunna bedöma och betygsätta elevens prestationer rättvist måste denne ta ställning till de styrdokument som finns. I dessa har begreppet *likvärdighet* en särställning. Likt andra positivt laddade värdebegrepp som demokrati och jämlikhet förs diskursiva strider om vad de skall innehålla och vilka konsekvenser som skall dras av en viss tolkning.

Tomas Englund har i sin forskning studerat begreppets användning på den högsta normgivande nivån i auktoritativa texter från 1970-talet och framåt.<sup>47</sup> Han identifierar en förskjutning av begreppets innebörd genom tre perioder. I detta sammanhang handlar det i första hand om i vilken mån en likvärdig utbildning kan uppnås i skolsystemet som helhet och inte direkt om betygsättning.

Under den första fasen gavs likvärdigheten en innebörd som handlade om utbildning som en social rättighet fördelad jämlikt i befolkningen. Att alla skulle fostras in i en gemensam referensram ansågs vara av central vikt. De som av olika skäl hade svårt att passa in eller hänga med i undervisningen skulle kompenseras för detta genom tillskjutande av extra resurser. Likvärdighetsbegreppet kommer under denna tid att successivt ersätta jämlikhetsbegreppet.<sup>48</sup> Tanken var under denna tid att ett starkt centraliserat skolsystem och ett normrelaterat betygssystem tillsammans skulle leda till att utbildningen på respektive linje blev så lika som möjligt. Skillnader i utbildningsresultat kan med detta synsätt sägas motsvara reella skillnader i begåvning.<sup>49</sup>

Den relativa konsensus som tidigare funnits bland de politiska partierna om denna jämlikhetsgrundade tolkning av likvärdighetsbegreppet utmanades under 1980-talet från höger. Kritiken grundades i uppfattningen att det på förhand bestämda kursinnehållet innebar att elevernas utveckling som individer hämmades. Likvärdighet uppstår enligt denna syn om varje elev ges förutsättningar att hitta den inriktning som bäst passar just denna individs intressen. Idealet handlar om att om alla får lika chans att utveckla sin komparativa fördel relativt andra i skolan uppstår lika möjligheter och därmed rättvisa.<sup>50</sup>

Under 1990-talet infördes både ett nytt kriterierelaterat betygssystem och ett nytt decentraliserat styrsystem för skolan. Likvärdighet skall i detta system garanteras av god uppföljning, utvärdering och tillsyn men också tydligt formulerade nationella målbeskrivningar i läroplanen, lärarutbildningens utformning och behörighetsregler för lärare. Englund identifierar en tyngdpunktsförskjutning under perioden fram till vår tid från en betoning av målen till ett fokus på resultaten. Innebörden av likvärdigheten begränsas till ett allmänt lägsta krav på godkänt. Likvärdighet ses som ett mått på i vilken mån enskilda skolor (med olika driftsform och/eller pedagogisk inriktning) lyckas uppnå detta lägsta krav.<sup>51</sup>

I denna beskrivning av hur innebörden av begreppet har förändrats över tid handlar det i första hand om *likvärdighet* på övergripande nivå som hela skolsystemet eller i en jämförelse mellan skolor. De olika tolkningarna framkallar ett spektrum av värderingsskillnader när det gäller den svenska skolans utformning. Ett sätt att förstå dessa olikheter är att analysera hur de

---

<sup>47</sup> Englund, Tomas & Ann Quennerstedt (red.) (2008) *vadå likvärdighet – studier i utbildningspolitisk språkanvändning*. Göteborg: Daidalos

<sup>48</sup> Ibid, s. 8

<sup>49</sup> Korp (2006) s. 49f

<sup>50</sup> Englund & Quennerstedt (red.) (2008) s. 19f

<sup>51</sup> Ibid, s. 24f

förhåller sig till det centrala liberala begreppet *lika möjligheter*.<sup>52</sup> Om elevernas olika förutsättningar inte skall ligga till grund för skillnader i resultat måste viss kompensation införas om man skall kunna hävda att systemet är rättvist menar vissa.<sup>53</sup> Andra hävdar att en individualiserad undervisning som fungerar är medlet för att erbjuda dessa lika möjligheter. Ytterligare en variant är att det är kvalitén på skolan som skall vara jämförbar så att eleverna skall ges lika chanser att lyckas.

Det är rimligt att anta att samtliga dessa olika uppfattningar finns representerade av enskilda lärare i det svenska skolsystemet. Eftersom tolkningen av likvärdigheten innebär olika uppfattningar av hur undervisningen skall anpassas och vilken roll betygen har kommer detta att bilda en viktig förutsättning för varje diskussion om rättvisa/jämförbara betyg. Till stöd för läraren när det gäller att sträva efter att följa statens intentioner för likvärdig utbildning och betygssättning finns ett antal styrdokument på olika nivåer.

I skollagens 1 Kap §2 står det: ”Alla barn och ungdomar skall, oberoende av kön, geografisk hemvist samt sociala och ekonomiska förhållanden, ha lika tillgång till utbildning i det offentliga skolväsendet för barn och ungdom. Utbildningen skall inom varje skolform vara likvärdig, varhelst den anordnas i landet.”<sup>54</sup> Denna paragraf talar om likvärdiga skolformer och överhuvudtaget inte om likvärdig bedömning. Den skiljer också tydligt mellan lika tillgång till utbildning och likvärdig utbildning inom varje skolform.

Den nu gällande läroplanen från 1994 stipulerar under rubriken ”Grundläggande värden” däremot att innebörden av likvärdighet inte innebär att undervisningen skall vara lika för alla utan snarare anpassad efter varje elevs olika ”förutsättningar, behov och kunskapsnivå”. Läroplanen ger också stort utrymme för olika innehåll och undervisningsformer. Här uppstår en svår fråga för läraren att ta ställning till. När man anpassar undervisningen efter elevernas olika förutsättningar och behov uppstår samtidigt en olikhet. Då kan man fråga sig var gränsen går för hur mycket olikhet som kan tolereras för att man skall kunna fortsätta hävda att likvärdighet råder? I ett system som bygger på tanken att olika vägar kan leda till samma mål blir tydligheten i målbeskrivningarna extra viktiga. Genom de riksgiltiga kursplanernas målbeskrivningar ges enligt läroplanen normerna som ligger till grund för likvärdigheten i skolsystemet som helhet.<sup>55</sup>

I Gymnasieförordningen står det ingenting uttryckligen om likvärdighet. De direktiv som man kan anta delvis syftar till en högre grad av jämförbarhet när det gäller betyg handlar om krav på att lokala betygsriterier skall finnas<sup>56</sup> samt att läraren inom kärnämnen svenska, svenska som andraspråk, engelska och matematik skall använda sig av nationellt fastställda prov för att få bedömningsgrunder som avser att leda till att dessa skall: ”bli så enhetliga som möjligt över landet.”<sup>57</sup> Här pekas de nationella proven ut som ett viktigt riktmärke men läraren får inga direktiv om hur de skall användas. Det är fortfarande en öppen fråga vilken relativ vikt resultat på det nationella provet skall tillmätas i förhållande till annat betygsunderlag.

---

<sup>52</sup> Kymlicha, Will (1995) *Modern politisk filosofi*. Nora: Bokförlaget Nya Doxa

<sup>53</sup> Frågan om hur denna kompensation skall se ut och hur långt man kan gå utan att andra värden hotas är en i högsta grad politisk fråga där olika synsätt står mot varandra.

<sup>54</sup> SFS 1985:1100

<sup>55</sup> Läroplan för de frivilliga skolformerna Lpf 94 1:2

<sup>56</sup> SFS 1992:394 Kap 7 §5 ändrad genom förordning 2001.56

<sup>57</sup> Ibid.



För att de kriterierelaterade betygen skall kunna bli jämförbara måste ett antal komponenter i systemet för målstyrning fungera på ett tillfredsställande sätt.<sup>58</sup> Lokala betygskriterier skall utformas på varje skola som sedan genom deltagande från elever konkretiseras till det stoff och de metoder som förverkligas i klassrummet.<sup>59</sup> I en avhandling från 2006 av Jörgen Tholin kunde konstaterats att en fjärdedel av de skolor han undersökte helt saknade lokala betygskriterier i ett eller flera ämnen.<sup>60</sup> 14% av skolorna hade samma betygskriterier 2005 som 1996. Detta menar Tholin kan ses som ett tecken på att betygskriterierna inte är levande dokument bland lärarna. Genom att jämföra de lokala betygskriterierna med tolkningar som görs av andra lärare samt resultat på nationella prov är det tänkt att graden av likvärdighet skall öka. Ytterligare en komponent är diskussioner mellan lärare från olika program, årskurser och skolor kring tolkningar av betygskriterier. Detta avser att som det står i Skolverkets allmänna råd ”skapa likvärdiga förutsättningar för undervisningen och bedömningen av elever i skolan och i olika skolor.”<sup>61</sup> Avsaknaden av en tydlig yttre måttstock gör att läraren i sin roll som professionell bedömare blir särskilt viktig. Om läraren verkar i en miljö där det saknas en levande diskussion bland kollegor om vad de lokala betygskriterierna innebär i respektive ämne riskerar rättvisan/jämförbarheten att bli lidande.

Enligt Ingrid Carlgren som ingick i läroplanskommittén och i den expertgrupp vid Skolverket som tog fram de första betygskriterierna bygger det nuvarande betygssystemet på det hon kallar ett *verksamt professionellt omdöme* och medverkan från eleverna i betydligt högre omfattning än tidigare.<sup>62</sup> Systemets legitimitetsgrund är enligt henne en annan än i det relativa betygssystemet där standardprov kunde utgöra en slags garant för jämförbarheten.

### **3.2 Handlingsutrymme**

De intentioner som kan utläsas ur styrdokumentet skall förverkligas av läraren i ett konkret organisatoriskt sammanhang. Ett sätt att beskriva lärarens roll i skolan är som tjänsteman.<sup>63</sup> I denna roll har läraren mycket stor betydelse för i vilken grad betygen i slutändan kan betraktas som rättvisa/jämförbara. Här kan man tänka sig att graden av frihet eller *handlingsutrymme* varierar. I följande avsnitt vill vi lyfta fram ett perspektiv på detta handlingsutrymme som i uppsatsen kan bidra till ytterligare förståelse av bedömningssituationens förutsättningar.

Trots senare tids decentralisering av skolan i form av kommunalisering och införande av resultat och målstyrning kan skolan fortfarande betraktas som en statligt styrd institution som i likhet med andra statliga institutioner ska agera i enlighet med statens beslut. Olika institutioner varierar dock i hur lätt- eller svårstyrda de är, där skolan hamnar långt ut på den

---

<sup>58</sup> Varianter av denna organisationsteori är mycket utbredd både i näringsliv och offentlig sektor men har också kritiserats för att inte fungera i praktiken se Rombach, Björn (1991) *Det går inte att styra med mål*. Lund: Studentlitteratur och Flaa. Paul m.fl. (1998) *Introduktion till organisationsteori*. Lund: Studentlitteratur 107-111

<sup>59</sup> Korp (2006) s. 57 och SKOLFS 2007:42

<sup>60</sup> Tholin, Jörgen (2006) *Att klara sig i ökad natur – En studie av betyg och betygskriterier – historiska betingelser och implementeringen av ett nytt system*. Avhandling. Skrifter från högskolan i Borås s. 119

<sup>61</sup> SKOLFS 2004:23

<sup>62</sup> Carlgren, Ingrid (2002) ”Det nya betygssystemets tankefigurer och tänkbara användningar” ur antologin: *Att bedöma eller döma? Tio artiklar om bedömning och betygsättning*. Skolverket s. 17f

<sup>63</sup> Lundquist, Lennart (2006) ”Läraren som demokratins väktare” ur *Ord & Bild 3 – 4 2006 Tema - Skolan* Detta är dock inte det enda sättet. Lärarens yrkesroll har i litteraturen beskrivits i termer av Profession, Samverkare och Säljare. se Fredriksson, Anders (2007) ”Läraryrket och den politiska styrningen av skolan” i Pierre, Jon (2007) *Skolan som politisk organisation*. Malmö: Gleerups Utbildning AB

svårstyrda kanten. De institutioner som är mer lättstyrda präglas av en byråkratisk och regelföljande organisationskultur. Exempel på sådana institutioner är skatteverket och försäkringskassan. På skatteverket finns det tydliga regler för vad som gäller i de flesta situationer och det är viktigt att alla människor möts och mäts i ljuset av dessa regler för att principen om lika behandling skall kunna upprätthållas. De mer svårstyrda institutionerna, till exempel skola och polis, lever i en nära förbindelse med medborgarna. Här uppkommer ständigt nya situationer som inte kunnat förutses av den mest ingående regelbok, vilket gör att reglerna för dess tjänstemän snarare får ses som ett ramverk att tillämpa på olika sätt i olika situationer.

Statsvetaren Michael Lipsky fångade in detta fenomen genom att studera tjänstemannautövandet hos polis, skola och socialtjänst. Han såg att tjänstemännen inom dessa tre institutioner arbetade så nära medborgarna att de hela tiden befann sig i nya situationer där de var tvungna att själva utifrån en tolkning av de politiska intentionerna på högre nivå fatta sina egna beslut i mötet med medborgaren. Ofta är situationen så komplicerad att det blir omöjligt att följa en given regel.<sup>64</sup> Han kallade dessa tjänstemän för frontlinjebyråkrater (street-level bureaucrats), eftersom de ska tillämpa byråkrati på frontlinjen mellan tjänstemannen och medborgaren.<sup>65</sup> Beslut som "egentligen" skulle fattats av politikerna överläts alltså på frontlinjebyråkraten och konsekvensen av detta blir att frontlinjebyråkraterna får ett stort ansvar och att deras handlingsutrymme växer.

Lipsky beskriver också hur organisationer kan försöka påverka graden av handlingsutrymme genom att införa vissa rutiner när det gäller särskilda uppgifter. Inom undervisningen kan exempelvis minskat handlingsutrymme för läraren uppnås genom införandet av standardiserat undervisningsmaterial eller datoriserade program för lärande.<sup>66</sup> På detta sätt kan man tänka sig att det reella handlingsutrymmet varierar från situation till situation. Den ökade graden av lika behandling leder samtidigt till minskade möjligheter för tjänstemannen att situationsanpassa.

Skolan styrs på olika nivåer. På Skolverkets hemsida i en övergripande förklaring av skolans styrsystem står det att beslut på dessa olika nivåer ska samverka.<sup>67</sup> Utbildningens inriktning, skolämnen, mål och riktlinjer ska bestämmas nationellt. Sedan är det på lokal nivå upp till kommun och skola att fördela resurser och ta beslut om organisation. Nere på lärarens nivå finns det dock mycket kvar att besluta och utforma, vilket enligt Skolverket i stor utsträckning ska göras tillsammans med eleverna. Denna nivå kallas *det lokala frirummet* och det är alltså här den konkreta, praktiska utformningen av kursen och lektionerna sker. Lärarna ska med sin yrkeskunskap och yrkeserfarenhet hjälpa eleverna att ta ställning till val av innehåll och arbetsformer som ligger i linje med elevernas "förutsättningar, erfarenheter, intressen och behov" och på så vis konkretisera de nationellt fastslagna kursplanerna.<sup>68</sup> Hela denna struktur på styrningen ger alltså stort utrymme för en positiv anpassning till situationen, men den ger även utrymme för orättvisor och minskar oundvikligen likvärdigheten.

Anders Fredriksson diskuterar Lipskys teorier om frontlinjebyråkrati som förklaringsmodell till en undersökning som genomförts bland gymnasielärare i Sverige.<sup>69</sup> I undersökningen ville

---

<sup>64</sup> Lipsky (1980) s. 15

<sup>65</sup> Ibid, s. 3

<sup>66</sup> Ibid, s. 16

<sup>67</sup> <http://www.Skolverket.se/sb/d/1597> 091217

<sup>68</sup> <http://www.Skolverket.se/sb/d/1598> 091217

<sup>69</sup> Fredriksson (2007)

man se hur lärare hanterar värdefrågor och resultatet visade att lärarna bemöter samma värdefråga på mycket skilda sätt. Eftersom styrdokumentet gäller för samtliga lärare kan de skilda sätten att lösa samma problem inte bero på att lärarna styrs på olika sätt av politikerna, utan svaret måste ligga någon annanstans.<sup>70</sup> Då lärarens handlingsutrymme ökar öppnas också fältet upp för andra påverkande faktorer. Fredriksson lyfter fram skolans problem med begränsade resurser i form av dålig ekonomi, tidsbrist och personalbrist. Dessa inverkar givetvis på tjänstemannens förmåga att utföra ett arbete i enlighet med styrdokumentet. Men även andra faktorer får möjlighet att påverka lärarens agerande, så som tjänstemannen själv, brukarna (eleverna) och tjänstemannens kollegor.

Lärarens roll som tjänsteman blir som mest tydlig när bedömning som ligger till grund för betygsättning skall göras. Eftersom betygen kan få stor betydelse för elevens framtida möjligheter måste kraven på lika behandling och rättvisa/jämförbara betyg ställas högt.

### **3.3 Sammanfattning**

Detta kapitel har i sin första del syftat till att visa hur innebörden i begreppet likvärdighet har förändrats över tid samt redogöra för den tolkning som de nu gällande styrdokumentet ger uttryck för. Begreppets mångtydiga karaktär kan skapa problem för läraren när det gäller att sträva efter rättvisa/jämförbara betyg. Om den ideologiska konflikten som begreppet rymmer decentraliseras till enskilda skolor kommer det vara upp till varje lärarlag att lokalt hitta de kompromisser som krävs för att kunna arbeta. I den andra delen har lärarens position i skolsystemet beskrivits med hjälp av begreppet handlingsutrymme. Systemets konstruktion skapar ett lokalt frirum där det är tänkt att lokala betygs-kriterier och diskussioner lärare emellan skall ge en högre grad av rättvisa/jämförbarhet när det gäller betygsättningen.

Den viktigaste komponenten för att sträva efter likvärdig utbildning och försöka upprätthålla rättvisa/jämförbara betyg i styr- och betygssystemet i den svenska skolan är den professionella läraren. Detta ställer mycket stora krav på att läraren använder sitt handlingsutrymme på ett klokt sätt. Forskningen visar dock på att mycket stora olikheter finns i det samlade lärarkollegiet när det gäller tolkningar av vad likvärdighet i en skola för alla egentligen betyder och hur rättvisa/jämförbara betyg skall kunna uppnås. Tillexempel lägger en del lärare stor vikt vid de nationella proven som riktlinje medans andra tonar ner denna betydelse.<sup>71</sup>

## **4. Design**

I vår undersökning bestämde vi oss för att genomföra ett experiment för att kunna studera en konkret faktor, nämligen lärarens vetskap om elevens programtillhörighet, och dess eventuella påverkan på bedömning. För att kunna slå fast ett samband mellan orsak och verkan är experimentet den bästa metoden.<sup>72</sup> Det som främst skiljer experiment från andra vetenskapliga metoder är möjligheten till specifika manipulationer av förklaringsfaktorn samt en god

---

<sup>70</sup> Om inte tolkningarna av styrdokumentens avsikter också uppfattas olika. Givet att värderingen är entydigt uttryckt i styrdokumentet borde lärarna också uppfatta dess innebörd på samma sätt.

<sup>71</sup> Korp (2006) s. 147

<sup>72</sup> Esaiasson, Peter m. fl. (2009) *Metodpraktikan – Konsten att studera samhälle, individ och marknad*. Stockholm: Norstedts Juridik AB s. 371

kontroll för bakomliggande faktorer.<sup>73</sup> I vårt experiment innebar detta möjligheten att manipulera stimuli så att lärarna fick specifik information, vilket gjorde att vi undvek andra stimuli som skulle påverkat i en verklig situation.<sup>74</sup> Den stränga kontrollen för bakomliggande faktorer åstadkommes genom randomisering av experimentdeltagarna. Detta innebär att den enda skillnaden mellan de två grupperna är själva stimulus, vilket får konsekvensen att man kan dra slutsatser om kausalitet. På grund av bedömningssituationens komplexitet är det svårt att genom intervjuer och enkäter gå till botten med orsakssamband. Experimentet fyller då en viktig funktion genom att renodla aspekten "lärarens vetskap om elevens programtillhörighet" och reducera påverkan från många andra håll, till exempel elevens uppförande, personlighet, etnicitet och kön.

## 4.1 Experimentet

Experimentet gick ut på att låta lärare bedöma samma skriftliga elevprestation men med olika information om elevens programtillhörighet. Läraren fick läsa igenom tre provfrågor och bedöma en elevs svar till dessa med ett samlat betyg på skalan IG - MVG. Vi använde alltså experimentmetoden endast-efter-design.<sup>75</sup> Denna metod innebär att man utsätter grupperna för en manipulation och sedan mäter beroendevariabeln. Den främsta nackdelen med metoden är att vi inte kan vara säkra på att grupperna verkligen är lika i den bemärkelsen att de hade bedömt uppgiften lika om de inte hade utsatts för stimuli. Hade vi däremot använt en före-efter-design för att komma runt detta problem hade det istället varit svårt att undvika att uppmärksamma lärarna på att stimuli var just stimuli, och lärarna skulle då genomskådat experimentet.

Som diskuterats tidigare hade vi bestämt oss för att använda naturkunskap A (NkA) som ram för vår undersökning. Vid val av ämnesområde inom kursen var det viktigt att välja ett område som inkluderas i kursplanen, dels för att öka frågornas trovärdighet och relevans för kursen och dels för att lärarna i sina bedömningar skulle kunna relatera till kriterierna för kursen enligt kursplanen. Utifrån detta valde vi att använda ämnesområdet växthuseffekten. Detta område knyter kanske främst an till målet att "Eleven skall kunna beskriva miljöproblem utifrån studieinriktning och aktivt delta i diskussioner om möjligheten att påverka utvecklingen", men även till ett av kriterierna för Godkänt "Eleven ger exempel på globala och regionala miljöproblem och beskriver lokala miljöproblem till följd av vardaglig och yrkesmässig verksamhet.", och ett kriterium för Väl godkänt "Eleven diskuterar konsekvenser av olika praktiska ställningstaganden i energi-, miljö- och resursfrågor."<sup>76</sup> Kursplanen tar inte upp växthuseffekten specifikt och det är därför upp till varje lärare om det ska ingå i kursen, men eftersom växthuseffekten är ett av våra mest omfattande och omskrivna miljöproblem utgår vi ifrån att den absoluta majoriteten av lärare i NkA anser att växthuseffekten är en viktig del.

Det stimulus som vi ville undersöka var elevens programtillhörighet. Vi valde att använda oss av två varianter av detta stimulus, vilket innebär att vi delade in lärarna i två experimentgrupper. För att få elever som generellt sett ligger långt ifrån varandra på

---

<sup>73</sup> Ibid, s. 103

<sup>74</sup> Experiment som metod blir ibland kritiserat för att experimentsituationen alltid är konstlad på ett eller annat sätt, och man menar då att den inte kan jämföras med en verklig situation. Men ofta kan en konstlad situation säga saker med extra skärpa om verkligheten, eftersom man kan lyfta ut vissa faktorer och granska dem närmare.

<sup>75</sup> Esaiasson m. fl. (2009) s. 376

<sup>76</sup> Kursplan för naturkunskap A (NK1201), Skolverket 2000-07

betygsskalan använde vi ett yrkesinriktat och ett studieinriktat program. Anledningen var att vi ville se om det fanns en skillnad; hade vi valt två program som legat närmare varandra så hade skillnaden eventuellt varit för liten för att kunna observeras i vårt experiment. Det var viktigt att stimulus varken var så tydligt att experimentdeltagarna genomskådade experimentet eller så svagt att det riskerade att inte uppfattas. Vi valde att skriva in programtillhörigheten i den inledande text som introducerar uppgiften till lärarna (se bilaga). På så sätt blev det klart och tydligt men inte alltför uppenbart. Vi använde oss av två manipulationskontroller för att visa att stimulus verkligen uppfattats av experimentdeltagarna – en direkt och en indirekt. Den direkta kommer från svaret från en av lärarna som vi använde oss av i en testomgång av experimentet.<sup>77</sup> Denna lärare förde ett resonemang om att elevsvaren inte kändes autentiska på grund av att nivån var för hög för att vara en representativ elev från det yrkesprogrammet, vilket tydligt visar att läraren tagit till sig av stimulus. Den indirekta manipulationskontrollen ser vi i Johanssons examensarbete där även han använder programtillhörighet som stimulus i en text på samma sätt som vi.<sup>78</sup> Hans resultat visade ingen skillnad mellan de två experimentgrupperna med olika programtillhörighet, men däremot mellan de två experimentgrupperna och kontrollgruppen som fått samma uppgift men utan kommentar om elevens programtillhörighet. Detta indikerar att lärarna lade märke till manipulationen, eftersom experimentgrupp och kontrollgrupp i övrigt var lika.

Valet av vilka två program som skulle få representera yrkes- respektive studieinriktat program var inte helt självklart. Det klassiskt studieinriktade och mest teoretiska programmet ansåg vi vara naturvetenskapsprogrammet (NV), men det var svårare att välja ut det yrkesinriktade programmet. Hotell- och restaurangprogrammet (HR) valdes framför andra program som fordonsprogrammet eller omvårdnadsprogrammet på grund av dess jämnare könsfördelning, för att kunna jämföras med NV utan att det istället blev en jämförelse mellan kön.<sup>79</sup> Dessutom var det viktigt att hitta en balans i hur mycket programmen skulle skilja sig åt i betygsgenomsnitt. För att få ett ordentligt stimulus krävdes en stor skillnad, samtidigt som en alltför stor skillnad skulle göra det omöjligt att konstruera ett elevsvar som kändes autentiskt för båda programmen. HR blev då valet vi gjorde för att få en passande medelnivå på elevprestationer.<sup>80</sup>

Urvalet av lärare skedde genom att leta efter lärares mejladresser och ämnesbehörigheter på skolors hemsidor, både kommunala och fristående skolor. Vi började med att leta i de tre större städerna Stockholm, Göteborg och Malmö, och fortsatte sedan med mindre städer som Örebro, Lund, Kalmar, Umeå med flera. Det mest ideala urvalet hade varit att ta med hela populationen svenska naturkunskapslärare, eller näst bäst ett statistiskt slumpmässigt urval. Detta hade vi dock inte resurser till inom ramen av ett examensarbete, men eftersom vi skickade mejl till lärare i flera städer ser vi det som att urvalet ändå är tämligen representativt

---

<sup>77</sup> För att försäkra oss om att uppgiften fungerade som vi tänkt innan vi genomförde experimentet använde vi oss av en testomgång, där vi lämnade ut uppgiften till några NO- eller NkA-lärare. De fick göra uppgiften och sedan svara på extra frågor om hur de upplevde uppgiften. Utifrån deras synpunkter gjorde vi några små justeringar.

<sup>78</sup> Johansson, David (2009) *Påverkas lärare av elevers programtillhörighet i betygprocessen? – En experimentell studie om huruvida lärares vetskap om elevers programtillhörighet påverkar betygsättningen*. Examensarbete i Samhällsvetenskap vid Göteborgs Universitet

<sup>79</sup> 91% är killar på Fordonsprogrammet enligt Skolverkets statistik. På hotell- och restaurang är killarna 38% och på Natur är de 53%. Uppgifterna gäller för antagna till nationellt program. Källa: <http://www.Skolverket.se/sb/d/1718/a/14862>, tabell 4C och 4D.

<sup>80</sup> Enligt Skolverket var den genomsnittliga betygspoängen för totala gymnasieskolan i Sverige 16,1 för NV och 12,8 för HR. Källa: <http://www.Skolverket.se/sb/d/1721>, tabell 2A.

i detta första skede. Vi ser inte någon större anledning till att anta att lärare i dessa städer/orter skiljer sig från resten av landet och anser därför att den externa validiteten är hög.<sup>81</sup>

Nästa steg var att skicka mejl till dessa lärare med förfrågan om de ville delta i undersökningen. För att begränsa urvalet till lärare som verkligen känner till kursen och dess betygsättning sattes kriteriet att lärarna skulle vara behöriga för NkA och/ eller undervisat minst ett år på denna kurs. Anledningen till att även lärare som är obehöriga för ämnet inkluderades är att det kan finnas många lärare, kanske speciellt äldre, som undervisar i NkA utan att ha läst in en behörighet för kursen men som istället har behörighet i liknande ämne som kemi, biologi eller fysik. Syftet med urvalet var att få en god bild av de lärare som faktiskt bedömer eleverna, inte bara de som har ämnesbehörighet och därmed ”bör” vara de som bedömer. För att däremot undvika att ta med lärare som bara hoppat in på en kurs utan att ha behörighet valde vi att kräva minst ett års undervisning i NkA (såvida de inte var behöriga i ämnet). Av de lärare till vilka vi skickade en förfrågan om deltagande var det 31 % (85 lärare) som ville delta. Ett par lärare svarade att de inte passade in på våra urvalskriterier och några svarade att de inte hann. Majoriteten av lärarna svarade alltså inte, vilket kan ha reducerat den externa validiteten om dessa lärare har en gemensam nämnare som gör att de skulle ha svarat liknande på uppgiften och därmed ha ändrat utfallet av experimentet. Ett minde bortfall fick vi i andra omgången, då uppgiften skickats ut till de lärare som tidigare tackat ja till att delta. Av dessa 85 lärare var det 65 som slutligen skickade in sitt svar till oss, alltså 76 %. Två av de lärare som utgjorde bortfallet svarade att de inte kunde delta eftersom de ansåg att uppgiften var omöjlig då den gick ut på att betygsätta alltför korta elevsvar. Ytterligare två uppgav att de inte hade tid. De andra lärarna svarade inte alls, vilket innebär att vi inte vet vad bortfallet berodde på. Sannolikt hade de liknande anledningar – åsikten att uppgiften var omöjlig eller brist på tid de förut trodde att de hade. Fördelningen i de två experimentgrupperna blev ändå någorlunda lika: 31 (NV) mot 34 (HR). Tabell 1 ger en övergripande bild av lärarna som deltog. Gruppen var förhållandevis jämnt fördelad mellan könen (45% män, 55% kvinnor), medelåldern var 41 år och det var en jämn fördelning över åldersspannet 24-65 år. Samtliga lärare fick svara på frågan om vilka program de huvudsakligen undervisar på, och ombads välja bland ett av tre alternativ: *teoretiska*, *praktiska* eller *ungefär lika mycket på båda* (benämns *lika* i resten av uppsatsen). Om de undervisade inom både teoretiska och praktiska program var det alltså upp till läraren själv att uppskatta förhållandet. Resultatet visade att majoriteten av lärarna undervisade huvudsakligen på teoretiska program (65%) och att av de andra var det lika många som undervisade på praktiska program (17%) som de som undervisade lika mycket på båda (17%). De flesta (60%) hade jobbat som lärare mindre än 10 år, men lärare fanns representerade ända upp till 45 år i tjänst.

---

<sup>81</sup> Ett slumpmässigt urval av undersökningsdeltagarna krävs inte för att genomföra ett experiment, men det är ändå önskvärt efter som det ger en hög extern validitet som gör det möjligt att dra slutsatser om hela populationen utifrån endast ett representativt urval.

**Tabell 1** Beskrivning av experimentgruppens sammansättning med avseende på faktorerna kön, huvudsakligt program läraren undervisar på, ålder samt tid i tjänst.

| Kön (%)        |    | Huvudsakligt program läraren undervisar på (%) |    | Ålder           |       | Tid i tjänst    |       |
|----------------|----|--|----|-----------------|-------|-----------------|-------|
| <i>Män</i>     | 45 | <i>Praktiska</i>                               | 17 | <i>Medel</i>    | 41 år | <i>Medel</i>    | 13 år |
| <i>Kvinnor</i> | 55 | <i>Teoretiska</i>                              | 65 | <i>21-30 år</i> | 26 %  | <i>0-10 år</i>  | 60 %  |
|                |    | <i>Lika mycket på båda</i>                     | 17 | <i>31-40 år</i> | 28 %  | <i>11-20 år</i> | 12 %  |
|                |    |  |    | <i>41-50 år</i> | 20 %  | <i>21-30 år</i> | 20 %  |
|                |    |  |    | <i>50-65 år</i> | 26 %  | <i>31-45 år</i> | 8 %   |

N för alla deltagare= 65.

För att behålla en av de absolut viktigaste fördelarna med experimentet som metod är det viktigt med en väl genomförd randomisering, vilket krävs för en god intern validitet. Vi använde oss av en slumplista för att avgöra vilken lärare som skulle placeras i vilken experimentgrupp, så att varje lärare skulle ha lika chans till båda grupperna.

Det klassiska experimentet inkluderar minst en experimentgrupp och en kontrollgrupp, så att man kan jämföra effekt av stimulus som använts i experimentgruppen men inte i kontrollgruppen.<sup>82</sup> I vårt experiment är vi dock inte intresserade av skillnaden mellan att introducera stimulus eller inte, utan istället skillnaden mellan två olika varianter av samma stimulus. Kontrollgruppen skulle i detta fall varit en lärargrupp som fick samma uppgift som de andra men utan information om elevens programtillhörighet, men eftersom vi var intresserade av skillnaden mellan olika program var det inte viktigt att ha med denna grupp. När man dessutom lägger till svårigheter att få tag i deltagare beslöt vi oss för att utesluta kontrollgruppen.

## 4.2 Uppgiftens utformning

Uppgiften som vi skickade ut bestod av tre delar (se bilaga). Först en inledning där vi presenterade oss och kort beskrev syftet med experimentet. Här informerades även deltagarna om deras anonymitet och rätt att när som helst avbryta experimentet.

Den andra delen bestod av tre provfrågor med ett elevsvar till varje. Detta var alltså själva materialet som lärarna skulle ta ställning till. Överst på sidan fanns dock en kort introduktion till elevsvaren, där läraren informerades om att kursen var NkA och att eleverna nyss gått igenom växthuseffekten. Det var i denna korta text som lärarna också informerades om vilket program eleven gick. Valet av frågor och svar var givetvis avgörande för experimentets framgång. Vi behövde frågor som täcktes av kursen NkA och som var utredande av sitt slag och inte bara rena faktafrågor. Dessutom var det viktigt att elevsvaren kändes autentiska så att det konstlade i experimentsituationen minimerades. Syftet med uppgiften var inte att hitta den "rätta" bedömningen utan istället att jämföra, och därför var vår ambition att lägga elevsvaren på vad vi anser motsvarar nivån för G-VG, så att det lättare fanns utrymme åt båda hållen. De första två frågorna och elevsvaren till dessa hämtade vi från Ämnesdidaktik i praktiken

<sup>82</sup> Esaiasson m.fl. (2009) s. 371

(2003).<sup>83</sup> Den tredje frågan och svaret konstruerade vi själva med motivet att ha en fråga som var mer öppen och som tillät kvaliteter för de högre betygsnivåerna. För att uppgiften skulle kännas autentisk måste elevsvaren upplevas som äkta elevsvar av lärarna, och därför gjorde vi en testomgång där den generella responsen vi fick var att elevsvaren verkade autentiska. Dock fick vi kritiken av en lärare som fått elevsvaret med en elev som gick det yrkesinriktade programmet att frågorna och svaren var för avancerade för den genomsnittliga nivån för dessa elever. Denna kritik gjorde att vi formulerade om några meningar för att sänka nivån en aning.

I sista delen fanns några frågor för lärarna att besvara, varav den första frågan uppmanade dem att bedöma elevsvaren med ett samlat betyg på skalan IG till MVG. Eftersom dessa betygssteg används i kursplaner med kriterier som baserar sig på en hel kurs, kan man misstänka att majoriteten av lärare inte är vana vid, och inte anser att det ligger i linje med kursplanerna, att bedöma tre korta elevsvar med ett betyg. Betyg anses av många lärare som något man sätter på slutet av en termin och inte förrän då. Möjligtvis på ett större prov eller arbete, men inte några enstaka frågor.<sup>84</sup> Detta var vi fullt medvetna om och önskar att vi hade kunnat undvika, men ett längre prov skulle tagit längre tid i anspråk av de deltagande lärarna, och det hade då varit svårt att få tag i experimentdeltagare. Vi anser därför att experimentet skulle ha lidit mer än tjänat på ett sådant upplägg. Det finns dessutom ännu en anledning till att välja just skalan IG - MVG som operationell indikator. Till skillnad från till exempel en skala med 0 - 3 poäng finns det en kursplan att falla tillbaka på, som sätter ramen. Hade lärarna fått poängsätta istället skulle man inte kunnat jämföra dessa poängbedömningar eftersom varje lärare har sin fulla rätt att lägga olika kriterier för de olika poängen. Det är just betygsättningen som har krav på sig om att vara likvärdig. Vi märkte också att många av lärarna använde sig av kursplanen för att motivera sin bedömning genom att lyfta fram kriterier och peka på hur de uppfylldes eller inte uppfylldes.

Sedan följde fyra frågor som var till för randomiseringskontroller (se Analysförberedelser) och slutligen en öppen punkt med utrymme för kommentarer. Drygt hälften av lärarna valde frivilligt att ge motivering till sin betygsättning under denna punkt, några kommenterade frågorna och undersökningen medan andra valde att inte lämna några kommentarer.

### **4.3 Analysförberedelser**

För att göra statistiska analyser av våra resultat använde vi oss av t-test i statistikprogrammet SPSS, vilket gav oss möjligheten att konstatera om eventuella skillnader i medelvärde berodde på en verklig skillnad mellan grupperna. Vi började med att göra en randomiseringskontroll, det vill säga en statistisk analys med avseende på några faktorer som bör vara jämt fördelade mellan grupperna om randomiseringen fungerat.<sup>85</sup> Lärarna ombads uppge kön, ålder, tid i tjänst och på vilket program de huvudsakligen undervisar (praktiska eller teoretiska). Detta är faktorer som vi tänker oss skulle kunna påverka bedömningen,

---

<sup>83</sup> Ämnesdidaktik i praktiken är en samling workshops framtagna av Enheten för ämnesdidaktik på Institutionen för pedagogik och didaktik vid Göteborgs universitet med syfte att fungera som fortbildning för lärare som undervisar inom naturvetenskap. För nedladdning av dessa och vidare information se <http://na-serv.did.gu.se/nordlab/se/se.html>. Frågorna har hämtats från verkliga situationer och elevsvaren är äkta, dock med några små justeringar som vi gjort för att de bättre skulle passa våra syften.

<sup>84</sup> Som förutspått var det flera lärare som kommenterade att det var svårt eller rent av omöjligt att sätta betyg. Två lärare som tidigare tackat ja till att delta hoppade av då de fått se uppgiften just med motiveringen att de inte kunde betygsätta materialet. Det stora flertalet lärare betygsatte dock utan att ta upp dessa svårigheter.

<sup>85</sup> Esaiasson m. fl. (2009) s. 376



utifrån tidigare studier.<sup>86</sup> I tabell 2 nedan finns resultatet från randomiseringskontrollen presenterat. Medelvärdeskillnaden redovisas med signifikansnivå i form av p-värde och F-värde. För att randomiseringen ska ha fungerat ska p-värdet hålla sig över 5 % -signifikansnivån, vilket innebär att det är mer än 5 % chans att skillnaden i medelvärde mellan grupperna bara beror på slumpen.

**Tabell 2** Randomiseringskontroll för faktorerna kön, ålder, tid i tjänst och huvudsakligt program läraren undervisar på.

|                           | Kön        |       | Ålder      |      | Tid i tjänst |      | Huvudsakligt program läraren undervisar på |      |
|---------------------------|------------|-------|------------|------|--------------|------|--|------|
|                           | Medelvärde | S.E.  | Medelvärde | S.E. | Medelvärde   | S.E. | Medelvärde                                 | S.E. |
| <b>HR (n=34)</b>          | 0,56       | 0,086 | 39,2       | 1,90 | 10,9         | 1,80 | 1,7  | 0,10 |
| <b>NV (n=31)</b>          | 0,55       | 0,091 | 43,1       | 2,30 | 14,5         | 2,14 | 1,3  | 0,17 |
| <b>Medelvärdeskillnad</b> | 0,01       |       | 3,9        |      | 3,6          |      | 0,4**                                      |      |

HR = hotell- och restaurangprogrammet, NV = naturvetenskapliga programmet, S.E. = standardfel. 1) Kön mäts med 0 = man och 1 = kvinna.  $p = 0,87$  och  $F = 0,03$ . 2) Ålder mäts i år och är ett medelvärde för gruppen.  $p = 0,15$  och  $F = 2,1$ . 3) Tid i tjänst mäts i år och är ett medelvärde för gruppen.  $p = 0,08$  och  $F = 3,3$ . 4) Huvudsakligt program läraren undervisar på mäts med 0 = praktiska, 1 = lika på båda och 2 = teoretiska.  $p = 0,00$  och  $F = 18,1$ . \* = signifikant på 5%-nivån, \*\* = signifikant på 1%-nivån.

Tabellen visar att enligt de flesta faktorer har randomiseringen fungerat. Vi ser att p-värdet håller sig klart och tydligt över 5 % -signifikansnivån för kön, ålder och tid i tjänst, vilket indikerar att det inte finns några skillnader mellan grupperna med avseende på dessa faktorer. Kontrollen av fördelningen av huvudsakligt program läraren undervisar på visar dock en signifikans redan på 1 % -nivån. När lärarna svarade på frågan om huvudsakligt undervisningsprogram kunde de välja mellan tre alternativ – *teoretiska*, *praktiska*, eller *ungefär lika mycket på båda* – vilket innebär att skillnaden i randomisering kan finnas mellan samtliga av dessa tre grupper eller endast mellan två av dem. För att ta reda på mellan vilka grupper skevheten fanns använde vi Tukey's post hoc-test i statistiktestet ANOVA, och kunde då konstatera att det fanns en statistiskt signifikant skillnad i tilldelning av experimentgrupp mellan lärarna som valt alternativet *praktiskt* och de som valt *teoretiskt* ( $p = 0,038$ ). Det fanns däremot ingen skillnad mellan grupperna *praktiskt* och *lika* ( $p = 0,079$ ) eller mellan *lika* och *teoretiskt* ( $p = 0,966$ ). Antal lärare i grupperna *praktiskt* och *lika* är dock endast 11, vilket egentligen är för litet för att göra en ordentlig statistisk analys på.

Ett sätt att hantera en skillnad i randomiseringen är att lyfta ut en av grupperna där randomiseringen inte var tillfredställande och utföra ett t-test på själva hypotesen inom denna grupp. Eftersom antal lärare som uppgett att de framför allt undervisar på teoretiska program är övervägande med 42 personer, valde vi att testa vår hypotes inom denna grupp.

<sup>86</sup> Klapp Lekholm (2008) och Korp (2006)

## 5. Resultat

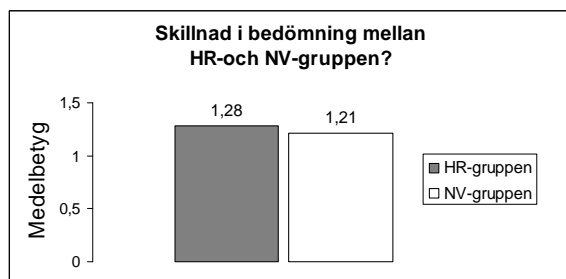
När vi analyserar experimentets utfall kan vi urskilja intressanta resultat på fyra olika områden. Det första och det andra ger svar på vår huvudfråga, nämligen huruvida det finns en skillnad i betygsmedelvärde mellan lärarna i HR-gruppen och NV-gruppen. Det tredje handlar om spridning i betygsättning, dels inom hela gruppen av experimentdeltagare och dels mellan de två experimentgrupperna. Slutligen tar det fjärde området upp skillnader i betygsättning mellan kvinnliga och manliga lärare.

I första delen ville vi svara på vår huvudsakliga fråga, alltså om det finns någon skillnad i bedömning mellan lärargruppen som fått information om att eleven går hotell- och restaurangprogrammet och lärargruppen som fått information om att eleven går naturvetarprogrammet. För att kunna jämföra bedömningen mellan grupperna omvandlade vi först de fyra betygsstegen IG, G, VG och MVG till respektive 0, 1, 2 och 3, och tog fram medelvärdet för hela experimentgruppen. Vid en jämförelse av de båda experimentgruppernas betygsmedelvärde såg vi att lärarna i HR-gruppen hade ett högre medelvärde än de i NV-gruppen, men när vi gjorde en statistisk analys hamnade värdet precis på 5 % -signifikansnivån ( $p = 0,05$ ). Tabell 3 redovisar siffrorna från analysen och figur 1 visar en mer lättöverskådlig jämförelse mellan grupperna. För att ett värde ska räknas som statistiskt signifikant ska  $p$ -värdet ligga *under* gränsvärdet (alltså under 0,05 för 5 % -signifikansnivån). Eftersom vårt värde hamnade precis *på* gränsvärdet är det så nära det kan komma utan att räknas som statistiskt signifikant, och följaktligen måste man ändå överväga att hantera skillnaden som en verklig skillnad. Signifikansnivåerna är trots allt uppskattningar beräknade på sannolikhet och inga absoluta värden. Vårt första resultat indikerar alltså att det finns en skillnad i betygsmedelvärde mellan experimentgrupperna men den är inte statistiskt signifikant.

**Tabell 3** Skillnad i bedömning mellan HR- och NV-gruppen?

|                          | Medelvärde | S.E. |
|--------------------------|------------|------|
| HR (n=34)                | 1,28       | 0,12 |
| NV (n=31)                | 1,21       | 0,10 |
| Medelvärdes-<br>skillnad | 0,070      |      |

HR = hotell- och restaurangprogrammet, NV = naturvetarprogrammet, S.E. = standardfel. För medelvärdesskillnaden var  $p = 0,050$ ,  $F = 4,0$ .  
\*signifikant på 5%-nivå, \*\*signifikant på 1%-nivå.



**Figur 1** En jämförelse av betygsmedelvärde mellan HR-gruppen och NV-gruppen visar en skillnad som inte är statistiskt signifikant ( $p=0,05$ ,  $F=4,0$ ).

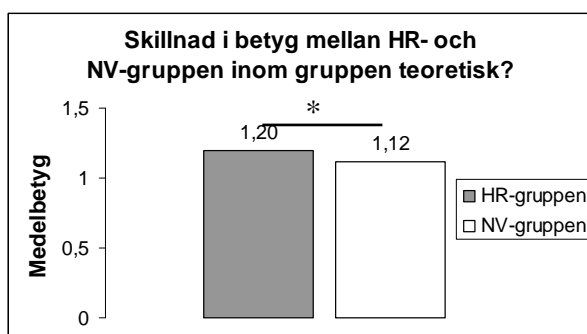
\*signifikant på 5%-nivå, \*\*signifikant på 1%-nivå.

Skillnaden mellan experimentgrupperna blev dock avsevärt större i andra delen när vi delade upp lärarna efter vilket program de huvudsakligen undervisar på. Som upptäcktes vid randomiseringskontrollen var fördelningen efter denna faktor skev mellan våra två experimentgrupper. Majoriteten av lärarna som arbetar huvudsakligen inom teoretiska program blev slumpmässigt placerade i HR-gruppen medan majoriteten av lärarna som arbetar huvudsakligen inom praktiska program blev slumpmässigt placerade i NV-gruppen. För att inte riskera att låta experimentet påverkas av denna skevhet valde vi att i denna andra del analysera skillnaden mellan HR-gruppen och NV-gruppen genom att endast jämföra de 42 lärare som undervisar huvudsakligen inom teoretiska program.<sup>87</sup> Resultatet redovisas nedan med siffror i tabell 4 och mer lättöverskådligt i figur 2, där det framgår av betygsmedelvärdena att lärarna i HR-gruppen har satt ett genomsnittligt högre betyg än de i NV-gruppen. Skillnaden är även klart statistiskt signifikant ( $p = 0,012$ ,  $F = 6,9$ ). I denna del två har vi alltså visat att då lärare som vanligtvis undervisar inom teoretiska program ombeds betygsätta ett provsvar från en elev på HR gör de det med ett genomsnittligt högre betyg än de ger en NV-elev under samma förutsättningar.

**Tabell 4** Skillnad i betyg mellan HR- och NV-gruppen inom lärargruppen teoretisk?

|                      | Medelvärde | S.E. |
|----------------------|------------|------|
| HR (n=25)            | 1,20       | 0,14 |
| NV (n=17)            | 1,12       | 0,08 |
| Medelvärdes-skillnad | 0,082*     |      |

HR = hotell- och restaurangprogrammet, NV = naturvetarprogrammet, S.E. = standardfel. För medelvärdes-skillnaden var  $p = 0,012$ ,  $F = 6,9$ .  
\*signifikant på 5%-nivå, \*\*signifikant på 1%-nivå.



**Figur 2** En jämförelse av betygsmedelvärde mellan HR-gruppen och NV-gruppen bland lärare som undervisar huvudsakligen inom teoretiska program visar en skillnad som är statistiskt signifikant ( $p=0,012$ ,  $F=6,9$ ).  
\*signifikant på 5%-nivå, \*\*signifikant på 1%-nivå.

<sup>87</sup> Vi valde gruppen av lärare som undervisar huvudsakligen på teoretiska program på grund av att den var överlägset störst (42 lärare jämfört med 11 lärare på ”praktisk” och 11 lärare på ”lika”). Eftersom denna grupp var så stor var det möjligt att genomföra statistiska analyser inom gruppen.

I analyserna ovan kan vi därmed se att våra resultat stämmer in på bilden som delar av tidigare forskning gett. Wikström har som diskuterats ovan (2.5) visat att elever på yrkesinriktat program får högre betygsavvikelse än andra då hon jämfört resultat från högskoleprovet med avgångsbetyg.<sup>88</sup> Vi kan nu styrka hennes resultat och utveckla det genom att hävda att om man tar bort all komplexitet och alla faktorer som påverkar i en verklig situation, så räcker bara vetenskapen om att eleverna går olika program för att de ska bedömas olika. Detta pekar på ett allvarligt problem i strävan efter rättvisa betyg i den svenska skolan. Skolverkets senaste rapport (2009) visar dock att det inte är någon skillnad i betygsavvikelse mellan yrkesinriktade och studieinriktade program då de jämför resultat från nationella prov med slutbetyg.<sup>89</sup>

I den tredje delen av analysen studerade vi spridningen i lärarnas bedömningar. Vi märkte tidigt att betygen som hade satts av lärarna varierade mycket inom hela gruppen av samtliga experimentdeltagare. Alla betygssteg, från IG till MVG, fanns representerade, även om majoriteten av lärarna bedömde uppgiften med G. Av 65 lärare satte 5 lärare IG, 40 lärare G, 2 lärare G/VG<sup>90</sup>, 16 lärare VG och 2 lärare MVG. Detta bekräftar vad som redan observerats i de tre examensarbetena av Anna-Maria Fjellman<sup>91</sup>, Henrik Gullberg<sup>92</sup> och David Johansson<sup>93</sup> som observerat påverkan av kön, etnicitet respektive programtillhörighet på lärares bedömning. Samtliga tre experiment visade liksom vårt en stor spridning i bedömning, vilket återigen är resultat som pekar på problem när det gäller rättvisa betyg.

Utifrån den stora spridningen som fanns i bedömning inom hela gruppen frågade vi oss sen om det fanns någon skillnad i spridning mellan de två experimentgrupperna. Var det så att lärarna i den ena gruppen var mer benägna att sätta ungefär samma betyg medan lärarna i den andra gruppen spretade mer i sina bedömningar? Så var mycket riktigt fallet. I tabell 3 kan man utläsa att spridningsmättet standardfel (S.E. i tabellen) är högre för HR-gruppen (0,12) än för NV-gruppen (0,10). Vi har inte kunnat beräkna den statistiska signifikansen för denna skillnad, men trenden visar på att lärarna som bedömde HR-eleven var mer oense i sin bedömning än lärarna som bedömde NV-eleven. I tabell 4, som till skillnad från tabell 3 endast inkluderar lärare med undervisning huvudsakligen på teoretiska program, visar standardfelet en ännu större skillnad i samma riktning. HR-gruppen skiljer sig med det höga värdet 0,14 från NV-gruppen med det lägre värdet 0,08.

Spridningsmättet ger en god uppfattning om den totala spridningen, men för att kunna studera effekten på en mer detaljerad nivå delade vi upp de två experimentgrupperna enligt vilket betyg de satt vilket redovisas i figur 3. Även här visas endast lärare som undervisar huvudsakligen på teoretiska program. Som man kan se i figuren så har den övervägande delen av lärarna inom båda grupperna satt betyget G, men HR-gruppen har sedan i större utsträckning använt även de andra betygsstegen.

---

<sup>88</sup> Wikström (2005) s. 35

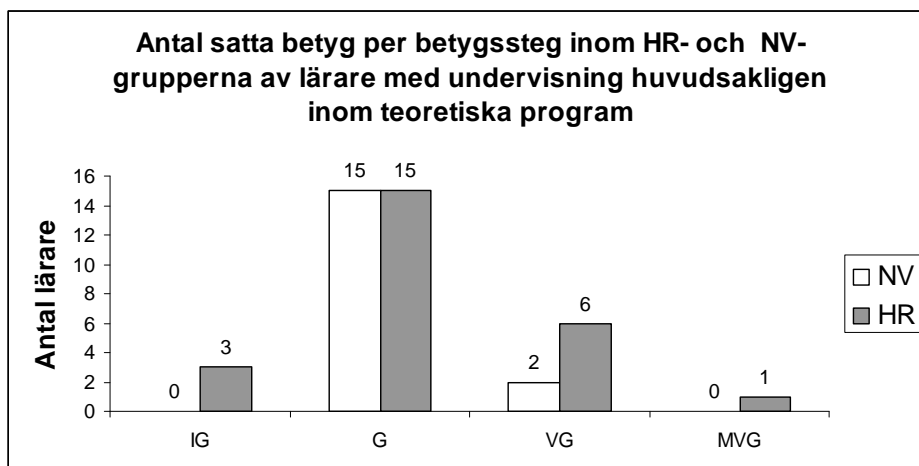
<sup>89</sup> Skolverket (2009) s 8.

<sup>90</sup> Detta kombinerade betygssteg fanns inte som alternativ, men två lärare skrev in det själva. Detta kan alltså inte tas som en verklig bild av gruppens bedömning; hade G/VG funnits som alternativ hade troligtvis fler lärare valt detta alternativ. För att inte snedvrider resultatet har G/VG hanterats som 1,5 i statistiken (G=1, VG=2).

<sup>91</sup> Fjellman (2008)

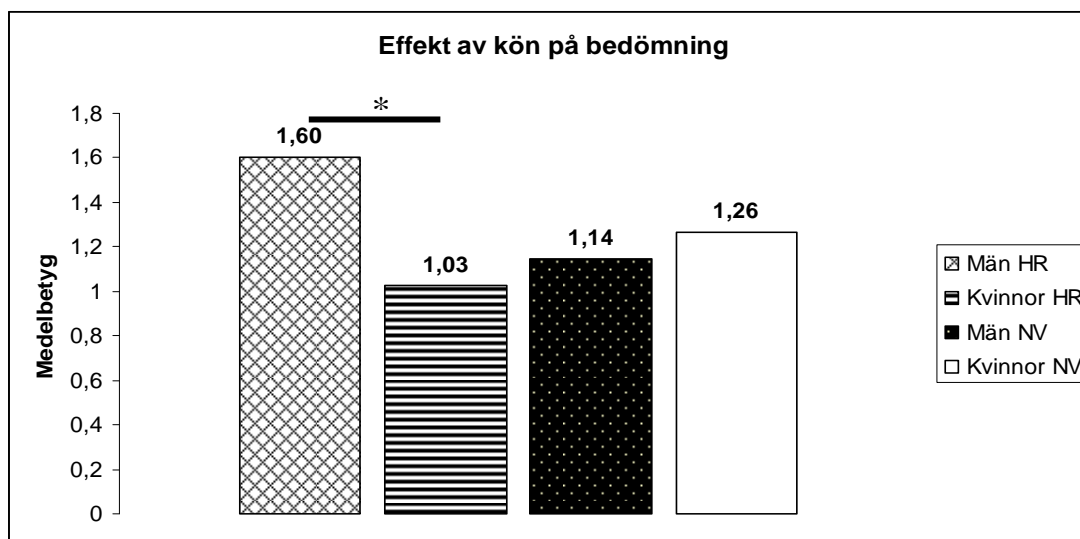
<sup>92</sup> Gullberg (2008)

<sup>93</sup> Johansson (2009)



**Figur 3** Visar antal lärare som har satt respektive betyg på uppgiften, med uppdelning i HR-gruppen och NV-gruppen. Majoriteten av de båda grupperna har satt G, men HR-gruppen har i större utsträckning än NV-gruppen använt även de andra betygsstegen. Endast de lärare som huvudsakligen undervisar inom teoretiska program har inkluderats i figuren. N för hela gruppen = 42, n för NV = 17, n för HR = 25.

I den fjärde och sista delen av resultatet vänder vi oss till en av randomiseringsfrågorna, nämligen den om kön. Figur 4 redovisar den skillnad man kan se i betygsättning mellan kvinnliga och manliga lärare. Eftersom det visat sig finnas en skillnad i betygsättning mellan HR-gruppen och NV-gruppen, jämförs kvinnliga och manliga lärare endast *inom* dessa två experimentgrupper. Inom HR-gruppen finns det en statistiskt signifikant skillnad mellan mäns och kvinnors bedömning ( $p = 0,04$ ), där männen har gett eleven högre betyg än kvinnorna. Det finns ingen statistiskt signifikant skillnad i bedömning mellan könen inom NV-gruppen.



**Figur 4** Medelvärde på satta betyg, där IG=0, G=1, VG=2, MVG=3 för grupperna män som fått HR-uppgiften (n=15), kvinnor som fått HR-uppgiften (n=19), män som fått NV-uppgiften (n=14) och kvinnor som fått NV-uppgiften (n=17). \* signifikant på 5%-nivån ( $p = 0,040$ ).

Eftersom vi inte fokuserat på skillnader i betygsättning mellan manliga och kvinnliga lärare i detta examensarbete får detta resultat stå relativt okommenterat, men det är ytterligare en aspekt som kan påverka det jämförbara betyget och skulle kunna undersökas vidare med andra metoder.

## 6. Avslutande diskussion

Utgångspunkten för vår studie har varit det system som omger bedömningsituationen genom ett stort handlingsutrymme för läraren som gatunivåbyråkrat i det lokala frirummet och olika tolkningar av vad likvärdig utbildning innebär. Med detta som grund ombads lärare i naturkunskap att betygsätta ett och samma elevsvar men utifrån olika bakgrundsinformation om elevens programtillhörighet. Tidigare forskning hade särskilt pekat ut denna faktor och visat på problem med orättvis betygsättning. Vi valde då att använda oss av ett experiment för att undersöka om lärarens vetskap om elevens programtillhörighet påverkar lärarens bedömning. I följande avsnitt avser vi att sammanfatta våra resultat och diskutera vad de betyder i ett vidare sammanhang.

Genom våra analyser av svaren från lärarna kunde vi påvisa effekter ur fyra olika aspekter. För det första såg vi att den grupp som hade fått information om att eleven gick hotell- och restaurangprogrammet (HR) fick ett lite, om än inte statistiskt signifikant, högre betygsmedelvärde än den grupp som hade fått information om att eleven gick naturvetenskapligt program (NV). För det andra ökade denna skillnad kraftigt och blev statistiskt signifikant då vi åter jämförde experimentgrupperna men endast inkluderade de lärare som undervisade huvudsakligen inom teoretiska program. För det tredje var spridningen i bedömning stor mellan samtliga lärare och trenden visade att spridningen var större inom HR-gruppen än inom NV-gruppen. För det fjärde och sista så upptäcktes det att betygsmedelvärdet som var satt av manliga lärare inom HR-gruppen var högre än det som var satt av kvinnliga lärare inom HR-gruppen. Sammantaget visar våra resultat att lärarna bedömer samma prestation med olika betyg vilket utgör ett hot för rättvisan och likvärdigheten i skolan. Dessa slutsatser stöds av tidigare forskning, bland andra av Wikström (2005) och Skolverkets rapport från i år (2009). Wikström ser i sin studie en större avvikelse mellan kursbetyg och högskolebetyg för elever på yrkesinriktade program än de på studieinriktade program vilket ligger helt i linje med våra resultat. Skolverket visar inte på samma trend mellan yrkes- och studieinriktade program, men ser en sammantagen bild av att betygavvikelsen skiftar beroende på lärare och drar efter det slutsatsen att bedömningen i gymnasieskolan inte är likvärdig.

De tendenser till olika bedömning som vårt experiment visar kan förklaras på olika sätt. En stor del av förklaringen vill vi hävda ligger i svårigheten att upprätthålla rättvisa betyg i ett system som syftar till att i första hand skapa utrymme för anpassning på den lokala nivån. Läraren befinner sig i ett mycket komplext sammanhang av att tolka betygsriterier och styrsystem i förhållande till de olika förutsättningar för undervisning som råder på yrkesinriktade och studieinriktade program. Även om varje lärare strävar efter att bedöma varje elevs individuella kunskaper mot betygsriterierna kan det vara svårt i denna situation att inte jämföra med den referensram som erfarenheten ger. Detta kan leda till att kategorierna som skapas styr de förväntningar som läraren har på elever med olika bakgrund.

Något som särskilt fångade vårt intresse i tredje delen av resultaten var att spridningen i betyg inom HR-gruppen verkade öka då vi endast inkluderade lärare som undervisar inom i huvudsak teoretiska program. Det skulle kunna tyda på att spridningen är större inom HR-gruppen på grund av att detta är ett program som läraren vanligtvis inte kommer i så stor kontakt med, och osäkerhet uppstår. Det innebär alltså att om vi hade låtit en övervägande del av lärare som undervisar huvudsakligen på praktiska program bedöma de två grupperna så skulle hypotetiskt sett resultatet kunna ha blivit det omvända – större spridning i bedömning inom NV-gruppen än inom HR-gruppen. Detta är tyvärr endast spekulationer eftersom gruppen av lärare som undervisade huvudsakligen inom praktiska program var för liten för att kunna analyseras, men som framtida experiment är det intressant. Det skulle i så fall kunna innebära att betygsättning inom ett visst programområde är något man inskolars i som lärare vilket gör att man är osäker i början vid mötet med det ”nya” programmet men sedan likriktar sin bedömning i linje med övriga lärare som undervisar inom samma inriktning.

Eftersom experiment går ut på att ta ut en liten del av hela verkligheten för att kunna undersöka den noggrant, blir följden att man måste vara försiktig med att låta delen säga för mycket generellt om helheten. Vi har klart och tydligt visat att lärare i naturkunskap bedömer samma elevsvar på mycket olika sätt, delvis beroende på vilket program eleven sägs gå på. Men detta innebär dock *inte* i förlängningen att vi har visat att dessa lärare skulle ha bedömt en och samma elev som deltagit i deras kurs med olika betyg. Anledningen till att man inte kan dra denna slutsats är delvis brister i vårt experiment, framför allt att lärare ombeds sätta betyg på tre korta elevsvar. Detta är en mycket svår uppgift eftersom underlaget är för litet, och därför är det inte konstigt att det råder viss osäkerhet inom lärarkåren med vilket betyg dessa svar bör bedömas. Många lärare har kommenterat i sina svar att det var svårt att bedöma med så lite underlag. En annan anledning till att denna slutsats inte kan dras är att olika lärare har olika sätt att bedöma och följaktligen kanske inte denna metod passar alla. Lärarens uppdrag är att skaffa sig en bild av elevens kunskapsnivå och bedöma den utifrån kriterier och mål. *Hur* läraren skaffar sig denna bild finns det dock inga regler om, vilket gör att läraren står tämligen fri att anamma den metod som läraren anser fungerar bäst. För en del lärare inkluderar det provfrågor liknande våra som eleven ska svara på, och dessa lärare är alltså väl förberedda för att svara på vår uppgift. Andra lärare jobbar inte med denna sorts prov och för dem kan det då bli svårare att svara på vår uppgift. Det gör att resultatet kanske skulle blivit annorlunda om lärarna fått använda sin egen metod.

Att olika lärare har olika metoder för att bedöma elevens kunskaper kan även kopplas till diskussionen om vad likvärdig utbildning och likvärdiga betyg egentligen innebär. När undervisningen anpassas efter olika elevers förutsättningar påverkas även möjligheterna att skapa former för bedömning som är jämförbara.

Efter att ha genomfört ett experiment med intressanta resultat vill vi slå ett slag för experimentets plats inom forskningsområdet betyg och bedömning. De olika metoder som har använts i tidigare forskning har kunnat svara på frågor som handlar om statistiska samband på aggregerad nivå mellan olika faktorer och avvikelser från det rättvisa betyget samt gå på djupet med lärares uppfattningar om betygssystemet genom intervjuer. Med dessa metoder är det dock svårt att visa på ett kausalt samband. Eftersom området är så komplext kan det behöva reduceras för att man ska kunna uttala sig om orsak och verkan och peka ut enskilda konkreta problemområden för vidare forskning. Det räcker inte att enbart konstatera att betygsättningen i Sveriges skolor är orättvis. Enskilda faktorer och orsakssamband behöver undersökas ytterligare för att en framåtsyftande diskussion skall kunna föras som kan leda till att vi kommer tillrätta med problemet. Användandet av experiment inom

utbildningsvetenskapen lyser med sin frånvaro och vi vill med vår uppsats visa på de möjligheter som finns att skapa kunskap på detta sätt. Om flera experiment genomförs med olika design och fokus på olika faktorer skulle en mer heltäckande bild kunna skapas som på ett bra sätt kompletterar de metoder som tidigare använts.

En viktig del i bekämpandet av orättvisa inom betygsättning är något som en del lärare och skolor redan bedriver, nämligen bedömningsdiskussioner mellan lärare. Handlingsutrymmet i det lokala frirummet skulle kunna hanteras med en professionell diskussion lärare emellan. När de centralt fastställda betygskriterierna skall översättas till den lokala nivån minskar risken för att alltför olika tolkningar uppstår om läraren har möjlighet att jämföra sin tolkning med andras. I Skolverkets rapport som diskuterats tidigare, trycker Skolverket på vikten av en dialog om betygsättning mellan lärare både på skolnivå och på nationell nivå.<sup>94</sup> De menar att det tolkningsutrymme som alltid kommer att finnas kring betygskriterierna gör det nödvändigt för lärare att utveckla gemensamma synsätt på betygsättning, och de betonar att alla parter är ansvariga för att skapa en fungerande diskussion på nationell nivå. Många av lärarna som deltog i vår undersökning visade att även de delade Skolverkets åsikt genom att uttrycka önskan om att ha mer diskussioner med sina kollegor. En del gav positiva exempel på att de redan har bedömningsdiskussioner på skolan, i vissa fall även med andra skolor. Våra resultat visar på ett särskilt behov av att skapa forum för diskussioner om bedömningspraxis mellan lärare på yrkesinriktade och studieinriktade program.

Inget system för bedömning av människors kunskaper kan utformas så att absolut rättvisa uppnås. Betygens dubbla funktion av att fungera som en återkoppling till eleven för att stimulera till vidare lärande och samtidigt vara ett urvalsinstrument för vidare studier kan i vissa fall hamna i konflikt med varandra. Våra resultat visar med tydlighet att det finns ett grundläggande problem när det gäller jämförbarhet och därmed i förlängningen med rättvis betygsättning i det nuvarande betyg- och styrsystemet. Utmaningen för den svenska skolan i framtiden blir att försöka förena en vilja till utvecklande av betygens kunskapsfrämjande funktion och samtidigt skapa förutsättningar för urval som kan uppfattas som legitima.

---

<sup>94</sup> Skolverket 2009 s 11.



## 7. Referenser

Andersson, Håkan (1999) *Varför betyg? Historiskt och aktuellt om betyg*. Lund: Studentlitteratur e-bok

Broady, Donald & Mikael Börjesson (2006) "En social karta över gymnasieskolan" ur *Ord & Bild 3 – 4 2006 Tema - Skolan*

Carlgrén, Ingrid (2002) "Det nya betygssystemets tankefigurer och tänkbara användningar" ur antologin: *Att bedöma eller döma? Tio artiklar om bedömning och betygsättning*. Skolverket

Dagens Nyheter (DN) 091130 "Gymnasieelevers betyg inte rättvisa"

Dagens Nyheter (DN) 090817 "Konkurrens om elever ger orättvisa betyg"

Englund, Tomas & Ann Quennerstedt (red.) (2008) *vadå likvärdighet – studier i utbildningspolitisk språkanvändning*. Göteborg: Daidalos

Esaiasson, Peter & Mikael Gilljam & Henrik Oscarsson & Lena Wängnerud (2009) *Metodpraktikan – Konsten att studera samhälle, individ och marknad*. Stockholm: Norstedts Juridik AB

Fjellman, Anna-Maria (2008) *Duktiga flickor och tuffa pojkar – en experimentell studie om könsskillnader i betygsättning*. Examensarbete inom lärarutbildningen i Göteborg. Göteborgs Universitet

Gipps, Caroline & Patricia Murphy (1994) *A fair test? – Assessment, achievement and equity*. Philadelphia: Open University Press

Gullberg, Henrik (2008) *Olika etnicitet, lika rätt? – En experimentell studie om etnicitetens påverkan på betyg och bedömning*. Examensarbete i Samhällskunskap Göteborgs Universitet

Flaa, Paul m.fl. (1998) *Introduktion till organisationsteori*. Lund: Studentlitteratur

Fredriksson, Anders (2007) "Läraryrket och den politiska styrningen av skolan" i Pierre, Jon (2007) (red.) *Skolan som politisk organisation*. Malmö: Gleerups Utbildning AB

Johansson, David (2009) *Påverkas lärare av elevers programtillhörighet i betygprocessen? – En experimentell studie om huruvida lärares vetskap om elevers programtillhörighet påverkar betygsättningen*. Examensarbete i Samhällsvetenskap vid Göteborgs Universitet

Klapp Lekholm, Alli (2008) *Grades and grade assignment - Effects of student and school characteristics*. Göteborg studies in educational science, Vol. 269. Göteborg: Acta Universitatis Gothoburgensis.

Korp, Helena (2003) *Kunskapsbedömning – hur, vad och varför* Stockholm: Myndigheten för skolutveckling.

Korp, Helena (2006) *Lika chanser i gymnasiet? En studie om betyg, nationella prov och social reproduktion*.

Kursplan för naturkunskap A (NK1201), Skolverket 2000-07

Kursplan för samhällskunskap A (SH1201), Skolverket 2000-07

Kymlicha, Will (1995) *Modern politisk filosofi*. Nora: Bokförlaget Nya Doxa

Lindahl, Erica (2007) *Systematiska avvikelser mellan slutbetyg och provresultat – Spelar elevens kön och etniska bakgrund roll?* Institutet För Arbetsmarknadspolitisk Utvärdering Rapport 2007:22

Lindberg, Viveca & Lars, Lindström (Red.) (2008) *Pedagogisk bedömning – Om att dokumentera, bedöma och utveckla kunskap*

Linde, Göran (2003) *Kunskap och betyg*. Lund: Studentlitteratur

Lindholm, Viveca (2005) "Svensk forskning om bedömning och betyg 1990-2005" *Studies in Educational Policy and Educational Philosophy* E-tidskrift 2005:1

Lipsky, Michael (1980) *Street-Level Bureaucracy – Dilemmas of the individual in public services*. New York: Russell Sage Foundation

Lundquist, Lennart (2006) "Läraren som demokratins väktare" ur *Ord & Bild 3 – 4 2006 Tema – Skolan*

Läroplan för de frivilliga skolformerna Lpf 94

SFS 1985:1100

SFS 1992:394

SKOLFS 2004:23

Rombach, Björn (1991) *Det går inte att styra med mål*. Lund: Studentlitteratur

Selghed, Bengt (2008) *Ännu icke godkänt – Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Malmö Studies in Educational Studies No 15 Malmö Högskola

Skolverket (2009) *Likvärdig betygsättning i gymnasieskolan – En analys av sambandet mellan nationella prov och kursbetyg*. Rapport 338

Tholin, Jörgen (2006) *Att klara sig i okänd natur – En studie av betyg och betygsriterier – historiska betingelser och implementeringen av ett nytt system*. Avhandling. Skrifter från högskolan i Borås

Wedman, Ingemar (1983) *Den eviga betygsfrågan: Historiskt och aktuellt om betygsättning i skolan*. Stockholm: Liber Utbildningsförlaget.

Wikström, Christina (2005) *Criterion-Referenced Measurement for Educational Evaluation and Selection*. Avhandling. Department of Educational Measurement Umeå University No 1

## **7.1 Elektroniska källor**

<http://www.skolverket.se/sb/d/1597> 091217

<http://www.skolverket.se/sb/d/1598> 091217

<http://www.skolverket.se/sb/d/1718/a/14862> 091220

<http://www.skolverket.se/sb/d/1721> 091220

<http://na-serv.did.gu.se/nordlab/se/se> 091217

## 8. Bilaga: uppgift till lärarna

*Hej!*

Vi, Robert Andersson och Jennie Gunnarsson, är lärarstudenter vid Göteborgs universitet och ska nu som sista steg i utbildningen skriva vårt examensarbete. Vi vore därför mycket tacksamma om Du i egenskap av naturkunskapslärare ville avsätta ungefär tio minuter för att delta i vår undersökning genom att svara på några frågor.

Vi är intresserade av bedömningssituationer och anser att det på grund av dess centrala plats inom skolan är ett viktigt forskningsområde. För att undersökningen inte ska påverkas negativt kan vi inte ge mer information än så just nu, men efter Ditt deltagande får Du närmare information om undersökningens syfte. Svaren är anonyma, deltagandet är naturligtvis frivilligt och Du kan när som helst avbryta Din medverkan.

För att experimentet ska fungera krävs det att Du genomför uppgiften *utan att diskutera den* med Dina kollegor. Tack för att Du respekterar detta!

Om du har frågor eller synpunkter är Du välkommen att höra av Dig till oss.

*Tack för Din medverkan!*

Jennie Gunnarsson  
gusgunjen@student.gu.se  
0731-835053

Robert Andersson  
roband@home.se  
076-84 84 396

## **Tänk dig att:**

Du undervisar en klass i Naturkunskap A. Klassen går första året på Naturvetenskapsprogrammet [alt Hotell- och restaurangprogrammet] och ni håller nu på med miljöfrågor. I slutet av avsnittet ger Du ut ett prov för att bedöma elevernas kunskaper. Följande tre frågor är ett utdrag ur provet, och Din uppgift är att läsa igenom ett elevsvar och sedan bedöma det på skalan IG - MVG. Vi förstår att det kan kännas svårt att bedöma några korta svar med betyg, men ber Dig ändå göra ditt bästa. Vi ber Dig också svara på några korta frågor på nästa sida.

### **Fråga 1**

*Förklara vad växthuseffekten är.*

Växthuseffekten är den effekt som blir på grund av att vi människor släpper ut för mycket t.ex. koldioxid. Gaserna som vi använder på jorden stiger upp till atmosfären där de bildar en hinna. Denna hinna gör att solstrålarna inte åker ut genom atmosfären efter det att de har reflekterats mot jorden. De studsar åter mot jorden och därför blir det duktigt varmt här. Genomsnittstempen höjs varje år.

### **Fråga 2**

*Vad blir konsekvenserna för hur vi lever om vi ska sänka utsläppen av koldioxid i Sverige?*

Ja levnadsstandarden måste sänkas betydligt. Vi måste satsa betydligt mer på miljövänliga artklar, bilar mm. Vi måste även hitta miljövänliga energikällor, använda oss av vatten, vind och solen. Fabrikers utsläpp måste också minskas betydligt.

### **Fråga 3**

*En del människor säger att växthuseffekten bara är en bluff. Hur skulle du svara på det påståendet?*

Nej vi vet att växthuseffekten finns för annars skulle vi frysa ihjäl. Man kan veta att vi människor släpper ut massa koldioxid i trafiken och industrier och när vi eldar. Sen går det upp i atmosfären och lägger sig som en hinna som stoppar värmen från att försvinna ut i rymden. De som säger att det är en bluff gör bara det för att de vill ha industrier och sånt, det är säkert mest dem som äger industrierna och som vill köra bil till jobbet. I filmen (den med al gore) kan man se att isbjörnarna dör och husen rasar så det händer faktiskt på riktigt. Om man själv springer för att springa från en översvämning vet man att det inte är en bluff. Så skulle jag svara.

**1. Vilket betyg ger Du svaren på uppgifterna (väg samman de tre svaren)?**

IG      G      VG      MVG

**2. Din könstillhörighet?**

MAN      KVINNA

**3. Ålder?**

\_\_\_\_\_ år

**4. Antal år som verksam lärare?**

\_\_\_\_\_ år

**5. Inom vilka program undervisar du mest?**

teoretiska                      praktiska                      ungefär lika mycket på båda

**6. Eventuella kommentarer**

*Tack för Din medverkan!*

*Robert och Jennie*