



Universitetspedagogiskt centrum

Målrelaterade betyg

Att arbeta med betygskriterier och bedömning i sju grader

Stefan Ekecrantz

UPC-rapport 2007:1

[Texten i sin helhet sidbryts för en dubbelsidig kopiering.]

Innehåll

Betygsättning med målrelaterade betyg	5
Förväntade studieresultat – examination – betygsbedömning	5
Målrelaterade betyg – vad är det, och vad är det inte?	5
Pedagogiska värden.....	6
Andra syften	6
Relationen mellan betygsriterier och förväntade studieresultat (lärandemål)	7
Problem eller utmaningar... ..	8
Betygsdebatten	8
Forskningen om mål- och kunskapsrelaterade betyg i grundskola och gymnasium.....	9
Vilka lärdomar kan vi dra?	9
Intuitiva betygsbedömningar – konsekvens och reliabilitet	10
Kan man inte jämföra studentarbeten i ett målrelaterat system?	10
Samma kriterier för allt och alla?	11
Graderade betyg, betygshets och studenters lärande	11
Är kejsaren naken? Betygsriterier som dimridåer och alibi.....	12
Att formulera betygsriterier	13
Betygsriterier = generella betygsmotiveringar	13
Den praktiska utformningen av betygsriterier	13
Kan betygsriterier uttryckas som poäng eller procent?	14
Två sätt att utforma kriterier: globalt eller per bedömningsområde.....	14
Hur detaljerat ska kriterietexterna formuleras?	15
Hur formulera? Kriterietexter om produkt, prestation eller lärande?.....	15
Två typer av kvalitetsskillnader – kontinuerliga och diskreta.....	16
Mer eller mindre tydliga gränsdragningar	17
Hur kategorisera i sju kvaliteter?	17
Hur kan man se om betygsriterierna är bra?	18
... och ett varnande finger.....	18
Metoder för att väga samman till ett betyg	19
Att välja bedömningsområden och kvaliteter	19
Flera lärandemål och bedömningsområden men bara ett betyg.....	19
Vad ska avgöra betyget?	20
Alternativ 1 – att låta det viktigaste avgöra.....	21
Alternativ 2 – att göra grövre graderingar i flera steg	21
Alternativ 3 – att låta lägsta eller högsta delprestationen avgöra	22
Alternativ 4 – att göra en helhetsbedömning.....	23

Lärandemål, bedömningsområden, kriterier och sammanräkningsmetod	24
Detaljerad feedback och betyg – en exkurs.....	26
Betyg som feedback? Två problemområden.....	26
Metoder för att göra återkopplingen mer relevant i tiden	26
Metoder för att ge en mer detaljerad återkoppling.....	27
Noter.....	29
Referenser	32

Bilaga A – Exempelsamling

Bilaga B – Checklista för förväntade studieresultat (lärandemål)

Bilaga C – Checklista för betygskriterier

Betygsättning med målrelaterade betyg

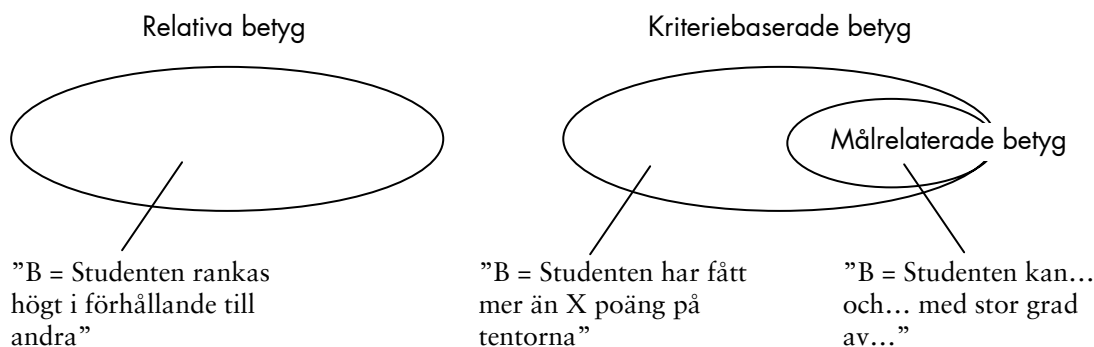
Den här framställningen är avsedd som en handledning till lärare och institutioner i deras arbete med att formulera skriftliga, målrelaterade betygskriterier i sju grader. Tanken är att olika delar ska kunna läsas separat och användas i olika moment av det arbetet. De två inledande avsnitten är av mer allmän karaktär och kan till exempel fungera som bakgrund till en diskussion vid en institution. De tre avslutande avsnitten är mer tekniska och torde främst vara relevanta för dem som håller på med det praktiska arbetet att formulera betygskriterier. I en bilaga presenteras en exempelsamling med förväntade studieresultat (lärandemål) och betygskriterier från olika ämnen.¹ Denna följs av två korta checklistor, en för förväntade studieresultat och en för betygskriterier. Texten i sin helhet kan ses som ett resultat av många och långa diskussioner med lärare vid Stockholms universitet under de senaste åren om hur betygskriterier kan eller bör utformas. De vanligast förekommande frågorna i de sammanhangen har tjänat som underlag för de olika underavsnitten.

Förväntade studieresultat – examination – betygsbedömning

Att sätta betyg är bara en liten del av en sammansatt process och bör inte behandlas som en isolerad företeelse. Innan det är meningsfullt att planera och utveckla metoder för betygsättning, formulera betygskriterier, etcetera, behöver ett antal andra saker ha avgjorts. I ett mål- och kunskapsrelaterat betygssystem anger de förväntade studieresultaten *vad* alla godkända studenter kan och betygskriterierna olika av grader *hur väl* studenterna kan detta. Att formulera betygskriterier före de förväntade studieresultaten är därför direkt olämpligt. Dessutom är betygsbedömning så intimt förknippad med examination att arbetsformerna för det ena inte kan utvecklas skilt från det andra. I praktiken är alltså betygsbedömning inte någonting enskilt, men det kan ändå vara bra att belysa ett antal frågor som har med betygsättning att göra – så länge bilden av detta som en länk eller sista led i någonting sammansatt upprätthålls.

Målrelaterade betyg – vad är det, och vad är det inte?

Figur 1. Relativa, kriteriebaserade och målrelaterade betyg



Med *relativa* betyg avses ett system där den ledande principen handlar om att rangordna studenterna inbördes, och ofta dessutom att betygen fördelas efter på förhand bestämda proportioner. (ECTS-betyg är i sin ursprungsversion ett sådant system, där 10 % av de godkända studenterna ska ha högsta betyget A, och de 10 % som har rankats lägst bland de godkända ska ha E, dvs. lägsta godkända betyg.² Övriga betyg fördelas efter andra procentsatser.) I Sverige och många andra länder tar man numera vanligtvis avstånd från relativa betyg, inte minst av pedagogiska skäl.³ En fördel med relativa betyg som ibland framförs är att betygsinflation per definition omöjliggörs eftersom det alltid är samma andel av en viss population som får ett visst betyg. Metodens motståndare hävdar å sin sida att den aspekten inte väger upp negativa effekter på undervisning och lärande.

I stället förespråkas ofta det som på engelska kallas "criterion referenced grading", vilket skulle kunna

översättas till *kriteriebaserade* betyg. (Relativa betyg kallas "norm referenced grading".) *Målrelaterade* betyg är en av flera möjliga varianter av ett sådant system. Man kan säga att målrelaterade betyg utgör en delmängd av kriteriebaserade betyg, och begreppen är därför inte synonyma. Kriteriebaserade betyg är i princip allt som inte är relativa betyg.⁴ Om det är möjligt för en student att få ett visst betyg helt oberoende av vad andra studenter presterar, så används ett kriteriebaserat betygssystem. Så också om det är hypotetiskt möjligt att alla studenter i en population får samma betyg. Med målrelaterade betyg räcker inte det. Då förutsätts därutöver att dessa kriterier representerar på förhand uppsatta kvalitativa mål för studenternas lärande. Med målrelaterade betyg kan man med ledning av betyget säga vad och hur väl en student med ett givet betyg kunde något vid kursens eller delkursens slut.⁵ Med kriteriebaserade betyg som är icke-målrelaterade är det inte möjligt på samma sätt. Forskningen om mål- och kunskapsrelaterade betyg i svensk grundskola och gymnasium visar också att många problem kan förklaras av att betygssättande lärare inte tillräckligt väl förmår skilja på dessa två alternativ.⁶

I svensk högre utbildning har vi aldrig behövt arbeta med relativa betyg, men å andra sidan har det inte heller funnits något tydligt regelverk som har fastslagit vad betygen ska representera. I högskoleförordningen anges bara vilka betygsgrader som ska användas (om inte det enskilda lärosätet fattar beslut om annat). Ibland har det lett till att man har arbetat med ett odefinierat kriteriebaserat system, där i princip varje lärare har att försöka uttolka vad betyg egentligen är för någonting. Dessa kriteriebaserade system var ibland relaterade till någon sorts kunskapsmål, ibland inte. Vid Stockholms universitet är det numera fastslaget att betygen ska vara just målrelaterade, och uttryckligen att dessa mål ska utgöras av de förväntade studieresultaten för varje delkurs eller kurs.⁷

Pedagogiska värden

Ur ett pedagogiskt perspektiv har målrelaterade betyg, oavsett antal grader, stora förtjänster jämfört med andra betygssystem. Att kunna uttala sig om hur väl en student kan någonting har betydelse för i princip alla som har anledning att ta del av betyget. Studenten själv kan använda betyget och dess motivering som en preciserad, kvalitativ återkoppling. Sådan kvalitativ feedback har visat sig vara mycket gynnsam för det fortsatta lärandet. Att få ett betyg som i princip enbart anger var någonstans studenten befinner sig på en allmän skala från bra till dåligt har en betydligt sämre pedagogisk potential.

Även läraren och institutionen kan ha stor nytta av kvalitativa, målrelaterade betyg i sin fortlöpande kursutveckling. Vad gick bra? Vad kan förbättras till nästa gång? Studenternas synpunkter i kursvärderingar brukar vara en viktig information, men det räcker inte. Ibland kan kursvärderingar och läranderesultat peka i samma riktning, ibland inte. Studenternas omdömen om en delkurs kanske förbättras utan att kunskaperna förbättrades i motsvarande grad, eller vice versa. Och, om det inte finns någon annan konkret, lätthanterlig data än kursvärderingsresultaten att tillgå riskerar frågan vad studenterna faktiskt lärde sig att hamna i skymundan. Målrelaterade betygsresultat kan då fungera som en lättillgänglig och kvantifierbar information om studentgruppens kvalitativa resultat.

Andra syften

När en examinator sätter ett betyg är det myndighetsutövning gentemot enskild person enligt förvaltningslagen. Sådana beslut måste hanteras rättssäkert och så konsekvent som möjligt. Betygskriterier kan i det sammanhanget fungera som på förhand kända grunder för hur betygsbedömningen kommer att ske.

För utomstående, som till exempel arbetsgivare eller andra utbildningar/lärosäten, kan det uppenbart vara värdefullt att via betyg och betygskriterier få veta examinatorns bedömning av hur väl den individuella studenten kan någonting. Särskilt viktigt är detta om dessa utomstående har andra intressen än att bara identifiera de enstaka individer som på något sätt har gjort bäst ifrån sig. (Sådana behov kan tillgodoses med nästan vilket betygssystem som helst.) Om det istället handlar om att en viss nivå på vissa specifika kunskaper behövs för en arbetsuppgift eller en utbildning, så kan kvalitativa, målrelaterade betyg och betygskriterier fungera som beslutsunderlag.

Relationen mellan betygsriterier och förväntade studieresultat (lärandemål)

Förväntade studieresultat anger vad alla studenter med godkänt betyg (dvs. betygen A–E) kan efter genomförd utbildning. Betygskriterierna kan med fördel skrivas så att de fungerar som preciseringar av dessa förväntade studieresultat. Därmed kan de förväntade studieresultatet skrivas i ganska allmänna termer, och överlåta konkretionen till betygsriterierna. Samtidigt behöver de förväntade studieresultatet vara så otvetydiga som möjligt eftersom de fungerar som övergripande styrmedel. Kvalitetskraven på de förväntade studieresultatets kvalitet måste ställas högt, och de är också på många sätt viktigare än betygsriterierna.

Figur 2. Relationen betygsriterier–förväntade studieresultat

Betygsriterier		Förväntade studieresultat (lärandemål)
A	Förklarar uttömmande med ett mycket självständigt resonemang möjliga orsakssamband... och tillämpar med stor säkerhet...	Efter genomförd delkurs förväntas studenten kunna: <ol style="list-style-type: none"> 1) förklara sambandet mellan X och Y 2) tillämpa metoden Z
B	Förklarar i stort sett uttömmande med eget resonemang möjliga orsakssamband... och tillämpar med stor säkerhet...	
C	Kan förklara hela processen, inklusive vissa av momenten som... Kan tillämpa...	
D	Kan förklara de viktigaste delarna i sambandet mellan... Kan tillämpa...	
E	Rudimentär förståelse av sambandet mellan... Kan med smärre missförstånd tillämpa...	
Fx	Ansats till förklaring med tyngdpunkt på mindre viktiga delar i processen och/eller tillämpar felaktigt metoden...	
F	Kan inte förklara sambandet... och tillämpar felaktigt metoden...	

Ett kanske enklare exempel på principen ovan kan vara tydligt kvantifierbara prestationer motsvarande omfattningen av en specifik kunskapsmängd. Betyget E på en delkurs kanske förutsätter att studenten kan översätta 250 av 500 vanligt förekommande kinesiska tecken. Betyget A förutsätter att studenten kan översätta 450 av dessa 500, med mellanliggande betyg i spannet däremellan. Det förväntade studieresultatet behöver då ha formulerats så att det anger vad samtliga godkända (dvs. alla med betyget A, B, C, D eller E) kan efter genomförd delkurs. Att då bara ange i lärandemålet att studenten "kan översätta 250 tecken" är visserligen korrekt, men det ger ett märkligt intryck. Dels har delkursen enbart presenterats i termer av den absolut lägst acceptabla prestationen, dels har en skev bild skapats av vad studenter med höga betyg kan.

En mer rättvisande beskrivning skulle rymmas i förväntade studieresultat som istället anger att "Efter genomförd delkurs förväntas studenten kunna översätta *minst* 250 av 500 vanligt förekommande kinesiska tecken". Eller, kanske ännu hellre, om den nivån kan beskrivas i mer allmänna termer: "Efter genomförd delkurs förväntas studenten kunna behärska ett grundläggande kinesiskt ordförråd." Om betygsriterierna är målrelaterade finns det ju också ett målrelaterat resonemang bakom att just att antalet 250 tecken har satts som tröskelvärde för godkänt. (Till exempel att det anses motsvara den grundläggande kunskap som krävs för att kunna läsa en viss typ av text.)

Problem eller utmaningar...

Betygsdebatten

Betyg i grundskola och gymnasium är ett kontroversiellt ämne, och debatter därom pågår ständigt, ofta med ideologiska undertoner.⁸ Betyg i högre utbildning är även det en källa till debatt. När ett ämne eller en utbildning har gått från en tvågradig (G/U) till en tregradig (VG/G/U) skala har det mötts av blandade reaktioner. Frågan om sjugradiga betyg (A–F) har varit än mer laddad.⁹ Den enskilda läraren/examinatorn har att anpassa sig till lokala föreskrifter och beslut om vilket betygssystem som ska användas vid det egna lärosätet. Det gör förstås inte den offentliga debatten mindre viktig. Tvärtom kan det vara bra att känna till de vanligast förekommande argumenten i den debatten:

Tabell 1. Betygsdebatten

Fråga	Argument för	Argument mot
Målrelaterade betyg	<ul style="list-style-type: none">• Mycket positiva effekter på studenters lärande• Betygskriterier fungerar som precision av de förväntade studieresultaten• Student samarbete gynnas• Urval grundat på kunskaper möjligt• Betyg kan fungera som återkoppling på lärandet	<ul style="list-style-type: none">• Svårt jämföra med relativa betygssystem• Arbetsintensivt i utvecklingsfasen• Betygsinflation möjlig
Kriteriebaserade betyg	<ul style="list-style-type: none">• Positiva effekter på studenters lärande• Samarbete gynnas	<ul style="list-style-type: none">• Svårt jämföra med relativa betygssystem• Betygsinflation möjlig
Relativa betyg	<ul style="list-style-type: none">• Betygsinflation omöjlig• Urval grundat på rangordning möjligt• Jämförbarhet med andra relativa betygssystem möjlig	<ul style="list-style-type: none">• Negativa effekter på studenters lärande• Student samarbete motverkas• Urval grundat på kunskaper ej möjligt• Svårt jämföra med kriteriebaserade betygssystem
Betyg i få grader	<ul style="list-style-type: none">• Fokus på lärande och formativa bedömningar• Undervisning/bedömning som fungerar väl med få grader kan behållas• Arbetsbesparande för läraren	<ul style="list-style-type: none">• Svaga incitament för studenten att engagera sig• Olika resultat ger samma betyg. Ingen åtskillnad mellan t ex G+ och G• Undervärdering av svenska studenter i internationella jämförelser
Betyg i många grader	<ul style="list-style-type: none">• Större precision och större rättvisa möjlig• Jämförbarhet med internationella flergradiga system underlättas• Möjlighet till förfinad och därmed bättre feedback på studentens lärande	<ul style="list-style-type: none">• Kan leda till betygshets och stress• Risk för försämrade examinationsformer• Ökad arbetsbörda för läraren• Risk för ökat fokus på yttre motivation ("piska/morot")

Det finns alltså en mängd argument för och mot betyg i olika former. Alla delar av de ideologiska och vetenskapliga debatterna om till exempel flergradiga betyg är inte nödvändigtvis relevanta för den enskilda lärarens faktiska betygsarbete. En del av argumenten är emellertid viktiga att beakta även i den konkreta utformningen av kriterier. Så även forskningen om motsvarande betygsreformer i gymnasium och grundskola.

Forskningen om mål- och kunskapsrelaterade betyg i grundskola och gymnasium

Under de senaste åren har flera avhandlingar publicerats som behandlar övergången från relativa till målrelaterade betyg i grundskola och gymnasium under 1990-talet, forskning som är högst relevant även för lärare i högre utbildning.¹⁰ Förutsättningarna för de olika utbildningsnivåerna skiljer sig förvisso åt på många sätt. Högre utbildning bygger inte på nationella läroplaner och tankar om likvärdighet på samma sätt som grundskola och gymnasium. Att arbeta med standardisering, nationella prov och nationella betygs-kriterier är ju inte aktuellt i högre utbildning, där vetenskaplig mångfald och beslut fattade av enskilda lärare är normen. Dessutom kan man argumentera för att det är lättare att formulera kravnivåer när kunskapsstoffet är enklare. Ju högre nivå desto högre abstraktionsnivå. Dessutom är universitetslärare i första hand ämnesexperter, medan lärare i grundskola och gymnasium i första hand är pedagoger och ämneslärare, etcetera. Samtidigt finns det många beröringspunkter. När det gäller själva bedömningsarbetet är den principiella problematiken ofta likartad på olika nivåer av utbildning. Nationella prov och kriterier finns dessutom bara i vissa ämnen och i vissa årskurser, vilket gör att skolans lärare sedan länge har ställts inför likartade uppgifter när det gäller utformning av målrelaterade betygs-kriterier.

Hur har det då gått med betygsreformerna? Skolverket konstaterade i en rapport 2004 att betygssättande lärare, skolor och kommuner hade vitt skilda föreställningar om de principiella grunderna för det betygssystem de var satta att implementera:

Granskningen visade att det fanns stora skillnader när det gäller förutsättningarna för en likvärdig bedömning av elevernas kunskaper. Skillnaderna gällde t.ex. att förståelsen av tankarna bakom betygssystemet i styrsystemet i dess helhet varierade, liksom kunskaperna hos lärarna om kursplanernas innehåll. [...] Dessutom varierade lärarnas syn på vad som skulle ligga till grund för betygssystemet, liksom huvudmännens och skolledningarnas engagemang och ansvarstagande för att skapa likvärdiga förutsättningar.¹¹

Även Bengt Selghed visar i sin avhandling *Ännu icke godkänt: lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*, att många problem med betygssystemet i skolan går att härleda till bristande förståelse av vad ett mål- och kunskapsrelaterat system innebär och hur det skiljer sig från andra alternativ.¹² Oreflekterade inslag av relativ bedömning är ett exempel på detta. Jörgen Tholin riktar i sin avhandling *Att klara sig i okänd natur – En studie av betyg och betygs-kriterier – historiska betingelser och implementeringen av ett nytt system*, en kanske ännu tyngre kritik mot hur en i och för sig lovvärd betygsreform har kommit att fungera i praktiken. Merparten av kritiken riktas mot Skolverket och mot en politisk nivå, men bilden av den enskilda lärarens praktik är ändå bekymmersam:

När man läser de styrdokument som vissa skolor har producerat är det lätt att drabbas av missmod. Talet om den professionella läraren känns mycket avlägset. [...]

Två bilder växer fram i denna studie. Den ena är bilden av läraren som inte är särskilt professionell eller analytisk i sin lärarroll, som omfattar en mycket traditionell kunskaps- och ämnessyn och som ibland använder betyg inte som ett mått på elevens kunskaper och färdigheter utan som ett sätt att disciplinera sina elever. Den andra bilden som framträder är den där lärare försöker göra sitt bästa trots att de[t] saknats tid för, information om och vana vid uppgiften.¹³

Helena Korps avhandling *Lika chanser på gymnasiet? En studie om betyg, nationella prov och social reproduktion* visar bland annat hur icke kunskapsrelaterade kriterier ibland tillåts inverka på betygssystemet.¹⁴ Elever som inte har uppnått uppsatta lärandemål godkänns ändå, och beteendekriterier (sic!) kan dessutom inverka på betygen – som om det som förr kallades ordning och uppförande vävs in i ämnesbetygen. Det sistnämnda är ett tydligt exempel på en tillämpning av kriterier som inte är mål- och kunskapsrelaterade.

Vilka lärdomar kan vi dra?

En samlad bild av forskningen om betygsreformen inom skolan visar att en dylik reform behöver genomföras väl för att inte en rad oacceptabla konsekvenser ska uppstå.¹⁵ Institutionella, sociala och

politiska faktorer inverkar starkt, men ett tydligt budskap är också att den enskilda läraren/examinatorn måste förstå vad ett målrelaterat system innebär, och hur det skiljer sig från icke-målrelaterade system (inklusive kriteriebaserade system som inte är målrelaterade). För vissa motsvarar det redan invanda tankesätt, för andra en betydligt större synvända.¹⁶

Frågan vi alla behöver ställa oss är hur vi kan undvika att reproducera de brister som har påvisats i grundskola och gymnasium? Ett svar kan vara att frågan om målrelaterad bedömning och betyg måste tas på största allvar på alla nivåer inom högskolan. Om reformen genomförs halvhjärtat, eller enkom som en byråkratisk anpassning till påbud uppifrån, är risken stor att motsvarande forskning om vår egen praktik kommer att ge en lika dystert bild framöver.¹⁷

Intuitiva betygsbedömningar – konsekvens och reliabilitet

I alla komplexa bedömningsprocesser är frågan om konsekvens och reliabilitet alltid viktig. Ett vanligt argument för skriftliga betygs-kriterier är just att de kan hjälpa läraren att göra mer konsekventa bedömningar. Om läraren har stor erfarenhet av att bedöma en viss typ av examinationsuppgift, och förmår bedöma efter samma principer över tid, kan detta i vissa fall åstadkommas även utan skriftliga kriterier. Men, frånvaron av på förhand formulerade kriterier också leda till en improviserad bedömning, där olika delbedömningar och olika perspektiv påverkar det sammantagna omdömet utan att läraren själv förmår överblicka processen. Situationen kan liknas vid en jurist som ska bedöma något utan tillgång till lagar och förarbeten, eller en forskare som klassificerar och kategoriserar ett material på okända grunder.

Oartikulerade överväganden som till del bygger på en tyst kunskap är en omistlig del av all bedömning av studenters kunnande.¹⁸ Men, graden av godtycke kan lätt nå oacceptabla nivåer om inte bedömningsgrunderna har artikulerats så väl som möjligt. Forskningen om faktorer som kan försämra examinatorers bedömningar är omfattande.¹⁹ Utan ett givet ramverk kan lärare i början av ett bedömningsarbete inleda med en tolerans för en viss typ av tillkortakommande, men bli allt mer frustrerad och omedvetet bedöma detta allt hårdare. Olika delbedömningar kan väga olika tungt vid olika tillfällen. En känsla halvvägs in i en hög med tentor av att ovanligt många i studentgruppen har bedömts som underkända, utan att gruppen i stort var svagare än tidigare grupper, kan skapa en tendens att godkänna fler av de kvarvarande än annars. Och så vidare. En bedömningsprocess som till övervägande del är intuitiv ökar dessutom risken för påverkan av helt ovidkommande faktorer. Kön, ålder och etnicitet är kända exempel, men det finns studier som visar att till och med studentens förnamn, personlighet, typsnittet i inlämnade uppgifter och papprets färg kan påverka betygsbedömningen.²⁰ I en studie vid Linköpings universitet där betygssättning simulerades kunde studentprestationernas kvalitet bara förklara 55–63% av givna betyg.²¹ Övriga faktorer som inverkade var bland annat systematisk lärarpåverkan och slumpen. Studien var begränsad, men bara möjligheten att ett betyg till så stor del kan påverkas av andra faktorer än studentens förståelse manar till eftertanke. Betygssättning är helt enkelt någonting som inte blir bra av sig självt.

Det är i det här sammanhanget viktigt att betona att examinatorns uppgift inte är att producera ett omdöme som är ”objektivt” oantastligt i ordets mer rigida mening. Det handlar om att en ämnesexpert efter bästa förmåga ska värdera hur väl en student har förstått något – varken mer eller mindre. Huruvida det bör kallas ”subjektivitet” eller inte kan diskuteras. Opartiskhet, transparens och intersubjektiv prövbarhet är likafullt värden som kan och bör värnas. Förväntade studieresultat och skriftliga betygs-kriterier som har förankrats i ett arbetslag och diskuterats ingående med studenterna kan fungera som ett stöd i detta.

Kan man inte jämföra studentarbeten i ett målrelaterat system?

Jo, att vända och vrida på olika studentprestationer och göra jämförelser är ofta en viktig del av bedömningsarbetet. Men, det är viktigt att klargöra vad det är för sorts jämförelser som görs, annars är det lätt att det smyger sig in relativa moment i processen. Om jämförelser mellan olika prestationer görs för att uppnå konsekvens i betygssättningen kan man säga att läraren egentligen jämför de egna bedömningarna med varandra. Hur såg jag på den här typen av resonemang i den andra tentan? Vilka betygs-kriterier matchade jag den här typen av missförstånd med igår? Jämförelserna syftar då i första

hand till att arbeta med exempel för att uppnå en konsekvens i bedömningen. Att göra en students betyg direkt beroende av vad andra har presterat är däremot att tillämpa ett relativ betygssystem. Även om det inte är en uttalad ambition kan en oreflekterad önskan att få "lagom" många godkända, eller "lagom" många med ett högt betyg, ändå influera betygssättningen i en grupp. Ett aktivt arbete med förväntade studieresultat och målrelaterade betyg kan motverka sådana tendenser.

Samma kriterier för allt och alla?

Vid vissa lärosäten formuleras betygs-kriterier generellt för alla ämnen och kurser. Exempelvis Mittuniversitet har beslutat att alla kan beskriva kraven för exempelvis betyget B som: "Studenten visar *framstående* resultat i förhållande till de krav som ställs i lärandemålen." Kraven för D skrivs som: "Studenten visar *tillfredsställande* resultat i förhållande till de krav som ställs i lärandemålen."²² Även internationellt förekommer den typen av gemensamma kriterier, ofta som markering att relativa betyg inte tillämpas. Vad sådana allmängiltiga kriterier får för konsekvenser beror till stor del på hur de hanteras på kurs- och delkursnivå. Ett alternativ är att de gemensamma ramarna ses som en övergripande karaktäristik som sedan ligger till grund för kursspecifika kriterier. Ett annat alternativ är att den typen av kriterier helt ersätter kursspecifika kriterier.²³

Vid Stockholms universitet ska betygs-kriterierna utformas med utgångspunkt i lärandemålen för varje kurs/delkurs som examineras separat, och utgöras av kvalitativa beskrivningar av hur väl de lärandemålen har uppnåtts.²⁴ Det är alltså inte möjligt att låta några gemensamma kriterier på institutions- eller fakultetsnivå ersätta kursspecifika kriterier. Frågan är ändå principiellt viktig, eftersom olika "tekniska" lösningar för att komma runt arbetet med kvalitativa kriterier ändå föreslås ibland, till exempel i formen av allmängiltiga kriterier. Gemensamt för den typen av lösningar är att de syftar till att komma runt uppgiften att formulera kursspecifika kvalitetsnivåer. Argumenten mot sådana kriterier knyter an till debatten om graderade betyg i stort: Risken är stor att man bara får möjliga nackdelar med fingerade betyg (betygshets, mm), men inte några av de viktigaste fördelarna (möjligheten till relevant återkoppling).

Graderade betyg, betygshets och studenters lärande

Det finns väl kända risker med införandet av flergradiga betyg: *Varför* man försöker förstå någonting har en avgörande betydelse för kvaliteten på lärandet.²⁵ När en student läser för att förstå tar den till sig kunskaperna på ett annat sätt än när den studenten enbart läser för att bevisa att den har förstått. I verkligheten tenderar individer att röra sig mellan de ytterligheterna, men tyngdpunkten kan ändå variera. Ett lärandeklimat som helt och hållet domineras av betyg och mätning av kunskap, snarare än kunskaperna som mål i sig, kan därför i vissa fall ge sämre förutsättningar för djuplärande och avancerad förståelse.²⁶ Vi kan aldrig bortse ifrån risken att ett införande av en månggradig skala kan leda till att individer och grupper i större utsträckning kraftsamlar på betygen istället för bästa möjliga förståelse.²⁷ Sådana krafter behöver alltid motverkas, vilket torde förutsätta att en betygsreform samtidigt behöver vara en pedagogisk reform.

En av de främsta pedagogiska fördelarna med kursspecifika, kvalitativa kriterier och mål är att de på ett enkelt sätt kan kopplas till arbetet med att ge studenter återkoppling på det egna lärandet. Fortlöpande feedback är en mycket effektiv undervisningsmetod som ofta är eftersatt. När en lärare genomför en genomgång efter en tenta och diskuterar olika typfel och olika nivåer av förståelse i gruppens prestationer görs precis detta i efterhand. Det finns mycket att vinna på att arbeta på liknande sätt även under kursens gång.²⁸ (Se avsnittet "Detaljerad feedback och betyg – en exkurs" nedan.) Det finns många metoder för att göra en sådan kontinuerlig återkoppling möjlig även i stora studentgrupper och med begränsade resurser, men de metoderna förutsätter att det finns en väl fungerande apparat för kvalitativa bedömningar.²⁹ Det går också att vända på resonemanget. Oavsett antalet betygssteg och typ av betygs-kriterier finns det stora pedagogiska vinster att göra om lärare utvecklar sin kompetens i att göra kvalitativa bedömningar av studenters förståelse. En lärare som har utvecklat metoder för detta har i princip gjort merparten av det arbete som krävs för att kunna konstruera väl fungerande, kursspecifika betygs-kriterier – det vill säga kvalitativa beskrivningar av hur väl lärandemålen har uppnåtts.

Är kejsaren naken? Betygskriterier som dimridåer och alibi

Även de mest utförliga och genomarbetade betygskriterietexter kan knappast vara mer än korta sammanfattningar av betydligt längre resonemang. De kan liknas vid isberg där betygskriterierna är topparna som syns ovanför vattenytan. Man ska därför inte ställa alltför höga krav på hur uttömmande kriterierna måste vara i sig själva.³⁰ Ett rimligt krav är emellertid att läraren kan svara på frågor om vad de olika formuleringarna egentligen betyder. Vad menas med ”mycket självständigt”? Vad betyder egentligen gränsen mellan ”refererande litteraturdiskussion” och ”recenserande litteraturdiskussion”? Läraren måste helt enkelt kunna exemplifiera och förklara vad som ligger bakom dessa uttryck när studenter undrar. Dessvärre är det vanligt med formuleringar som inte alls svarar mot något utvecklat resonemang hos den som har författat kriterietexten. Den typen av problem kan lätt uppstå om författandet av kriterietexter enbart betraktas som en slags formuleringsexercis. Sådana betygskriterier tenderar också att förbli meningslösa.

I praktiken kan examinatoren då komma att bedöma studenternas prestationer helt efter en intuitiv, allmän skala mellan något slags ”bra” och något slags ”dåligt”. Vad detta ”bra” och ”dåligt” representerar är då okänt och kan variera över tid. Kriterietexterna kan bestå av diverse tomma superlativ för A och en diffus, sjunkande skala av allt mindre smickrande adjektiv och adverb därunder. Det går ganska lätt att utarbeta texter som på det sättet ”ser ut” som betygskriterier utan att vara det. Om de inte motsvarar de faktiska värderingar som examinatoren kommer att använda sig av kan de på sin höjd fungera som alibi för den som har ålagts att formulera dem.

I ett sådant fall har alltså betygskriterier formulerats, och examinationsuppgifter bedömts. De två processerna hänger däremot inte samman. Ett arbete som istället inleds med att författaren verkligen försöker skapa sig en bild av kvalitativa skillnader har betydligt bättre förutsättningar att bli bra. Skrivandet handlar då om att sätta ord på professionella värderingar, inte om att producera en formellt oantastlig text. Ett effektivt sätt att därefter komma oavsiktliga dimridåer på spåren är att tillsammans med andra pröva kriterierna på riktiga examinationsuppgifter, som ett par gamla tentor eller projektarbeten. Då uppdragas obönhörligen vilka formuleringar som verkligen är meningsbärande och vilka som inte är det.

Att formulera betygskriterier

Betygskriterier = generella betygsmotiveringar

Att utforma skriftliga målrelaterade betygskriterier uppfattas ibland som någonting nytt och komplicerat. För att det ska bli bra krävs visserligen ett grannliga arbete, men det är egentligen en typ av uppgift som många har ägnat sig åt under lång tid – om än i en annan form. Lärare lägger mycket tid och möda på att skriva kommentarer till studenter på skriftliga examinationsuppgifter. Dessutom förs ofta diskussioner med studenter där läraren muntligen motiverar sina betygsbeslut. Om dessa skriftliga och muntliga motiveringar är kopplade till de förväntade studieresultaten (lärandemålen) för delkursen kan de betraktas som en slags individuellt formulerade ”kriterier”, meddelade i efterhand. Betygskriterier kan betraktas som en slags *generaliseringar* av just sådana motiveringar, meddelade i förväg:

Figur 3. Betygskriterier som generella motiveringar

Individuell feedback/motivering efter examination	Betygskriterier
<ul style="list-style-type: none">• "Sara, ditt projektarbete fick betyget VG därför att..."• "William du får VG därför att du i din tentamen visar att du kan... och... med en god förmåga att..."	} "För VG krävs att studenten kan ... och ... med kvaliteten ..."
<ul style="list-style-type: none">• "Reza, din uppsats är... och får därför G. För ett högre betyg skulle ... ha behövts, jämte..."• "Lisa du får G därför att du i ditt arbete visar att du kan ... Därtill har du..."	
<ul style="list-style-type: none">• "Ditt projektarbete fick betyget U därför att..."• "Kim, ditt arbete har inte..., vilket motiverar betyget U i det här läget. Vid omprovet bör du..."	} "U ges vid ett eller flera av följande tillkortakommanden: ..."

Om en lärare ska utarbeta betygskriterier för en kurs som har getts flera gånger tidigare kan det vara relativt enkelt att i allmänna ordalag försöka beskriva vad som brukar känneteckna prestationerna för olika betygsgrader. Om kriterier ska skrivas för en helt ny kurs, om ett betygssystem med fler grader än tidigare har införts eller om motiveringar till betyg tidigare inte gavs med utgångspunkt i kvalitativa mål, krävs ett större arbete. Många gånger kan även det arbetet underlättas om man börjar med att försöka föreställa sig typfall av olika prestationer och hur dessa skulle kunna kommenteras, för att därefter börja med generella beskrivningar.

Den praktiska utformningen av betygskriterier

Målrelaterade betygskriterier kan rent praktiskt utformas på väldigt många olika sätt. Alla sätt har sina för- och nackdelar, och det kan vara bra att noga tänka igenom vilket format man vill arbeta med, då ett sådant val kan få konsekvenser för verksamheten i stort. En institution kan behöva överväga om man behöver ett gemensamt arbetssätt, till exempel i utbildningar där flera olika lärare samarbetar, och där studenterna passerar igenom en serie olika delkurser och kurser inom ramen för en och samma utbildning. I samband med valet av utformning av kriterierna är det dock oerhört viktigt att det arbetet sätts i rätt perspektiv: vilka funktioner ska betygskriterierna fylla, vad ska de användas till? Det är vanligt att arbetet inleds med en alltför stor fokusering på mallar, regler och formalia. Det kan vara naturligt när man ställs inför en ny typ av uppgift, på ungefär samma sätt som studenter kan inleda sitt arbete med uppsatser och projektarbeten med uppmärksamheten riktad mot sidantal och formalia. Ibland är det ett sätt att skapa sig ett nödvändigt ramverk för det fortsatta tankearbetet, men ofta skymmer det de mer kvalificerade och prioriterade delarna av uppgiften.

Betygskriterierna ska i ett målrelaterat system motsvara ett omdöme om vad och hur väl en individuell

student kan något. I ett sjugradigt system: vad kan till exempel den som får betyget B som den som får betyget C inte kan? Vad kan den som får betyget D? I ett tregradigt system: vad kan den som får betyget VG och hur väl? Betygskriterierna måste fylla den funktionen och formerna för hur detta uppnås är av helt underordnad betydelse. Om betygen däremot inte fyller en sådan målrelaterad funktion, så kommer de förr eller senare att behöva omarbetas i grunden. Allt arbete med dessa kriterier kan då ha varit förgäves. Det är mot den något allvarstygda bakgrunden nedanstående diskussion om praktisk utformning, alternativ och exempel ska ses.

Kan betygs-kriterier uttryckas som poäng eller procent?

Inte om dessa poängtal representerar okända kvaliteter, som ofta är fallet. Att arbeta med kvantitativa sammanräkningar av exempelvis antal poäng på tentor, kan förvisso vara förenligt med ett målrelaterat betygssystem – men bara om dessa siffervärden på något sätt motsvarar på förhand bestämda kvaliteter (som då ändå behöver formuleras som betygs-kriterier). Vad kan den som får mer än X rätt på tentan? Vad kan *inte* den som får mindre än X rätt på tentan? Om det inte finns några svar på den typen av frågor kan poängen/betygen enbart användas till att differentiera mellan studenter. I ett målrelaterat system är examinatorns uppgift mer kvalificerad än så.

Frågan berör en central problematik som i vissa fall kan ha långtgående konsekvenser – förhållandet mellan målrelaterade betyg och kriteriebaserade betyg som inte är målrelaterade.³¹ Om poäng på tentor används som betygsunderlag *utan* att dessa poäng är kopplade till på förhand formulerade kvaliteter, så kan det visserligen vara förenligt med ett kriteriebaserat betygssystem. Alla studenter kan ju i princip få samma betyg, och man har därmed inte använt sig av ett relativt betygssystem. Om varken examinatorn, studenten eller någon utomstående med ledning av betyget kan sluta sig till vad och hur väl studenten kunde något, kan betyget däremot inte ses som målrelaterat. I ett målrelaterat system är examinationen ett verktyg som används för att göra det möjligt att matcha studentens prestationer mot förutbestämda nivåer av förståelse. Sådana betygs-kriterier av hög kvalitet kan fungera som utgångspunkt för utvecklingen examinationsuppgifterna. Omvänt kan väl fungerande examinationsformer fungera som utgångspunkt för utformningen av betygs-kriterierna. Det får i vilket fall som helst inte uppstå ett glapp mellan beskrivna kvalitetsnivåer och de verktyg som används för att pröva dem.

Observera att gränsdragningen mellan kriteriebaserade betyg som är respektive inte är målrelaterade inte alls går mellan kvantitativ och kvalitativ metod i vanlig mening. Kvantitativa tekniker kan mycket väl användas i ett målrelaterat betygssystem, men det ställer särskilda krav på hur detta görs, inte minst vad gäller examinationsinstrumentens validering. Om till exempel 80% av maxpoäng på en examinationsuppgift ska ses som ett tröskelvärde för ett högt betyg, måste examinationsuppgifterna utformas så att en bestämd kvalitet på studentens kunnande ”orsakar” just dessa 80%. Det omfattande arbete som kan krävas för att åstadkomma examinationsuppgifter som är tillräckligt valida och reliabla för att fylla den funktionen kan ibland vara motiverat, till exempel med mycket stora studentgrupper och då examinationsuppgifterna kan behållas oförändrade över tid.

Två sätt att utforma kriterier: globalt eller per bedömningsområde

I de flesta fall arbetar läraren med ett antal områden eller perspektiv när studenternas lärande och prestationer bedöms. Ofta är lärandemålen eller examinationsformerna grund för dessa delbedömningar, ibland konstruerar man helt andra bedömningsområden. Betygs-kriterier kan i princip utformas på två olika sätt: (1) som globala/holistiska kriterier, där varje betygsgrad beskrivs med en sammanhängande text, respektive (2) uppdelade per bedömningsområde, ibland kallat analytiska eller atomistiska kriterier.³² Båda formerna har sina för- och nackdelar. Globala betygs-kriterier gör det möjligt att ge en samlad bild av vad som kännetecknar olika kvalitetsnivåer. Det kan vara en stor fördel inte minst när betygs-kriterierna ska kommuniceras med andra. Dessutom ger de en tydlig signal om att helhetsbedömningen är viktigast. Nackdelen är att det ibland är svårt att skriva dessa globala texter så att gränserna mellan de olika kvalitetsnivåerna framgår tillräckligt tydligt. Kriterier uppdelade efter bedömningsområden, ofta skrivna i olika typer av matriser, visar tydligt hur olika aspekter bedöms. Nackdelen är att det ofta är svårt att se hur dessa delbedömningar beskriver ett samlat betyg för kursen/delkursen.

Figur 4. Globala kriterier och kriterier uppdelade i bedömningsområden

	Globala kriterier	Bedömningsområde 1	Bedömningsområde 2	Bedömningsområde 3
A	"Text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text..."	"Text, text, text..."	"Text, text, text..."	"Text, text, text..."
B	"Text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text..."	"Text, text, text..."	"Text, text, text..."	"Text, text, text..."
C	"Text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text..."	"Text, text, text..."	"Text, text, text..."	"Text, text, text..."
D	"Text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text..."	"Text, text, text..."	"Text, text, text..."	"Text, text, text..."
E	"Text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text..."	"Text, text, text..."	"Text, text, text..."	"Text, text, text..."
Fx	"Text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text..."	"Text, text, text..."	"Text, text, text..."	"Text, text, text..."
F	"Text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text, text..."	"Text, text, text..."	"Text, text, text..."	"Text, text, text..."
"Kriterierna bygger på följande bedömningsområden:... 1.... 2... 3..."		"Delbedömningarna ovan omvandlas till ett samlat betyg på delkursen genom en samlad bedömning, där bedömningsområde 1 är viktigast, och..."		

Hur detaljerat ska kriterietexterna formuleras?

Det är något som författaren av kriterierna själv måste avgöra. Svaret på frågan beror bland annat på hur detaljerat de förväntade studieresultaten är formulerade, och hur detaljerat olika kvaliteter av kunnskap i just det aktuella ämnet bör eller ens kan beskrivas. Dessutom behöver hela den sammantagna kommunikationen kring förväntade studieresultat och betygskriterier tas i beaktande. Det handlar trots allt om att synliggöra och artikulera ett antal kvalitativa gränsdragningsregler. Om detta dessutom diskuteras ingående i olika typer av kompletterande texter, i kursbeskrivningar, muntligt i samband med undervisningen, med mera, påverkar det förstås hur uttömmande dessa små kriterietexter behöver vara. För att kriterierna ska kunna fylla sin funktion kan de emellertid inte vara så allmänna och oprecisa att de blir meningslösa för utomstående.

Hur formulera? Kriterietexter om produkt, prestation eller lärande?

Även om betygskriterierna alltid ska fylla funktionen att beskriva hur väl en student kan något, kan det rent språkligt uttryckas på olika sätt. Man kan exempelvis välja att formulera kriterierna så att de syftar på examinationsuppgifternas kvalitet för ett visst betyg: "Uppsatsen kännetecknas av...". Beskrivningen av den kvaliteten blir då som en indirekt beskrivning av studentens kunnskap. Ett annat alternativ är att formulera det direkt som beskrivningar av vad studenten kan: "Studenten kan... med en hög grad av...". Dessa två uttrycksätt kan kombineras, med studenten som subjekt i förhållande till prestationen: "Studenten visar en god förmåga att...". Ytterligare ett alternativ kan vara att formulera texterna kort och utan explicita subjekt: "Nyanserat resonemang. Stor grad av självständighet." Alla skrivsätt ovan är förenliga med ett målrelaterat betygssystem. Av tydlighetsskäl kan det vara bra att utforma alla sina kriterier språkligt efter en och samma princip. Om kriterierna för B uttrycks på ett sätt och kriterierna för C på ett helt annat blir lätt gränsen däremellan onödigt svårbegriplig för läsaren. På den här detaljnivån är valet av formuleringalternativ oftast upp till den eller dem som har till uppgift att författa kriterierna, men det kan hända att man på en institution vill förenkla för läsarna och författarna genom att enas om gemensamma riktlinjer även för den språkliga utformningen.

Två typer av kvalitetskillnader – kontinuerliga och diskreta

När en lärare formulerar sina mångfacetterade betygsriterier kan arbetet många gånger förenklas om hon/han först tar ställning till vilka typer av bedömningar det egentligen handlar om. Då framgår det tydligare hur det går att arbeta vidare med själva formuleringarna. En typ är *kontinuerliga* kvalitetskillnader, där en och samma företeelse kvalitetsbedöms efter en skala mellan ”mycket” och ”lite” av något, som i ett kontinuum mellan ljus och mörkt.³³ Om ett moment av en kurs handlar om att kunna översätta ett antal ord på ett främmande språk kan man lätt tänka sig ett kontinuum av antalet korrekta översättningar, från hundratals felfria översättningar ner till noll rätt. I andra sammanhang kan ett sådant kontinuum till exempel handla om empirins eller forskningslägets omfattning. Det är då en och samma företeelse som gradvis kan ökas eller minskas ungefär som ett volymvred på en radio. Kontinuerliga skillnader kan också vara fråga om kontinuerliga värderingar, till exempel ett kontinuum mellan ”bra” och ”dåligt” av en viss aspekt.

Figur 5. Kontinuerliga kvaliteter/bedömningar

Mycket empiri	Lite empiri
Stark teoriansknytning	Svag teoriansknytning
Nyanserad analys	Onyanserad analys
x x x x x x x x x x x	

I andra fall kan olika kvaliteter beskrivas som mer eller mindre *diskreta* skillnader. Det kan handla om att skillnader av en och samma aspekt av studenters prestationer i praktiken tenderar att klumpa ihop sig, även om de i princip skulle kunna vara kontinuerliga. I vissa sammanhang kan exempelvis graden av självständighet i studentarbeten ha den karaktären, där de flesta prestationerna ganska tydligt kan karaktäriseras som antingen ”osjälvständiga”, ”självständiga” eller ”mycket självständiga”. Att den typen av kvalitetskluster ibland uppstår kan förklaras med att de olika diskreta nivåerna svarar mot olika tröskelvärden i studenters förståelse. Varje högre kvalitetsnivå kan då representera en ny plåtå eller ”aha-upplevelse” hos studenten.³⁴

Figur 6. Diskreta kvaliteter/bedömningar

Mycket självständigt... Jämförande litt. genomgång... Text som ”diskussion”	Självständigt... Recenserande litt. genomgång... Text som ”synpunkt”	Osjälvständigt... Refererande litt. genomgång... Text som ”återgivning”
x x x x x	x x x x x x	x x x x

En annan sorts diskreta kvaliteter motsvarar mer innehållsliga eller analytiska skillnader. I en känd studie av Dai Hounsell identifierades tre distinkt åtskilda föreställningar bland studenter om vad examinationsuppgiften uppsats var för någonting.³⁵ En uppfattning var att uppsatsen var en ordnad ”återgivning” av idéer och fakta. En annan att det skulle vara en ”synpunkt”, eller tolkning av ett problem. Den sista kategorin som identifierades var studenter som såg uppsatsen som en organiserad, tolkande ”diskussion”, med en argumentation underbyggd av data. Hounsell studie handlade främst om studenters föreställningar, men dessa föreställningar fick tydliga konsekvenser för uppsatsernas kvalitet. Ett annat exempel kan vara en litteraturkurs med förväntade studieresultatsom handlar om att studenterna ska kunna diskutera och ta ställning till ett forskningsläge. Om litteraturen enbart återges kan det kallas ”refererande” – i många sammanhang ett rimligt kriterium för underkänt. I en ”recenserande” text förs istället relevanta resonemang om styrkor och svagheter i olika framställningar. I en text som för samman sådana recenserande resonemang i utvecklade ”jämförelser” uppvisas en ännu

större analytisk förmåga. Sett i radiometaforen skulle de skillnaderna kunna motsvara distinkta stationer, snarare än en kontinuerlig volym.

Mer eller mindre tydliga gränsdragningar

Kategoriseringen handlar både om att finna så rättvisande formuleringar möjligt och om att därefter kunna göra bedömningar med dessa texter som riktlinjer. Med en sjugradig skala är det uppenbart att vissa kvalitetsskillnader kommer att kunna bedömas enklare än andra. Den viktigaste gränsen går alltid mellan godkänt och underkänt, mellan E och Fx. De övriga fem gränserna behöver sedan vara så otvetydiga som möjligt. Ibland finns det väl inarbetade former för att identifiera kvalitetsskillnaderna mellan det gamla systemets G och VG. Den gränsen kan med fördel fungera som utgångspunkt för en mer fingeraderad bedömning, till exempel mellan betygen B och C. Det är inte givet att alla gränsdragningar som tillkommer därutöver kan skrivas och bedömas lika säkert. Att i praktiken behöva arbeta med både skarpa och luddiga gränser kan ses som en naturlig konsekvens av ett fingeraderat betygssystem. (Med färre grader ökas säkerheten i bedömningarna, till priset av en minskad precision.)

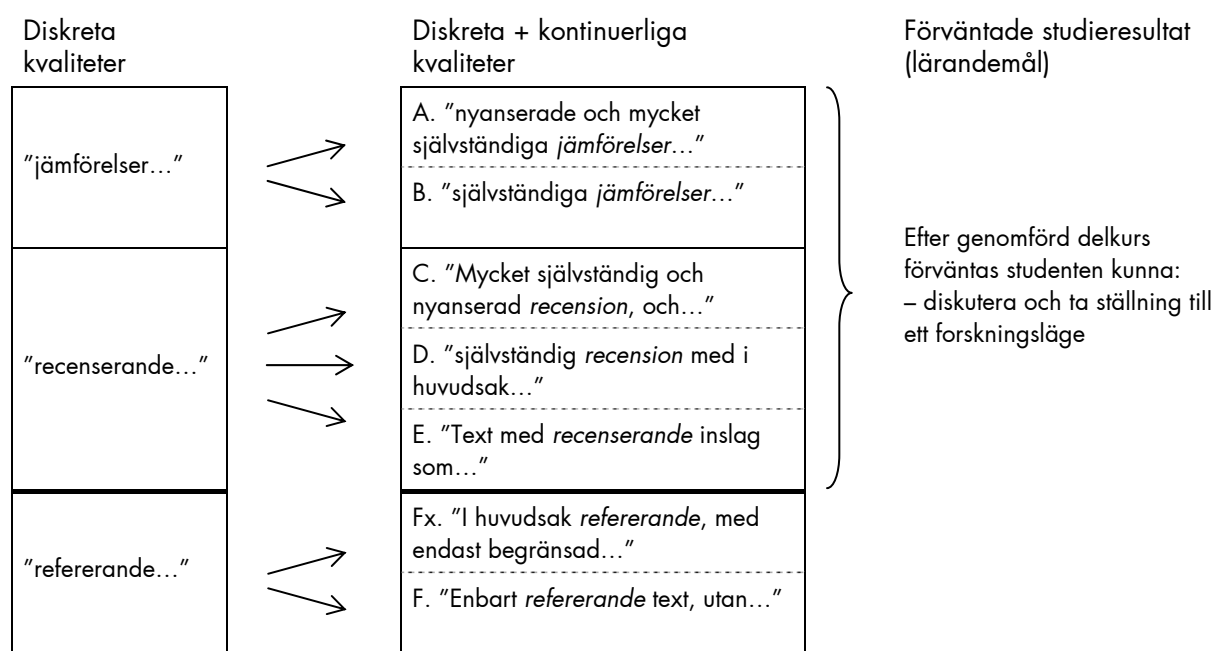
På ett sätt kan man säga att all betygssättning innebär att någonting flerdimensionellt (studentens förståelse) översätts till en endimensionell skala (betyg). I den enklaste av världar skulle studenters lärande och prestationer alltid motsvara endimensionella, tydligt diskreta kvaliteter i precis samma antal grader som rådande betygssystem. I ett sjugradigt, målrelaterat betygssystem skulle det då vara tämligen enkelt att utforma otvetydiga betygskriterier i just sju steg. I verkligheten är det sällan så. Bedömningar av de komplexa fenomen som kännetecknar lärande i högre utbildning kräver oftast att olika delbedömningar och olika typer av bedömningar kombineras till en helhet. Den samlade bilden av dessa delbedömningar behöver anpassas till att passa just en sjugradig skala så bra som möjligt.

Hur kategorisera i sju kvaliteter?

När det gäller kontinuerliga kvaliteter väcks ofta frågan hur dessa ska kunna beskrivas i separata grader. Hur ska gråzonen mellan ”mycket” och ”lite” empiri kunna skiktas på ett meningsfullt sätt? Hur ska ett kontinuum av mer eller mindre korrekta översättningar i ett språkämne kunna uttryckas i just sju steg? Om ett sådant kontinuum motsvarar kvantiteter kan de ganska enkelt delas upp i det antal man önskar ungefär som när en demograf delar upp ett kontinuerligt åldersspann i ett konstruerat antal åldersgrupper. Kontinuerliga skillnader som inte kan kvantifieras på det sättet kan även de delas upp i ett konstruerat antal grupper med markörer som ”mycket...” ”viss...” ”begränsat...” och liknande. Sådana skillnader kan emellertid vara svåra att formulera i många steg. Att formulera och sedan urskilja skillnader i två-tre nivåer kan däremot ofta gå ganska lätt.

När det gäller diskreta kvalitetsskillnader uppstår problemet bara när de skillnaderna inte motsvarar antalet betygssteg. I exemplen ovan på olika diskreta skillnader var ju utgångspunkten tre sådana nivåer, inte sju. Ett möjligt sätt att arbeta är att kombinera både diskreta och kontinuerliga skillnader, för att på så sätt åstadkomma en skiktning i tillräckligt många grader. Det kan då motsvara en stegvis process i både formuleringsarbetet och examinatorns bedömningar av studentprestationerna: först kan en grovsortering göras där gruppen godkända prestationer skiljs från dem som inte uppnår de förväntade studieresultaten. Därefter kanske en relativt tydlig sortering mellan två eller tre diskreta kvaliteter kan göras. Inom dessa kategorier är det möjligt att det går att finna bra beskrivningar av mer kontinuerligt slag. I principskissen nedan framgår hur detta kan kombineras till en sjugradig skala. Diskreta skillnader markeras i figur 7 nedan med heldragna linjer, kontinuerliga skillnader med streckade linjer:

Figur 7. När kriterietexter skrivs som både diskreta och kontinuerliga kvaliteter



Hur kan man se om betygskriterierna är bra?

Det kan man inte. Ibland går det däremot att se om de är dåliga. Texter som inte handlar om en precisering av de förväntade studieresultaten, som inte beskriver grader av förståelse och betygskriterier som är obegripliga för utomstående är uppenbart problematiska. Huruvida de är "bra" avgörs till stor del om gränsdragningarna verkligen motsvarar examinatorns värderingar, och det ledet framgår inte av kriterietexterna i sig. Motsvarar de toppen av ett isberg av bakomliggande resonemang eller är det någonting annat som ligger och flyter i vattnet? Svaret på den frågan står ofta att finna i hur kriterietexterna används. Hur vidareutvecklas betygskriterierna? Många styrkor och svagheter i kriterietexter kan i princip bara uppdragas efter en praktisk tillämpning, när de har ställts mot verkliga studenters verkliga examinationsuppgifter. Är det en skrivbordsprodukt som bara plockas fram när en student begär omprövning av ett betyg, som *post hoc*-motivering till intuitivt gjorda bedömningar? När ett lärarlag diskuterar relationen mellan kvalitetskrav och genomströmningskrav, relaterar de då till kriterietexterna? När en student frågar om vad som kommer på tentan, utgör kriterietexterna ett viktigt underlag för lärarens svar? Svaren på dessa frågor avgör om betygskriterierna är "bra" eller inte.

... och ett varnande finger

Det är viktigt att bevaka så att det som var avsett som ett hjälpmedel inte får oönskade effekter. Det gäller såväl feedbackmallar som betygskriterier och förväntade studieresultat.³⁶ Vad händer om en student tar sig an detaljerade anvisningar med avsikten att "vara till lags"?³⁷ Vad blir de eventuella styreffekterna? Hur stor är risken att studierna blir alltför instrumentella eller alltför fragmenterade? Hur förhåller sig tydlighet från lärarens sida till studentens självständiga och kritiska tänkande?³⁸ Har det som skulle ge en ökad transparens gjort utbildningen helt "färdigtuggad"?³⁹ Ibland står det mellan förutsägbarhet och godtycke. Andra gånger mellan tydlighet och övertydlighet. Den här typen av frågor har inte några enkla svar, men det gör dem inte mindre viktiga nu när betygsreformen ska implementeras. Till väldigt stor del kan den enskilda läraren påverka hur tungt vågskålen med fördelar väger i förhållande till vågskålen med nackdelar. Med det följer både handlingsutrymme och ansvar. Om det blir bra eller dåligt bestämmer vi alltså själva.

Metoder för att väga samman till ett betyg

Att välja bedömningsområden och kvaliteter

Så långt arbetet med att gradera kvaliteter efter en dimension, i figur 7 uttryckt som kvaliteter av hur väl ett enda lärandemål har uppnåtts. I verkligheten är de flesta delkurser och kurser sammansatta av ett antal förväntade studieresultat (lärandemål), vilket kan ses som ett uttryck för att vi vill bedöma prestationernas kvalitet efter mer än en parameter. Ofta är utgångspunkten i formulerandet av förväntade studieresultat just en analys av vad som är möjligt att behärska olika väl. När det gäller arbetet med att formulera kriterier i sju steg är detta både en komplikation och en lösning på ett antal problem. Mångfacetterade bedömningar är komplexa, men de öppnar samtidigt för ett antal alternativ när det gäller att fånga en flergradig bedömning.

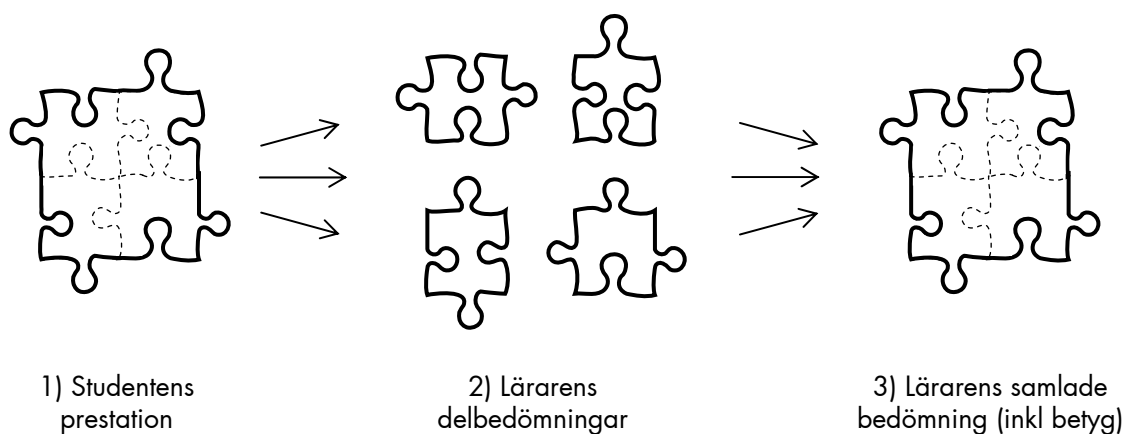
I ett första steg kan en lärare välja att formulera de förväntade studieresultaten för en kurs/delkurs som motsvarar de perspektiv som anses fruktbara att skilja på analytiskt. Studentens faktakunskaper kanske behöver bedömas skilt från samma students förmåga att tillämpa en viss metod, och så vidare. Med ett sådant betraktelsesätt kan man tänka sig en serie av frågor som leder fram till en fingraderad bedömningsmetod med tillräckligt god validitet:

- Vilka bedömningsområden *kan* och *bör* beaktas var för sig? (Lärandemålsvis, per examinationsform eller annat?)
- Hur *kan* olika kvaliteter av dessa bedömningsområden bedömas? (Hur många kontinuerliga och/eller diskreta gränsdragningar *går* att urskilja?)
- Hur många meningsbärande gränsdragningar *bör* göras inom dessa bedömningsområden? (Vilka kontinuerliga och/eller diskreta gränsdragningar är *relevanta* att urskilja i det här sammanhanget?)
- Hur ska dessa delbedömningar vägas samman till ett betyg för delkursen? Vad väger tyngst? Tröskelvärden? Andra principer?

Flera lärandemål och bedömningsområden men bara ett betyg...

Om delbedömningar görs, oavsett själva utformningen av kriterierna, ställs läraren alltid inför frågan hur olika aspekter ska vägas ihop till ett enda betyg. (Även om en global helhetsbedömning görs ingår ofta mer eller mindre explicita delbedömningar i processen, då olika aspekter av en prestation analyseras separat.) Schematiskt kan bedömningsarbetet beskrivas som att (1) studenten gör en komplex prestation, som sedan (2) analyseras med hjälp av olika delbedömningar, vilka renderar ett samlat omdöme (3) uttryckt som ett betyg.

Figur 8. Komplex prestation – delbedömningar – samlad bedömning (inklusive betyg)



Pusselmetaforen ovan har begränsningar som kan vara viktiga att beakta innan vi för det resonemanget vidare. Delbedömningar utgörs sällan av distinkt åtskilda ”pusselbitar”, utan kan ibland lika gärna betraktas som skilda perspektiv på ett och samma fenomen. Oavsett vilket, ligger det i examinatorns uppgift att inte förlora greppet om helhetsbilden när bedömningsarbetet delas upp på något sätt. I avsnittet nedan diskuteras olika metoder för att väga samman sådana delbedömningar till ett betyg. För tydlighetens skull illustreras det i exemplet och i figurerna 9–12 med betygskriterier uppdelade per bedömningsområde. Det principiella resonemanget är dock precis lika giltigt för betygskriterier som skrivs i ett globalt format. Observera att presenterade illustrationer syftar på principiella överväganden med avseende på sammanvägning. De är alltså inte i första hand avsedda som förslag till utformning av kriteriemallar.

Vad ska avgöra betyget?

Om pusselbitarna i figur 8 ovan får symbolisera skilda lärandemål kan betygskriterier skrivas i en matris där dessa lärandemål utgör bedömningsområden. Om vi tänker oss att en lärare utgår ifrån tre sådana lärandemål och att dessa både kan och bör bedömas i lika många steg som nivåer i betygsskalan, så blir det $3 \times 7 = 21$ fält som ska fyllas med meningsfulla texter. Det kan låta mycket i sig – men det som skapas är en långt mer fingeraderad mall än så. Antalet möjliga studentprestationer som kan översättas till matrisens 21 texter uppgår faktiskt till $7 \times 7 \times 7$, det vill säga 343 kombinationsmöjligheter...

Figur 9. Tre delbedömningar i sju steg (Ex: betyg=?)

	Bedömnings- område 1	Bedömnings- område 2	Bedömnings- område 3
A	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
B	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
C	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
D	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
E	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
Fx	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
F	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
”Delbedömningarna ovan omvandlas till ett samlat betyg på delkursen genom att...???”			

Uppgiften är att kategorisera studenters prestationer i sju grader men studenternas möjliga prestationer är uppenbart mer variationsrik än så. (Med fem bedömningsområden uppgår antalet möjliga kombinationer i en matris som ovan till 16807!) Frågan är då hur man ska hantera denna extrema variationsmöjlighet? Vad ska vi göra med markeringarna A-D-C i betygskriteriematrisen ovan? Vad hade hänt om studenten istället hade presterat C-D-A, eller D-A-C? Den svårigheten får många att börja överväga matematiska lösningar, där olika delbedömningar tillskrivs ett numeriskt värde. Problemet är att effekten lätt blir att examinatorn då överlåter steget mellan delbedömningar och betyg åt miniräknaren. I ett slag kan examinatorn ha frånhänt sig möjligheten att kunna svara på frågan vad den som får betyget B kan som den som får betyget C eller D inte kan. ”Jasså, det hela var en B-prestation! Det hade jag ingen aning om.” I värsta fall kan den matematiska sammanräkningen producera ett betyg som strider mot examinatorns helhetsbedömning. ”Jaha, det blev ett B! Vad synd, det är ju inte rättvisande...” Om examinatorns helhetsbedömning får stå tillbaka för en godtyckligt utformad

Sammanräkningsteknik är det inte ens säkert att kvalitetsgränsen för godkänt kan garanteras. Därför behöver metoderna utformas så att de både är förutsägbara och producerar betyg som svarar mot examinatorns faktiska uppfattning.

Alternativ 1 – att låta det viktigaste avgöra

I vårt exempel ovan hade läraren valt att göra en sjugradig bedömning av alla bedömningsområden. Det skapade som vi såg ett långt mer fingeraderat system än sju grader. En metod för att komma fram till ett samlat betyg kan vara att enbart fingeradera det viktigaste bedömningsområdet. (Se exempel II, IV och VI i bilaga A nedan för två praktiska tillämpningar.) På så sätt kan också antalet gränsdragningar som vi behöver göra hållas nere till ett mer rimligt antal.

Figur 10. Det viktigaste bedömningsområdet avgör betyget (Ex: betyg=B)

	Bedömningsområde 1	Bedömningsområde 2	Bedömningsområde 3
A	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
B	"Text, text, text, text, text, text, text ..."		
C	"Text, text, text, text, text, text, text ..."		
D	"Text, text, text, text, text, text, text ..."		
E	"Text, text, text, text, text, text, text ..."		
Fx	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
F	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."

"Lärandemål 1 avgör betyg A–F på delkursen. För A krävs högsta omdöme även under lärandemål 2. (I övrigt beaktas lärandemål 2 i förstone vid eventuella gränsfall.) För godkänt på kursen krävs minst E på samtliga lärandemål."

Metoden att låta ett bedömningsområde skrivas som det viktigaste och helt avgöra betyget underlättar på många sätt. Det förutsätter förstås att just det viktigaste lärandemålet (eller annat bedömningsområde) är möjligt att fingeradera i sju grader. Ibland är det inte så. Om bedömningsområdena är indelade per examinationsform och exempelvis ett större projektarbete bedöms med få grader och en mindre viktig tenta med sju grader kan det få oönskade effekter. Det fingeraderade uppfattas lätt som viktigast.

Alternativ 2 – att göra grövre graderingar i flera steg

Ett snarlikt sätt att arbeta kan vara att gradera det viktigaste bedömningsområdet i tre grader, och sedan ta hjälp av det näst viktigaste för att sammantaget åstadkomma en fingeradering. I figur 11 nedan har lärandemål 1 graderats i tre kvaliteter ovanför gränsen för godkänt. För A krävs högsta omdöme under både lärandemål 1 och 2. I alla andra fall är omdömet under lärandemål 1 utgångspunkt och omdömet under lärandemål 2 faller sedan avgörandet om det ska bli det högre eller lägre av de två möjliga betygen. (B-C) under lärandemål 1 kan alltså utmynna i ett B om sedan lärandemål 2 ges omdömet (A-B) och i ett C om lärandemål 2 ges omdömet (C-E).

Det här alternativet är relativt enkelt att arbeta med och framstår dessutom som lättbegripligt. Som alla andra alternativ kan det dock bara fungera om bedömningsområden och kriterietexter är utformade för att fungera tillsammans med sammanräkningsmetoden. Att det skapas en behändig teknik räcker inte.

Betygskriterierna måste ställas mot alla möjliga utfall. Är verkligen bedömningsområde 2 i det här exemplet så underordnat bedömningsområde 1 att det bara ska vara utslagsgivande? Vill läraren verkligen bara bedöma lärandemål 3 i godkänt och underkänt? Om svaret på någon av frågorna är nej kan metoden inte användas, med mindre än att den kan leda till betyg som strider mot examinatorns värderingar.

Figur 11. Grövre bedömningar i två led (Ex: betyg=C)

	Bedömningsområde 1	Bedömningsområde 2	Bedömningsområde 3
A	(A) "Text, text, text, text, text, text, ..."	(A-B) "Text, text, text, text, text, text, text ..."	
B	(B-C) "Text, text, text, text, text, text, text ..."		"Text, text, text, text, text, text, text ..."
C		(C-E) "Text, text, text, text, text, text, text ..."	
D	(D-E) "Text, text, text, text, text, text, text ..."		
E			
Fax	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
F	"Text, text, text, text, text, text, text ..."		

"Lärandemål 2 är sekundärt i förhållande till lärandemål 1. För A på kursen krävs högsta omdöme under både lärandemål 1 & 2. I övrigt fungerar lärandemål 2 somutslagsgivande. (B-C) och (A-B) ger B. (B-C) och (C-E) ger C. (D-E) och (A-B) ger D. (D-E) och (C-E) ger E. För godkänt på kursen krävs godkänt på samtliga lärandemål."

Alternativ 3 – att låta lägsta eller högsta del prestationen avgöra

Ytterligare en metod går ut på att skriva betygskriterierna så att varje bedömningsområde rymmer beskrivningar av en lägsta acceptabel nivå för varje betygsgrad. (Se exempel III och V i bilaga A nedan för praktiska tillämpningar.) Det lägsta omdömet ger då betyget på kursen/delkursen.

Figur 12. En students lägsta delbedömning avgör betyget (Ex: betyg=D)

	Bedömningsområde 1	Bedömningsområde 2	Bedömningsområde 3
A	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
B	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
C	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
D	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
E	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
Fx	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."
F	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text ..."

"Betygskriterierna syftar på lägsta acceptabla prestation för varje betygsgrad, uppdelat per bedömningsområde. Det lägsta delomdömet ger betyget på kursen."

Om bedömningsområdena är lika viktiga och ämnet som kursen behandlar dessutom kräver en hög lägstanivå kan det vara ett alternativ som kan övervägas. I normalfallet torde det förutsätta att bedömningsområdena verkligen beskriver aspekter av ämnet som måste behärskas i sin helhet, och att just lägstanivån är det som är intressant. I väldigt många fall är det förstås inte så. Vill vi verkligen att en prestation som skulle motsvara A-A-E i figur 12 ovan ska ge samma betyg som prestationen E-E-E? (Båda prestationerna skulle ge betyget E på kursen...)

En närbesläktad metod är att istället låta den *högsta* prestationen avgöra betyget. Med utgångspunkt i ett antagande att det är den bästa prestationen som motsvarar studentens ”verkliga” kapacitet, kan ett antal sämre prestationer kanske betraktas som ovidkommande så länge samtliga lärandemål kan garanteras. Det kan till exempel vara aktuellt om betygskriterierna är uppdelade per eximationsform, och samtliga examinationsformer täcker in samtliga lärandemål. Det kanske inte är så vanligt, men om breddkunskaper och jämna prestationer inte har någon betydelse kan det vara ett alternativ att överväga. Samma typ av fråga måste emellertid ställas till den här varianten: Vill vi verkligen att prestationen A-A-A ska ge samma betyg som A-E-E? (Båda prestationerna skulle ge betyget A på kursen...)

Alternativ 4 – att göra en helhetsbedömning

Figur 13. Helhetsbedömningen avgör (Ex: betyg=C)

	Bedömnings- område 1	Bedömnings- område 2	Bedömnings- område 3
A	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
B	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
C	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
D	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
E	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
Fx	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
F	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”	”Text, text, text, text, text, text, text ...”
<p>”Betyg för kursen bygger på en helhetsbedömning med utgångspunkt i delbedömningarna ovan. Lärandemål 1 och 2 väger tyngst i den bedömningen. Betygskriterierna ska ses som typprestationer. Examinatorn avgör om enstaka styrkor och svagheter kan kompensera varandra.”</p>			

Det säkraste sättet att garantera att examinatorns samlade bedömning inte övertrumfas av någon sammanräkningsteknik är att låta just helhetsomdömet avgöra betyget. (Se exempel I i bilaga A nedan för en praktisk tillämpning.) Delbedömningarna ses då mer som ett beslutsunderlag inför helhetsomdömet än några definitivt fastställda ingångsvärden. Metoden är i princip den som används i de flesta sammanhang där någonting komplext ska bedömas och värderas: När sakkunniga skriver ett utlåtande om en ansökan till en tjänst; när en vetenskaplig tidskrift väljer mellan refusering och publicering; när Högskoleverket bedömer examensrättsprövningar; när Vetenskapsrådets bedömaregrupper tar ställning till projektansökningar... På förhand bestämda kriterier, vare sig de har nedteknats eller inte, ligger till grund för ett samlat omdöme: placering i en tåggrupp, publicering, avslag på examensrätt, beviljade forskningsmedel. Få skulle i de sammanhangen argumentera för att olika matematiska sammanräkningar skulle kunna ersätta metoden att sakkunniga gör helhetsbedömningar med stöd av förutbestämda kriterier.

En invändning som kan göras mot metoden är att den kan framstå som mindre tydlig och mindre förutsägbar för utomstående. Vad som i detalj ligger bakom examinatorns helhetsomdöme är på ett sätt mindre synligt. Har då en ökad validitet skett på bekostnad av en minskad reliabilitet?⁴⁰ Kanske till viss del, men ett motargument är att de komplexa bedömningar som görs inte kan beskrivas på annat sätt. I så fall skulle en minskad ”tydlighet” i själva verket motsvara en mer rättvisande bild av den verkliga bedömningsprocessen.⁴¹

Kriterier som i förstone bygger på helhetsbedömningar kan ofta med fördel skrivas i ett globalt format. På så sätt kan det framgå tydligare att de inte är avsedda att läsas som någon checklista utan som beskrivningar av *typiska* prestationer för varje betygsgrad. Betygskriterierna i figur 13 ovan kan skrivas ihop till ett globalt format som brödtext eller i punktform. Det förra signalerar tydligare att det är en samlad kvalitet som beskrivs. Dessutom är det då inte lika nödvändigt att konsekvent behandla alla bedömningsområden explicit i alla kriterietexter. Det kan ha fördelar, men om tanken är att läsaren ska kunna jämföra enskildheter i kriterierna och de har utformats mer systematiskt, kan det vara tydligare att skriva dem i punktform. Det blir då något av ett mellanting mellan globala kriterier och kriterier uppdelade per bedömningsområde.

Figur 14. Globala betygs-kriterier som brödtext respektive i punktform

A	”Text, text...”	A	1. ”Text, text, text, text, text, text, text...” 2. ”Text, text, text, text, text, text...” 3. ”Text, text, text, text, text, text...”
B	”Text, text...”	B	1. ”Text, text, text, text, text, text, text...” 2. ”Text, text, text, text, text, text...” 3. ”Text, text, text, text, text, text, text...”
C	”Text, text...”	C	1. ”Text, text, text, text, text, text, text, text...” 2. ”Text, text, text, text, text, text...” 3. ”Text, text, text, text, text, text, text...”
D	”Text, text...”	D	1. ”Text, text, text, text, text, text, text, text...” 2. ”Text, text, text, text, text, text...” 3. ”Text, text, text, text, text, text, text...”
E	”Text, text...”	E	1. ”Text, text, text, text, text, text, text, text...” 2. ”Text, text, text, text, text, text...” 3. ”Text, text, text, text, text, text, text...”
Fx	”Text, text...”	Fx	1. ”Text, text, text, text, text, text, text, text...” 2. ”Text, text, text, text, text, text...” 3. ”Text, text, text, text, text, text, text...”
F	”Text, text...”	F	1. ”Text, text, text, text, text, text, text, text...” 2. ”Text, text, text, text, text, text...” 3. ”Text, text, text, text, text, text, text...”
<p>”Betygskriterierna ska läsas som typprestationer. I det enskilda fallet utgår examinatorn ifrån ett helhetsomdöme, där enstaka styrkor och svagheter kan kompensera varandra. Kriterierna bygger på följande bedömningsområden: Lärandemål 1... Lärandemål 2... Lärandemål 3... Lärandemål 1 och 2 väger tyngst i bedömningen.”</p>		<p>”Betygskriterierna ska läsas som typprestationer. I det enskilda fallet utgår examinatorn ifrån ett helhetsomdöme, där enstaka styrkor och svagheter kan kompensera varandra. Kriterierna bygger på följande bedömningsområden: Lärandemål 1... Lärandemål 2... Lärandemål 3... Lärandemål 1 och 2 väger tyngst i bedömningen.”</p>	

Lärandemål, bedömningsområden, kriterier och sammanräkningsmetod

Vid där laget torde det ha framgått tydligt att valet av sammanräkningsmetod har konsekvenser för hur kriterietexterna ska skrivas. Om till exempel sammanräkningsmetoden bygger på att den lägsta prestationen avgör betyget så måste förstås kriterietexterna utformas så att de beskriver just lägstanivåer för att fungera. Lika klart är det att valet av utformning av betygs-kriterierna kan få konsekvenser för

lärandemålen utformning om man väljer att arbeta med lärandemålen som bedömningsområden. Det kan till exempel bli nödvändigt att hålla nere antalet lärandemål för att inte få ohanterligt komplexa betygskriterier. Ibland går det att göra på ett bra sätt, ibland innebär det att lärandemålen försämras. Då gäller det att göra en klok prioritering. Lärandemålen kvalitet är i högsta grad överordnad! Väl fungerande lärandemål fungerar som stöd och styrmedel för hela utbildningens upplägg och genomförande. Att kompromissa med väl fungerande lärandemål bara för att de ska kunna pressas in i en behändig mall för betygskriterier, tillsammans med en viss sammanräkningsmetod, kan vara att göra sig själv en björntjänst. I sådana fall kan det finnas anledning att utforma andra bedömningsområden specifikt för kriterierna. (Se exempel V och VI i bilaga A nedan för två praktiska tillämpningar.) Det viktiga är då att de är härledda ur lärandemålen, och inte talar emot dessa.

Låt säga att en lärare har kommit fram till sex väl fungerande och uttömmande lärandemål. Samtidigt vill läraren arbeta med betygskriterier i en matris uppdelat per bedömningsområde, och använda sammanräkningsmetoden i alternativ 2 ovan där grövre bedömningar i två steg ska åstadkomma en sjugradig bedömning. (Det kan vara ett gemensamt format som ett arbetslag eller en institution har enats om.) Att hålla fast vid de sex lärandemålen som separata bedömningsområden förutsätter då att fyra av dem endast bedöms tvågradigt, godkänt respektive underkänt, vilket kanske inte alls är vad läraren vill. Ett alternativ är att minska antalet lärandemål och göra dem mer övergripande, men i det här fallet skulle det troligen innebära en försämring. Ett bättre alternativ är då att behålla de sex lärandemålen, och utarbeta tre nya bedömningsområden för betygskriterierna. Ett ganska smidigt sätt att göra det på är att se om lärandemålen går att samla ihop i tre olika grupper som fungerar ihop med sammanräkningsmetoden: De viktigaste lärandemålen som graderas tregradigt ovanför godkänt, ett eller ett par lärandemål som bedöms tvågradigt och ett eller ett par som inte graderas alls vid godkänt.

Figur 15. Exempel på lärandemål som samlas ihop till färre bedömningsområden (Ex: betyg=D)

	Bedömningsområde 1 (Lärandemål 1, 2 & 3)	Bedömningsområde 2 (Lärandemål 4)	Bedömningsområde 3 (Lärandemål 5 & 6)
A	(A) "Text, text, text, text, text, text, ..."	(A-B) "Text, text, text, text, text, text, text ..."	
B	(B-C) "Text, text, text, text, text, text, text ..."		
C		(C-E) "Text, text, text, text, text, text, text ..."	
D	(D-E) "Text, text, text, text, text, text, text ..."		"Text, text, text, text, text, text, text ..."
E			
Fx	"Text, text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text, text ..."	"Text, text, text, text, text, text, text, text ..."
F	"Text, text, text, text, text, text, text, text ..."		
"Bedömningsområde 2 är sekundärt i förhållande till bedömningsområde 1. För A på kursen krävs högsta omdöme under både bedömningsområde 1 & 2. I övrigt fungerar bedömningsområde 2 utslagsgivande. (B-C) och (A-B) ger B. (B-C) och (C-E) ger C. (D-E) och (A-B) ger D. (D-E) och (C-E) ger E. För godkänt på kursen krävs godkänt på samtliga lärandemål."			

Det ovan beskrivna exemplet är bara en av ett närmast oändligt antal varianter på tekniska lösningar. Det som är svårt och viktigt har dock väldigt lite att göra med teknikaliteter. Med sunt förnuft går det ofta att komma fram till bra lösningar, så länge det som är viktigast inte kommer bort i processen: att betygskriterierna ska motsvara de värderingar som examinatorn de facto använder i sin betygssättning, och att utformning/sammanräkningsmetod inte får kunna leda till ett betyg som strider mot examinatorns helhetsbedömning.

Detaljerad feedback och betyg – en exkurs

Betyg som feedback? Två problemområden

Som diskuterades i de inledande avsnitten kan ett målrelaterat betyg och dess motivering fylla funktionen av återkoppling på studentens lärande. Sådan återkoppling eller feedback kan jämföras med undervisning och förbättrar kvaliteten på lärandet. (Kriteriebaserade betyg som inte är målrelaterade har en betydligt sämre potential i det avseendet, för att inte tala om relativa betyg.) För optimal effekt bör målrelaterad återkoppling dels vara tillräckligt detaljerad, dels ske i nära anslutning till lärtillfället.⁴² Betyg som feedback brister en del i båda dessa avseenden, brister som behöver hanteras.

Betyget är ett samlat omdöme om komplexa prestationer. Det gör det svårt att enbart med ledning av betyget förstå precis vad som gick rätt och vad som gick fel. En doktorand som i en recension får veta att avhandlingen är stark i teoridelen, har ett bra språk men uppvisar smärre brister i empirin, kan lära sig betydligt mer av detta än av ett ”godkänt” från betygsnämnden. En forskare som får veta att en projektansökan har brister i redovisningen av det internationella forskningsläget kan lära sig betydligt mer av detta än av ett okommenterat avslag. På samma sätt behöver våra studenter en detaljerad återkoppling utöver det samlade omdöme som betyget ger.

Om studenten får en utförlig kommentar/motivering från läraren tillsammans med betyget är mycket vunnet men problemet med tidsaspekten kvarstår.⁴³ Föreställ dig att en av dina studenter i mitten av en femveckorskurs kämpar med att förstå en viss princip inom ämnet. Under efterföljande veckor övergår studierna till att handla om andra delar av kursen. I slutet av kursen ges en salstenta som studenten får tillbaka ett par veckor därefter, med ditt beslut om betyg och tillhörande motivering. Till att börja med finns det ett stort problem med själva avståndet i tiden. Det har nu gått över en månad sedan studenten försökte förstå den aktuella principen. (Tentasvaret i den delen var kanske inte mycket mer än ett försök att rekonstruera tankarna från delkursens mitt.) När betyget och motiveringen kommer studenten till del är hon/han dessutom mitt uppe i nästa delkurs, vilket ytterligare minskar relevansen och därmed effekten. En lärare kan på det sättet ha ägnat stor möda åt att ge återkoppling som inte får någon betydelse eftersom den kommer vid fel tidpunkt.

Metoder för att göra återkopplingen mer relevant i tiden

Återkoppling förutsätter att någon har gjort en individuell bedömning av ett lärande. Sådana pedagogiska bedömningar kan delas in med utgångspunkt i dess syften, respektive vem som utför dem. Summativa bedömningar syftar till att ge omdömen om en avslutad prestation. Det kan ske i slutet av en kurs eller vara slutomdömen om kontinuerligt inlämnade examinationsuppgifter. Formativa bedömningar har ofta en diagnostiserande karaktär och syftar till att skapa omdömen som studenten kan använda sig av under lärprocessens gång. Vad har jag begripit? Vad har jag missförstått? Vad ska jag kraftsamla på för att förstå bättre?

Tabell 2. Pedagogisk bedömning i olika former – exempel på återkoppling på det egna lärandet

	Läraren	Studenten	Studiekamraterna
Formativ Bedömning	<ul style="list-style-type: none">• Kommentarer på övningsuppgifter•Handledning	<ul style="list-style-type: none">• Självstudier• Strukturerad självvärdering	<ul style="list-style-type: none">• Informella samtal• Strukturerade kommentarer
Summativ Bedömning	<ul style="list-style-type: none">• Betygsättning• Kommentarer på prov	<ul style="list-style-type: none">• Strukturerad självvärdering	<ul style="list-style-type: none">• Opposition och likn.

När det gäller vem som utför bedömningen ser vi det traditionellt som lärarens uppgift att ensam stå för all återkoppling, ofta i form av kommentarer på en avslutad tentamen. Betygsättning som summativ bedömning är visserligen enbart examinatorns ansvar, men studenten själv (”self assessment”) och

studiekamrater ("peer assessment") är ofta underutnyttjade resurser när det gäller bedömning i en vidare bemärkelse.⁴⁴ Om en mängd olika vägar används för att möjliggöra en återkommande feedback skapas inte heller samma uppdämda behov av väldigt utförliga kommentarer i efterhand. De potentiella pedagogiska vinsterna med ett vidgat perspektiv på bedömning och återkoppling är alltså stora. Dessutom kan det innebära att lärarens begränsade resurser används på ett bättre sätt.

Metoder för att ge en mer detaljerad återkoppling

Alla dessa typer av bedömning kan både underlättas och förbättras om läraren kan skapa väl fungerande arbetsformer.⁴⁵ En utvecklad kompetens i att göra målrelaterade bedömningar, och utveckla tillhörande betygskriterier, kan vara en bra utgångspunkt i det arbetet. De kriterierna ska ju emellertid fylla en mängd behov och är därför inte alltid optimerade just för en detaljerad återkoppling till studenten. En väg att gå är att komplettera kriterierna med separat utformade feedbackmallar. Sådana mallar kan även användas som ett underlag för läraren i arbetet med att matcha en prestation med en viss nivå i betygskriterierna. I sådana mallar kan sedan läraren (eller andra) göra fria kommentarer med hänvisningar till studentens inlämnade uppgifter.

Feedbackmallar och rättningsmallar kan utformas betydligt friare än formella betygskriterier, inte minst därför att de inte behöver generera ett enda samlat omdöme. Bedömningsområden behöver inte vägas på guldvåg, och de behöver inte heller fungera som en direkt precisering av de förväntade studieresultaten. Givetvis måste även dessa mallar utformas så att de inte ger en missvisande bild av vad läraren ser som viktigt, men begränsningarna är ändå få i jämförelse. Om läraren vill markera att feedbackmallen inte ska läsas som en direkt förlängning av betygskriterierna utan mer som ett sätt att ge struktur på återkopplingen, kan det vara nödvändigt att utforma mallen så att den inte ens påminner om kriterier. Om siffervärden eller olika kvalitativa texter används kan det till exempel vara klokt att använda ett annat antal gränsdragningar än antalet betygsgrader. (Annars är risken stor att studenter och andra förväntar sig att feedbackomdömena på något sätt ska kunna räknas ihop till ett betyg.) Teknikerna som kan användas i utformningen av feedbackmallar är till exempel: dikotoma checklistor, kontinuerliga VAS-skolor, likertskalor och kvalitativa beskrivningar (se tabell 3 nedan)⁴⁶.

Tabell 3. Exempel på tekniker för att utforma feedbackmallar

Dikotoma checklistor

"Redovisar egna resultat tydligt": ja/nej

"Kan med säkerhet utföra beräkningar av typen X": ja/nej

"Korrekt formalia": ja/nej

Kontinuerliga skalor

"Arbetet visar på mycket goda faktakunskaper"

Instämmer helt Instämmer inte alls

← X →

"Resonemanget är..."

Mycket självständigt Osjälvständigt

← X →

Likertskalor

"Arbetet visar på mycket goda faktakunskaper"

Instämmer helt Instämmer inte alls

"Resonemanget är..."

Mycket självständigt Osjälvständigt

Kvalitativa beskrivningar

Språk och formalia

Mycket välskrivet. Stor stilistisk medvetenhet	Korrekt skrivet. Viss stilistisk medvetenhet	Vissa brister i språklig framställning.	Stora språkliga brister. Svårt att följa diskussionen.
Stor noggrannhet i akribi	Noggrannhet i akribi	Fullgod akribi	Bristar i akribi
Mycket god beläggsteknik, felfritt	God beläggsteknik, tydligt	Fullgod beläggsteknik, vissa brister	Inkonsekvent, bristfällig beläggsteknik
Oförvitlig käll- och litteraturförteckning	Tydlig käll- och litteraturförteckning	Bristfällig käll- och litteraturförteckning	Käll- och litteraturförteckning saknas

Noter

¹ Begreppen förväntade studieresultat och lärandemål är synonyma. Den formella beteckningen vid Stockholms universitet är förväntade studieresultat. Begreppet lärandemål har fått stor spridning och används bland annat av Höskoleverket.

² *ECTS User's Guide. European Credit Transfer and Accumulation System.* (2005, s. 14 ff).

³ Ett av få undantag är en internationell utbildning vid Malmö högskola, IMER, där relativa ECTS-betyg har införts. En oberoende utvärdering pekar på flera problem med systemet, och förespråkar istället ett målrelaterat betygssystem. Åkesson (2006 s. 31 f).

⁴ Sadler (2005, s. 175–194).

⁵ Kurs avser den storhet som har en egen kursplan. (HF 6 kap. 6, 7 §§). En kurs kan vara kort eller lång, från någon enstaka vecka upp till ett helt läsår. Delkurs avser en del av en kurs som examineras och poängsätts separat men som inte ges som fristående kurs och därför inte har en egen kursplan. (HF 6 kap. 6, 7 §§). En 20 veckor (30 hp) lång kurs kan t ex bestå av fyra delkurser om 5 veckor (7,5 hp) vardera. *Allt resonemang som förs i den här framställningen syftar på den nivå som examineras separat, och därmed ska ha betygskriterier.* Ofta sker detta på delkursnivå och på fristående kortare kurser, men det är även möjligt att examinera större kurser och utbildningar. I sådana fall ska betygskriterier skrivas på den nivån.

⁶ Tholin (2006), Korp (2006), Selghed (2004).

⁷ Rektorsbeslut Stockholms universitet 2007-01-18. *Riktlinjer och tidsplan för Bolognaarbetet vid Stockholms universitet, reviderad version januari 2007.* Dnr SU 50-1075-05.

⁸ Skolverket (2002).

⁹ Se t ex Dahlgren & Fejes (2005), Milberg, m fl (2005), Lindberg-Sand (2003). Frågan har komplicerats av att avgörande nationella och lokala beslut har fattats kontinuerligt, vilket har förändrat förutsättningarna för debatten och forskningen. Innebörden i centrala begrepp som ”Bologna-processen”, ”ECTS-betyg” mm har också förändrats och tycks i dagsläget ha närmast lokala betydelser.

¹⁰ För en historik, se t ex Lundahl (2006, s. 359–404).

¹¹ Skolverket (2004, s. 5).

¹² Selghed (2004).

¹³ Tholin (2006, s. 189).

¹⁴ Korp (2006, s. 211 ff).

¹⁵ En del av den bedömnings- och betygsrelaterade skolforskningen fokuserar på läroplansforskning, nationell kontra lokal styrning, storskaliga prov och liknande. Se t ex Lundahl (2006) och Wikström (2005). Även om många intressanta problemområden lyfts fram torde bara en del av de frågorna vara angelägna för den enskilda universitetslärarens betygsarbete.

¹⁶ Tan & Prosser (2004, s. 267–282). Om examination och bedömning som socialt fenomen, se t ex Filer (2000) och Korp (2006).

¹⁷ Kent Löfgren vid Umeå universitet bedriver ett forskningsprojekt om Bologna-processen och ”ECTS-betyg” ur ett svenskt perspektiv. Projektet planeras bli färdigt under 2007 och omfattar således inte longitudinella studier.

¹⁸ O'Donovan, m fl (2004, s. 325–335).

¹⁹ För en forskningsöversikt, se Korp (2003, s. 126–134).

- ²⁰ Se t ex Carrell (1995, s. 185 ff), Greatorex & Bell (2004 s. 31 ff) och Klein & Taub (2005).
- ²¹ Milberg, m fl (2005, s. 174).
- ²² Rektorsbeslut Mittuniversitetet 2006-10-03, Dnr MIUN 2005/1086. [Min kursiv.]
- ²³ ”Rektor beslutar att: [...] Kursspecifika betygskriterier inte behövs i de fall de ämnes- och/eller nivåspecifika betygskriterierna ger tillräcklig information för varje kurs.” Rektorsbeslut Mittuniversitetet 2006-10-03, Dnr MIUN 2005/1086.
- ²⁴ Rektorsbeslut Stockholms universitet 2007-01-18. *Riktlinjer och tidsplan för Bolognaarbetet vid Stockholms universitet, reviderad version januari 2007*. Dnr SU 50-1075-05.
- ²⁵ Marton & Säljö (1986, s. 74–80), Marton, m fl (1999), Marton & Booth (2000, s. 54–80).
- ²⁶ Biggs (2003, s. 65 ff), Norton (2004, s. 689).
- ²⁷ I en debattartikel mot flergradiga betyg menar Milberg, Bergman & Dahlgren (2006) att ”pedagogisk forskning visar att månggraderade betyg snarast leder till kvalitativt sämre utbildningar” (s. 6). Anförda referenser är dock bara delvis relevanta, där Dahlgren & Fejes (2005) kan ses som den enda empiriska studien som berör just försämrad kvalitet på lärandet som effekt av flergradiga betyg.
- ²⁸ Hounsell, m fl (2005, s. 2 f), Ramsden (1992, s. 167 f), Laurillard (2002, s. 55 ff).
- ²⁹ Se t ex Angelo & Cross (1993) och Juwah m fl (2004).
- ³⁰ Ecclestone (2001, s. 331 f), O’Donovan, m fl (2004, s. 325–335), Wyatt-Smith & Castleton (2005, s. 150 f).
- ³¹ Om olika sätt att arbeta med kriteriebaserad betygssättning, se t ex Knight (2002) & Sadler (2005).
- ³² I litteraturen syftar begreppen “global/holistic” respektive “analytic/atomistic scoring/assessment” ofta på själva bedömningsprocessen, om helhetsbedömning eller delbedömningar är utgångspunkten för värderingen. Se t ex Hout (1990), Bacha (2001) Matzen & Hoyt (2004) och Korp (2003). Begreppen används i det här sammanhanget närmast för att beskriva två principiella sätt att utforma betygskriterierna rent tekniskt.
- ³³ Falchikov (2004, s. 144).
- ³⁴ Meyer & Land (2005, s. 373–388).
- ³⁵ Hounsell (1986, s. 150 ff).
- ³⁶ O’Donovan, m fl (2004).
- ³⁷ Korp (2003, s. 128), Hedin (2006, s. 371).
- ³⁸ Andrade (2005, s. 27 ff).
- ³⁹ Yorke (2002, s. 155–171), Hussey & Smith (2002, s. 220–233), Högskoleverket (2006, s. 63 f).
- ⁴⁰ För en diskussion om validitet-reliabilitet apropå examination i högre utbildning, se t ex Slomp & Fuite (2005), Moss (2003) och Moss (1994).
- ⁴¹ Om valida bedömningsgrunder i högre utbildning över huvud taget går att formulera detaljerat, se t ex Baume (2004), Johnston (2004), Norton (2004) och Knight (2002).
- ⁴² Chickering & Gamson (1987, s. 4), Juwah (2004, s. 11), Hounsell, m fl (2005, s. 2 f).
- ⁴³ Om svårigheten att formulera feedback som uppfattas som läraren har avsett, se Chanock (2000, s. 103 f).
- ⁴⁴ Se t ex Race (2001) och Falchikov (2001).
- ⁴⁵ Andrade (2000, s. 13 ff).

⁴⁶ För en översikt över forskning om bl a kamratvärdering där dessa och andra tekniker har använts, se Falchikov (2005, s. 145 f).

Referenser

- Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57 (5), 13–18.
- Andrade, H. (2005). Teaching with rubrics: the good, the bad, the ugly. *College Teaching*, 53 (1), 27–30.
- Angelo, T. & Cross, P. (1993). *Classroom assessment techniques – A handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.
- Bacha, N. (2001). Writing evaluation: what can analytic versus holistic essay scoring tell us? *System*, 29, 371–383.
- Baume, D., Yorke, M. & Coffey, M. (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment & Evaluation in Higher Education*, 29, 451–477.
- Biggs, J. (2003). *Teaching for quality learning: what the student does* (2nd ed.). Buckingham, UK: SRHE & Open University Press.
- Carrell, P. L. (1995). The effect of writers' personalities and raters' personalities on the holistic evaluation of writing. *Assessing Writing*, 2, 153–190.
- Chanock, K. (2000). Comments on Essays: do students understand what tutors write? *Teaching in Higher Education*, 5(1).
- Chickering, A. & Gamson, Z. (1987). Seven Principles for Good Practice in Undergraduate Education. *AAHE Bulletin*, 7.
- Dahlgren, L. O. & Fejes, A. (2005). ECTS-skalan: Att mäta eller mota lärande? *Pedagogisk forskning i Sverige*, 10, 24–34.
- Ecclestone, K. (2001). 'I know a 2:1 when I see it': understanding criteria for degree classifications in franchised university programmes. *Journal of Further and Higher Education*, 25 (3).
- ECTS User's Guide. European Credit Transfer and Accumulation System*. Bryssel, 14 february (2005).
- Falchikov, N. (2001). *Learning together. Peer tutoring in higher education*. London & New York: RoutledgeFalmer.
- Falchikov, N. (2005). *Improving Assessment Through Student Involvement: Practical Solutions For Aiding Learning in Higher and Further Education*. London: Routledge.
- Filer, A. (2000). Introduction. In Filer, A. (Red.) *Assessment. Social practice and social product*. London: RoutledgeFalmer.
- Greatorex, J. & Bell, J. F. (2004). Does the gender of examiners influence their marking? *Research in Education*, 71, 25–36
- Hedin, A. (2006). *Lärande på hög nivå. Idéer från studenter, lärare och pedagogisk forskning som stöd för utveckling av universitetsundervisning*. Uppsala: Uppsala universitet.
- Hounsell, D. (1986). Inläring och uppsatsskrivande. I Marton, Hounsell & Entwistle (Red.) *Hur vi lär*. Stockholm: Rabén & Sjögren.
- Hounsell, D., Hounsell, J., Litjens, J. & McCune, V. (2005). Enhancing guidance and feedback to students: Findings on the impact of evidence-informed initiatives. Conference paper, European Association for Research on Learning and Instruction (EARLI) 11th Biennial Conference.
- Hout, B. (1990). Reliability, validity and holistic scoring: what we know and what we need to know. *College Composition and Communication*, 41, 201–213.
- Hout, B. (1996). Toward a new theory of writing assessment. *College Composition and*

- Communication*, 47, 549–566.
- Hussey, T. & Smith, P. (2002). The trouble with learning outcomes. *Active Learning in Higher Education*, 3, 220–233.
- Hård af Segerstad, H. (Red.). (2005). *Nya villkor för lärande och undervisning – 9:e Universitetspedagogiska konferensen vid Linköpings universitet 17 oktober 2005*. CUL-rapport nr 10.
- Härnqvist, K. (1999). *En akademisk fråga – en ESO-rapport om ranking av C-uppsatser*. Ds 1999:65 Högskoleverket. (2006). *Examination med kvalitet – en undersökning av examinationsförfarandet vid några svenska högskolor*. Rapport 2006:45 R.
- Johnston, B. (2004). Summative assessment of portfolios: an examination of different approaches to agreement over outcomes. *Studies in Higher Education*, 29, 395–412.
- Juwah C., m fl. (2004). *Enhancing student learning through effective formative feedback*. June 2004 - The Higher Education Academy Generic Centre.
- Klein, J. & Taub, D. (2005). The effect of variations in handwriting and print on evaluation of student essays. *Assessing Writing*, 10, 134–148.
- Knight, P. T. (2002). Summative assessment in higher education: practices in disarray. *Studies in Higher Education*, 27.
- Korp, H. (2003). *Kunskapsbedömning – vad, hur och varför*. Stockholm: Myndigheten för skolutveckling.
- Korp, H. (2006). *Lika chanser på gymnasiet? En studie om betyg, nationella prov och social reproduktion*. Diss. Malmö: Malmö Högskola.
- Laurillard, D. (2002). *Rethinking University Teaching – a Conversational Framework for the Effective Use of Learning Technologies* (2nd ed.). London: RoutledgeFalmer.
- Lindberg-Sand, Å. (2003). *Kompetensbedömning eller prestationskontroll? Examination som praktiserad kunskapssyn*. Utvärderingsenheten, Lunds universitet, Rapport nr 2003: 222.
- Lundahl, C. (2006). *Viljan att veta vad andra vet : kunskapsbedömning i tidigmodern, modern och senmodern skola*. Diss. Stockholm: Elander Gotab.
- Marton, F. & Booth, S. (2000). *Om lärande*. Lund: Studentlitteratur.
- Marton, F. Dahlgren, L.O. Svensson, L. & Säljö, R. (1999). *Inläring och omvärldsuppfattning. En bok om den studerande människan*. Stockholm: Prisma.
- Marton, F., Hounsell, D. & Entwistle, N. (Red.). (1986). *Hur vi lär*. Stocholm: Rabén & Sjögren.
- Marton, F. & Säljö, R. (1986). Kognitiv inriktning vid inläring. I Marton, Hounsell & Entwistle (Red.), *Hur vi lär* (s. 56–80). Stocholm: Rabén & Sjögren.
- Matzen, R. N. & Hoyt, J. E. (2004). Basic writing placement with holistically scored essays: research evidence. *Journal of Developmental Education*, 28, 2–34.
- Meyer, J. & Land, R. (2005). Threshold concepts and troublesome knowledge: epistemological considerations and a conceptual framework for teaching and learning. *Higher Education*, 49(3), 373–388.
- Milberg P., Bergman K-O. & Dahlgren L-O. (2006). ECTS-betyg och pedagogisk evidens. *Universitetsläraren*, (1).
- Milberg, P., m fl. (2005). Precision i poängsättning av essäfrågor och examensarbeten: Implikationer för betyg enligt ECTS. I Hård af Segerstad, H. (Red.), *Nya villkor för lärande och undervisning – 9:e Universitetspedagogiska konferensen vid Linköpings universitet 17 oktober 2005*. CUL-rapport nr 10.

- Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23, 5–12.
- Moss, P. A. (2003). Reconceptualizing validity for classroom assessment. *Educational Measurement, Issues and practice*, 22, 13–25.
- Norton, L. (2004). Using assessment criteria as learning criteria: A case study in psychology. *Assessment & Evaluation in Higher Education*, 29 (6).
- O'Donovan, B., Price, M. & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, 9(9).
- Race, P. (2001). A Briefing on Self, Peer and Group Assessment. *LTSN Generic Centre, Assessment Series No 9*.
- Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30, 175–194.
- Selghed, B. (2004). *Ännu icke godkänt: lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Diss. Lund: Lunds universitet.
- Skolverket. (2004). *Likvärdig bedömning och betygssättning*. Stockholm: Statens skolverk.
- Skolverket. (2002). *Att bedöma eller döma. Tio artiklar om bedömning och betygssättning*. Liber: Stockholm.
- Slomp, D. H. & Fuite, J. (2005). Following Phaedrus: alternate choices in surmounting the reliability/validity dilemma. *Assessing Writing*, 9, 190–207.
- Tan, K. H. K. & Prosser, M. (2004). Qualitatively different ways of differentiating student achievement: a phenomenographic study of academics' conception of grade descriptors. *Assessment & Evaluation in Higher Education*, 29.
- Tholin, J. (2006). *Att klara sig i ökad natur – En studie av betyg och betygskriterier – historiska betingelser och implementeringen av ett nytt system*. Diss. Borås: Högskolan i Borås.
- Wikström, C. (2005). *Criterion-referenced measurement for educational evaluation and selection*. Diss. Umeå: Umeå universitet.
- Wyatt-Smith, C. & Castleton, G. (2005). Examining how teachers judge student writing: an Australian case study. *Journal of Curriculum Studies*, 37, 131–154.
- Yorke, M. (2002). Subject benchmarking and the assessment of student learning. *Quality Assurance in Education*, 10, 155–171.
- Åkesson, E. (2006). ECTS-betyg? En studie av erfarenheter inom området IMER vid Malmö högskola. Centre for Educational Development (CED) Lunds universitet, augusti 2006.



Bilaga A – Exempelsamling

[Uppdaterad 2007-01-31]

När betygskriterier diskuteras är behovet av exempel ofta stort. Många principiella resonemang förblir ofta tomma ord utan tillgång till konkreta, autentiska tillämpningar. Samtidigt är det något av ett dubbeleggat svärd. Exempel kan ibland läsas extensivt och tolkas som riktlinjer, vilket inte alltid är avsikten. Dessutom riskerar exempel som ligger nära det egna ämnet att alltför lättvindigt accepteras som förlaga, trots att formatet som valts i exemplet bara är ett av många alternativ när det gäller utformning, bedömningsområden och metod för sammanvägning. Ibland kopieras kriterietexter från exempel utan att ett tillräckligt arbete har lagts på att anpassa dem till den egna kursen/delkursen. I sådana fall kan exempel göra mer skada än nytta.

Nedanstående exempelsamling är en mindre sammanställning avsedd att fungera som enstaka illustrationer snarare än som en systematisk täckning av alla möjliga ämnen och alla möjliga tillvägagångssätt när det gäller utformning av betygskriterier. I Stockholms universitets kurskatalog för 2006/07 listas 1238 olika kurser. Tusentals kurser och delkurser av en hisnande mångfald kommer således att behöva skriftliga betygskriterier i linje med universitetets tidsplan. Om ett par år kommer förhoppningsvis standardiserade former för att tillgängliggöra kriterier och kursbeskrivningar ha utvecklats vid universitetet, så att det blir enkelt för lärare att ta del av alla andras arbete. Först då blir det kanske möjligt att tala om någon fullständig täckning av exempel.

Innehåll

Exempel I) Kandidatuppsats engelsk litteratur (15 hp).....	A-3
Exempel II) Modern kinesiska I (7,5 hp)	A-6
Exempel III) Samtidshistoria ”Teori och kronologi” (7,5 hp)	A-9
Exempel IV) Kandidatuppsats historia (15 hp).....	A-12
Exempel V) Fysik – Grundläggande mekanik (12 hp)	A-15
Exempel VI) Formell språkinläring, läromedel och bedömning av språkfärdighet (7,5 hp)	A-18

Exempel I) Kandidatuppsats engelsk litteratur (15 hp)

Förväntade studieresultat (lärandemål):

Efter genomförd delkurs förväntas studenten kunna...

1. ... genomföra en självständig litteraturvetenskaplig undersökning, med högt ställda krav på analysförmåga och forskningsanknytning
2. ... presentera egna resultat skriftligt med högt ställda krav på argumentation, språklig framställning och formalia.
3. ... presentera egna resultat muntligt och kunna argumentera för dessas giltighet.
4. ... analysera och värdera andras vetenskapliga resultat.

Examinationsformer:

- Egen uppsats, inklusive ventilerung
- Opposition av annans uppsats
- Seminariediskussioner om andras uppsatser

Undervisningsformer:

Seminarieundervisning och individuell uppsatshandledning

Kommentar till betygskriterier och bedömningsområden:

Betygskriterierna är utformade som globala beskrivningar av helhetsomdömen. Med delkursens förväntade studieresultat som övergripande ram används följande bedömningsområden för att avgöra uppsatsens kvalitet:

- Analytisk förmåga
- Forskningsanknytning
- Teoretisk medvetenhet
- Metodval och genomförande
- Språklig utformning (stil, grammatik och vokabulär)
- Formalia (enligt MLA:s standard)

Betygskriterierna nedan ska förstås som samlade beskrivningar av olika kvalitetsnivåer. Helhetsomdömet är viktigast, vilket gör att en enstaka svaghet inom ett bedömningsområde kan vägas upp av en styrka inom ett annat. I den samlade bedömningen väger den egna uppsatsen avgjort tyngst. Opposition och seminariediskussion kan emellertid fälla avgörandet i eventuella gränsfall.

[Exemplet utarbetat av Elias Schwieler, UPC & Engelska institutionen, SU.]

I) Betygskriterier – "Kandidatuppsats engelsk litteratur" (15 hp)

A	Nyskapande problemval och analys genomförd med stor grad av självständighet i förhållande till den litterära texten och tidigare forskning. Stort värde. En mycket övertygande vetenskaplig argumentation. Väl motiverade avgränsningar. Relevant och vittomfattande teorianknytning. Stor tydlighet och logik i disposition. Språklig stringens. Oförvitlig formalia; stor noggrannhet i akribi, beläggsteknik, notapparat, käll- och litteraturförteckning som följer MLA:s standard. Noggrann genomgång av annans C-uppsats, då både förtjänster och tillkortakommanden diskuteras. Mycket relevant prioritering av centrala kontra perifera problem i den uppsatsen.
B	Intresseväckande problemval och analys genomförd med stor grad av självständighet. En i huvudsak övertygande vetenskaplig argumentation. Väl motiverade avgränsningar. Relevant teorianknytning. Stor tydlighet och logik i disposition. Språklig stringens. Oförvitlig formalia; stor noggrannhet i akribi, beläggsteknik, notapparat, käll- och litteraturförteckning som följer MLA:s standard. Noggrann genomgång av annans C-uppsats, då både förtjänster och tillkortakommanden diskuteras. Mycket relevant prioritering av centrala kontra perifera problem i den uppsatsen.
C	Självständigt genomförd analys där frågeställningar, problematiseringar och analys hänger samman. Relevant metod- och teorival och koppling till tidigare forskning. Ett gediget arbete utan betydande tillkortakommanden. Endast ett fåtal mindre språkliga brister. Formalia utan brister; noggrannhet i akribi, beläggsteknik, notapparat, käll- och litteraturförteckning som följer MLA:s standard. Noggrann genomgång av andras C-uppsatser, då både förtjänster och tillkortakommanden diskuteras.
D	Frågeställningar, problematiseringar och analys hänger samman. Relevant metod, men begränsad teorianknytning. Viss koppling till tidigare forskning. Språkliga brister kan förekomma. Formalia utan större brister; akribi, beläggsteknik, notapparat, käll- och litteraturförteckning som följer MLA:s standard. Noggrann genomgång av andras C-uppsatser, då både förtjänster och tillkortakommanden diskuteras.
E	Frågeställningar, problematiseringar och analys hänger mestadels samman. Relevant metod, men begränsad teorianknytning. Svag koppling till tidigare forskning. Språkliga brister kan förekomma. Formalia utan större brister; akribi, beläggsteknik, notapparat, käll- och litteraturförteckning som följer MLA:s standard. Noggrann genomgång av andras C-uppsatser, då både förtjänster och tillkortakommanden diskuteras.
Fx	Lärandemålen uppnås inte. Ett eller flera av följande tillkortakommanden: Frågeställningar och analys hänger inte samman. Irrelevant metod. Bristande teorianknytning. Koppling till tidigare forskning saknas. Otydlig i argumentationen. Stora språkliga brister. Formaliafel. Bristfällig genomgång av andras C-uppsatser: Uppenbara brister och förtjänster förbises. Direkt olämplig prioritering av centrala kontra perifera problem.
F	Lärandemålen uppnås inte. Uppgiften framstår som omöjlig att genomföra inom ramen för handledarledd tid.

Diskussion ang ex I)

En delkurs som kandidatuppsatsen förutsätter en komplex prestation, sammansatt av komponenter som var och en dessutom befinner sig på en hög abstraktionsnivå. I ett sådant fall kan globala helhetsomdömen som noga förankras i lärandemål och förutbestämda bedömningsområden vara ett attraktivt betygskriteriealternativ. Examinatorns arbete motsvarar de kvalitativa bedömningar som görs i många andra liknande sammanhang: När forskningsråd bedömer projektplaner, när sakkunniga bedömer sökande till en tjänst, när en redaktör bedömer ett forskningsmanus eller när forskaren klassificerar ett komplext material. Som i dessa exempel vill man i regel bibehålla den övergripande bilden med ledning av på förhand bestämda bedömningsområden. Dessa bedöms visserligen analytiskt var för sig, men utan att skiljas från helhetsbilden. (Att mekaniskt dela upp delbedömningarna var för sig för att sedan adderas ihop till en slags matematisk slutsumma är ett alternativ som sällan ens övervägs i de sammanhangen.)

Metodens styrka, att behålla fokus på helheten, är samtidigt dess svaghet. Det kan vara svårt att utforma kriterierna så att en lagom stor grad av tydlighet och precision kommuniceras. Kriterietexterna kan lätt bli så allmänna att de inte ger mycket vägledning om vad som egentligen skiljer olika kvalitetsnivåer åt. Å andra sidan bör bedömningar som i verkligheten innehåller många led av svåra avvägningsfall inte beskrivas som någonting enkelt. Största möjliga tydlighet och precision utan att göra våld på en komplex verklighet är en princip som är lätt att acceptera. Att omsätta det i praktiken kan dock kräva ett grannlaga arbete och ett antal revideringar innan det blir tillräckligt bra. Globala betygskriterier formulerade som helhetsomdömen behöver, som alla andra alternativa format, revideras och prövas mot olika prestationer för att kunna bli bra nog. Det arbetet görs med fördel inom ett lärarlag eller genom att den enskilda läraren ber enstaka kollegor om hjälp i utvecklingsfasen.

Ett sätt att hantera ett stort mått av allmängiltighet i globala betygskriterier är att arbeta med mer detaljerade format i andra led av undervisning och bedömning. På så sätt kan man försöka utnyttja det bästa av två världar genom att även arbeta med en mer utvecklad feedbackmall där olika kvaliteter av de olika bedömningsområdena anges. Om helhetsomdömet är överordnat delbedömningarna, är det särskilt viktigt att en sådan feedbackmall inte uppfattas som ett mekaniskt betygssunderlag. Ett sätt att understryka detta kan vara att använda en annan skala än den aktuella betygsskalan. Omdömen som uttrycks som siffror leder ofta tanken till att dessa på något sätt kan räknas samman till ett betyg. Om det inte är avsikten kan det vara klokt att istället arbeta med begrepp som "Utmärkt", etcetera.

Ex: Feedbackmall som komplement till formulerade betygskriterier (markeras med X)

	Utmärkt	Otillräckligt
Analytisk förmåga	←	→
Forskningsanknytning	←	→
Teoretisk medvetenhet	←	→
Metodval och genomförande	←	→
Språklig utformning	←	→
Formalia	←	→

Exempel II) Modern kinesiska I (7,5 hp)

Förväntade studieresultat (lärandemål):

Efter genomförd delkurs förväntas studenten kunna...

1. ... muntligt och skriftligt översätta texter på modern kinesiska till svenska med hjälp av ordbok.
2. ...skriva och översätta ca 500 skrivtecken utöver de ca 1300 tecken som ingick i Kinesiska I-II.

Examinationsformer:

- Salstentamen
- Glostest

Undervisningsformer:

Lärlarledda föreläsningar och seminarier, själv- och kamratbedömning.

Kommentar till betygskriterier:

Procentsatserna avseende salstentamen syftar på grad av korrekthet, där 100 % motsvarar en helt felfri översättning. Slutbetyget A–E på delkursen avgörs av resultatet på salstentamen. För betyget godkänd (A–E) krävs godkänt på båda examinationsuppgifterna. Rättningsmall utdelas vid kursstart.

[Exemplet utarbetat av Lena Rydholm, Institutionen för orientaliska språk, SU.]

II) Betygskriterier – "Modern kinesiska I" (7,5 hp)

	Examinationsuppgift 1 (salstentamen)	Examinationsuppgift 2 (glostest)
A	Texten översatt till minst 90 % rätt (med avseende på att ord och meningar översatts, och översatts korrekt vad gäller betydelse och grammatik och syntax). Ordvalet vid översättningen speglar en djupare förståelse av ordens konnotationer. Översättningen är noga genomtänkt och det svenska språket i översättningen flyter naturligt och avviker inte från normal svensk meningsbyggnad och språkbruk.	Minst 35 av 50 rätt på samtliga fem glostest.
B	Texten översatt till minst 90 % rätt (se A) Ordvalet vid översättningen speglar en djupare förståelse av ordens konnotationer alt. det svenska språket i översättningen flyter naturligt och avviker inte från normal svensk meningsbyggnad och språkbruk.	
C	Texten översatt till minst 80 % rätt (se A) Det svenska språket i översättningen flyter någorlunda naturligt och avviker inte nämnvärt från normal svensk meningsbyggnad och språkbruk	
D	Texten översatt till minst 70 % rätt (se A) Det svenska språket i översättningen flyter någorlunda naturligt och avviker inte alltför mycket från normal svensk meningsbyggnad och språkbruk.	
E	Texten översatt till minst 60 % rätt (se A) Det svenska språket i översättningen flyter någorlunda naturligt och avviker inte alltför mycket från normal svensk meningsbyggnad och språkbruk.	
Fx	Texten översatt till mindre än 60 % rätt (se A). Det svenska språket i översättningen avviker mycket från normal svensk meningsbyggnad och språkbruk.	Mindre än 35 av 50 rätt på ett eller flera av fem glostest.
F	Texten översatt till mindre än 50 % rätt (se A). Det svenska språket i översättningen avviker mycket från normal svensk meningsbyggnad och språkbruk.	

Diskussion ang ex II)

Alternativet att utforma betygskriterier per examinationsuppgift kan vara ett enkelt sätt att koppla dem till lärarens praktiska arbete. Det finns fördelar med detta, bland annat att det kan framstå som tydligt och lättöverskådligt. För många lärare motsvarar det dessutom ett invariant arbetssätt. Nackdelen är att examinationsformerna så att säga låses fast över tid. Det arbete som ofta krävs för att få betygskriterierna att fungera väl förutsätter ett grannliga utvecklingsarbete, följt av ett antal revideringar efter de första genomförandena. Det innebär att det kan vara förenat med ett stort merarbete att byta bedömningsområden, i det här fallet delkursens examinationsuppgifter. Vad händer om läraren finner att examinationsformerna behöver förändras i grunden? Risken är uppenbar att illa fungerande examinationsformer behålls på grund av att ett byte skulle innebära att stora delar av betygskriteriearbetet behöver göras om. Om det finns anledning att tro att examinationsformerna inte kommer att behöva förändras under överskådlig tid är det en risk som man kanske är villig att ta.

I exemplets båda bedömningsområden anges olika siffervärden. Just dessa storheter representerar principiellt viktiga överväganden. I fallet med översättningarna anger procentsatserna graden av korrekthet. Gränsen för C på salstentan anges som en översättning som till 80% är korrekt. Det är någonting i grunden annorlunda än att ange 80% av antalet möjliga poäng på en tenta, det vill säga 80% av en okänd storhet. (En sådan angivelse förutsätter i princip att tentan bifogas betygskriterierna för att det ska vara möjligt att skapa sig en bild av vilka kvaliteter som avses.)

I det andra bedömningsområdet, examinationsform 2, anges siffervärden som absoluta tal. Inte heller i det här fallet ska det blandas samman med metoden att bara ange antalet poäng på en tenta. De absoluta talen motsvarar i det här fallet en viss omfattning av gloskunskaer som det är möjligt att få en bild av. En utomstående kan ha användning för den här uppgiften eftersom det är en siffra som representerar en känd kvalitet. Siffran anger alltså hur väl studenten kan kinesiska glosor, inte bara hur många poäng på en okänd tenta som uppnåddes.

Exempel III) Samtidshistoria "Teori och kronologi" (7,5 hp)

Förväntade studieresultat (lärandemål):

Efter genomförd delkurs förväntas studenten kunna...

1. ... identifiera olika typer av orsaksförklaringar, ändamålsförklaringar, funktionalistiska förklaringar, klassificering, kolligation, beskrivning och karakteristik i historiska framställningar
2. ...analysera och diskutera användandet av dessa former av beskrivningar och förklaringar i olika typer av historiska framställningar, som översiktverk och dokumentära skildringar.

Examinationsformer:

- Individuella inlämningsuppgifter
- Skriftliga gruppuppgifter
- Skriftliga delexaminationer

Undervisningsformer:

Lärlarleda föreläsningar, internetbaserade gruppövningar, filmvisning och seminarier.

Kommentar till betygsriterier och examinationsuppgifter:

Betygskriterierna anger minimikrav för olika betygsnivåer. (För att denna bedömning ska ske förutsätts att kurskraven är uppfyllda.) Kriterierna tar sin utgångspunkt i delkursens lärandemål. Dessa två lärandemål är lika viktiga och om en student presterar olika väl i dessa två avseenden avgörs betyget av den lägsta prestationen. (Den som bedöms ha uppnått A under lärandemål 1 och C under lärandemål 2 får således betyget C på delkursen.)

I ett första led bedöms om prestationen visar att studenten har uppnått lärandemålen, dvs uppnått något av de godkända betygen A–E. I ett andra led avgörs om kvaliteten motsvarar kriterierna för de högre betygen A–B eller betygen C–E. Gränsen mellan A–B respektive C–E kan beskrivas som både kvalitativ och kvantitativ: För de högre betygen krävs *både* en högre grad av analytisk förmåga *och* en större kunskapsbredd, jämfört med vad som krävs för betygen C–E. I ett tredje led bedöms vilket betygsriterium enligt skalan A–E som bäst svarar mot individens prestation.

För A–B krävs att studenten med stor säkerhet kan identifiera ett stort antal teoretiska perspektiv, och föra ett nyanserat och analytiskt resonemang om dessa. Resonemangen kännetecknas av självständighet. Brister och förtjänster i tillämpningar analyseras ingående. För C–E krävs att studenten kan identifiera flertalet av de under kursen behandlade teoretiska begreppen, och föra ett eget resonemang om dessa i olika tillämpningar. Egna resonemang visar på en grundläggande förståelse, ev med mindre missförstånd. Fx–F ges om lärandemålen inte uppnås, dvs att studenten inte kan identifiera flertalet av de under kursen behandlade teoretiska begreppen och inte heller kan föra ett eget resonemang om dem.

Störst vikt vid bedömningsarbetet läggs vid delexaminationerna. De övriga individuella och gruppvisa uppgifterna ingår också i den sammantagna betygsbedömningen, men då främst som komplement till delexaminationerna, t ex i olika gränsfall. Restuppgifter medges inte. Vid betygen Fx eller F ges möjlighet till omprov vid tre tillfällen per läsår.

[Exemplet utarbetat av Stefan Ekecrantz, UPC, SU.]

III) Betygskriterier – Samtidshistoria "Teori och kronologi" (7,5 hp)

A	<p><i>Lärandemål 1</i> Felfri identifikation av teoretiska perspektiv, med mycket relevanta val av exempel.</p> <p><i>Lärandemål 2:</i> Mycket nyanserad analys, med ett resonemang som kännetecknas av stora mått av självständighet. Styrkor och svagheter i tillämpningar diskuteras insiktsfullt.</p>
B	<p><i>Lärandemål 1:</i> Felfri identifikation av teoretiska perspektiv.</p> <p><i>Lärandemål 2:</i> Nyanserad analys, med ett självständigt resonemang. Ett antal styrkor och svagheter i tillämpningar diskuteras.</p>
C	<p><i>Lärandemål 1:</i> I huvudsak korrekt identifikation av samtliga teoretiska perspektiv.</p> <p><i>Lärandemål 2:</i> Analys som visar på en god förståelse, med ett visst självständigt resonemang.</p>
D	<p><i>Lärandemål 1:</i> Mestadels korrekt identifikation av flertalet teoretiska perspektiv.</p> <p><i>Lärandemål 2:</i> Analys som visar på en god förståelse, med ansatser till självständigt resonemang.</p>
E	<p><i>Lärandemål 1:</i> Mestadels korrekt identifikation av ett antal teoretiska perspektiv.</p> <p><i>Lärandemål 2:</i> Analys som visar på tillräcklig förståelse. Eget resonemang, ev med mindre missförstånd.</p>
Fx	<p>Ett eller flera av följande tillkortakommanden presteras:</p> <p><i>Lärandemål 1:</i> Studenten kan identifiera endast ett begränsat antal teoretiska perspektiv.</p> <p><i>Lärandemål 2:</i> Resonemangen om dessa är begränsade och/eller alltför onyanserade. Vissa missförstånd.</p>
F	<p><i>Lärandemål 1:</i> Studenten kan inte identifiera teoretiska perspektiv.</p> <p><i>Lärandemål 2</i> Resonemangen är begränsade. Betydande missförstånd.</p>

Diskussion ang ex III)

De två lärandemålen utgör bedömningsområden och är utgångspunkt för varje kriterietext, även om de inte har skrivits i separata kolumner. Betygskriterierna med tillhörande övergripande beskrivning är utformade så att de återspeglar examinatorns bedömningsprocess. Först görs en bedömning om prestationerna visar att studenterna har uppnått de förväntade studieresultaten och har klarat godkänt (A–E) på delkursen. Därefter gör en grovsortering med utgångspunkt i de diskreta skillnaderna mellan A–B respektive C–E. Först därefter görs en fingeraderad bedömning, som utgörs av en kombination av kontinuerliga och diskreta bedömningar inom dessa två grupper av prestationer.

En fördel med att både arbeta och presentera kriterierna på det här sättet är att det är ganska lätt att kommunicera vad som är viktigast för studenterna. (Väl fungerande betygskriterier tillsammans med övrig dokumentation och fortlöpande diskussion ska ju i bästa fall ersätta studenternas behov av information om lärarens värderingar från mer eller mindre pålitliga källor, som gamla/inaktuella tentor eller studenter från tidigare årskullar.) Under delkursens gång kan läraren med utgångspunkt i ett begränsat underlag ge återkoppling till studenterna var någonstans i den grövre indelningen de befinner sig, till exempel att inlämningsuppgifterna som lämnats in halvvägs in i delkursen tycks ligga någonstans inom spannet C–E. Den grövre skalan kan också användas i olika övningar där studenterna under kursens gång försöker bedöma sina egna eller kamraternas prestationer ("self-" och "peer assessment").

Betygskriterierna är utformade så att det är lägstaprestationen för respektive förväntat studieresultat (lärandemål) som avgör betyget på delkursen. (Jmf alternativ 3 i huvudtexten.) Metoden fungerar i det här fallet av två skäl: För det första täcker båda förväntade studieresultaten in hela kursinnehållet, om än från två skilda perspektiv ("identifiera" respektive "analysera"). För det andra har de bedömts som lika viktiga och den som utformade dem ville inte att en svaghet i det ena perspektivet ska kunna kompenseras av en styrka i det andra. (I praktiken är de dessutom så nära relaterade med varandra att det är ovanligt att en student till exempel är mycket bra på att identifiera men dålig på att analysera samma stoff, även om det i sig är tänkbart.)

Exempel IV) Kandidatuppsats historia (15 hp)

Förväntade studieresultat (lärandemål):

Efter genomförd delkurs förväntas studenten kunna...

1. ... genomföra och presentera en självständig historisk undersökning på grundnivå, med högt ställda krav på empiriskt hantverk, samt teoretisk och metodologisk medvetenhet.
2. ... analysera och värdera andras vetenskapliga resultat.
3. ... presentera egna resultat muntligt och kunna argumentera för dessas giltighet.

Examinationsformer:

- Egen uppsats, inklusive ventilerings
- Opposition av annans uppsats
- Seminariediskussioner om andras uppsatser

Undervisningsformer:

Seminarieundervisning, handledning i smågrupp och individuell uppsatshandledning

Kommentar till betygsriterier och bedömningsområden:

Betygskriterierna är utformade som delbedömningar uppdelade per examinationsuppgift. För betyget A–F väger uppsatsen avgjort tyngst. För betyget A krävs högsta angivna kvalitet på både uppsats och opposition. Vid övriga uppsatsbetyg kan oppositionens kvalitet avgöra delkursens betyg i förstone vid eventuella gränsfall. För godkänt betyg (A–E) krävs godkänt på alla tre examinationsuppgifterna. Med delkursens förväntade studieresultat som övergripande ram har följande bedömningsområden använts för att utforma betygsriterierna för uppsatsmomentet, och kommer tillika att utgöra grund för bedömningen:

- De vetenskapliga resultatens värde (ex. relevans i förhållande till forskningsläge)
- Teoretisk och metodologisk medvetenhet (självständighet och relevans)
- Empirins genomförande och omfattning (källanalys, noggrannhet, akribi)
- Framställningens språkliga kvalitet och formalia (argumentation och stilistisk medvetenhet)

De tre första bedömningsområdena viktas lika. Vad gäller det fjärde området har vetenskaplig argumentation störst betydelse. (Språklig kvalitet och formalia kan i vissa fall påverka ett omdöme negativt, men däremot inte kompensera andra typer av tillkortakommanden.) Samma bedömningsområden ligger till grund för kriterierna för opposition och seminariedeltagande, men syftar då på studentens förmåga att diskutera och bedöma dessa värden i andras arbeten.

[Exemplet utarbetat av Stefan Ekecrantz, UPC, SU.]

IV) Betygskriterier – "Kandidatuppsats historia" (15 hp)

	Examinationsuppgift 1 (Egen uppsats)	Examinationsuppgift 2 (Opposition)	Examinationsuppgift 3 (Seminarier)
A	Nyskapande problemval och/eller tolkning. Nya och värdefulla resultat. En mycket övertygande vetenskaplig argumentation. Omfattande och väl genomförd empiri med väl motiverade avgränsningar. Relevant och vittomfattande teoriansknytning. Stor tydlighet och logik i disposition. Språklig stringens. Oförvitlig formalia; stor noggrannhet i akribi, beläggsteknik, notapparat, käll- och litteraturförteckning.	Mycket noggrann genomgång av annans uppsats. Själständigt genomförd uppgift, med viktiga förbättringsförslag. Relevant prioritering av centrala kontra perifera problem i uppsatsen. Grundlig kontroll av akribi genomförd.	Ett antal andra seminariedeltagares resultat diskuteras självständigt. Förberedda frågor och analyser med god relevans.
B	Intresseväckande problemval och/eller tolkning. En övertygande vetenskaplig argumentation. Nya resultat. Väl genomförd empiri med väl motiverade avgränsningar. Relevant teoriansknytning. Stor tydlighet och logik i disposition. Språklig stringens. Oförvitlig formalia; stor noggrannhet i akribi, beläggsteknik, notapparat, käll- och litteraturförteckning.		
C	Frågeställningar, undersökning och analys hänger samman. Delvis nya resultat. Relevant metodval. Viss koppling till tidigare forskning. Nyanserad källkritisk analys. Formalia utan brister; noggrannhet i akribi, beläggsteknik, notapparat, käll- och litteraturförteckning.	Noggrann genomgång av annans uppsats, då både förtjänster och tillkortakommanden diskuteras. Eventuellt viss sammanblandning av centrala kontra perifera problem. Kontroll av akribi genomförd.	
D	Frågeställningar, undersökning och analys hänger samman. Delvis nya resultat, men av osäkert värde. Relevant metod, men empirins omfattning begränsad. Viss koppling till tidigare forskning. Grundläggande källkritisk analys. Formalia utan större brister; akribi, beläggsteknik, notapparat, käll- och litteraturförteckning.		
E	Frågeställningar, undersökning och analys hänger mestadels samman. Delvis nya resultat, men av begränsat värde. Relevant metod, men empirins omfattning begränsad. Svag koppling till tidigare forskning. Grundläggande källkritisk analys. Formalia utan större brister; akribi, beläggsteknik, notapparat, käll- och litteraturförteckning.		
Fx	Ett eller flera av följande tillkortakommanden: Frågeställningar, undersökning och analys hänger inte samman. Inga nya resultat. Irrelevant metod. Bristande omfattning på empirin. Koppling till tidigare forskning saknas. Grundläggande källkritisk analys saknas. Otydlig i argumentationen. Formaliafel.	Bristfällig genomgång av annans C-uppsats. Uppenbara brister och förtjänster förbises. Direkt olämplig prioritering av centrala kontra perifera problem i uppsatsen. Akribi ej kontrollerad.	Inga eller få andra seminariedeltagares resultat diskuteras, och/eller frågor och analyser med svag relevans.
F	Uppgiften framstår som omöjlig att genomföra inom ramen för handledarledd tid.		

Diskussion ang ex IV)

Delkursen i fråga är snarlik exempel I – kandidatuppsats engelsk litteratur. Betygskriterierna är annorlunda uppbyggda i exempel IV och de två kan tjäna som illustration av olika sätt att arbeta med två relativt likartade delkurser. Betygskriterierna i exempel IV är uppdelade med examinationsformerna som bedömningsområden. Nackdelen med uppdelningen per examinationsform är ju annars att man riskerar att försvåra viktig kursutveckling genom att befästa själva formerna för examinationen i betygskriterierna. I princip skulle lika gärna de förväntade studieresultaten ha kunnat fungera som bedömningsområden, men i det här fallet kan det finnas en poäng med att behålla indelningen per examinationsform. För det första är examinationsformerna i just den här kursen synnerligen stabila. Det är nästan svårt att tänka sig en uppsatskurs som inte examineras med uppsats, opposition och seminarier, och därmed framstår inte detta som ett problem. För det andra motsvarar uppdelningen ett väl fungerande arbetssätt. Uppsatsen utgör den absolut största delen av kursen och är det som i princip avgör betyget. En mycket stark eller mycket svag opposition kan i vissa fall komma att påverka betyget, men det kommer oftast i fråga om uppsatsen står och väger mellan två betyg.

Själva uppsatsen i sig är en högst komplex uppgift som analyseras med hjälp av ett antal bedömningsområden: De vetenskapliga resultatens värde, teoretisk och metodologisk medvetenhet, empirins genomförande och omfattning, samt framställningens språkliga kvalitet och formalia.

Alla dessa delbedömningar kan också utvecklas vidare, för att studenten ska kunna få en än mer detaljerad återkoppling, både formativt under uppsatsarbetets gång och summativt under ventileringen. Hur långt man vill gå i detaljbeskrivningar varierar, men ofta utvecklas olika typer av mallar eller checklistor angående formalia för att undervisningen så mycket som möjligt ska kunna ägnas åt annat. Även språkliga aspekter kan behöva differentieras. (Det är ju t ex en stor skillnad på disposition och språklig korrekthet.) Metodvalets relevans för uppgiften kan vara något som behöver skiljas från själva tillämpningen av metoden. Och så vidare.

Exempel V) Fysik – Grundläggande mekanik (12 hp)

Förväntade studieresultat (lärandemål):

Efter genomförd delkurs förväntas studenten kunna...

1. ...lösa grundläggande mekaniska problem.
2. ...förklara grundläggande mekaniska fenomen med hjälp av kraftbegreppet och Newtons lagar.
3. ...redogöra för fenomenen i den speciella relativitetsteorin.
4. ...bedöma när den klassiska mekaniken är tillämpbar.

Examinationsformer:

- Skriftligt prov, uppdelat i sex bedömningsområden

Undervisningsformer:

Föreläsningar, självstudier och övningar.

Kommentar till betygskriterier, examinationsuppgift och rättningsmall:

Betygskriterierna är utformade som globala beskrivningar av hur väl de förväntade studieresultaten uppnås. Examinationsuppgiften och rättningsmallen är utformade så att poängsättningen svarar mot dessa kriterier. Mallen och examinationens olika delar utgår ifrån sex bedömningsområden, utvecklade från kursens fyra lärandemål. Lärandemål 1 är uppdelat i tre bedömningsområden. (I spannet B–D finns ett visst utrymme för kompensation inom bedömningsområdena 1a och 1b. Ett lägre poängtal i det ena bedömningsområdet kan kompenseras av ett högre i det andra.) Bedömningsområde/lärandemål 2–4 bedöms enbart tvågradigt: tillräckligt (ja) eller otillräckligt (nej). Rättningsmallen anger lägsta poäng för respektive betyg.

Rättningsmall för examinationsuppgiften.

Bedömningsområdet	1) ...lösa grundläggande mekaniska problem. a) Icke-relativistiska problem. Enbart translationsrörelse. (0-12p)	1) ...lösa grundläggande mekaniska problem. b) Icke-relativistiska problem inkl rotationsrörelse. (0-8p)	1) ...lösa grundläggande mekaniska problem. c) Relativistiska problem. (0-4p)	2) ...förklara grundläggande mekaniska fenomen med hjälp av kraftbegreppet och Newtons lagar. (ja/nej)	3) ...redogöra för fenomenen i den speciella relativitetsteorin. (ja/nej)	4) ...bedöma när den klassiska mekaniken är tillämpbar. (ja/nej)
A	10p	6p	3p	ja	ja	ja
B	6p/10p	6p/3p	3p	ja	ja	ja
C	3p/6p	6p/3p	3p	ja	ja	ja
D	3p/6p	6p/3p	0p	ja	ja	ja
E	3p	3p	0p	ja	ja	ja
Fx	3p	3p	0p	ja	nej	nej
F	0p	0p	0p	nej	nej	nej

[Exemplet utarbetat av Henning Schmidt, UPC & Fysikum, SU.]

V) Betygskriterier – Grundläggande mekanik (12 hp)

A	Studenten kan med hjälp av kraftbegreppet och Newtons lagar förklara grundläggande fenomen och lösa ickerelativistiska problem som kräver stor kreativitet och en välutvecklad helhetsbild av kursinnehållet. Kunskaperna om speciell relativitet räcker till att redogöra för relativistiska fenomen och lösa enklare problem och härunder göra en adekvat bedömning av ifall den relativistiska kinematiken är nödvändig för en tillräckligt noggrann lösning av ett givet problem.
B	Studenten kan med hjälp av kraftbegreppet och Newtons lagar förklara grundläggande fenomen och lösa ickerelativistiska problem som kräver en välutvecklad helhetsbild av kursinnehållet. Kunskaperna om speciell relativitet räcker till att redogöra för relativistiska fenomen och lösa enklare problem och härunder göra en adekvat bedömning av i fall den relativistiska kinematiken är nödvändig för en tillräckligt noggrann lösning av ett givet problem.
C	Studenten kan med hjälp av kraftbegreppet och Newtons lagar förklara grundläggande fenomen och lösa vissa ickerelativistiska problem som kräver en välutvecklad helhetsbild av kursinnehållet. Kunskaperna om speciell relativitet räcker till att redogöra för relativistiska fenomen och lösa enklare problem och härunder göra en adekvat bedömning av i fall den relativistiska kinematiken är nödvändig för en tillräckligt noggrann lösning av ett givet problem.
D	Studenten kan med hjälp av kraftbegreppet och Newtons lagar förklara grundläggande fenomen och lösa vissa ickerelativistiska problem som kräver en välutvecklad helhetsbild av kursinnehållet. Kunskaperna om speciell relativitet räcker till att redogöra för relativistiska fenomen och göra en adekvat bedömning av i fall den relativistiska kinematiken är nödvändig för en tillräckligt noggrann lösning av ett givet problem
E	Studenten kan med hjälp av kraftbegreppet och Newtons lagar förklara grundläggande fenomen och lösa enkla ickerelativistiska. Kunskaperna om speciell relativitet räcker till att redogöra för relativistiska fenomen och göra en adekvat bedömning av i fall den relativistiska kinematiken är nödvändig för en tillräckligt noggrann lösning av ett givet problem.
Fx	Studenten behärskar kraftbegreppet och Newtons lagar tillräckligt väl för att kunna förklara grundläggande fenomen och lösa enkla problem, men kunskaperna om speciell relativitet är bristfälliga och räcker varken för redogöra för relativistiska fenomen eller för att göra en adekvat bedömning av i fall den relativistiska kinematiken är nödvändig för en tillräckligt noggrann lösning av ett givet problem.
F	Studenten behärskar inte kraftbegreppet och Newtons lagar tillräckligt väl för att såväl kunna förklara grundläggande fenomen samt lösa enkla problem.

Diskussion ang ex V)

Betygskriterierna är utformade i ett globalt format, men de är direkt kopplade till de bedömningsområden som ligger till grund för examinationsuppgiftens rättningsmall. Sammanräkningsmetoden kan ses som en kombination av alternativ 2 och alternativ 3 i huvudtexten. Ett bedömningsområde skiktas i fyra nivåer (1a), ett annat i tre nivåer (1b) och övriga i två nivåer (jmf alternativ 2). Kraven för respektive betygsnivå uttrycks i termer av lägstanivåer där samtliga lägstakrav ska vara uppfyllda för varje givet betyg (jmf alternativ 3).

Betygskriterier

Rättningsmall

Av rättningsmallen framgår tydligt att poängsättningen förutsätter att olika bedömningsområden hålls isär när ett betyg ska tilldelas. I många fall representerar ett sådant arbetssätt en viktig reformering av examinationsprocessen. Väsentliga delar av kursinnehållet samlas inte ihop till en stor pott med poäng som sedan mekaniskt skiktas i betygsnivåer. Äpplen och päron blandas inte samman. Det gör det möjligt att verkligen kunna uttala sig om vad en student kan. (Det finns inte längre en risk att ett i sig oacceptabelt missförstånd i en del automatiskt kompenseras av en fullständig förståelse av någonting helt annat.)

Bakom utformningen av den aktuella examinationsuppgiften ligger dessutom mycket viktiga överväganden angående betygssättningen som inte framgår av rättningsmallen. Poängsättningen inom respektive bedömningsområde representerar skilda svårighetsnivåer kopplade till betygskriterierna för kursen. Poängen är alltså uttryck för att studenten har klarat att lösa problem på en viss svårighetsnivå, inte bara ett visst antal uppgifter på en och samma nivå. Rättningsmallen kan dessutom väldigt enkelt utvecklas till en väl fungerande feedbackmall, där de kvalitativa skillnader som poängen representerar förklaras för studenterna i ord. Det är tveksamt om ett dylikt arbetssätt kan benämnas som en kvantitativ metod även om kvalitetsskillnader uttrycks i rättningsmallen som poäng och till viss del antal korrekta lösningar och förklaringar. I vilket fall som helst ligger det grundliga, kvalitativa överväganden bakom processen, vilket gör det möjligt att koppla examinationen till de målrelaterade betygskriterierna.

Avslutningsvis kan konstateras att de två underkända betygen Fx och F representerar specifika kvalitetsskillnader. (Annars är det vanligt att Fx bara kännetecknas av ett mindre antal av de fel som anges under F.) Den som får betyget Fx har visat att den har grundläggande kunskaper inom vissa områden, men att det inte räcker för godkänt. På så vis skapas en stor tydlighet vad gäller gränsen mellan E och Fx. Gränsen för godkänt – den viktigaste – blir därmed lättare att kommunicera och bevaka.

Exempel VI) Formell språkinläring, läromedel och bedömning av språkfärdighet (7,5 hp) – Delkurs 3 i Svenska som andraspråk II, GN, (30 hp)

Förväntade studieresultat (lärandemål):

Efter genomförd delkurs förväntas studenten kunna...

1. ... redogöra för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll i språkinläringen
2. ... resonera kring principer för produktion av språktest och hur olika typer av språktest kan användas för olika syften och i olika sammanhang
3. ... diskutera principer för innehåll och grammatisk progression i läromedel med utgångspunkt från teorier om språkinläring

Examinationsformer:

- Salstentamen
- Muntlig och skriftlig redovisning av gruppuppgift

Undervisningsformer:

Undervisningen består av lektioner som ägnas åt genomgång och diskussion av de olika delmomenten språkinläring, läromedelsanalys och språktest. Närvaron är ej obligatorisk.

Kommentar till examinationsuppgifter, betygskriterier och bedömningsområden:

Lärandemål 1 och 2 examineras genom salstentamen och betygsdifferentieras i en sjugradig skala.

Lärandemål 3 examineras genom en gruppuppgift (en analys av ett läromedel i svenska som andraspråk) som redovisas muntligt och skriftligt och betygsdifferentieras i tregradig skala.

[Exemplet utarbetat av Gunilla Jansson, Institutionen för nordiska språk, SU.]

VI) Betygskriterier – Formell språkinläring, läromedel och bedömning av språkfärdighet (7,5 hp)

A	Studenten redogör självständigt och på ett sätt som genomgående stämmer med vedertagna vetenskapliga uppfattningar för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll i språkinläringen samt resonerar på ett genomgående adekvat, kreativt och problematiserande sätt kring principer för produktion av språktest och hur olika typer av språktest kan användas för olika syften och i olika sammanhang. Dessutom diskuterar studenten principer för innehåll och grammatisk progression i läromedel med utgångspunkt från teorier om språkinläring. Diskussionen förankras genomgående i relevanta resonemang från litteraturen.
B	Studenten redogör självständigt och på ett sätt som i huvudsak stämmer med vedertagna vetenskapliga uppfattningar för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll i språkinläringen samt resonerar med anknytning till relevant litteratur och på ett delvis problematiserande sätt kring principer för produktion av språktest och hur olika typer av språktest kan användas för olika syften och i olika sammanhang. Dessutom diskuterar studenten principer för innehåll och grammatisk progression i läromedel med utgångspunkt från teorier om språkinläring. Diskussionen förankras genomgående i relevanta resonemang från litteraturen.
C	Studenten redogör delvis självständigt och på ett sätt som i huvudsak stämmer med vedertagna vetenskapliga uppfattningar för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll i språkinläringen samt resonerar med anknytning till relevant litteratur och på ett någorlunda problematiserande sätt kring principer för produktion av språktest och hur olika typer av språktest kan användas för olika syften och i olika sammanhang. Dessutom diskuterar studenten principer för innehåll och grammatisk progression i läromedel med utgångspunkt från teorier om språkinläring. Diskussionen förankras i viss mån i relevanta resonemang från litteraturen.
D	Studenten redogör på ett sätt som i huvudsak stämmer med vedertagna vetenskapliga uppfattningar för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll i språkinläringen samt resonerar med anknytning till relevant litteratur kring principer för produktion av språktest och hur olika typer av språktest kan användas för olika syften och i olika sammanhang. Dessutom diskuterar studenten principer för innehåll och grammatisk progression i läromedel med utgångspunkt från teorier om språkinläring. Diskussionen förankras i viss mån i relevanta resonemang från litteraturen.
E	Studenten redogör på ett sätt som någorlunda stämmer med vedertagna vetenskapliga uppfattningar för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll i språkinläringen samt resonerar med anknytning till relevant litteratur kring principer för produktion av språktest och hur olika typer av språktest kan användas för olika syften och i olika sammanhang. Dessutom diskuterar studenten principer för innehåll och grammatisk progression i läromedel med utgångspunkt från teorier om språkinläring. Diskussionen förankras i viss mån i relevanta resonemang från litteraturen.
Fx	Studenten redogör för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll i språkinläringen på ett sätt som inte stämmer med vedertagna vetenskapliga uppfattningar. Studenten resonerar kring hur läsförståelse och skriftlig kommunikation är beroende av kulturellt betingad förståelse och hur läromedelstexter är uppbyggda på ett sådant sätt som visar på bristande förståelse för resonemang i litteraturen.
F	Studenten dokumenterar knappast någon förståelse för hur några teoretiska modeller inom andraspråksforskningen förklarar interaktionens, lärarens och inlärarens roll och visar knappast någon förmåga att föra principdiskussioner kring innehåll och grammatisk progression i läromedel.

Diskussion ang ex VI)

Betygskriterierna är inskrivna i den globala mall som Institutionen för nordiska språk använder. Den enhetliga formen gör det lätt för utomstående att ta del av betygskriterier från många olika delkurser och kurser. Betygskriterierna för delkursen i exempel VI är inte desto mindre de facto uppdelade i två olika bedömningsområden. Den första utgörs av lärandemål 1 och 2, som bedöms sjugradigt. Den andra utgörs av lärandemål 3 som bedöms tregradigt. De ihopskrivna betygskriterierna bygger således på en tydlig uppdelning i delbedömningar:

Betygskriterier	Bedömningsområde 1 (lärandemål 1 & 2)	Bedömningsområde 2 (lärandemål 3)

Bilaga B – Checklista för förväntade studieresultat (lärandemål)

Lärandecentrerat?

Är de förväntade studieresultaten formulerade i termer om vad studenten ska kunna/kunna göra efter avslutad kurs?

- JA: "Efter genomförd kurs förväntas studenten kunna analysera primärkällor källkritiskt i förhållande till en given problemställning."
- NEJ: "Kursen handlar om att ge studenterna kunskaper i källkritik."

Mål eller medel?

Syftar det styrande verbet verkligen på ett mål, eller uttrycker det vad studenten ska ha gjort?

- JA: "Efter genomförd kurs förväntas studenten kunna analysera primärkällor källkritiskt i förhållande till en given problemställning."
- NEJ: "Efter genomförd kurs har studenten genomfört ett antal källkritiska analyser av primärkällor, och läst kurslitteraturen X, Y och Z."

Mål eller mätmetod?

Syftar det styrande verbet verkligen på ett mål, eller uttrycker det hur lärandet ska "bevisas"?

- JA: "Efter genomförd kurs förväntas studenten kunna analysera primärkällor källkritiskt i förhållande till en given problemställning."
- NEJ: "Efter genomförd kurs förväntas studenten kunna redogöra för hur en källkritisk analys av primärkällor kan gå till."

Hur pröva?

Mål skiljs från mätmetod, men det är väldigt viktigt att varje förväntat studieresultat följs upp med frågan "Hur ska detta kunna prövas?", dvs. vilka examinationsformer ska användas? Om ett förväntat studieresultat inte låter sig prövas behöver det omarbetas, eller ses som ett mer övergripande syfte med kursen. Om en viss kunskap eller färdighet *de facto* inte examineras så är det definitionsmässigt inte ett förväntat studieresultat för just den kursen.

Är de förväntade studieresultaten tillräckligt begripliga och tydliga?

Att skriva förväntade studieresultat handlar om att kommunicera med andra vad alla som har klarat kursen kan. Därför måste lärandemålen (som läses tillsammans med betygskriterier, examinationsform, litteratur, mm) vara begripliga. Är det möjligt att skapa sig en rättvisande bild? Hur är det med tydligheten? Vissa styrande verb är så allmängiltiga att de betyder allt och inget. "Förstå" är ett sådant. Försök formulera förväntade studieresultat som är så otvetydiga som möjligt.

Rätt prioritet?

En sekundär kunskap (t ex enklare former av faktakunskaper), kan ibland vara en förutsättning för det egentliga målet (t ex någon sorts tillämpning av dessa kunskaper.) Det viktigaste målet bör då ges en mer framträdande plats. Beakta att det som står överst kan tolkas som det viktigaste av läsaren.

Lagom krav?

När samtliga förväntade studieresultat är formulerade är det viktigt att se över helheten. Är kraven för lågt eller för högt ställda? Två frågor som behöver besvaras är: (1) Om studenterna kan göra allt som anges, bör det räcka för att ha klarat kursen? Om svaret är nej behöver de förväntade studieresultaten kompletteras eller utvidgas. (2) Bör alla som inte kan göra *allt* som anges underkännas? Om svaret är nej är kraven för höga. Kanske några av formuleringarna kan passa bättre som betygskriterier för höga betyg?

Bilaga C – Checklista för betygskriterier

Vilket format, vilka bedömningsområden?

Finns det riktlinjer eller mallar vid din institution eller motsvarande som anger hur betygskriterierna ska skrivas och utformas? Ska de skrivas globalt eller per bedömningsområde? Vilka bedömningsområden i så fall? Per lärandemål, per examinationsform, eller annat? Om det inte finns några sådana riktlinjer behöver du ta ställning till dessa formfrågor själv. Är lärandemålen skrivna så att de kan fungera som bedömningsområden? Om inte, finns det andra perspektiv som kan förbättra och förenkla bedömningen? Även betygskriterier skrivna i ett globalt format kan många gånger fungera bättre om de utgår ifrån väl definierade bedömningsområden.

Målrelaterat?

Formuleras betygskriterierna i termer av *hur väl* studenten har uppnått de förväntade studieresultaten (lärandemålen) för den aktuella kursen/delkursen? Fungerar kriterierna för godkänt (A–E) som kvalitativa preciseringar av de förväntade studieresultaten (lärandemålen)?

Fungerar kriterierna och examinationen?

Kan du använda kriterierna till att besvara frågor från utomstående om vad studenter med olika betyg kan? Vad kan studenter med betyget C på just den här kursen som de med D och E inte kan? Fungerar betygskriterierna som beskrivning av de skillnaderna?

Garanteras en lämplig lägsta gräns?

Gränsen mellan godkänt och underkänt är viktigast. Att det är svårt att formulera meningsfulla kriterietexter kan ibland göra att kraven för E skrivs för lågt, (och alltför högt för A). Inled gärna arbetet med att formulera kraven för E, och därefter för A. Lärandemålen omfattar spannet A–E. Att det är svårt att fingradera mellanliggande kvaliteter får inte innebära att de gränserna ruckas på.

Skiljs kurskrav från kvaliteter?

Obligatoriska moment, deadlines, närvarokrav och liknande listas med fördel separat som ”kurskrav”. Betygskriterierna bör renodlas så att de enbart beskriver olika nivåer av förståelse, dvs syfta på en bedömning av dem som har uppfyllt alla formella krav. Om kurskrav omnämns i betygskriterierna kan de uppfattas som självändamål – och de kvalitativa kraven framstå mindre tydligt.

Hur blir det ett samlat betyg?

Om studentens prestation och lärande analyseras i delbedömningar, hur vägs det sedan samman till ett enda betyg för kursen/delkursen? Ett alternativ är att i en berättande text ange vilka aspekter som väger tyngst. Det är av största vikt att sammanvägningsmetoden inte kan leda till ett betyg som strider mot examinatorns helhetsbedömning. En helt mekanisk metod kan ofta te sig tydlig och rättvis, men kan innebära att examinatorn förlorar kontrollen över helhetsomdömet.

Är kriterierna ”på riktigt”?

Kan varje formulering motiveras, förklaras och exemplifieras? Spela djävulens advokat med texten och försök avslöja delar som inte svarar mot några verkliga bedömningsgrunder. Huruvida kriterierna är ”på riktigt” visas också i den efterföljande processen. Används de verkligen som stöd i arbetet med att betygssätta studenternas examinationsuppgifter, eller ligger de kvar i skrivbordslådan?

Används rätt examinationsform och examineras rätt saker?

Uppgiften att differentiera sju nivåer av förståelse måste ses i sitt rätta perspektiv. Fingraderingen i sig är inte överordnad alla andra värden. Att öka andelen examinationsformer som mäter enklare former av kunskap enbart för att resultaten ska gå lättare att kvantifiera innebär en kvalitetssänkning. Examinationsformer ska alltid utformas för att motivera studenterna att lära sig rätt saker på rätt nivå, även om det innebär att efterföljande betygsbedömningar blir mindre tvärsäkra.