

Beteckning: _____



Institutionen för Pedagogik, Didaktik och Psykologi

Att döma framåtsyftande

En problematisering av bedömningsbegreppet
i aktuell svensk och anglosaxisk pedagogisk forskning

Elisabeth Eriksson

Juni 2009

Uppsats, 15 högskolepoäng
Didaktik

Vetenskapligt arbete

Handledare: Jan Grannäs och Elisabeth Björklund

Examinator: Bengt Nilsson

Elisabeth Eriksson (2009): *Att döma framåtsyftande – en problematisering av bedömningsbegreppet i aktuell svensk och anglosaxisk pedagogisk forskning*. Uppsats i didaktik C. Institutionen för Pedagogik, didaktik och psykologi. Högskolan i Gävle.

Sammanfattning

Bedömning inom skolan är något komplext och bedömningsbegrepp likaså, hur de definieras, tolkas och används. Hur bedömning förmedlas till en elev är av stor betydelse för det fortsatta lärandet. Rätt utformad kan denna förmedling resultera i ett ökat engagemang hos eleven och en förbättring av elevens lärande. Behovet av att förstå bedömningsprocesser och olika former av bedömning är idag högaktuellt inom svensk grundskola i och med att landets skolor och rektorer givits i uppdrag att utforma skriftliga omdömen från skolår 1. Även om det gått snart ett år sedan uppdraget gavs är det i kontakter med lärare och rektorer tydligt att utformandet av skriftliga omdömen i många fall är en process som är långt ifrån klar. För att kunna utforma dessa omdömen på tillfredsställande sätt behövs djupare kunskaper om bedömningsbegrepp.

Denna studie är en forskningsöversikt, en analytisk undersökning av aktuell forskning kring bedömningsbegreppet. I översikten problematiseras bedömningsbegreppet inom skolans värld för att tydliggöra hur olika former av bedömning definieras och används inom svensk och anglosaxisk forskning. Undersökningen omfattar 48 källor, varav 27 svenska. Dessa har analyserats genom en studie av definitioner, argument, formuleringar, författarnas perspektiv och eventuella personliga ställningstaganden, samt om de är att betrakta som problematiserande eller enbart konstaterande.

Resultatet visar att det inom den anglosaxiska forskningen, som dominerar den internationella forskningen kring bedömning, finns en betydligt högre grad av problematisering och diskussion kring bedömningsbegrepp än inom svensk forskning. Inom svensk forskning är det istället lätt att tro att det råder konsensus i hur man tolkar och förstår bedömningsbegrepp. Denna villfarelse skapas av en brist på definitioner och en avsaknad av diskussioner kring begrepp. Istället problematiseras summativ bedömning i form av betyg utifrån frågan om likvärdighet, en diskussion som saknas i anglosaxisk forskning.

Vid en jämförelse av svensk och anglosaxisk forskning kring bedömning framträder en tydlig skillnad i fokus, även om svensk forskning börjar närma sig ett anglosaxiskt perspektiv. Detta genom att inta ett övergripande perspektiv på bedömning istället för att se formativ och summativ bedömning som separata företeelser. Att tala övergripande i termer av kunskapsbedömning innebär dock inte att svårigheterna för bedömning att uppfylla kraven för att tillskrivas såväl formativ som summativ karaktär försvinner. Detta är något som den anglosaxiska forskningen tydligt vittnar om.

Nyckelord: bedömning, formativ, kunskapsbedömning, lärande, summativ,

Keywords: assessment, assessment for learning, assessment of learning, educational assessment, formative, learning, summative

Innehållsförteckning

INLEDNING	1
DISPOSITION	2
BEGREPPSFÖRKLARINGAR.....	3
<i>Bedömning</i>	3
BAKGRUND	4
<i>Svensk bedömnings- och betyghistorik</i>	4
<i>Synen på kunskapsbedömning i ett historiskt perspektiv</i>	5
<i>Vägen till individuella utvecklingsplaner med skriftliga omdömen</i>	7
SYFTE	7
<i>Frågeställningar</i>	7
<i>Avgränsningar</i>	8
METOD	9
OM FORSKNINGSÖVERSIKTEN SOM METOD.....	9
METODVAL	9
<i>Bakgrund till metodval</i>	9
<i>Val av metod</i>	10
SYSTEMATISKT URVAL OCH ANALYSMETODER	10
<i>Urval</i>	11
<i>Gradering gällande relevans</i>	11
<i>Teser och argument</i>	12
<i>Bedömning av argument</i>	13
RESULTAT	14
BEDÖMNINGSBEGREPP I SVENSK FORSKNING	14
<i>Definierande och tolkande av bedömningsbegrepp</i>	14
<i>Användande av bedömningsbegrepp</i>	15
<i>Sammanfattning av svensk forskning kring bedömning</i>	19
BEDÖMNINGSBEGREPP I ANGLOSAXISK FORSKNING	20
<i>Definierande och tolkande av bedömningsbegrepp</i>	20
<i>Användande av bedömningsbegrepp</i>	22
<i>"Inside the black box" och "bibeln" för formativ bedömning</i>	26
<i>Sammanfattning av anglosaxisk forskning kring bedömning</i>	28
FOKUS PÅ BEDÖMNING I SVERIGE OCH INTERNATIONELLT – EN JÄMFÖRELSE	29
<i>En förskjutning i fokus</i>	29
<i>Olika sorters forskning och perspektiv</i>	30
DISKUSSION	33
BEGREPP – DEFINITIONER OCH AVSAKNAD AV DEFINITIONER	33
<i>Inbillad konsensus</i>	33
<i>Otydlighet i begreppsfloran</i>	34
<i>Syfte eller inte</i>	34
FORSKNINGSFOKUS	35
<i>Skillnad i fokus, frågor och ickefrågor</i>	35
AVSAKNAD AV PROBLEMATISERING	35
<i>Att döma framåtsyftande</i>	36
METODDISKUSSION.....	36
TANKAR OM FORTSATT FORSKNING	37
<i>Hur har skriftliga omdömen utformats?</i>	37
<i>Hur definierar och använder lärare bedömningsbegrepp?</i>	37
<i>Vad innebär en digitalisering av skriftliga omdömen?</i>	38
REFERENSLISTA	39

Inledning

När det talas om bedömning på skolor, i arbetslag eller i ämnesgrupper för att nå någon form av samstämmighet och med likvärdig bedömning som mål och ambition, kan det vid ett ytligt samtal verka som om samsynen redan finns där. Men vid en djupare diskussion är möjligheterna stora att det visar sig att det inte är så. Begreppet bedömning är i skolans värld något komplext och hur det tolkas och förstås beror på en mängd skilda faktorer, där syftet med bedömningen är en central faktor. Christian Lundahl (2006), som i sin doktorsavhandling studerar kunskapsbedömning i ett historiskt perspektiv, berör detta på följande sätt i avhandlingens avslutande diskussion:

Hur skolan ska bedöma barns och ungdomars kunskaper kan betraktas som ett evigt dilemma. Olika aktörer har under olika perioder skilda förslag, behov och intressen av att veta vad eleverna lär sig i skolan. Mellan dessa olika viljor att veta uppstår spänningar relaterade till skilda perspektiv på skolans uppdrag, på kunskap, undervisning och kanske inte minst på lärarna och eleverna. (Lundahl, 2006, s. 407)

I och med Utbildningsdepartementets pressmeddelande 10 juni 2008 gällande skriftliga omdömen för elever i grundskolan har den nu sittande regeringen som aktör i detta dilemma målat upp sin bild av hur elevernas kunnande och kunskapsinhämtande ska tydliggöras (Utbildningsdepartementet, 2008a). I pressmeddelandet klargjordes att regeringen beslutat att samtliga elever från och med höstterminen 2008 ska ges skriftliga omdömen gällande kunskapsutvecklingen i varje ämne. Denna information ska ges varje termin och ansvaret för utformningen av dessa omdömen åläggs varje enskild rektor. Det uppdrag landets rektorer givits är inte alldeles enkelt, vilket blir tydligt när direktiven klargörs. De skriftliga omdömena får vara av betygsliknande karaktär, vilket river upp ett tidigare förbud mot detsamma (ibid). Att de skriftliga omdömena får vara betygsliknande innebär att det nu är tillåtet att skriva om en elev har nått målen i ett ämne eller inte (Utbildningsdepartementet, 2008b). I den promemoria som ligger till grund för beslutet framgår att de skriftliga omdömena även bör vara framåtsyftande för att främja elevens fortsatta lärande (Utbildningsdepartementet, 2008c). Utifrån dessa direktiv ska varje enskild elevs resultat beskrivas i förhållande till den lokala pedagogiska planeringen, kursplaner och läroplan. Frågan är hur rektorerna väljer att utforma dessa omdömen och på vilken grund de fattar sina beslut. Som hjälp till rektorer och lärarlag har Skolverket på uppdrag från regeringen omarbetat sina allmänna råd om den individuella utvecklingsplanen att nu gälla individuell utvecklingsplan (IUP) med skriftliga omdömen (Skolverket, 2008). Då oron bland landets lärare ändå tycks vara stor kring detta uppdrag, har olika texter för att tolka de allmänna råden skrivits (Tholin, 2008; Moreau, 2008; Wretman, 2008).

Det är, utifrån Lundahls (2006) påpekande gällande tolkningen av kunskapsbedömningsbegreppet, en rätt komplex uppgift som delegerats landets rektorer. Jag hade en ambition att studera det vetenskapliga underlag som låg till grund för den ovan beskrivna lagändringen. Detta i avsikt att jämföra de ställningstaganden som kan avläsas där med innebörden av de skriftliga omdömena och eventuellt göra jämförelser med annan aktuell forskning rörande bedömning. Då varken promemorior eller andra dokument direkt kopplat till detta beslut anger några referenser tog jag kontakt med Utbildningsdepartementet via en e-postlänk på deras hemsida och det svar jag fick från Hedwig Kastenholm i Departementsledningen var följande:

Generellt sett så redovisas inte forskningsunderlaget i regeringens propositioner och beslut. Däremot finns det mycket intressant forskning om bedömning och utvärdering av elever som Utbildningsdepartementet tagit del av under utvecklandet av propositionen. Det finns dock inget från Utbildningsdepartementets sida publicerat material angående denna forskning. Jag skulle rekommendera att du istället tar del av den forskning som publicerats på ämnet i vetenskapliga skrifter och avhandlingar. (H. Kastenholm, Utbildningsdepartementet, personlig kommunikation, 9 mars 2009)

Detta svar och en artikel av Hans-Åke Scherp (2008) fick mig att välja att göra precis det som brevet från Utbildningsdepartementet uppmanar till, nämligen att ta del av den aktuella forskningen i ämnet, något som departementet säger sig ha gjort, men Scherp ifrågasätter genom att ställa en fråga om "vart" forskningen tog vägen (Scherp, 2008, s.12). Scherp har granskat regeringens skolpolitik i relation till forskningsresultat inom relevanta områden och menar att det inte finns mycket forskning som stöder den kursomläggning som regeringen genomför gällande ämneskunskaper, nationella prov och betyg. Även Håkan Andersson (2007) ställer sig frågande till vetenskaplig grund för de förändringar inom skolan som den sittande regeringen nu vill genomdriva. Han refererar till att utbildningsminister Jan Björklund vid den presskonferens där han presenterade sitt förslag till nytt betygssystem klargjorde att det handlade om politiska ståndpunkter och löften och att det inte byggde på forskning. Utifrån detta uttalande är det intressant att ställa sig frågan vilken forskning som Kastenholm refererar till och vad "tagit del av" innebär. Det är dock inte vad jag tänker göra här. Medan Scherp (2008) använder vassa ordalag i sin kritik mot den skolpolitik som förs har jag inte några sådana avsikter med min studie. Jag ämnar inte kritisera en lag som är genomdriven och för oss inom skolan att anpassa oss efter. Istället är jag intresserad av att studera vad det är för uppdrag som landets skolor och rektorer givits i och med denna lagändring, och eftersom det i sig är ett stort område tänkte jag börja med att studera kärnbegreppet bedömning. Detta gör jag genom att undersöka hur bedömningsbegrepp definieras och används inom svensk och anglosaxisk forskning.

Disposition

I det inledande kapitlet redogörs för de centrala begreppen i studien. Här ges även en historisk presentation av hur bedömningssystemet i den svenska skolan har sett ut sedan införandet av "det absoluta betygssystemet" för folkskolan i slutet av 1800-talet, följt av en redovisning av hur de idag gällande skriftliga omdömena växt fram. Den historiska tillbakablicken sammanfattas med en kort analys av mönster i utvecklingen av betyg och omdömen inom den svenska skolan. Inledningen avrundas med att läsaren får ta del av syftet med studien, med tillhörande frågeställningar och avgränsningar. I det andra kapitlet beskrivs forskningsöversikten som metod. Metodval och avväganden som lett till dessa presenteras tillsammans med en beskrivning av urvalsproceduren och analysmetod. Metodkapitlet följs av en resultatredovisning, där svensk och anglosaxisk forskning kring bedömningsbegreppet inledningsvis presenteras åtskilt. Här redogörs vilka begrepp och definitioner som används inom forskning kring bedömning, samt hur dessa begrepp används. Resultatkapitlet avslutas med en jämförelse av svensk och internationell forskning kring bedömning. I det avslutande diskussionskapitlet ges där en friare analys av och reflektion kring resultaten. Kapitlet omfattar även en metoddiskussion, där frågor rörande undersökningens trovärdighet behandlas. Allt avslutas med några tankar kring fortsatt forskning inom området bedömning.

Begreppsförklaringar

Några av de begrepp som är centrala i denna studie lämnar tolkningsutrymme till läsaren och därför kan det vara på sin plats att redan tidigt presentera dessa och de definitioner av dem som kommer att användas i denna studie.

Bedömning

Då bedömning är det centrala begreppet som problematiseras i denna studie kommer här inte någon färdig definition. Vad som istället presenteras är en beskrivning av några olika typer av bedömning inom skolans värld. Inom dessa olika genrer av bedömning finns olika begrepp som används mer eller mindre frekvent. Några forskare nöjer sig inte med de redan etablerade bedömningsbegrepp som finns, utan skapar nya som de anser bättre beskriver det de vill belysa (se t.ex. Carless, 2007). Medan en del forskare sätter likhetstecken mellan ett svenskt och ett anglosaxiskt begrepp, väljer andra att särskilja dem av olika skäl. Detta är anledningen till att jag i denna studie valt att både använda mig av svenska och anglosaxiska begrepp, samt föra diskussioner kring de definitioner och gradskillnader av definitioner som lyfts i det behandlade materialet.

Newton (2007) har delat in bedömning i tre olika användningsområden och det är de som jag har valt att presentera här. Valet baseras på att det handlar om övergripande användningsområden för bedömning, inom vilka de bedömningsbegrepp som denna studie omfattar passar väl. De begrepp vilka Newton använder sig av är inte frekvent förekommande i vetenskapliga artiklar inom området, men kommer att vara de som används som övergripande bedömningsbegrepp i denna studie. Detta i en ambition hos mig som författare att distansera mig till de val av begrepp som används av författarna till de olika vetenskapliga texter som studeras.

Om bedömningen är av *dömande karaktär* handlar det om den typ av bedömning som resulterar i betyg, eller kanske betygsliknande omdömen av siffer- eller bokstavskaraktär, eller av typen uppnått och uppnått väl (Newton, 2007). Det handlar om en avstämning av kunskaper som finns hos den man bedömer. Vanligt förekommande bedömningsbegrepp som beskriver denna form av bedömning är *summativ bedömning* och *assessment of learning*.

Det finns även bedömning av *beslutskaraktär* (Newton, 2007). Oftast handlar detta om bedömningar man gör utifrån resultatet av en dömande bedömning. I skolans värld handlar denna typ av bedömning om en avvägning gällande framtida möjligheter inom ett område. Det kan röra sig om bedömningar som direkt påverkar små eller stora val, fortsatta satsningar inom idrott, ämnesval eller val av inriktning för fortsatta studier (ibid). Denna studie kommer inte att fokusera på beslutande bedömning, men när den typen av bedömning behandlas är det under begreppet *beslutande bedömning*.

Den tredje formen av bedömning är den *framåtsyftande*, med avsikt att påverka och utöva inflytande över en pågående process (Newton, 2007). Vanliga bedömningsbegrepp för att beskriva denna typ av bedömning är *formativ bedömning* och *assessment for learning*.

Bakgrund

Här ges en presentation av svensk betygshistorik från införandet av ”det absoluta betygssystemet” 1897 fram till införandet av målrelaterade betyg 1996. Det åskådliggörs även hur synen på kunskap förändrats med tiden. Bakgrunden avslutas med en beskrivning av införandet av skriftliga omdömen.

Svensk bedömnings- och betygshistorik

I och med införandet av ”det absoluta betygssystemet” för folkskolan 1897 börjar man tillämpa en enhetlig betygsskala i Sverige. Skalan som då används är en sjugradig bokstavsskala från A till C, där varje betygssteg har benämningar som exempelvis berömlig (A), icke utan beröm godkänd (Ba) och icke fullt godkänd (BC). Betygen ges i alla årskurser och deras funktion är inte primärt av urvalskaraktär, utan de används främst för att påvisa i vilken grad eleven uppnått vissa i förväg bestämda kunskaper (Lundahl, 2006). Urval till vidare studier görs av mottagande skola, ofta genom antagningsprov (Tholin, 2006). Då intresset för vidare studier ökar växer behovet av att använda betygen för urval fram och 1949 sker en ändring som ger bokstavsskalan urvalskaraktär och vi får ”det relativa betygssystemet”. Detta sker genom att betygen i bokstavsskalan fördelas enligt en normalfördelningskurva, där procentangivelser anger fördelningen av betyg. Denna normalfördelning ska motsvara betygsfördelningen på nationell nivå (ibid). För att ytterligare öka likvärdigheten i betygssystemet införs standardprov, vilka ska hjälpa lärare att skaffa en bild av hur klassen ligger till kunskapsmässigt i ett nationellt perspektiv. 1962, i och med införandet av läroplanen för grundskolan Lgr 62, byts den sjugradiga bokstavsskalan ut mot en femgradig sifferskala, fortfarande med en normalfördelningskurva som rekommenderas motsvara betygsfördelningen på nationell nivå (Tholin, 2006). Betyg ges nu på vårterminen från och med årskurs 2 och både höst- och vårtermin från och med årskurs 6. I och med att läroplanen för grundskolan Lgr 69 införs minskar betygstillfällena ytterligare till att delas ut på våren i årskurs 3, 6 och 7, samt höst och vår i årskurs 8 och 9. När läroplanen för grundskolan Lgr 80 införs är det endast i årskurs 8 och 9 som det ges betyg.

Det relativa betygssystemet fyller sin funktion som urvalsinstrument, men blir tidigt kritiserat och ett flertal utredningar tillsätts för att utvärdera det. I början av 70-talet läggs ett förslag fram om slopande av betyg, vilket inte går igenom (Tholin, 2006). I början av 90-talet står det klart att det relativa betygssystemet ska bytas ut, men frågan är till vad. Två olika förslag läggs fram, det ena handlar om ett slopande av betyg och det andra förslaget förespråkar ett målrelaterat system. Målrelaterade system har då diskuterats i åtskilliga år, men problemet med att lösa frågan om likvärdig bedömning har bidragit till att ett konkret förslag dröjt (Wedman, 2003). Det målrelaterade betygssystemet genomförs 1996. Det tidigare problemet med likvärdighet antar utredarna kommer att lösas av sig själv, med motiveringen att om bara lärarna får lite tid så kommer deras tolkningar av styrdokumentet i slutändan se ganska lika ut (ibid). Detta antagande saknar helt stöd i forskning, och problemet med lärarnas uppdrag att ge likvärdig bedömning trots avsaknad av betygskriterier har studerats och debatterats (Tholin 2003; 2006; Wedman, 2003).

Synen på kunskapsbedömning i ett historiskt perspektiv

När Lundahl (2006) sammanfattar sin studie av kunskapsbedömning i svenskt skolväsende från 1500-tal till nutid är begreppen ”traditioners makt” och ”kulturell tröghet” framträdande. Den kulturella trögheten gällande utveckling och ändring av betygssystem är något som även Tholin (2006) framhåller i den betyghistoriska delen av sin avhandling:

Att betrakta betygssystemet som *ett slags frusen ideologi*¹ tillför ett intressant perspektiv. Den sjugradiga skala, ibland kallad ”det absoluta betygssystemet”, som infördes för folkskolan 1897 kom med vissa justeringar att fungera i ungefär femtio år. 1962 infördes den femgradiga skalan i grundskolan som i princip behölls intakt till 1994. Betygssystem fungerar alltså länge, betydligt längre än till exempel läroplaner, och sätter rimligtvis så djupa avtryck i vårt sätt att tänka kring och tala om betyg att man kan tala om att avlagringar av de tidigare systemen ligger under de nya. (Tholin, 2006, s. 61, egen kursivering)

Dessa avlagringar refereras ibland till i debatten kring dagens skolpolitik (se Tholin, 2008; Scherp, 2008). Vad Tholin (2006) och Lundahl (2006) menar är att förändringar sker långsamt, då vi är styrda av de traditioner och den kultur som formar det samhälle vi lever i. Det relativa betygssystemet, vilket ersatte det absoluta betygssystemet 1949, blev tidigt ifrågasatt, men trots åtskilliga betygsutredningar dröjde det ända till 1996 innan vi började tillämpa målrelaterade betyg (Tholin, 2006). Då hade målrelaterade betyg som alternativ diskuterats i närmare 30 år, men genomförandet väntade, då det var svårt att se hur betygens urvalsfunktion skulle kunna behållas (Wedman, 2003).

70- och 80-talet

Utifrån en studie av drygt 80 artiklar publicerade i *Kritisk utbildningstidskrift (Krut)* under åren 1977 till 1986 konstaterar Lundahl (2006) följande:

Krut-skribenternas beskrivningar av kunskapsbedömning och dess konsekvenser i termer av dolt och öppen sortering skulle kunna sammanfattas som en kritisk strategi vilken å ena sidan gjorde tidigare dolda praktiker synliga genom språket, och å andra sidan påvisade hur tidigare öppna praktiker gjorts dolda genom språket. /---/ Beskrivningarna om dolt och öppet utgår nästan undantagslöst från föreställningar om klassamhällets reproduktion. Betyg och test möjliggör en sortering inom skolan och i samhället. Bedömningarnas validitet och bedömnarna, diskuteras inte i denna sorteringstematik. Över huvud taget är kunskapsbedömning som vetenskaplig och pedagogisk praktik relativt osynlig i *Krut*. (Lundahl, 2006, s. 311)

Lundahl påtalar att fokus gällande kunskapsbedömning i *Krut* läggs på den sorterande funktionen. Kunskapsbedömningen behandlas som ett urvalsinstrument som beskrivs, snarare än diskuteras. Fokus ligger på dömande bedömning och på betygens sorterande funktion och den framåtsyftande bedömningen diskuteras knappt. Kanske är det så just för att artiklarna i *Krut* under den tid Lundahl studerat är så tydligt kritiska och det är lätt att kritisera ett system, men svårare att kritisera den dagliga framåtsyftande bedömningen, då det indirekt skulle innebära kritik riktad mot det egna arbetet. Forsberg och Lundahl (2006) uttrycker följande om samma tid:

¹ Tholin (2006) refererar till Liedman, S.-E. (1997). *I skuggan av framtiden. Modernitetens idéhistoria*. Stockholm: Bonnier Alba.

Under 70- och 80-talen blir det tydligt att de politiska löftena inte kunnat infrias. Staten hade styrt men inte parerat för att likhet i chans inte nödvändigtvis leder till likhet i resultat. Elevers prestationer kunde användas för konkurrens om arbetstillfällen och studieplatser på ett sätt som snarare ledde bort från visionerna om jämlikhet än till dess förverkligande. Skolan uppfattades vara ett kugghjul i det kapitalistiska maskineriet. (Forsberg & Lundahl, 2006, s. 19)

Detta, och Lundahls (2006) reflektioner kring *Krut* får här representera bilden av den bedömningssyn som genomsyrar skola och skolforskning under sent 70- och tidigt 80-tal. Hela den del av Lundahls avhandling som behandlar *Krut* under denna tid osar av socialism, en skildring av den tidsanda som tidskriften föddes ur.

90-talet och några år till

Den kursändring som övergången till ett målrelaterat betygssystemet innebär sätter sina spår i synen på kunskap och kunskapsbedömning inom skolan. Sorteringsfunktionen ses enbart som viktig i vidare urval, istället är lusten att lära central (Lundahl, 2006). Betygsberedningen 1992 definierar enligt Tholin (2006) kunskapsutveckling som en individuell process, relaterad till såväl ämnesinnehåll som olika situationer och sammanhang. Innebörden i en sådan definition och eventuella konsekvenser för ett sådant synsätt är däremot ingenting som de klargör (ibid). Lundahl (2006) uttrycker 90-talets anda gällande kunskap och bedömning med följande ord:

Valfrihet, rörlighet, flexibilitet och lust att lära är ideal utifrån vilka å ena sidan mer öppna och tydliga bedömningskriterier kan informera eleverna om deras behov av mer fördjupade kunskaper. Å andra sidan, i ett diskursivt klimat där ingen kan misslyckas, växer psyvetenskaperna² tillbaka in i skolan med förnyad kraft /.../. I diskursen blir exempelvis bristande vilja en oacceptabel ursäkt när samhällsordningen bygger på individuellt ansvarstagande. (Lundahl, 2006, s. 411-412)

Elevens eget ansvar och ökade inflytande över det egna lärandet under 1990-talet är något som Lundahl tydligt lyfter fram. Enligt Tholin (2006) framgår i ett nyhetsbrev från Skolverket³ att läraren ska sätta betygskriterier i samråd med eleverna. Detta samtidigt som man brottas med problemet att skapa likvärdighet i bedömningen, då egenkontroll snart visar sig fungera dåligt (Forsberg & Lundahl, 2006). Detta problem löses genom införandet av nationella prov och diagnoser, individuella utvecklingsplaner, kvalitetsredovisningar och inspektioner (ibid).

I och med övergången till ett målrelaterat betygssystem slutar skolpolitiska beslut vara forskningsrelaterade (Tholin, 2006; Wedman, 2003). Från att skolpolitik och forskning tidigare haft en tydlig koppling, hänvisas vid skolpolitiska beslut nu endast till att man tagit del av forskning, något flera ställer sig frågande till (Andersson, 2007; Scherp, 2008). Visserligen var en hel del av den tidigare forskningen så tydligt knuten till skolpolitiken att resultat ibland ansågs vara beställda, men den senare tidens frånvaro av forskningsanknytning i skolpolitiska beslut ses på med större oro (Tholin, 2006).

² Med psyvetenskaper menas psykologiska vetenskaper. Enligt Forsberg och Lundahl (2006) karaktäriseras dessa inom skolans värld av att elever bedöms med hjälp av test, siffror och tabeller. "Psykologin erbjöd en språkdräkt för olikhet. Denna var uppbyggd på naturvetenskapliga metoder och kodifierade olikhet som objektiva fakta om individen. Utifrån dessa fakta kunde styrningen – organiserandet av individen, skolan och samhället – anses säker och rationell." (Forsberg & Lundahl, 2006, s.18)

³ Skolverket (1996). *Nyhetsbrev*, Nr. 19.

Vägen till individuella utvecklingsplaner med skriftliga omdömen

Sedan 1 juli 2005 används individuella utvecklingsplaner, där insatser för elevens utveckling mot läroplans- och kursplanemål tydliggörs (Utbildningsdepartementet, 2008c). I regeringens budgetproposition för 2007 påtalas en brist i informationen till elever och föräldrar i samband med utvecklingssamtal och det betonas att de individuella utvecklingsplanerna därför behöver struktureras och förbättras (Finansdepartementet, 2006). Det framhålls att betyg och skriftliga omdömen underlättar utvärderingen av elevens kunskapsresultat (ibid). Regeringen menar att det finns en inre konflikt i tidigare lagstiftning då den anses förhindra lärare att ge föräldrar den information de efterfrågar och att man i och med ett borttagande av förbudet gällande omdömen av betygsliknande karaktär tillmötesgår dessa önskemål från föräldrahåll (Utbildningsdepartementet, 2008c). I propositionen 2007/07:86 föreslås en lagändring med innebörden att det inte längre ska finnas ett förbud mot skriftlig information av betygskaraktär i samband med utvecklingssamtal (Utbildningsdepartementet, 2007). Att ge lärarna möjlighet att uttrycka sig i betygsliknande former ses som ett sätt att öka tydligheten i informationen till elever och föräldrar. Det framhålls att denna lagändring möjliggör för elever och föräldrar att redan från det att eleven börjar sin skolgång följa elevens utveckling mot läroplans- och kursplanemål (ibid). I promemorian gällande den förordning som trädde i kraft 1 juli 2008 är kraven om skriftliga omdömen än mer tydliga (Utbildningsdepartementet, 2008c). Där klargörs att utvecklingsplanen ska, förutom att vara framåtsyftande, innehålla skriftliga omdömen om elevens kunskapsutveckling i varje enskilt ämne eleven undervisas i. De skriftliga omdömena bör i sin tur vara framåtsyftande och får ha betygsliknande karaktär. Enligt Utbildningsdepartementets (2008c) definition av betygsliknande omdömen är dessa inte att blanda ihop med betyg, vilka värderar avslutade kurser och betecknas därför som summativa. Betygsliknande omdömen däremot bör enligt regeringen ha en framåtsyftande karaktär, vilket innebär att det handlar om bedömning som ska främja elevens fortsatta lärande (ibid). Märk väl ordet *bör* i förra meningen och den innebörd som ordet ger meningen. Det är nämligen inte så att de skriftliga omdömena ska vara av formativ karaktär, utan just bör, en rekommendation. Detta innebär att även om den individuella utvecklingsplanen ska vara framåtsyftande och främja elevens fortsatta lärande, så kan de skriftliga omdömena ha en annan utformning, även om detta inte rekommenderas (Skolverket, 2008).

Syfte

Syftet med denna studie är att problematisera bedömningsbegreppet inom skolan i svensk och anglosaxisk forskning och genom detta bidra till att förtydliga det uppdrag som har givits landets rektorer och skolor i och med införandet av skriftliga omdömen från skolår 1. Valet att ställa svensk och anglosaxisk forskning i relation till varandra gjordes utifrån ett tidigt konstaterande att Storbritannien och USA dominerar den internationella forskningen om bedömning i skolan.

Frågeställningar

Den övergripande frågan jag ställer mig i denna studie är vilka aspekter gällande bedömningsbegreppet som är av relevans för utformandet av skriftliga omdömen. För att ta reda på detta studeras aktuell forskning gällande bedömning utifrån följande frågeställning:

Hur definieras olika former av bedömning?

I denna frågeställning ryms följande funderingar:

- Vilka olika former av och syften för bedömning finns?
- Hur definieras dessa? Vilka specifika begrepp används?
- Vad styr författarens val av bedömningsbegrepp och definitioner?
- Hur används bedömningsbegrepp?
- Vilka möjligheter och/eller svårigheter gällande bedömning med såväl framåtsyftande som dömande karaktär påtalas, och hur motiveras dessa ställningstaganden?

Avgränsningar

Då skriftliga omdömen görs av läraren, är det dennes bedömning i undervisningen som studien fokuseras på. Det innebär att den bedömning som elever själva gör inte undersöks. Undersökningar som enbart behandlar bedömning av elever i högre utbildningar faller också utanför ramen för denna studie, då de bedöms ha lägre relevans i och med att skriftliga omdömen ges elever i grundskolans tidigare år.

Metod

Det behövs ingen längre studie av metodlitteratur för att konstatera att forskningsöversikt som metod inte är särskilt vanligt förekommande i uppsatsskrivande. Exempel på upplägg av en sådan studie är både få och fåordiga, kanske beroende på det faktum att metoden inte är särskilt vanlig, något som konstateras av Stukát (2005). I metodavsnittet redogörs för hur denna studies upplägg slutligen blev, samt de val som ledde dit.

Om forskningsöversikten som metod

Forskningsöversikten är enligt Wigblad (1997) en lågstrukturerad, indirekt metodik. Med lågstrukturerad menar han att oväntade svar kan förekomma. Detta är givet i en forskningsöversikt, då svaren på undersökningsfrågorna inte är begränsade till ett fåtal i förväg specificerade svarsalternativ. Det är även rimligt att anta att underlagen för de olika vetenskapliga texter som är av intresse för denna studie ser olika ut, liksom författarnas frågeställningar och perspektiv. Då forskningsöversikten är en studie av andras forskningsresultat och alltså en hantering av sekundärinformation handlar det om en indirekt metodik. Backman (1998) menar att forskningsöversikten utan en analytisk strategi riskerar att bli en ytlig litteraturinventering istället för en vetenskaplig analys. Samma noggranna arbete och systematik som används vid granskning av primärdata ska användas vid denna typ av studie (ibid). Bryman (2001) berör inte forskningsöversikten som metod, men för en diskussion kring innehållsanalys, vilket enligt honom är mer ett angreppssätt än en forskningsmetod, då det snarare handlar om att analysera än att generera data. Objektivitet och systematik är två begrepp som Bryman lyfter fram som centrala i en innehållsanalys, och det är enligt honom just tydligheten gällande analysen som gör att innehållsanalysen ändå är att betrakta som forskningsmetod. Objektiviteten i innehållsanalysen ligger i ett tydligt klagörande av metod för kategorisering av materialet för studien, distinkt formulerad för att minimera personliga värderingar, medan systematiken i en observationsanalys ligger i ett konsekvent tillämpande av de regler som satts upp för analysen (ibid).

Metodval

För att hantera och strukturera upp information krävs en strategi. Här presenteras val av metod och strategi, samt metodanalysen som ledde dit.

Bakgrund till metodval

Hantering av vetenskapliga dokument kräver en analytisk strategi som grundar sig i valet av synsätt, perspektiv och ansats. Samtidigt som författarna för de olika metodböcker jag studerat presenterar sina tankar kring detta är det enbart Backman (1998) som ger ett konkret förslag till hur en uppläggning av en forskningsöversikt kan se ut. Wigblad (1997) visar genom några exempel olika perspektiv man som författare kan ta vid denna typ av studie. Wigblads bakgrund är dock inom industriell ekonomi, och det är även inom det området han presenterar sina exempel, en jämförande analys av fyra olika studier kring en specifik händelse. Jag studerar ett begrepp, hur det tolkas och används och har svårt att se hur något av

Wigblads angreppssätt ska kunna vara användbart i min studie. Att låta Backmans (1998) modell styra hade varit enkelt, särskilt som han är väldigt enkel och lättfattlig i sitt uttryck, till skillnad från Wigblad (1997). Men Bryman (2001) lyfter i sitt resonemang möjligheten att skapa sin egen modell, bara man har en tydlig strategi, och detta lockar mig, nu när jag ändå gör en typ av studie där det inte ges särskilt många tydliga direktiv. Detta resonemang stärks av Sohlberg och Sohlberg (2001), vilka påpekar att de i varierande grad standardiserade regler som styr uppsatsskrivandet inte har något egenvärde i sig. De lyfter noggrannhet och systematik som viktiga komponenter för en god studie, men framhåller transparens som det viktigaste begreppet gällande kunskapssökande inom vetenskap och säger att "All god forskningskultur kännetecknas av transparens och öppenhet" (Sohlberg & Sohlberg, 2001, s. 14). De menar att en läsare som är insatt i ämnet utan problem ska kunna följa författarens resonemang och att det tydligt ska framgå när det handlar om tolkning av andras forskning och när det är egna reflektioner. Denna transparens är enligt Sohlberg och Sohlberg överordnad procedurreglerna.

Val av metod

Vad är det då som styr mina val kring strategi för hantering av information? En problematisering av ett begrepp innebär att slutsatser och resultat i olika studier och artiklar jämförs med varandra. Det ligger i undersökningens analys att väga in vem som författat texten och det underlag som resultaten eller slutsatserna som framhålls bygger på. Enligt Ödman (2007) är det viktigt att tydliggöra sina tankar, tolkningar och hur man förstår en text och att fokusera på sakförhållanden i förståelsen. Bryman (2001) lyfter tolkningsperspektivet som typiskt för en kvalitativ forskningsstrategi. Utgångspunkten i denna studie är en frågeställning, inte ett påstående. Jag har ingen teori som jag ämnar pröva, utan analyserar materialet utifrån noggrant avgränsade frågeställningar med intentionen att undersökningen i slutändan kan komma att resultera i en teori, vilket stämmer väl med Brymans beskrivning av induktion. Mitt primära syfte är dock inte att generera en egen teori, utan att göra en noggrann analytisk studie av aktuell forskning gällande bedömningsbegreppet. Det handlar om jämförande analys där, förutom en jämförelse av begreppstolkning och begreppsanvändning, en redovisning och analys av eventuella argument för dessa görs. Även om min studie är tolkande, vill jag inte kalla den hermeneutisk⁴ som metod. Den problematisering av tolkningar som ett hermeneutiskt arbetssätt innebär (Ödman, 2007), är säkert möjlig i denna studie och att rekommendera gällande de texter i studien som är av argumentativ karaktär. Det som lockar och kan ses som eftersträvansvärt i ett hermeneutiskt förhållningssätt är öppenheten till det som tolkas och till förståelsen. Även om underlagen för resultaten i de vetenskapliga artiklarna kommer att läggas under lupp, är de många till antalet och av skiftande karaktär, vilket gör att en hermeneutisk metod för att analysera dem känns som ett svåruppnåeligt ideal. Detta är något som bekräftas av Ödman, som lyfter fram svårigheten att arbeta hermeneutiskt och menar att även om ambitionen finns, saknas ofta den problematisering av och argumentation för tolkningar, som är karaktäristiskt för den hermeneutiska metoden.

Systematiskt urval och analysmetoder

För att tydliggöra resultaten och lyfta fram de begreppsval som olika forskare gör väljer jag att använda de mindre vanliga bedömningsbegreppen *övergripande*, *dömande* och *framåtsyftande bedömning* i min analys och diskussion. Dessa begrepp presenteras under

⁴ Hermeneutik är en vetenskaplig metod, med ursprung i teologin, som används för att tolka och förstå text genom att förstå personen bakom texten och den kontext inom vilken texten producerats (Bryman, 2001).

rubriken "Begreppsförklaringar". Jag ser en poäng med att själv inte använda de bedömningsbegrepp som vanligtvis förekommer i forskningen. Detta då ett eget användande av sådana begrepp kan ge en skev bild av hur pass vanligt förekommande de olika begreppen är inom forskningen och hur de används. Här följer en presentation av övriga praktiska överväganden och val som gjordes gällande urval och analys.

Urval

Urvalsproceduren handlar inledningsvis om att finna de vetenskapliga texter som är av relevans för studien. För att göra detta gjordes sökningar i bibliotek och databaserna ERIC och Academic Search Elite. Begreppen *bedömning*, *skriftliga omdömen*, *betyg*, *summativ*, *formativ*, *lärande*, *målrelaterad bedömning* och de engelska begreppen *assessment*, *formative*, *summative*, *learning*, *criterion-referenced measurement*, *grading* och *marks* användes, samt kombinationer och varianter av dessa begrepp. Normativ bedömning valdes bort som sökbegrepp, då vi idag har ett målrelaterat betygssystem och det inte finns några uttalade planer om att återinföra ett normativt betygssystem. Ganska snart urskiljdes ett mönster i referenslistorna hos artiklar som var av intresse för studien. Detta mönster var gemensamma referenser, ibland specifika återkommande texter, men oftare handlade det om författare som återkom. Då fortsatte ett sökande utifrån referenslistor, för att skaffa tillgång till dessa högst sannolikt relevanta texter och författare. Dessa ledde vidare till andra texter och författare och till slut fann jag mig i läget där sökresultaten ständigt visade material jag redan hade. Jag insåg då att jag ringat in området väl och att jag ur det digra material jag hade skulle ha råd att kassera en betydande mängd och ändå ha tillräckligt med underlag för denna studie. Under arbetet med insamlingen av underlag för studien gjordes en enklare kategorisering av materialet utifrån en analys av vilka aspekter av bedömning som texten ifråga berörde.

Gradering gällande relevans

Nästa steg i urvalet handlar om analyserande av text, där en gradering av relevans för studierna ytterligare görs bland det utvalda materialet och en del av det insamlade materialet sorteras bort som irrelevant för studien, medan annat material ges olika grad av relevans. Detta urval görs utifrån frågeställningar om, hur och i vilken grad texten svarar på någon eller några av analysfrågorna i studien. Frågeställningarna är öppna, vilket ger mig utrymme till analys av sådant som formuleringar, perspektiv och eventuella värderingar författaren förmedlar via sin text. Bryman (2001) talar om "latent innehåll", intryck man får av författaren, genom exempelvis hur tonvikt lagts på olika teman i text och påvisar även den möjlighet till oförutsedda svar som öppna frågor ger, det Wigblad (1997) kallar lågstrukturerad metodik. Graderingen skedde genom att jag inledningsvis studerade artiklars abstrakt och böckers innehållsförteckningar och sammanfattningar för att skapa mig en bild av om bedömningsbegreppet diskuterades. De artiklar som föreföll intressanta lästes igenom och sorterades utifrån hur de behandlade bedömningsbegrepp, om de var problematiserande eller enbart konstaterande och vilka bedömningsbegrepp som behandlades.

Underlag för studien

Då jag ville fokusera på aktuell forskning hade jag inledningsvis föreställt mig att använda material som inte var mer än fem år gammalt, det vill säga från 2004 och framåt. Gällande den anglosaxiska forskningen fann jag inte några större svårigheter att hålla detta, med undantag för enstaka artiklar som var ständigt återkommande som referenser i en betydande del av materialet och som jag efter närmare granskning bedömde relevanta att ha med. Detta utifrån att det handlade om större undersökningar som flera andra forskare byggt vidare sina studier utifrån. Att avstå från att använda dessa kändes som att välja bort primärkällor. När det kom till den svenska forskningen blev det betydligt svårare att hålla femårsgränsen, detta då

det är väldigt lite svensk forskning som diskuterar bedömningsbegreppet, utan fokus inom detta område ligger främst på betyg. Visserligen är betyg en högst relevant del av bedömningen inom skolan, men utan en diskussion kring bedömningsbegreppet var det en del av materialet inom denna genre som inte uppfattades som relevant för studien. Jag konstaterade snabbt att jag skulle få med en del intressant material om jag valde att utöka ramen för det svenska materialet med två år, och gjorde därför så. Vid den första insamlingen av material fanns drygt hundra källor som bedömdes kunna vara av relevans för studien. Efter genomläsningar och ytterligare urval inom detta material fanns till slut 48 källor, varav 27 svenska, och 21 anglosaxiska. Av det anglosaxiska materialet var en övervägande del, 16 källor, brittiska, följt av fyra amerikanska och en australiensisk. Anledningen till denna till synes sneda fördelning grundar sig i svårigheter i att finna amerikanska artiklar som inte fokuserade på betyg, i kombination med en rik bedömningsdiskussion inom brittisk forskning. När det gäller den anglosaxiska forskningen består underlaget för studien av 15 artiklar, tre kapitel ur ett par antologier, två böcker och ett paper. Gällande svensk forskning består materialet av fyra böcker, 16 kapitel ur antologier och artikelsamlingar, tre artiklar, tre doktorsavhandlingar och en licentiatuppsats. Utöver dessa källor refereras i resultatet även till ett par debattartiklar, vilka inte tillskrivs vetenskaplig grund, men som visar de diskussioner som pågår kring bedömning och vilka känslor ämnet kan väcka hos lärare.

Teser och argument

I nästa led genomförs ytterligare genomläsningar av materialet, där analys av författare och underlaget för uttalanden görs, för att kunna värdera textens tillförlitlighet. Under genomgången av texterna förs även anteckningar utifrån intryck, vad tesen, det författaren hävdar är, och med vilka argument författaren styrker sina resonemang och slutsatser. Boréus och Bergström (2005) talar om tre olika sorters teser, normativa, vilka beskrivs som värderande, preskriptiva, vilka manar till handling och faktuelle, ickevärderande, vilka ger intryck av att presentera fakta eller sanningar. Det känns rimligt att anta att den senare sorten är den som är mest vanligt förekommande i vetenskapliga texter. För att göra en argumentationsanalys behöver teserna synliggöras och argumenten för eller mot dem granskas. Genom att studera författarens skäl för eller mot argumenten fördjupas analysen med argument av andra ordningen och ibland även argument i ännu fler led (se Sohlberg & Sohlberg, 2001; Boréus & Bergström, 2005). Boréus och Bergström uttrycker följande gällande forskningsartiklar och contra-argument:

Beroende på genre är det mer eller mindre troligt att man finner både pro- och contra-argument i samma text. /---/ Många forskningsartiklar är inte argumenterande, men i de som har sådana inslag är det vanligt att man finner en behandling även av motargument mot det författaren själv hävdar. Så ser nämligen idealet för detta slags text ut: även de starkaste argumenten mot den egna uppfattningen bör redovisas och man bör kunna bemöta dem om man står fast vid sin egen tes. (Boréus & Bergström, 2005, s. 103)

Även Sohlberg och Sohlberg (2001) lyfter användandet av motargument som en metod för att övertyga andra och öka sin egen trovärdighet. Det faktum att pro- och contra-argument kan finnas i flera led innebär att kodningen kan bli tidsödande. Många forskningsartiklar är dock inte av argumenterande karaktär (Boréus & Bergström, 2005). Analysen genomfördes genom en enkel kodning innehållande tes och tillhörande pro- och contra-argument. Vilka argument talar för respektive mot tesen? Sohlberg och Sohlberg (2001) påtalar att teser ibland kan vara svåra att få grepp om och rekommenderar i sådana fall att härleda tesen genom att studera argumenten, alltså genom att gå bakvägen. Den avsaknad av argument som Boréus och Bergström (2005) påtalar blir tydlig i analysen, då de argument som hittades var få och sällan i mer än ett eller två led. Att finna teser via argument visade sig vara desto svårare, då

bedömningsbegreppen ofta inte var det centrala, särskilt i de svenska källorna. Jag väljer att kalla denna analys av enklare slag. Detta då jag inledningsvis ställde mig tveksam till hur många argumentled jag skulle finna och sedan i arbetet med analysen fick mina misstankar om en betydande avsaknad av argument bekräftad.

Bedömning av argument

Hur ska då de argument som finns bedömas? Gällande osaklighet i argumentation lyfter Sohlberg och Sohlberg (2001) upp fyra knep som brukar användas. Det första är utnyttjande av auktoritet, att ge sken av vetenskaplighet. I debattartiklar handlar det om att lyfta fram författarens meriter och titlar, för att påvisa status. Det andra är att undergräva en alternativ ståndpunkt. När argumenten för den egna ståndpunkten inte räcker till attackerar man istället en ståndpunkt man vill ta avstånd från för att hävda sin rätt. Utnyttjande av värderande ord är ett tredje sätt att förvränga sanningen enligt Sohlberg och Sohlberg. Att framställas som snål låter inte lika bra som att bli kallad ekonomisk. Det fjärde och sista knepet för osaklig argumentation som de lyfter är det enkla i att undanhålla obekväma fakta.

Ytterligare sorteringar och kategoriseringar av begreppsdefinitioner och begreppsanvändande i de olika vetenskapliga texterna genomförs utifrån anteckningar i form av kategoriseringar gällande jämförelser kring hantering och definiering av begrepp och eventuella personliga ställningstaganden hos författarna. Utifrån svaren på de undersökningsfrågor som ställs byggs bilden av bedömningsbegreppet upp.

Resultat

För att ge resultatredovisningen en tydlig struktur och för att göra den lättöverskådlig har följande uppdelning gjorts: Inledningsvis presenteras en sammanställning av de definitioner och diskussioner gällande bedömningsbegrepp som identifierats i studier av aktuell forskning inom området. Först redovisas den svenska forskningen och därefter den anglosaxiska. Avslutningsvis görs en jämförelse gällande användandet av begrepp och fokus inom forskning om bedömning.

Bedömningsbegrepp i svensk forskning

Definitioner av bedömningsbegrepp är inte särskilt vanligt förekommande i den svenska forskningen. De begreppsdefinitioner som identifierats i studien presenteras här, följt av en beskrivning av hur bedömningsbegreppen används. Presentationen avslutas med en kort sammanfattning.

Definierande och tolkande av bedömningsbegrepp

Presentationen av bedömningsbegrepp är uppdelad i kategorierna övergripande, dömande och framåtsyftande bedömning, vilka är de begrepp Newton (2007) använder för att särskilja olika typer av bedömning och även de begrepp som jag valt att använda mig av i denna studie, och som presenterades i begreppsdefinitionen.

Övergripande...

Två olika begrepp används som övergripande gällande bedömning i skolan. Forsberg och Lundahl (2006) använder sig av begreppet *kunskapsbedömning* i en intention att närma sig det internationella perspektivet inom pedagogisk forskning, som de menar skiljer sig från det svenska perspektivet, vilket de beskriver med följande ord:

Bedömningar av olika slag har i svensk pedagogisk forskning ofta betraktats som åtskilda fenomen. Diagnoser, läxor, prov, individuella utvecklingsplaner, utvecklingssamtal, betyg, nationella prov, uppföljningar, utvärderingar, kvalitetsredovisningar, inspektioner, internationella kunskapsmätningar etcetera har därför sällan studerats och analyserats inom ramen för ett sammanhållet perspektiv. (Forsberg & Lundahl, 2006, s. 7)

Det internationella perspektivet, som Forsberg och Lundahl väljer att ta, innebär att alla dessa olika former av bedömning ses som olika aspekter av samma sak, kunskapsbedömning, vilket Forsberg och Lundahl ger samma innebörd som det anglosaxiska *assessment in education*. Lundahl (2006) menar att:

/.../ kunskapsbedömningar kan förstås som produktionen av ett vetande om ett barn och deras relation till sina kunskaper. (Lundahl, 2006, s.407)

Även Korp (2003) och Lindström och Lindberg (2005) väljer begreppet kunskapsbedömning när de intar ett övergripande perspektiv på bedömning i skolan, men ingen av dem definierar begreppet. Vid genomläsning av den engelska sammanfattningen av Korp (2003) används *educational assessment* som uttryck för kunskapsbedömning, vilket tolkas som att Korp

likställer dessa begrepp med varandra. Det anglosaxiska *assessment in education* översätts av Nyström (2004) med uttrycket *bedömning i skolan*. Han definierar begreppet som bedömning i alla former i utbildningssammanhang.

...dömande...

Summativ bedömning summerar elevens kunskaper vid en viss tidpunkt (Björklund Boistrup, 2005). Det finns enligt Korp (2003) alltid ett värderande moment i *summativa bedömningar*. Hon menar att det handlar om bedömning för rangordning av elevens kunskaper eller betygsättning av någon form och saknar pedagogiskt syfte och uttrycker gällande summativa bedömningar att:

Deras grundläggande syfte är /.../ att generera en så tillförlitlig bild som möjligt av sådana kvalitativa och kvantitativa aspekter av elevernas lärande som är relevanta i förhållande till undervisningsmålen, och värdera och tillskriva dessa någon form av betyg eller omdöme. (Korp, 2003, s. 77)

Ur detta grundläggande syfte urskiljer Korp fyra olika syften för summativ bedömning. Dessa är *certifikation*, att intyga att en person har en viss kompetens, *differentiering*, att skilja sökande åt i en urvalsprocedur, *kvalifikation*, att garantera att eleverna får den utbildning de har rätt till och slutligen *kvalitetssäkring* av utbildning (Korp, 2003, s. 78-79).

...och framåtsyftande bedömningsbegrepp

Genomgående inom svensk forskning är begreppet *formativ bedömning* det frekvent använda när det gäller framåtsyftande bedömning. Pettersson (2005) definierar formativ bedömning som "en bedömning som ska forma lärandet" (s. 32). Även Lundahl (2007) lyfter den starka kopplingen till själva lärandet som grundläggande för formativ bedömning. Syftet med formativ bedömning är att verka som ett stöd för lärandeprocesser, enligt Björklund Boistrup (2005), som också talar om formativ bedömning som en del av undervisningen. Detta överrensstämmer väl med Korp (2003) som talar om de dagliga utvärderingar som lärare gör i sin undervisning som en form av formativ bedömning. Ett liknande resonemang för Tholin (2007), som lyfter att formativ bedömning ges under pågående arbete och att syftet oftast är att ge information som möjliggör förbättringar under den arbetsprocessen. Diagnoser av olika slag har alltså formativa syften (ibid). Lindström och Lindberg (2005) fokuserar på kunskapsbedömning med pedagogiskt syfte och beskriver det som bedömning som görs utifrån frågeställningar kring syftet med bedömningen, vad som ska bedömas och hur bedömningen ska ske. I resonemanget kan vikten av en medvetenhet hos bedömaren kring dessa frågeställningar utläsas som central för *pedagogisk bedömning* (ibid). Lindström (2005a) framhåller återkoppling som det viktigaste inslaget i formativ bedömning.

Användande av bedömningsbegrepp

Under senare år har bedömning börjat användas för både fler och andra syften än vad tidigare varit vanligt i skolan (Giota, 2006). Dessutom har fokus för bedömningen flyttats från att kontrollera elevens detalj- och faktakunskaper till att kontrollera deras kunskaper om processer (Lindberg, 2002). Här ges en inblick i hur diskussionerna kring användandet av bedömningsbegrepp ser ut inom svensk forskning. Frågan om huruvida summativ och formativ bedömning kan närma sig varandra, och om de bör det, diskuteras flitigt enligt Giota (2006). Hon menar att det inte finns några enkla svar i denna fråga, men att oavsett hur läraren löser bedömningen har det konsekvenser både för undervisningen och för eleven som blir bedömd.

Övergripande

Kunskapsbedömning har, enligt Selghed (2006), oavsett syfte skett på ungefär samma sätt tidigare. Han menar att det behövs en anpassning av bedömningsformen utifrån vad som ska bedömas för att bedömningarna ska spegla sanningen. Uppgiften att kvalitetsbedöma elevers kunskaper ställer enligt Selghed dessa krav på variation av bedömningsformer. Även Nyström (2004) menar att:

Bedömning utgår /.../ från en strävan att skapa bästa möjliga förutsättningar för alla att visa vad de kan. (Nyström, 2004, s. 3)

Nyströms uppfattning om bedömning är väl överensstämmande med Selgheds (2006). En kvalitetsbedömning av elevers kunskaper som är så sann som möjligt borde rimligtvis vara en där eleven givits förutsättningar för att prestera sitt bästa. Lundahl (2006) menar att det finns olika sätt att förstå kunskapsbedömningar:

Ur ett reformperspektiv kan kunskapsbedömningar förstås som en politisk eller administrativ funktion inom skolan, som kan justeras och ändras allt eftersom samhällets förväntningar och behov ändras. Med ett sociologiskt perspektiv framstår kunskapsbedömningar som ett symboliskt kapital, med vilket privilegierade grupper återskapar sin privilegierade position i samhället. Skolan sägs vara rättvis, men betyg och dylikt bidrar till att alla ändå inte får samma chans i samhället. Med ett psykometriskt perspektiv⁵ hamnar ofta mättekniska frågor i förgrunden. Tekniker avlöser varandra genom historien i jakt på det fulländade mätinstrumentet eller betygssystemet som garanterar objektivitet, rättvisa och tillförlitlighet i bedömningarna. Att förstå kunskapsbedömningar i termer av vetande, är ett sätt att frigöra sig från dessa typer av dominerande historiebetyg, och återge kunskapsbedömandets historia ett pedagogiskt perspektiv. (Lundahl, 2006, s. 407)

Det psykometriska perspektivet som Lundahl talar om har inte haft särskilt starkt fäste inom svensk skola (Nyström, 2004). Lundahl (2006) menar att samtidigt som kunskapsbedömningar är typiska för skolan präglas de av förändringar av kunskapssyn i samhället. Han intar själv ett vetandeperspektiv, vilket han menar möjliggör synliggörande av "relationen mellan skolans produktion av eget vetande och reproduktion av andras vetande" (Lundahl, 2006, s. 407).

Dömande

När man tänker på bedömning av dömande karaktär är betygen det tydligast framträdande exemplet på denna form av bedömning. Betyg kan ha olika funktion, de kan fungera som urvalsinstrument inför högre utbildningar, de kan ha en informativ funktion till elever och föräldrar, en kontrollfunktion för nationella och internationella jämförelser, en disciplinerande funktion och de kan ha en motiverande funktion (Nordgren, 2008). Nordgren framhåller att bedömning har en ideologisk dimension samtidigt som det handlar om makt. Han uttrycker detta i följande termer:

Skolan är en del av ett samhälle och påverkas av dess politiska, sociala, kulturella och ekonomiska strukturer. Det finns ingen neutral "betygslära", ingen objektiv bedömningsmetod som saknar koppling till ideologi, makt och moral. Det går inte att komma förbi att alla bedömningssystem är uttryck för vissa kunskapssyner som premierar vissa kunskaper och lärostrategier men diskriminerar andra. (Nordgren, 2008, s. 9)

⁵ Ett psykometriskt perspektiv utgår från psy-vetenskaperna (se fotnot 2, s. 6).

Nordgren lyfter också det etiska åtagande som ligger i värderandet av en annan persons prestationer och kunnande. Han talar om en bedörmakt hos den som utför bedömningen, en makt som medför stort ansvar då elevers framtid och självkänsla kan påverkas av denna. Även Wehner-Godée (2005) påtalar den effekt bedömning kan ha på en elevs självkänsla och självbild. Hon framhåller att det inte går att undvika bedömningar, då vi alla bedömer varandra dagligen, och menar att bedömningar kan vara en fantastisk hjälp för oss i vår utveckling, men att negativ bedömning kan ha motsatt effekt. Om bedömning av kunskaper tolkas som bedömning av person kan det sätta tydliga spår hos den vars kunskaper bedöms (ibid). Detta kan kopplas till Andersson (2007) som menar att många elever uppfattar betyget Godkänd som knutet till personligheten och att betygen för en del elever är en bekräftelse på misslyckanden. Emanuelsson (2002) anser även han att det finns anledning att fundera över varför man kom att använda Godkänd som betygsform. Han menar att elever redan tidigare lätt definierat betyg som knutet till deras person och att denna uppfattning stärks med denna betygsformulering. Idag är ett nytt betygssystem för grundskola under utarbetande, där detta ifrågasatta begrepp försvinner, men istället kommer eleverna troligtvis inte att möta ett, utan två betyg som motsvarar den tidigare frånvaron av betyg, när eleven inte nått kursplanemålen i ett ämne (Andersson, 2007). Vilken påverkan det har på våra elever får framtiden utvisa. Wehner-Godée (2005) ställer sig frågan:

Hur känns det för ett barn att bli bedömd och få det på pränt från förskolan upp till mellanstadiet? Det är säkert oerhört betydelsefullt och stärkande att bli uppmärksammas för det man är bra på och kan. Men hur påverkar det barnet att bli uppmärksammas för det som lärarna/pedagogerna anser vara brister och som man inte kan? Och hur påverkar det föräldrarnas syn på sitt barn? (Wehner-Godée, 2005, s. 94)

Några år efter att Wehner-Godée ställt denna fråga har skriftliga omdömen från första skolåret införts i svensk grundskola, vilket aktualiserar hennes undran.

Framåtsyftande

Björklund Boistrup (2005) talar om vikten av att förstärka den framåtsyftande bedömningen och hänvisar till en sammanställning av över 20 studier gjord av Black och Wiliam (1998b), vilken tydligt påvisar att en höjning av kvaliteten på den formativa bedömningen bidrar till förbättrat lärande hos eleverna. Som metoder för att förstärka den framåtsyftande bedömningen lyfts portföljmetodik och användande av matriser som en del av undervisningen fram som främjande för elevernas utveckling (Dysthe, 2002; Lindström & Lindberg, 2005). Detta genom att eleverna ges en större delaktighet och inblick i både lärande- och bedömningsprocessen (Kjellström, 2005). Listan med metoder för framåtsyftande bedömning kompletteras av Lundahl (2007) med bland annat traditionella uppgifter, observationer, självvärderingar, problemlösningar, essäer och intervjuer. Han uttrycker att:

I praktiken kan det handla om att få både lärare och elever att anamma ett förhållningssätt till bedömningar där till exempel läraren inte ges rätten att ställa frågor och bedöma svaren utan eleverna ska involveras i både frågandet och bedömandet. (Lundahl, 2007, s. 63)

Den aktiva elevmedverkan som påtalas kräver motivation hos eleverna, ett område som enligt Giota (2006) är eftersatt inom pedagogisk forskning.

Möjligheter och svårigheter med att döma framåtsyftande

Skolans framåtsyftande uppdrag och betygens dömande funktioner framställs enligt Nordgren (2008) ibland som något som står i motsättning till varandra. Han menar att:

Det förra uppdraget knyts till en kunskapssyn där lärandet ses som en komplex process, medan det senare framställs som en enkelspårig kunskapssyn som förlitar sig på mättekniker. Det kan onekligen tyckas vara motsägelsefullt att å ena sidan framhålla kunskap som en konstruktion där individen skapar mening i en social kontext och å andra sidan gradera det eleven uttrycker vid en viss tidpunkt. (Nordgren, 2008, s. 12)

Även Björklund Boistrup (2005), menar att det är vanligt att särskilja de två typerna av bedömning, med motiveringen att summativ bedömning görs som en mätning av kunskaper vid en viss tidpunkt, medan formativ bedömning är integrerad i undervisningen. Som exempel på summativ bedömning framhåller hon de betyg som ges elever i Sverige från och med skolår 8. Till skillnad från Nordgren (2008) påtalar Björklund Boistrup (2005) inga starka motsättningar mellan bedömningsformerna, vilket inte heller Selghed (2006) gör. Han för följande resonemang kring förhållandet framåtsyftande – dömande bedömning:

Distinktionen mellan formativ och summativ bedömning kan vara svår att upprätthålla. Begreppens innebörder är relativa. Samma bedömning kan i ett sammanhang ses som summativ, men i ett annat som formativ. I vårt nuvarande betygssystem uppfattas vanligtvis betygen som ett resultat av summativ bedömning. Men terminsbetygen i skolår 8 och vid höstterminens slut i skolår 8 i grundskolan kan också ses som formativa, eftersom vad eleven uppnått i slutet av grundskolan är det viktigaste i en målstyrd skola. (Selghed, 2006, s. 204)

I Selgheds resonemang framställs begreppen som svåra att särskilja, då bedömning i vissa fall kan ha både dömande och framåtsyftande funktion. Till skillnad från Björklund Boistrup (2005) tillskriver Selghed (2006) de tidigare betygen en formativ funktion, med motiveringen att de inte är definitiva. Detta är ett ställningstagande som jag inte har funnit stöd för någonstans i resterande underlag för denna studie. Linde (2003) talar om bedömningsformer i termerna informell, då lärare lyssnar in elevers förståelse och tankar under pågående verksamhet, och formell, genom exempelvis prov. Det syfte för bedömning han fokuserar på är bedömning av kunskaper. Han uttrycker gällande detta att:

För att kunna ge en adekvat återkoppling till eleven behöver läraren på något sätt diagnostisera elevens uppgiftslösningar. Dessa diagnoser och den nyanserade återkopplingen kan inte ersättas med betyg, vilka inte kan betraktas som nyanserade omdömen. Betyget "Icke godkänd" [gymnasienivå] i engelska /.../ säger inte *vad* det är eleven bör ägna mer kraft åt i detta fall. (Linde, 2003, s. 115)

Den avsaknad av nyanser i återkoppling med betyg som Linde lyfter talar mot Selgheds (2006) en bild av formativa betyg. Andersson (2007) menar att betygens roll som morot och motivationsskapande endast fungerar för de elever som ligger på den övre delen av betygsskalan, medan de för elever med sämre skolprestationer har motsatt effekt och att de därmed inte kan ges status som formativa. Även Jansdotter Samuelsson och Nordgren (2008) resonerar kring en eventuell framåtsyftande roll hos betyg och ställer sig frågan om betygssystemet är en tillgång eller ett hinder när det gäller elevens lärande. De konstaterar att tillämpningen av ett betygssystem borde ha någon form av inverkan, men att det behövs forskning inom området.

Giota (2006) påvisar en brist på pedagogisk forskning gällande elevers motivation och hur motivationen påverkas och påverkas av bedömning. Hon summerar de forskningsresultat hon funnit med följande ord:

De få studier som gjorts inom området konstaterar emellertid att strävan efter höga betyg och att bry sig om sitt eget lärande inte behöver vara oförenliga mål, utan i vissa situationer till och med kan vara ömsesidigt förstärkande. Dessa studier indikerar med andra ord att summativ och formativ bedömning inte behöver utesluta varandra utan vara komplement till varandra. (Giota, 2006, s. 110)

Giota kommer dock inte med några praktiska exempel på hur denna komplettering ska ske. Detta är även fallet för Wikström (2006), som även om hon främst behandlar bedömning för betygsättning, påtalar att bedömning och betygsättning bör integreras med undervisningen och att bedömningen bör ha en formativ såväl som en summativ funktion. Några som däremot ger exempel på bedömning som fungerar framåtsyftande såväl som dömande är Lindström och Lindberg (2005). Kjellström (2005) menar att användande av bedömningsmatriser som en del i undervisningen kan fungera formativt, eftersom elevernas kunskapsutveckling och vidare steg tydliggörs. Men att detta förutsätter att eleverna själva har en aktiv del i användandet av och bedömandet kopplat till matriserna, ett resonemang som stöds av Dysthe (2002). Portfolio som bedömningsmetod är annars den metod som inte ifrågasätts när det gäller bedömning som fungerar framåtsyftande såväl som dömande (Lindström, 2005b; Jönsson, 2005).

Att bedömningsbegreppet är något som väcker diskussioner även i lärarrummet tydliggörs i en debattartikel av Habbe och Björkman (2007). De anser att lärare har svårt att skilja bedömningsbegreppet från betygsättning och därför tenderar att betygsätta enskilda uppgifter och glömma den viktiga kommunikation som bedömning innebär. De uttrycker sin oro med följande ord:

När betygsättning används i samband med bedömning tenderar det att överskugga och ibland helt ersätta vägledningen att utveckla de kunskaper som så tydligt definieras i styrdokumentet. En bedömning av elevernas prestationer ska innehålla information om vad de har uppnått utifrån styrdokumentets mål, vilket på fackspråk kallas summativ bedömning, samt en vägledning om hur eleverna ska gå till väga för att uppnå de kompetenser, kunnande och vetande som hon eller han har missat eller kan förbättra, det kallas på fackspråk för formativ bedömning. Bedömningen av elevernas kunskaper är en mycket viktig del i lärprocessen, ja till och med nyckeln till lärandet. All bedömning innebär kommunikation och kräver ett språk om hur man talar om kunskaper i specifika ämnen och det är med hjälp av det språket som vi vidareutvecklar elevernas kunskaper. Betygsättning däremot innebär, brutalt uttryckt, att läraren avsäger sig viljan att kommunicera. (Habbe & Björkman, 2007, s. 73)

Habbe och Björkman uttrycker även en oro för att man vid införande av betyg i lägre åldrar riskerar att försämra bedömningen. Så som jag tolkar deras inlägg i bedömnings- och betygsdebatten är det den framåtsyftande bedömningen som riskerar att försämrans i och med att den dömande bedömningen ges större utrymme.

Sammanfattning av svensk forskning kring bedömning

Den sneda fördelningen av utrymme de olika formerna av bedömning ges motsvarar den fördelning de olika formerna av bedömning har inom forskningen. Att betrakta bedömning med ett övergripande begrepp är någonting relativt nytt inom svensk forskning. I och med detta är begreppen inte heller inarbetade och klara. De forskare som ger sig i kast med att tala övergripande om bedömning tar även på sig uppgiften att tydliggöra begrepp och

begreppsdefinitioner. Det mest frekvent använda övergripande bedömningsbegreppet inom svensk forskning är kunskapsbedömning. Summativ och formativ bedömning är de dominerande begreppen gällande dömande och framåtsyftande bedömning. Av dessa ges summativ bedömning betydligt större utrymme i diskussioner kring dess innebörd än formativ bedömning. Detta kan bero på ett ifrågasättande av summativ bedömning som är lätt att avläsa i ett flertal texter (se t.ex. Jansdotter Samuelsson & Nordgren, 2008; Linde, 2003). Ifrågasättandet handlar om användandet av summativ bedömning i form av betyg i relation till främjandet av elevers fortsatta lärande, där ren summativ bedömning kan ha en negativ effekt på elevers lärande, självförtroende och självkänsla om den uppfattas som en bedömning av personen (ibid). Det finns en avsaknad av definitioner av formativ bedömning inom svensk forskning. Flera forskare tydliggör att formativ bedömning är viktig och talar om hur och varför utan att tydliggöra vad det är de menar med begreppet (se t.ex. Linde, 2003, Giota, 2006). Å andra sidan framstår det som att de forskare som definierar begreppet formativ bedömning är enade kring begreppsdefinition.

Bedömningsbegrepp i anglosaxisk forskning

Det finns en större mängd bedömningsbegrepp inom anglosaxisk forskning än inom svensk forskning. Detta innebär också en större diskussion kring bedömningsbegrepp, vilket redovisas under denna rubrik.

Definierande och tolkande av bedömningsbegrepp

Här presenteras anglosaxiska forskares tolkning av bedömningsbegrepp. Taras (2005) menar att det finns brister i hur bedömning definieras:

Assessment is of central importance in education, and yet there is a lack of commonality in the definition of the terminology relating to it. Development of both theoretical and practical applications will suffer unless there is coherence and agreement in the definition of the terms. Assessment for learning or for formative assessment /.../ is increasingly being emphasised, yet its relationship to summative /.../ has been little explored. (Taras, 2005, s. 466)

Dessa skillnader och eventuella oklarheter synliggörs i såväl resultatdelen av detta arbete, som i diskussionen.

Övergripande...

Educational assessment, vilket är det begrepp svensk forskning främst relaterar till, används av Newton (2007) som inte definierar begreppet, utan låter det tala för sig själv. Någon som också använder sig av *educational assessment* som övergripande begrepp är Popham (2008), som ger följande definition:

Educational assessment is a formal attempt to determine students' status with respect to educational variables of interest. (Popham, 2008, s. 6)

I klargörandet av denna definition lyfter Popham fram att det handlar om formell och medveten bedömning samt att "variabler av intresse" varierar, beroende på vad det är som ska bedömas. Black och Wiliam (1998b) använder begreppet *assessment* som ett övergripande begrepp omfattande all bedömning i skolan som tillhandahåller information användbar som feedback för att ändra eller vidareutveckla undervisningen. De är av åsikten att all bedömning i undervisningssammanhang ska tjäna syftet att stödja lärandet (Black & Wiliam, 2006),

vilket överensstämmer med Harlens (2005) definition av assessment. Även Tomlinson (2005) använder begreppet assessment som övergripande bedömningsbegrepp och definierar det som insamlade av information som underlag för instruktioner.

...dömande...

Harlen (2005) använder begreppet "*assessment of learning*" när hon talar om dömande bedömning och likställer det med begreppet *summativ bedömning*. Hon definierar "*assessment of learning*" som bedömning med syfte att summera lärande. Den summativa bedömningen kan enligt Harlen delas in i intern och extern summativ bedömning, där den interna omfattar rapportering till elev och föräldrar och intern kontroll inom skolan, medan den externa bedömningen omfattar slutbetyg och muntliga rekommendationer vid tjänsteansökan. Hayward och Hedge (2005) använder sig av begreppet "*assessment for accountability*" när de talar om dömande bedömning, vilket är ett val baserat på att dömande bedömning ofta registreras och protokollförs. Summativ bedömning klargör vad eleven vet och kan göra och ska enligt Welsh och D'Agostino (2009) användas för betygsättning.

... och framåtsyftande bedömningsbegrepp

När bedömningen är av sådan karaktär att den faktiskt används för att anpassa undervisningen efter elevernas behov är den att betrakta som *formativ* enligt Black och Wiliam (1998b). De använder även begreppet "*assessment for learning*", när de talar om framåtsyftande bedömning och menar att den framåtsyftande bedömningens främsta syfte gällande både utformning och användande är att främja elevernas lärande (Black, Harrison, Lee, Marshall & Wiliam, 2004). Harlen (2005) och Hayward och Hedge (2005) använder också begreppet *assessment for learning*, när de talar om framåtsyftande bedömning och likställer det med *formativ bedömning*. Harlen definierar "*assessment for learning*" som bedömning med syfte att hjälpa lärandet. Gardner (2006) menar att skillnaden mellan *formativ bedömning* och "*assessment for learning*" handlar om hur länge begreppen använts. *Formativ bedömning* är ett äldre begrepp, medan *assessment for learning* började användas i slutet av 80- och början av 90-talet. Begreppen står enligt Gardner för bedömning som främjar lärandet. Hargreaves (2007) menar att även om all bedömning påverkar elevens lärande direkt eller indirekt är förbättrad inlärning syftet med och resultat av *formativ bedömning*. Men det finns även andra, mer utförligt sätt att definiera *formativ bedömning*:

Formative assessment is all about sharing information. Teacher-to-student communication – teachers showing students where teachers believe learning should be headed and what students need to do to get there – is important in formative assessment, just as in conventional assessment. But the power of formative assessment comes from the addition of student-to-teacher communication. Each student shows the teacher all along the way where his or her understanding is deep, shallow or stalled. (Brookhart, Moss & Long, 2008, s. 52)

Här sätts tvåvägskommunikationen i fokus som det utmärkande för *formativ bedömning* och Brookhart, Moss och Long (2008) lyfter fram elevens aktiva deltagande i bedömningsprocessen. De menar att denna typ av kommunikation stärker eleven och gynnar elevens intresse av att lära sig, samt gör läraren effektivare. Den *formativa bedömningen* blir en strategi för instruktioner (ibid).

När det gäller framåtsyftande bedömning skiljer Welsh och D'Agostino (2009) på *diagnostisk bedömning* och *formativ bedömning*. *Diagnostisk bedömning* definierar de som att ta reda på vad eleven kan och vilka kunskapsluckor som finns innan man ger instruktioner, medan *formativ bedömning* sker under instruktion eller lektion för att avgöra vad som varit bra med instruktionen eller lektionen och vad som behöver förbättras. De definitioner Welsh och

D'Agostino använder lägger fokus gällande den diagnostiska bedömningen på elevens lärande, medan fokus gällande formativ bedömning läggs på lärarens förmåga att lära ut. Till skillnad från Brookhart et al. (2008) fokuserar Welsh och D'Agostino (2009) på lärare som bedömare och förmedlare av bedömningsinformation.

Användande av bedömningsbegrepp

Flera forskare (t.ex. Black et al., 2004) använder begreppen "assessment for learning" och "assessment of learning" såväl som formativ och summativ bedömning och växlar mellan begreppen i sina texter. Harlen (2006) menar att det i begreppen "assessment for learning" och "assessment of learning" tydligare framkommer att det handlar om samma bedömning, men att syftet är olika, medan formativ och summativ bedömning lätt tolkas som olika typer av bedömning. När det finns behov av att tydliggöra att det är bedömningens syfte som är kriteriet för om bedömningen är dömande eller bedömande väljs begreppen "assessment for learning" och "assessment of learning" (ibid).

I en debattartikel kritiserar Petty (2008) användandet av begreppet *assessment for learning* istället för *formativ bedömning*. Han hänvisar till the King's-Medway-Oxfordshire Formative Assessment project [KMOFAP] som startades 1999 utifrån de resultat Black och Wiliam redovisade 1998 om hur elever lär och där den framåtsyftande bedömningen lyftes fram som central för elevers lärande (Black & Wiliam 1998a; 1998b). Projektet fick sedan en efterföljare med namnet Assessment for Learning [AfL] (se Black et al. 2004), och det är detta namn Petty (2008) ställer sig kritisk till. Han menar att även om formativ bedömning kanske inte är särskilt spännande begrepp som del i ett projektnamn, så uttryckte det ändå klart syftet med projektet, vilket inte är fallet med AfL. Enligt Petty kan vilken betydelse som helst som rör bedömning och lärande läggas in i begreppet "assessment for learning". Han tydliggör det genom att hänvisa till att den brittiska regeringens uttalade syfte med AfL är att genom tydliga mål och *frekventa summativa bedömningar* hjälpa eleverna effektivisera sin inläring, något som även kritiskt kommenterats av Paul Black som en misstolkning av det avsiktliga syftet med projektet (ibid). Även Sadler (2007) uttrycker tveksamhet till begreppet "assessment for learning", men också till "assessment of learning", och av annan anledning än Petty (2008). Sadler (2007) menar att det finns en otydlighet kring begreppet "learning" och menar att det kan referera till en meningsfull aktivitet, en process eller till en mängd av något som lärts in, en produkt. Denna otydlighet bidrar enligt Sadler till att begreppen "assessment for learning" och "assessment of learning" kan ges olika innebörd beroende på hur man definierar learning. En misstolkning och felaktig begreppsanvändning, liknande den Petty (2008) påtalar gällande begreppet "assessment for learning", menar Gardner (2006) har gjorts av begreppet formativ bedömning. Han säger att begreppet har använts i samband med regelbundna mätningar inför en slutlig summativ bedömning, vilket inte bidrar till elevernas lärande och därmed inte kan ses som "assessment for learning" eller formativ bedömning. Detta menar Gardner har skett på grund av att ett flitigt användande av ett populärt begrepp inom bedömning bidragit till en urvattning av begreppets egentliga mening. Även Newton (2007) talar om bedömningsbegrepp som feltolkas och uttrycker sig så här:

The fact that a system which is fit for one purpose will not necessarily be fit for all purposes is a fundamental consideration when evaluating the legitimacy of proposals. It is one of the most important messages for policy-makers to understand. Unfortunately, it can be quite hard to convey the significance of this message; and part of the reason for this is a lack of clarity within the professional discourse of educational assessment. (Newton, 2007, s. 149)

Newton menar att innebörden av dömande, beslutande och framåtsyftande användningsområden för bedömning behöver tydliggöras var för sig, för att en bedömningsdesign som inte är diffus ska kunna utformas. Han uttrycker att det faktum att det finns olika typer av bedömningar, vilka används olika beroende på vilket syfte det finns med bedömningen, medför en risk för misstolkning av syftet (ibid). Den reflektion Perry (2008) gör angående misstolkningar av AfL kan ses som ett tecken på detta.

Taras (2005; 2009), är precis som Newton (2007), kritisk till hur bedömningsbegrepp definieras och används. Newton (2007) menar att begreppen *assessment* och *judgement* lätt blandas ihop och likställs, men menar att det inte finns ”formative judgements”⁶ och inte heller ”summativa syften”. Han motiverar detta ställningstagande med dessa ord:

If you get to be a summative purpose simply by being based on summative judgements, then the definition is hollow: the idea of a summative purpose is conceptually redundant. Likewise, what if a ‘summing up’ judgement is used for formative purpose? Would it simultaneously be both a summative judgement and a formative judgement? Again, if you get to be a formative judgement simply by being used for a formative purpose, then the definition is hollow: the idea of a formative judgement is conceptually redundant. Obviously, formative uses do require assessment judgements. The point is simply that we risk confusing ourselves, and others, when we call them ‘formative judgements’ (since there is nothing about judgements, *per se*, that is formative). It is an assessment judgement used for a formative purpose, not a formative judgement. To avoid getting ourselves confused, and to avoid confusing others, we need to use the language of assessment with greater precision. We may talk of a formative purpose, to indicate the use to which a result is put, but we ought not to talk of a summative purpose. Likewise, we may talk of a summative judgement, but we ought not to talk of a formative one. (Newton, 2007, s. 157)

Både Newton (2007) och Taras (2005) menar att bedömning i grunden alltid är summativ. Den grundläggande bedömningen som görs refererar Newton (2007) till som judgement. Resultatet av bedömningen kan sedan användas på olika sätt (Newton, 2007, Taras, 2005; 2009). Hur detta val görs och hur bedömningen används är det som Newton (2007) kallar assessment. Taras (2005) för inga resonemang kring begreppen judgement och assessment, utan håller sig till det sistnämnda. Hennes kritik gällande hur bedömningsbegrepp används handlar mer om hur de definitioner som görs tillämpas och eftersöker en mer enhetlig syn på hur bedömningsbegrepp definieras och används. Taras menar exempelvis att forskningen som följt i kölvattnet av Black och Wiliam (1998a; 1998b) har haft en kursändring i och med tillägg som skapat inre motsägelser (se Taras, 2009). Detta presenteras mer utförligt under rubriken ”*Inside the black box*” och ”*bibeln*” för *formativ bedömning*.

Möjligheter och svårigheter med att döma framåtsyftande

Black och Wiliam (1998b) menar att det inte går att hålla summativ och formativ bedömning isär, då det handlar om bedömning som läraren har ett ansvar för. De menar dock att det inte är någon lätt uppgift att förlika sig med de två roller som summativ och formativ bedömning innebär och påtalar att denna svårighet kan utgöra ett hinder för läraren att förbättra undervisningen. Wiliam (2006) uttrycker:

⁶ Begreppen *judgement* och *assessment* översätts i svenskan båda till bedömning, men innebörden skiljer sig åt. Judgement är ett mer dömande begrepp än det mer neutrala assessment. För att undvika misstolkning har begreppet ”formative judgement” inte översatts.

Assessment that serves the need of teachers and their students is seen as completely separate from, and indeed, incompatible with, assessment that serves the need of parents, administrators, policy makers and other stakeholders. (Wiliam, 2006, s. 1)

Wiliam menar att denna osammanhängande bild av bedömning är vanlig bland lärare i hela världen. Gällande hur man bedömer elever nämns prov och andra formella metoder, medan informella metoder som samtal och observationer i klassrum är övervägande vid frågor om hur de vet att elever lärt sig vad de lärt dem. Denna bild av bedömning menar flera forskare är ett hinder i utvecklingen av en bedömning som fungerar såväl dömande som framåtsyftande (se t.ex. Black, Harrison, Lee, Marshall & Wiliam, 2003). Harlen (2005) anser att summativ och formativ bedömning inte är två olika typer av bedömning, utan att det istället rör sig om bedömning med olika syften, något som Taras (2005) instämmer i. Harlen (2005) menar att en och samma information, insamlad på ett och samma sätt beroende av hur det används tillskrivs antingen en summativ eller en formativ funktion (se tabell 1). Newton (2007) särskiljer dömande och framåtsyftande bedömning på så vis att han, till skillnad från Harlen (2005) och Taras (2005; 2009), ser det som olika användningsområden istället för syften för bedömningar. Däremot ser Newton (2007) precis som Harlen (2005) syftet med bedömningen som central, medan Taras (2005; 2009) önskar se ökat fokus på bedömningsprocessen.

Tabell 1. En möjlig dimension av bedömningssyften och -praktik.

	<i>Formative</i>		<i>Summative</i>	
	Informal formative	Formal formative	Informal summative	Formal summative
Major focus	What are the next steps in learning?		What have been achieved to date?	
Purpose	To inform next steps in learning	To inform next steps in teaching	To monitor progress against plans	To record achievements of individuals
How is evidence collected?	As normal part of class work	Introduced into normal class work	Introduced into normal class work	Separate task or test
Basis of judgement	Student referenced	Student and criterion referenced	Criterion referenced	Criterion referenced
Judged by	Student and teacher	Teacher	Teacher	Teacher or external marker
Action taken	Feedback to students and teacher	Feedback into teaching plans	Feedback into teaching plans	Report to student, parent other teachers, etc.
Epithet	Assessment for learning	Matching	Dip stick	Assessment of learning

(Harlen, 2006, s. 114)

När det gäller framåtsyftande användande av resultat av dömande bedömning anser Harlen (2006) att man behöver undersöka resultaten utifrån frågeställningar som dessa:

- Does it focus on how students learn?
- Is it sensitive and constructive?
- Does it foster motivation?
- Does it promote understanding of goals and criteria?
- Does it help learners to know how to improve?
- Does it develop the capacity for self-assessment?
- Does it recognize all educational achievements?

(Harlen, 2006, s. 108)

Frågeställningarna är fokuserade kring eleven och dennes medvetande om det egna lärandet och delaktighet i bedömningsprocessen. Dessa är kriterier Harlen menar behövs för att bedömning verkligen ska fungera framåtsyftande. Även Tomlinson (2005) talar om elevers inställning till lärande och bild av sig själva som lärande är av betydelse för utvecklandet av deras engagemang i lärandeprocessen. När det gäller att ge respons på genomfört arbete menar Black et al. (2004) att mängden respons är av mindre betydelse, men att utformningen av den är desto viktigare. De styrker detta påstående med att hänvisa till egen forskning som visat att medan elevens lärande kan förbättras genom respons på uppgifter i form av kommentarer, uteblir den positiva effekten om responsen ges i form av betyg eller poäng. Om eleven ges både betyg och kommentar läggs fokus på betyget och kommentaren ignoreras, vilket innebär att kommentarens positiva effekt uteblir. Black et al. (2004) menar att lärare ofta förvånas över dessa resultat, men att de som övergår till att enbart kommentera elevarbeten bekräftar dessa uppgifter, elevernas engagemang i att förbättra sitt eget arbete ökar. Summativa prov, som stämmer av och mäter kunskaper, kan enligt Black et al. (2003) ges en formativ funktion om eleverna engageras i processen kring proven, utformning, bedömning av resultat och analys. Detta då eleverna ges möjligheten att se den egna nyttan med proven, istället för att se dem som ett nödvändigt ont (ibid). Portfolio lyfter Harlen (2005) fram som en form av bedömning som är utformad så att det fungerar framåtsyftande såväl som dömande. Prov, och särskilt externa prov som till exempel nationella prov, kan enligt Harlen (2006) bidra till att identifiera vad som behöver arbetas vidare med, men hon tillskriver dem ingen framåtsyftande funktion med motivationen:

- The collection of summative evidence does not occur sufficiently frequently;
 - The information is not sufficiently detailed to be diagnostic;
 - It only occurs in reality in relation to tasks chosen, administered and marked by teachers;
 - There is a danger of mistaking frequently collected summative evidence for evidence that can be used formatively and therefore neglecting genuine assessment for learning;
 - Using external tests in this way risks teaching to the tests;
 - It rarely matches the principles of assessment for learning.
- (Harlen, 2006, s. 108-109)

Detta motsägs till viss del av Black et al. (2003), som menar att provresultat kan ges en framåtsyftande funktion genom att involvera eleverna i utformandet och utvärderandet av uppgifterna. Inom den dömande bedömningen är betygssättande en central och frekvent förekommande del och även om bedömning kan fungera framåtsyftande såväl som dömande, så är relationen mellan betyg och framåtsyftande bedömning en annan (Tomlinson, 2005; Black & Wiliam, 2005). Tomlinson (2005) menar att:

Grading is making an end-point judgement about students' achievement. Formative, or on-going, assessment should not all be graded. Its purpose is to help both teacher and student see how learning is progressing and to adjust as necessary to make sure learning stays on course. Grading on-going assessment can undermine student willingness to take mental risks, absorb an undue amount of teacher time, and contribute to a skewed view of what a student actually learns... (Tomlinson, 2005, s. 264)

Detta kan kopplas till att ökat testande av elever i kontrollerande syfte har visat sig motarbeta effektiv formativ bedömning (Black & Wiliam, 2005).

När det kommer till framåtsyftande bedömning som kan användas i dömande syfte talar Harlen (2006) om vikten av att sätta upp tydliga korta delmål som eleverna arbetar mot. Det blir då lätt för lärare såväl som för elev att se när målen nåtts och vad som ska arbetas vidare

med, då målen som följer också är tydligt specificerade. Hur dessa mål ser ut och hur de följer på varandra tydliggörs i listor, utan nivågraderingar, utan bara som en hjälp för hur man arbetar vidare, vilket överensstämmer med syftet med formativ bedömning. För att bedömningen ska kunna användas även summativt behöver det enligt Harlen tydliggöras hur pass väl en elev klarar ett mål. Detta görs genom att sammanställa bedömningen från flera olika lektioner med samma mål, där eleven ges möjlighet att påvisa sitt kunnande. Det finns svårigheter även med att få resultat av formativ bedömning uppfylla kriterierna för summativ bedömning. Dessa svårigheter är enligt Harlen:

- It is essential to reinterpret the evidence in relation to the same criteria for each student;
 - It is important for teachers to be very clear about when level-based criteria are appropriate and not to use them for grading students when the purpose of assessment is formative;
 - Since the formative use of evidence depends on teachers' judgements, additional quality assurance procedures will be needed when the information is used for a different purpose;
 - Teachers may need to supplement evidence from regular classroom events with special tasks to ensure that all necessary evidence is collected for all students;
 - The difficulty of dealing with ephemeral evidence from regular classroom events could lead to a tick-list approach or a series of summative tasks;
 - This may well change the nature of formative assessment, making it more formal;
- (Harlen, 2006, s. 113)

Dessa svårigheter kan ställas i relation till Taras (2005; 2009) och Newton (2007) som menar att all bedömning är summativ i grunden och att det därför inte finns någon renodlad formativ bedömning. Enligt dem skulle det som Harlen (2006) redogör för istället betraktas som bedömning först använd för ett formativt syfte och därefter använd även summativt.

”Inside the black box” och ”bibeln” för formativ bedömning

Black och Wiliam presenterar 1998 en sammanställande analys av 250 publicerade undersökningar kring bedömning och lärande (Black & Wiliam, 1998a⁷). Resultatet visar att formativ bedömning gynnar elevens lärande och höjer undervisningens standard, men att de arbetssätt som då används inom den brittiska skolan inte gör det. Det presenteras dock inget förslag på hur man skulle kunna förändra undervisningen för att höja standarden i undervisningen (ibid). Lite senare samma år publiceras en till artikel av Black och Wiliam (1998b⁸), en artikel populärt kallad ”Inside the black box”, i vilken de ställer sig frågorna:

- Is there evidence that improving formative assessment raises standards?
- Is there evidence that there is room for improvement?
- Is there evidence about how to improve formative assessment?

(Black & Wiliam, 1998b, s. 140)

De svar de finner på sina frågor är alla jakande, även om svaret på den sista frågan inte är lika tydligt som för de andra frågorna. Detta då det ges idéer om förbättringsåtgärder, men saknas detaljerade beskrivningar på hur lärare skulle kunna åstadkomma denna förbättring. De idéer som presenteras handlar om feedback, elevens självbedömning, elevens förståelse för kriterier och lärares sätt att ställa frågor, och Black och Wiliam (1998b) uttrycker tydligt intentioner att fortsätta arbetet inom detta område för att bidra till att höja undervisningsstandarderna. Gällande frågan om förbättrad formativ bedömning höjer kunskapsnivån hos eleverna

⁷ Utifrån att en stor del av dagens forskning kring bedömning bygger på resultat av denna studie tillskrivs den sådan relevans att den ingår i studien trots att den inte håller femårsgränsen uppsatt för studien.

⁸ Se föregående not.

konstaterar Black och Wiliam att resultaten visar att det hjälper lågpresterande elever i högre utsträckning än högpresterande och därigenom bidrar till att minska kunskapsspannet, samtidigt som det sker en genomgående höjning av elevernas kunskande. Resultaten som Black och Wiliam (1998a; 1998b) presenterade har resulterat ett i kraftigt ökat fokus på framåtsyftande bedömning. Hargreaves (2005) uttrycker:

Black and Wiliam's (1998 [b]) pamphlet called *Inside the black box* has been one of the most powerful vehicles for urging teachers, and policymakers, that assessment for learning is beneficial for our students. (Hargreaves, 2005, s. 213)

Ett tydligt tecken på detta är att en betydande del av den forskning som idag bedrivs i Storbritannien kring bedömning baserar sig på de resultat som Black och Wiliam (1998a; 1998b) presenterade. Black och Wiliam utgjorde tillsammans med forskarkollegorna Harrison, Lee och Marshall vid King's College en forskargrupp som fokuserade på framåtsyftande bedömning utifrån Black och Wiliams resultat. Detta samarbete resulterade bland annat i boken *Assessment for Learning – Putting it into Practice* (Black et al., 2003)⁹, en bok som Taras (2009) refererar till som "the 'bible' of assessment for learning" (s. 57). I ett vidare samarbete i Storbritannien är Black och Wiliam tillsammans med forskare som Daugherty, Ecclestone, Gardner, Harlen, James, Pedder, Sebba och Stobart engagerade i något som kallas *Assessment Reform Group*. De arbetar bland annat med ett skolutvecklingsprojekt kallat *Assessment for Learning [AfL]*, ett projekt som även det är ett resultat av forskningsresultaten presenterade av Black och Wiliam 1998 (Gardner, 2006; Black & Wiliam, 1998a; 1998b). I samband med att AfL startade inleddes även ett samarbete med Stanford University i Kalifornien, där man var intresserade av Black och Wiliams (1998a; 1998b) resultat och ville genomföra ett liknande projekt i skolor i Kalifornien (se Black et al. 2003).

Kritiska röster...

Även om resultatet i den forskning Black och Wiliam presenterade 1998 gällande bedömningens effekt på elevers lärande inte har ifrågasatts nämnvärt har ändå kritiska röster höjts gällande en del av det efterföljande arbetet (Black et al., 2003).

... gällande 'bibeln' för formativ bedömning

Den kritik Taras (2009) riktar mot Black et al. (2003) baseras på en uppfattning om att fokus gällande bedömningen har förskjutits över tid. Som Taras (2009) tolkar Black och Wiliam (1998a) låg fokus ursprungligen på bedömningsprocessen, medan det hos Black et al. (2003) snarare är det bedömningens funktion som framställs som central. Taras (2009) ser detta som en inkonsekvens i det teoretiska ramverket hos Black och Wiliam och menar att problem har skapats genom att bedömningsprocessen hamnat i skuggan av bedömningens funktion. Problemet är, enligt Taras, att Black et al. (2003) presenterar ett par olika bedömningsprocesser, vilka enligt henne varken kompletterar eller stödjer varandra. Den första processen Taras (2009) pekar på är inom kontexten klassrumsinteraktion och fokuserar på läraren som den som styr bedömningsprocessen. I detta fall framstår formativ bedömning som någonting som endast kan ske i kontexten klassrumsinteraktion (ibid). Dessutom menar Taras att den jämbördighet och balans gällande ansvar för bedömning mellan lärare och elev, som Black och Wiliam (1998a) talade om, hos Black et al. (2003) har ersatts av fokus på läraren som bedömare. I den andra bedömningsprocessen som Taras (2009) beskriver jämförs information

⁹ Den centrala roll denna bok har i Storbritannien, och det faktum att Taras (2009) refererar till det som en "bedömningsbibel" gör att den tillskrivs sådan relevans att den ingår i studien trots att den inte klarar det angivna tidsspannet för studien.

om faktiskt kunnande med önskade kunskapsnivåer för att identifiera eventuella luckor. Informationen används sedan för att påverka dessa luckor. Skillnaden mellan processerna är att den första fokuserar på lärarens arbete medan den andra fokuserar på elevernas lärande, samt att bedömningsprocessen i det första exemplet handlar om lärarens metodik i klassrummet, medan det i det andra handlar om själva produkten (ibid). Taras menar att:

What Black et al. seem be doing is to have their own definition of FA [formative assessment] which is based on the distinction of functions (taken from Bloom, Hastings, and Madaus, 1971¹⁰) which they relate to classroom interaction and which is under the control of the teacher, but also, seeing the necessity to address FA as a learner /.../ they adopt Sadler's¹¹ definition. (Taras, 2009, s. 63)

Sadler definierar, enligt Taras (2009), formativ bedömning som en tilläggssteg till summativ bedömning, vilket kräver feedback som visar vilka luckor som behöver fyllas för att nå uppsatta mål och kriterier. Taras, som själv utgår från Scriven¹² och Sadler i sina definitioner av bedömningsbegrepp, ser motsättningar i den kombination av Sadler och Bloom, Hastings och Madaus som Black et al. (2003) använder sig av för att definiera formativ bedömning. Taras menar att Assessment Reform Group i det inledande arbetet med AfL var tydliga med att lärarna skulle skilja på formativ och summativ bedömning, men att detta mötte motstånd bland lärare. Detta motstånd innebar att de valde att omformulera sig i hur arbetet med framåtsyftande bedömning skulle fungera i praktiken (ibid). Taras anser att det i och med detta omformulerande uppstått tvetydigheter. Hon finner det anmärkningsvärt att det, trots att Black et al. (2003) har tillskrivits så stor tyngd i brittisk forskning att den nu är obligatorisk läsning för alla blivande lärare, är det förvånansvärt lite diskussion kring såväl paradigmen som definitionerna för detta. Taras (2009) anser att en lösning på motsättningarna hos Black et al. (2003) är ett paradigmskifte innebärande att definitionerna av formativ och summativ bedömning baseras på bedömningsprocesser istället för funktioner.

...gällande AfL-projektet

Ifrågasättande röster har riktats mot AfL-projektet (t.ex. Petty, 2008). Petty menar att man tappat fokus. Trots att projektets grundtanke är att använda formativ bedömning för att stimulera elevernas lärandeprocesser, så görs betygslänkande omdömen. De regelbundna externa prov som eleverna ska göra under fastslagna skolår motarbetar enligt Petty (2008) intentionen med projektet genom att få lärarna, och därmed även eleverna, att fokusera på lärande inför dessa dömande prov.

Sammanfattning av anglosaxisk forskning kring bedömning

Inom anglosaxisk forskning har en övergripande syn på bedömning varit det vanliga sättet att se på bedömning, medan fokus på framåtsyftande bedömning är något relativt nytt. Detta uttrycker sig i forskningen med att definitionerna av sammanfattande bedömningsbegrepp är få. Det vanligast förekommande är det enkla *assessment*. Dömande bedömning behandlas antingen i samband med betyg och då är det betygsbegrepp och inte andra begrepp för dömande bedömning som används, eller i samband med att framåtsyftande bedömning behandlas och då i relation till framåtsyftande bedömning. Detta kan ses som en anledning till

¹⁰ Bloom, B.S., Hastings, J.T. & Madaus, G.F. (Red.) (1971). *Handbook on the formative and summative evaluation of student learning*. New York: McGraw-Hill.

¹¹ Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), s. 145-165.

¹² Scriven, M. (1967) The methodology of evaluation. I R. Tyler, R. Gagne & M. Scriven (Red.), *Perspectives on Curriculum Evaluation* (AERA Monograph Series – Curriculum Evaluation). Chicago: Rand McNally.

att definitionerna för dömande bedömningsbegrepp är få. Det faktum att bedömning inom skolan ses och länge har setts som en helhet tydliggörs i diskussioner kring relationen mellan framåtsyftande och dömande bedömning. Ingen uttrycker att den ena bedömningen utesluter den andra, däremot att det är komplext och att det är hur man väljer att använda bedömningen som avgör om den är framåtsyftande eller dömande eller om bedömningen kan tillskrivas såväl dömande som framåtsyftande karaktär. Den framåtsyftande bedömningen har kommit i fokus sedan Black och Wiliam (1998a; 1998b) presenterade sina forskningsresultat och tydliggjorde den formativa bedömningens betydelse för elevers lärande. De nyare bedömningsbegreppen *assessment for learning* och *assessment of learning* i relation till de äldre, inarbetade begreppen *formativ* och *summativ* bedömning, samt diskussioner kring tolkningar och risk för misstolkningar av dessa begrepp sätter sin prägel på den anglosaxiska forskningen. Dessa resultat har även påverkat skolpolitiken i Storbritannien, där satsningar på projekt som AfL har gjorts för att främja elevers lärande. Det har dock skett en politisk förvanskning genom att de presumtiva resultaten av dessa satsningar kontrolleras med externa prov, vilka motverkar det framåtsyftande arbetet (Black & Wiliam, 2005). Det fokus på dömande bedömning i form av test och betyg som präglat och fortfarande präglar skolväsendet i såväl Storbritannien som USA (Black & Wiliam, 2005; Popham, 2008) har bidragit till ett fokus på dömande bedömning i skolan som inte alltid är kompatibel med ett formativt syfte (Black & Wiliam, 2005).

Fokus på bedömning i Sverige och internationellt – en jämförelse

Följande presentation är en sammanställning av de jämförelser mellan svenskt och internationellt perspektiv på bedömning som görs i den svenska delen av underlaget för studien, med inslag av jämförelser med anglosaxisk forskning.

En förskjutning i fokus

Både Lindström (2005a) och Nyström (2004) reflekterar kring ändrat fokus inom internationell forskning kring bedömning. Nyström (2004) talar om detta i termerna paradigmskifte från ett psykometriskt synsätt, där lärarens uppgift gällande bedömning var att administrera test utformade av andra, till ett mer pedagogiskt perspektiv, där läraren och lärandet är centralt för bedömningen. Han menar dock att:

Det finns anledning till att ifrågasätta om denna beskrivning stämmer för svenska förhållanden. Svenska lärare har aldrig varit reducerade till förmedlare av test som utformas av någon annan och psykometriska utgångspunkter har knappast varit särskilt framträdande när det gäller bedömningar i skolan. (Nyström, 2004, s. 2)

Detta är en reflektion som inte Lindström (2005a) gör i sin redovisning av bedömningsfokus inom svensk och internationell forskning. Hans fokus ligger istället på att se hur man kan förstå det nya paradigmet ur ett svenskt perspektiv. Lindström beskriver den ändrade synen på bedömning inom internationell forskning som en förskjutning mot ökat fokus på framåtsyftande bedömning (se tabell 2, s. 30). Den förskjutning av fokus som Lindström påtalar kan man också utläsa hos Popham (2008), som särskiljer gamla och nya syften för bedömning. Följande syften benämner Popham som gamla anledningar till bedömning:

- To diagnose students' strengths and weaknesses
 - To monitor students' progress
 - To assign grades to students
 - To determine instructional effectiveness
- (Popham, 2008, s. 11)

Medan han ser följande anledningar till bedömning som nyare:

- Test results determine public perceptions of educational effectiveness.
 - Students' assessment performances are increasingly seen as part of the teacher evaluation process.
 - As clarifiers of instructional intentions, assessment devices can improve instructional help.
- (Popham, 2008, s. 16)

Popham påtalar, precis som Lindström (2005a), att nyare typer av och anledningar till bedömning är ett komplement till och inte en ersättning för äldre anledningar, vilka fortfarande är lika viktiga inom bedömning i skolan.

Tabell 2. Pedagogisk bedömning

<i>En förskjutning från...</i>	<i>i riktning mot:</i>
Bedömning används främst för att kontrollera vad eleverna lärt sig	Bedömning används även för att befrämja lärande och diagnostisera lärande
Bedömning och lärande hålls isär	Bedömning av och för lärande sker fortlöpande
Läraren bedömer på egen hand elevernas kunskaper	Lärare och elev bedömer tillsammans var eleven befinner sig och hur hon kan gå vidare
Ateoretisk bedömning	Bedömning grundad på teori om hur man lär sig inom ett bestämt kunskapsområde
Bedömning av kunskaper och färdigheter	Bedömning av förståelse och förmågor som kritiskt tänkande, kreativitet, kommunikation och problemlösning i realistiska situationer
Produkter i fokus	Processer i fokus
Tonvikten ligger vid de rätta svaren	Tonvikten ligger vid fruktbara frågor och konsten att lära av erfarenheter
Normrelaterad bedömning	Mål- och kunskapsrelaterad bedömning
Resultat redovisas i form av en totalpoäng	Visar på starka och svaga sidor, uppmärksammar framsteg
Elever arbetar var för sig utan tillgång till hjälpmedel	Elever ger varandra gensvar och kan utnyttja olika former av minnestöd och tankeverktyg
Huvudsakligen skriftliga prov	Dokumentation av olika slag, t.ex. loggar, portföljer, gestaltning, utställning, cd-rom

(Lindström, 2005a, s. 12)

Olika sorters forskning och perspektiv

Selghed (2006) utläser två riktningar inom internationell bedömningslitteratur utifrån de sista fyra raderna i Lindströms (2005a) tabell, vilka behandlar hur bedömning genomförs (se tabell 2, s. 31). Dessa inriktningar är "assessment" och "testing", om vilka Selghed (2006) uttrycker följande:

Konflikten mellan de olika riktningarna består alltid i vad som anses viktigast i samband med bedömning. "Testningsriktningen" fokuserar på *hur* bedömningen ska genomföras, genom krav på objektivitet, distans, standardiserade förhållanden, lika villkor vid mätningen, oberoende neutrala bedömare, allt med syftet att göra bedömningen tillförlitlig och därmed rättvis och likvärdig. "Assessmentriktningen" ser innehållet, *vad* som ska bedömas som viktigast. Detta kan endast göras om den som ska bedömas får möjlighet att under naturliga eller realistiska

förhållanden visa vad han/hon behärskar. Detta omöjliggör standardiserade situationer med på förhand hårt strukturerade och avgränsade provuppgifter. Bedömningssituationen måste vara dynamisk, varför det blir svårt med jämförelser mellan olika bedömare på grund av större inslag av subjektivitet. (Selghed, 2006, s. 210)

Kritiken mot assessmentriktningen inom internationell bedömningsforskning handlar om bristen på objektivitet och svårigheten att nå likvärdighet, vilket Selghed uttrycker med följande frågeställningar:

Tolkar olika bedömare samma prestationer på samma sätt? Går det att språkligt utforma kraven så att olika bedömare gör likvärdiga tolkningar? Är lärare tillräckligt kunniga i bedömningsfrågor, för att klara att utföra bedömningarna på ett tillförlitligt sätt? (Selghed, 2006, s. 209)

Utifrån frågeställningar som dessa kan man se tydliga skillnader i svenskt och anglosaxiskt perspektiv. Lundahl (2006) ställer sig frågande till hur lärare inte verkar klara av att bedöma elevers lärande utan nationella prov, eller att någon utifrån, exempelvis forskare talar om för dem hur de skulle göra. Black och Wiliam (2005) däremot uttrycker att det i Storbritannien finns en tveksamhet till den ökade subjektivitet som det innebär om lärarna skulle ges större inflytande över den dömande bedömningen. Här blir det tydligt att det psykometriska synsättet, som så tydligt finns i det resonemang Black och Wiliam redogör för, saknas hos Lundahl (2006), vilket också gäller jämförelser i ett större perspektiv (Nyström, 2004). Lundahl (2006) menar att framåsyftande bedömning i ett internationellt perspektiv inte står i motsats till dömande bedömning på samma sätt som inom svensk forskning. Han menar att man kan urskilja två sorters forskning kring kunskapsbedömning: forskning *i* respektive forskning *om* bedömningar. Den förstnämnda typen handlar enligt Lundahl om utvecklandet av metoder för mätning, beskrivning och sammanställning av elevens lärandeutveckling. Den andra handlar om att försöka förstå kunskapsbedömningar som ett socialt fenomen och de intressekonflikter och dilemman som kan uppstå. Svensk forskning kring bedömning har främst fokuserat på den senare sorten, forskning om bedömning (ibid). Med intention att förklara dessa skillnader i fokus för Lundahl följande resonemang:

Den senare typen av forskning kommer ofta från England och USA. Att denna typ av forskning inte är lika vanlig i Sverige kan nog delvis förstås utifrån att läroplan, kursplan och bedömning länge betraktats som separata styrinstrument där experter på läroplaner haft ringa kunskap om kunskapsbedömning och experter på kunskapsbedömning haft ringa kunskap om läroplaner. En annan tolkning kan vara att externa kunskapsbedömningar har hög legitimitet i Sverige och sett till de senaste åren använts med försiktighet. I England och USA där *high stake*-bedömningar är mycket vanligare har så att säga hela yrkeskåren av psykometriker kommit att ifrågasättas [...]. Ett sätt att återvinna legitimitet kan då vara att närma sig den pedagogiska praktiken och peka på hur kunskapsbedömningar är ett didaktiskt redskap som främjar en ny typ av kunskaper; dessutom en "högre ordningens tänkande". (Lundahl, 2006, s. 397)

Forsberg och Lundahl (2006) lyfter skillnaden mellan svensk och internationell forskning kring bedömning i skolan genom att synliggöra dessa olika fokus. De uttalar en intention att bryta detta mönster genom att själva inta ett internationellt sett mer vedertaget perspektiv. Detta genom att inte betrakta de olika bedömningsformerna separat, utan gemensamt. Skillnaderna i perspektiv påtalas även av Wedman (2003), som efter att ha studerat den svenska skoldebatten konstaterar att:

Inte bara den allmänna skoldebatten har uppmärksammat betyg och betygsfrågor. Ett tiotal statliga utredningar har behandlat betygen sedan 60-talet. Intressant att nämna i förbigående är att frågan inte på långt när haft samma framskjutna plats i den internationella skoldebatten. (Wedman, 2003, s. 307)

Wedman konstaterar att en skillnad gällande betyg är att de krav på jämförbarhet som vi har här, inte verkar vara lika uttalad i ett internationellt perspektiv.

Diskussion

Det finns inom svensk forskning en tydlig brist på och även ofta avsaknad av definitioner av bedömningsbegrepp. Vid en jämförelse av de uttalade definitioner av bedömningsbegrepp som görs i underlaget för denna studie kan det verka som att svenska forskare nått konsensus. Det är dock en inbillad konsensus, vilket klart framgår när man tar i beaktande de skillnader i användande av bedömningsbegrepp som finns och den bristande begreppsproblematisering och -diskussion som präglar svensk forskning. Inom anglosaxisk forskning är diskussionen betydligt livligare. Även om begreppsdefinitionerna vid en yttlig översikt även där ter sig relativt lika förs en diskussion kring definitioner och begreppsval som saknas i svensk forskning.

Begrepp – definitioner och avsaknad av definitioner

I svensk forskning är det väldigt lite diskussion kring bedömningsbegrepp och deras innebörd. Det finns en avsaknad av diskussion kring vilka begrepp som används, hur och varför. De frekvent använda begreppen formativ och summativ bedömning upplevs därför som ganska statiska. Detta är inte fallet i anglosaxisk forskning kring bedömning, där nyanserade diskussioner kring begreppsskillnader förs och forskare använder olika sätt att definiera och därmed särskilja eller likställa begrepp. Taras (2005; 2009) upplevs dock som exceptionell med ett närmast statistiskt fasthållande av äldre definitioner av bedömningsbegrepp, medan andra menar att det sker en utveckling i hur man definierar bedömningsbegrepp (Hayward & Hedge, 2005). Taras (2005) har även en betydligt mer teoretisk ansats i sina diskussioner kring bedömning än många andra forskare, vilka i mycket större utsträckning fokuserar på det praktiska användandet av bedömning.

Inbillad konsensus

De svenska definitioner som ges av bedömningsbegrepp med framåtsyftande eller dömande innebörd är få, fåordiga, och väldigt lika. Det saknas variation i begrepp, formativ och summativ är de allmänt vedertagna begreppen. I och med att inga andra begrepp används, förs heller inte några resonemang kring begreppsval. En allmän tro på att det råder konsensus i vad man menar med summativ respektive formativ bedömning kan förklara avsaknaden av definitioner. Detta antagande om inbillad konsensus stärks av att det inte finns någon påtaglig diskussion eller problematisering inom svensk forskning gällande framåtsyftande och dömande bedömningsbegrepp. Att det inom anglosaxisk forskning är tydligare definitioner och mer diskussion kring bedömningsbegrepp och dess användning än inom svensk forskning innebär inte att definitioner är tydliga, vilket Sadler (2007) uttrycker följande om:

.../ many of the terms we use in discourse on assessment and grading are used loosely. By this I mean that we do not always clarify the several meanings a given term may take even in a given context, neither do we necessarily distinguish various terms from one another when they occur in different contexts. (Sadler, 2007, s. 388)

Det vore anmärkningsvärt om man i Sverige nått så långt att man är helt överens om begreppsinnbörd, när man i den anglosaxiska forskningen konstaterar att det vore bra om man var överens (Newton, 2008; Sadler, 2007). Detta med tanke på att man i anglosaxisk

forskning faktiskt skapar möjligheter att bli överens genom att föra diskussioner om vad man menar, något som jag inte sett tecken på att det görs inom svensk bedömningsforskning.

Otydlighet i begreppsfloran

Kunskapsbedömning är det mest använda övergripande bedömningsbegreppet inom svensk forskning. Jag ställer mig frågande till varför det är så av samma anledning som Sadler (2007) ställer sig tveksam till begreppen "assessment for learning" och "assessment of learning", nämligen risken för misstolkning. När Selghed (2007) reder ut kunskapsbegreppet står det klart att detta är ett begrepp lätt att tolka och misstolka. Selghed menar:

Kunskap uppfattas vara av antingen/eller-karaktär. Kunnigheten består i att man har ett stort förråd eller lager av kunskaper, som direkt kan tas fram och återges. Detta sätt har under lång tid dominerat skola och utbildning. I samhället i stort är denna kunskapssyn fortfarande dominerande. (Selghed, 2007, s. 39)

Även om denna kunskapssyn fortfarande är dominerande i samhället har kunskapssynen i skolan förändrats och kunskap definieras inte längre som något statiskt, utan något som sker i interaktionen mellan lärare och elev (Selghed, 2007). För mig som gick lärarutbildningen när läroplanen hette Lgr80 är de fyra kunskapsformerna *fakta*, *förståelse*, *färdighet* och *förtrogenhet* ordentligt inbakade i mitt medvetande. Säkert är det så även för alla andra som arbetar inom skolans värld. Men i samhället i stort är definitionen av kunskapsbegreppet en annan, där är kunskap någonting faktiskt och oföränderligt (ibid). Begreppet kunskapsbedömning kan lätt tolkas som att det är bedömning av faktakunskaper som är primärt och att bedömning av lärprocesser och färdigheter är sekundärt. Detta gör att jag ställer mig tveksam till användandet av kunskapsbedömning som svensk form av övergripande bedömningsbegrepp och istället föredrar Nyströms (2004) *bedömning i skolan*.

Begreppet *kunskapsbedömning* är ett resultat av svenska strävanden efter att finna någon bra översättning till det anglosaxiska *educational assessment*, ett begrepp väl förankrat inom anglosaxisk forskning, men som inte alls används lika ofta som det enkla *assessment*. *Assessment* är det övergripande bedömningsbegrepp som jag uppfattar som det mest frekvent förekommande inom anglosaxisk forskning. Av de 20 anglosaxiska texter som låg som underlag för denna studie är det enbart Newton (2007) och Popham (2008) som använder sig av begreppet *educational assessment*. Det är intressant att ställa sig frågan varför man inom svensk forskning inte nöjt med att bara översätta begreppet *assessment*. Kanske beror det på att begreppet *bedömning* motsvarar såväl *assessment* som *judgement*, en definition som kan ses som alltför vid.

Syfte eller inte

Det visade sig vara inte alltför enkelt att svara på frågan om vilka former av och syften för bedömning som finns. I ett traditionellt svenskt perspektiv ses framåtsyftande och dömande bedömning som separata former av bedömning, medan man i ett anglosaxiskt perspektiv ser dem som delar av en helhet. Aktuell svensk forskning närmar sig ett anglosaxiskt perspektiv genom att tala övergripande om bedömning och inom anglosaxisk forskning finns ökat fokus på hur framåtsyftande bedömning kan utvecklas. Det finns inom anglosaxisk forskning även en problematisering av syftesbegreppet, då Newton (2007) menar att eftersom all bedömning i grunden är dömande så kan bedömningen inte också ha ett dömande syfte. Dessa skillnader i förhållningssätt redovisas i resultatet.

Forskningsfokus

Redan vid sökningen av artiklar som skulle vara underlag för denna studie stod det klart för mig att det var stort fokus på *framåtsyftande bedömning* i anglosaxisk forskning. Detta tog sig uttryck i att artiklar om formativ bedömning utgjorde en stor del av träffarna när jag sökte på *summative assessment*, medan samma effekt gällande summativ bedömning inte uppkom vid sökning på *formative assessment*. Det mesta av den anglosaxiska forskningen kring bedömning som inte behandlar framåtsyftande bedömning, enskilt eller i relation till dömande bedömning, tycks fokusera på betygssystem eller på utformandet av teknik för underlättande av bedömning.

Skillnad i fokus, frågor och ickefrågor

Det är anmärkningsvärt att skillnaderna mellan fokus på bedömning i Sverige och internationellt är så påtagliga. Gällande *framåtsyftande bedömning* inom svensk pedagogisk forskning är det inte mycket skrivet. Framåtsyftande bedömning lyfts genomgående fram som centralt och viktigt i lärarnas bedömningsuppdrag. Ändå decimeras den i forskningen till att främst nämnas i samband med *dömande bedömning*, i en problematisering eller avproblematisering gällande integrering av dessa bedömningsformer (Jansdotter Samuelsson & Nordgren, 2008; Selghed, 2006). Detta till skillnad från bedömningsforskning i ett anglosaxiskt perspektiv, eller främst brittiskt, där *assessment for learning* ges betydande utrymme. Det svenska fokus på vilken form av betyg som är att föredra, problem med likvärdighet i målrelaterade betygssystem och betygens vara eller icke vara lyser med sin frånvaro i anglosaxisk forskning. Där är betyg en ickefråga, något som bara är, det är målrelaterat och betygsdiskussionerna handlar istället om hur man kan utveckla system för betygsbedömning. Även om diskussionen kring dömande och framåtsyftande bedömning pågår inom anglosaxisk forskning precis som inom svensk forskning är fokus olika. Medan betygens roll ges stort fokus i Sverige, ser man ur anglosaxiskt perspektiv på den framåtsyftande bedömningen för att där finna lösningarna till bedömningen i stort. Visserligen reflekterats det över att de nationella prov man har inte gynnar den framåtsyftande bedömningen, men det är inte särskilt mycket mer än ett konstaterande (Harlen, 2005; Popham, 2008).

Avsaknad av problematisering

Jag upplever en avsaknad av diskussion kring den framåtsyftande bedömningen i svensk forskning, något som Lindberg (2007) bekräftar. Det är inte så att jag menar att den framåtsyftande bedömningen ska särskiljas från annan bedömning och sedan analyseras. Istället menar jag att den trend jag kan avläsa i att även svenska forskare försöker se på bedömning inom skolan som en helhet, också innebär att vi inte kan nöja oss med att konstatera att framåtsyftande bedömning är viktigt. Det finns behov av att närmare studera hur denna bedömning görs. Genom att synliggöra den framåtsyftande bedömningen kan den utvecklas. *Världens bästa fröken* och *Vygotskij i praktiken* är båda exempel på böcker där lärarnas arbete studeras, med fokus på hur bra lärararbete bedrivs i praktiken (Steinberg, 2004; Strandberg, 2006). Morberg (2002) uttrycker följande gällande svårigheter och hjälp för nyexaminerade lärare att utveckla den egna lärarrollen samt läraryrkets komplexitet:

Inga uppräddade expertlösningar på komplexa problem! Att enbart lita till insamlandet av goda exempel/modeller går idag inte längre. *Mångfaldiga* exempel, dåliga och goda, och *mångfaldiga* modeller, goda och dåliga, är viktigare än 'goda exempel och undervisningsmodeller'. Av samma skäl är nyansrika variationer viktigare än tydliga förebilder /.../ (Morberg, 2002, s. 134)

Taras (2009) kritik mot att Black et al. (2003) fokuserar på att presentera goda exempel på hur formativ bedömning kan göras, övertygar mig om att det Morberg (2002) uttrycker är applicerbart även inom forskning kring formativ bedömning. I internationell forskning diskuteras framåtsyftande bedömning med samma intensitet som dömande bedömning. Den ifrågasätts inte, alla är överens om dess vikt. Problematiseringen gällande framåtsyftande bedömning handlar snarare om hur man får den att fungera på bästa sätt i praktiken och vilka hinder som finns för detta. I Sverige har den framåtsyftande bedömningen sedan länge en naturlig del i skolan, medan den i exempelvis Storbritannien vinner mark i ett system som länge byggt på ett psykometriskt synsätt med stark tilltro till bedömning av dömande karaktär. Denna skillnad i syn på bedömning kan förklara det bristande problematisering av framåtsyftande bedömning som präglar svensk forskning. Fransson (2002) talar om att identifiera "...institutionaliserade eller förgivettagna fenomen vad gäller förskola, skola, undervisning och lärande" (s. 112). Lärares bedömning är inget fenomen. Dock kan det ses som ett fenomen att bedömningen som ingår i läraruppgiften är så pass tagen för givet att diskussionerna kring hur lärarna genomför den ständigt pågående bedömningen av elevernas kunskaper är anmärkningsvärt få (William, 2006).

Att döma framåtsyftande

Kan man då döma framåtsyftande? Resultaten i denna studie ger inte ett enkelt ja eller nej som svar på den frågan. Vad som framkommer är dock att genom ett medvetet bedömande av elevers kunskaper kan bedömning tillskrivas framåtsyftande såväl som dömande karaktär. Hur detta görs finns inte en enkel modell för, utan snarare lika många tankar om som forskare som studerat området, plus ytterligare en mängd, utformade av lärare som praktiserar denna form av bedömning. Det minst omdebatterade och ifrågasatta sättet att använda bedömning i såväl framåtsyftande som dömande syfte är att tillämpa portfolio. Det finns dock flera olika sätt att lägga upp en portfolio, baserat på vilket det primära syftet är (Popham, 2008).

Metoddiskussion

Utökningen av svenska källor för studien från en gräns på fem till sju år innebar att underlaget för studien till slut kom att omfatta fler svenska än anglosaxiska källor, vilket kan få valet att verka tveksamt. I resultatredovisningen framkommer dock tydligt att det finns en betydligt rikare diskussion kring bedömningsbegrepp inom anglosaxisk forskning, vilket bland annat yttrar sig i att de svenska källorna mer sällan behandlar fler än enstaka av de frågeställningar studien utgår från, medan en högre andel av de anglosaxiska källorna berör flera av frågeställningarna. I de anglosaxiska referenserna kan även en högre grad av diskussion och problematisering kring bedömningsbegrepp urskiljas än i de svenska referenserna. Tunnsåddheten inom svensk forskning kring bedömning jämfört med motsvarande anglosaxiskt material innebär att jag hårdare har kunnat skära inom det anglosaxiska materialet, medan jag i hanterandet av det svenska materialet nästan varit tvungen att ta med det jag kunnat finna som på något sätt diskuterar bedömningsbegrepp inom skolan. Jag är av uppfattningen att resultatet hade blivit detsamma även om utökningen av det svenska materialet inte hade gjorts, men att konsekvensen blivit ett något tunnare underlag.

Är det rimligt att tänka sig att denna studie kan genomföras igen med samma resultat? Ja, det är högst sannolikt. I urvalsproceduren uppstod ganska snart ett läge där ett antal artiklar och författare var ständigt återkommande och det vore högst osannolikt att tänka sig att dessa inte skulle dyka upp i en liknande studie. Däremot har ett urval gjorts bland andra artiklar, där en del har valts bort baserat på att de då inte uppfattats som relevanta för studien. Trots att detta urval skett utifrån uppsatta kriterier för studien påverkas det av subjektiva uppfattningar av materialets innehåll och relevans. Jag finner det därför högst osannolikt att någon annan som skulle genomföra denna studie skulle göra exakt samma urval som jag gjort. Det är även högst troligt att jag själv idag skulle ha bedömt vissa källor annorlunda, baserat på de kunskaper om bedömningsbegrepp som jag har utvecklat under det pågående arbetet. Jag hade under insamlingen av material svårigheter att hitta amerikanska bedömningsartiklar som inte enbart behandlade betyg. Detta resulterade i en kraftig dominans av brittiskt material i det anglosaxiska underlaget för studien. Den sneda fördelningen mellan brittisk och amerikansk forskning i underlaget för denna studie borde tas i beaktande och försöka åtgärdas vid en ny studie. Den prägel de brittiska forskarna kring Paul Black har satt på brittisk forskning och de influenser William spridit i USA under sina år där gör dock att jag är tämligen säker på att de anglosaxiska resultaten skulle likna varandra. Detta bekräftas till viss del av Popham (2008) som ofta återkopplar till forskning av Black och William. Min enda reservation gällande resultatet rör begreppet *educational assessment* och hur pass frekvent förekommande det är inom amerikansk forskning. Av de fyra amerikanska källorna var det endast en som använde sig av begreppet. Men då urvalet är så litet säger det oss inget om vilket av begreppen *assessment* och *educational assessment* som är vanligast förekommande som övergripande begrepp inom amerikansk forskning. När det gäller svensk forskning kring bedömningsbegrepp och dess användning är det som sagt tunnsått, vilket gör att jag finner det är svårt att komma fram till så väldigt skiftande resultat.

Tankar om fortsatt forskning

I början av denna studie var mina intentioner att studera utformandet av skriftliga omdömen. Denna tanke finns kvar, men även andra funderingar har väckts av den analys av material som jag gjort.

Hur har skriftliga omdömen utformats?

Skolor och rektorer har givits uppdraget att själva utforma de skriftliga omdömen som alla elever från skolår 1 ska ges. De individuella utvecklingsplanerna *ska* vara framåtsyftande, de skriftliga omdömena *bör* vara framåtsyftande, men de *kan* även ges betygsliknande karaktär (Utbildningsdepartementet, 2008c). Här är det intressant att fråga sig hur rektorerna har tolkat uppdraget att utforma skriftliga omdömen och hur skolornas dokument för skriftliga omdömen för yngre elever (skolår 1-5/6) är utformade? Vilka avväganden har legat till grund för dessa beslut om utformandet av skriftliga omdömen? I hur stor utsträckning kan de dokument som används på skolorna anses vara av formativ karaktär och i hur stor utsträckning är de av betygsliknande karaktär? Detta har jag tankar om att studera genom en dokumentanalys av skriftliga omdömen, kompletterad med intervjuer med rektorer och lärare.

Hur definierar och använder lärare bedömningsbegrepp?

Införandet av skriftliga omdömen från skolår 1 innebär att lärare för de tidigare skolåren inom grundskolan nu ska ge eleverna tydligare dömande bedömning och detta i interaktion med en väl fungerande framåtsyftande bedömning. Detta gör det intressant att ställa sig frågan om hur

lärare definierar bedömning och hur de använder bedömningsbegrepp. I en strävan efter en likvärdig bedömning är det inte bara vad som ska bedömas som är av relevans utan även hur lärare ser på bedömning. En brittisk undersökning av hur lärare definierar bedömningsbegrepp påvisade ett spektra av definitioner och perspektiv (Hargreaves, 2005). Det skulle vara intressant att genomföra en liknande studie inom svensk skola för att se vilka resultaten skulle bli och se om de skiljer sig från brittiska lärares uppfattning om bedömning. Denna undersökning skulle även kunna göras longitudinellt genom att inledningsvis göra en studie av nyexaminerade lärares bild av bedömning och sedan återkomma till samma lärare efter ett antal år i yrket.

Vad innebär en digitalisering av skriftliga omdömen?

Samtidigt som uppdraget om utformandet av skriftliga omdömen gavs tydliggjordes att det var upp till de enskilda rektorerna att utforma hur dessa omdömen skulle se ut. I Stockholm fattades trots detta ett gemensamt utformande av underlag för skriftliga omdömen, något som mötte kritik. Idag kan vi se att det sker en digitalisering av skriftliga omdömen i en del av landets kommuner, troligen för att omdömena ska kunna ha en kontrollfunktion i en strävan efter likvärdighet i bedömning. Frågan är vad detta leder till. Jag kan inte låta bli att tänka på vad Håkan Andersson (2007) uttrycker angående tidigareläggande av betyg:

Eftersom ungdomarna förändras och utvecklas olika finns ingen anledning att tidigt cementera genom skriftliga omdömen eller tidiga betyg. /---/ Grundskolan är till för elevernas allsidiga utveckling. Ökad disciplinering av eleverna, ökad makt för lärarna, ökad betygsättning och provverksamhet riskerar att få motsatt effekt. (Andersson, 2007, s. 63-64)

Andersson menar att han i egen forskning inte funnit något stöd för denna utveckling av bedömning inom skolan, varken gällande gynnande av elevers utveckling eller lärarrollen som yrkesval. Merparten av de forskningsresultat som redovisats i detta arbete stödjer Anderssons resonemang. Utifrån detta är det svårt att låta bli att undra vad ett införande av digitaliserad kontroll av skriftliga omdömen innebär för elevernas utveckling i de tidigare skolåren. Hur en sådan studie skulle göras har jag ännu ingen klar bild av.

Efter att ha genomfört denna studie kan jag inte göra annat än att hålla med Wikström (2006) när hon säger att den svenska bedömningsarenan kommer att vara ett spännande och intensivt område framöver, för forskare såväl som för de som arbetar med bedömning i skolan.

Referenslista

- Andersson, Håkan (2007). Betyg är ett trubbigt instrument. *Pedagogiska magasinet*, (3), s. 60-64.
- Backman, Jarl (1998). *Rapporter och uppsatser*. Lund: Studentlitteratur.
- Björklund Boistrup, Lisa (2005). Att fånga lärandet i flykten. I L. Lindström & V. Lindberg (Red.), *Pedagogisk bedömning* (s. 111-129). Stockholm: HLS Förlag.
- Black, Paul, Harrison, Christine, Lee, Clare, Marshall, Bethan & Wiliam, Dylan (2003). *Assessment for Learning – Putting it into Practice*. Maidenhead: Open University Press.
- Black, Paul, Harrison, Christine, Lee, Clare, Marshall, Bethan, & Wiliam, Dylan (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86(1), s. 9-21.
- Black, Paul & Wiliam, Dylan (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), s. 7-74.
- Black, Paul & Wiliam, Dylan (1998b). Inside the Black Box. *Phi Delta Kappan*, 80(2), s. 139-148.
- Black, Paul & Wiliam, Dylan (2005). Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices. *The Curriculum Journal*, 16(2), s. 249-261.
- Black, Paul & Wiliam, Dylan (2006) Assessment for Learning in the Classroom. I J. Gardner (Red.), *Assessment and Learning* (s. 9-25). London: SAGE Publications.
- Boréus, Kristina & Bergström, Göran (2005). Argumentationsanalys. I G. Bergström & K. Boréus (Red.) *Textens mening och makt – Metodbok i samhällsvetenskaplig text- och diskursanalys* (2:a uppl.) (s. 89-147). Lund: Studentlitteratur.
- Brookhart, Susan, Moss, Connie & Long, Beverly (2008). Formative Assessment That Empowers. *Educational Leadership*, 66(3), s. 52-57.
- Bryman, Alan (2001). *Samhällsvetenskapliga metoder*. Malmö: Liber.
- Carless, David (2007). Conceptualizing pre-emptive formative assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), s. 171-184.
- Dysthe, Olga (2002). Om förhållandet mellan individ och grupp i portföljprocessen. I Skolverket, *Att bedöma eller döma – Tio artiklar om bedömning och betygssättning* (s. 91-108). Stockholm: Liber .
- Emanuelsson, Ingemar (2002). I behov av särskilt stöd – och ändå godkänd?. I Skolverket, *Att bedöma eller döma – Tio artiklar om bedömning och betygssättning* (s. 27-38). Stockholm: Liber.
- Finansdepartementet (2006). *Budgetpropositionen för 2007*. Prop. 2006/07:1, Stockholm 16 oktober 2006.
- Forsberg, Eva & Lundahl, Christian (2006). Kunskapsbedömningar som styrmedia. *Utbildning och Demokrati* 2006 vol.15 nr 3, s. 7-29.

- Fransson, Göran (2002). Om institutionella förväntningar på lärare. I G. Arfwedsson (Red.) *Mellan praktik och teori – Tio didaktiska berättelser om undervisning* (s. 100-112). Stockholm: HLS Förlag.
- Gardner, John (2006) *Assessment and Learning: An Introduction*. I J. Gardner (Red.), *Assessment and Learning* (s. 1-5). London: SAGE Publications.
- Giota, Johanna (2006). Självsbedöma, bedöma eller döma? Om elevers motivation, kompetens och prestationer i skolan. *Pedagogisk Forskning i Sverige 11(2)*, s. 94-115.
- Habbe, Peter & Björkman, Joachim (2007). Bedömning är nyckeln till lärande. *Pedagogiska magasinet* (3) s. 72-73.
- Hargreaves, Eleanore (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education*, 35(2), s. 213-224.
- Hargreaves, Eleanore (2007). The validity of collaborative assessment for learning. *Assessment in Education*, 14(2), s. 185-199.
- Harlen, Wynne (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *Curriculum Journal*, 16(2), s. 207-223.
- Harlen, Wynne (2006) On the Relationship Between Assessment for Formative and Summative Purposes. I J. Gardner (Red.), *Assessment and Learning* (s. 103-117). London: SAGE Publications.
- Hayward, Louise, & Hedge, Nicki (2005). Travelling towards change in assessment: policy, practice and research in education. *Assessment in Education: Principles, Policy & Practice*, 12(1), s. 55-75.
- Jansdotter Samuelsson, Maria & Nordgren, Kenneth (Red.) (2008) *Betyg i teori och praktik*. Malmö: Gleerups.
- Jönsson, Anders (2005). Att bedöma naturvetenskapligt arbetssätt. I L. Lindström & V. Lindberg (Red.), *Pedagogisk bedömning* (s. 219-239). Stockholm: HLS Förlag.
- Kjellström, Katarina (2005). Bedömningsmatriser – en metod för analytisk bedömning. I L. Lindström & V. Lindberg (Red.), *Pedagogisk bedömning* (s. 193-217). Stockholm: HLS Förlag.
- Korp, Helena (2003). *Kunskapsbedömning – hur, var och varför* (Forskning i fokus, nr 13). Stockholm: Myndigheten för skolutveckling.
- Linde, Göran (2003). *Kunskap och betyg*. Lund: Studentlitteratur.
- Lindberg, Viveca (2002). Införandet av godkändgränsen – konsekvenser för lärare och elever. I Skolverket, *Att bedöma eller döma – Tio artiklar om bedömning och betygssättning* (s. 39-56). Stockholm: Liber .
- Lindberg, Viveca (2007). Långtgående slutsatser trots få lärare och elever i studierna. I A. Pettersson (Red.), *Sporre eller otyg – om bedömning och betyg* (s. 131-153). Stockholm: Lärarförbundets förlag.
- Lindström, Lars (2005a). Pedagogisk bedömning – Inledning. I L. Lindström & V. Lindberg (Red.), *Pedagogisk bedömning* (s. 11-27). Stockholm: HLS Förlag.
- Lindström, Lars (2005b). Portföljmetodik i estetiska ämnen. I L. Lindström & V. Lindberg (Red.), *Pedagogisk bedömning* (s. 159-192). Stockholm: HLS Förlag.

- Lindström, Lars & Lindberg, Viveca (2005). *Pedagogisk bedömning*. Stockholm: HLS Förlag.
- Lundahl, Christian (2006) *Viljan att veta vad andra vet – Kunskapsbedömning i tidigmodern, modern och senmodern skola*. Uppsala: Uppsala Universitet.
- Lundahl, Christian (2007). Kunskapsbedömningens historia. I A. Pettersson (Red.), *Sporre eller otyg – om bedömning och betyg* (s. 51-68). Stockholm: Lärarförbundets förlag.
- Morberg, Åsa (2002) Ett dilemma för lärarutbildning och lärare. I G. Arfwedsson (Red.), *Mellan praktik och teori – Tio didaktiska berättelser om undervisning* (s. 113-134). Stockholm: HLS Förlag.
- Moreau, Helena (2008). IUP med skriftliga omdömen - tolkat utifrån Skolverkets allmänna råd, *Grundskoletidningen* (6), s. 7-12.
- Newton, Paul E. (2007). Clarifying the purpose of educational assessment. *Assessment in Education* 14(2) s. 149-170.
- Nordgren, Kenneth (2008). Varför betyg?. I M. Jansdotter & K. Nordgren (Red.), *Betyg i teori och praktik* (s. 6-16). Malmö: Gleerups.
- Nyström, Peter (2004). *Rätt mått på prov: Om validering av bedömningar i skolan*. (Pedagogiska rapporter från Pedagogiska institutionen, Nr 71) Umeå: Umeå universitet, Pedagogiska institutionen.
- Pettersson, Astrid (2005). Bedömning – varför, vad och varthän?. I L. Lindström & V. Lindberg (Red.), *Pedagogisk bedömning – Om att dokumentera, bedöma och utveckla kunskap* (s. 31-42). Stockholm: HLS Förlag.
- Petty, Stephen (2008). Like bikers, we're victims of a marketing blunder that has caused a stink. *Times Educational Supplement*, 7 november 2008, s. 44-45.
- Popham, W. James (2008). *Classroom Assessment – What Teachers Need to Know* (5:e uppl.). Boston: Pearson/Allyn & Bacon.
- Sadler, D. Royce (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education*, 14(3) s. 387-392
- Selghed, Bengt (2006). *Betygen i skolan – kunskapssyn, bedömningsprinciper och lärarpraxis*. Stockholm: Liber.
- Selghed, Bengt (2007). Oförtjänt kritik mot lärare. I A. Pettersson (Red.), *Sporre eller otyg – om bedömning och betyg* (s. 37-49). Stockholm: Lärarförbundets förlag
- Scherp, Hans-Åke (2008). Vart tog forskningen vägen? – regeringens skolpolitik under luppen. *Pedagogiska Magasinet* (3) s. 12-19.
- Skolverket (2008). *Allmänna råd och kommentarer – Den individuella utvecklingsplanen med skriftliga omdömen*. Stockholm: Fritzes.
- Sohlberg, Peter & Britt-Marie (2001). *Kunskapens former – Vetenskapsteori och forskningsmetod*. Malmö: Liber.
- Steinberg, John (2004). *Världens bästa fröken: när modern pedagogik fungerar*. Solna: Ekelunds förlag.
- Strandberg, Leif (2006). *Vygotskij i praktiken: bland plugghästar och fuskklappar*. Stockholm: Norstedts akademiska förlag.

- Stukát, Staffan (2005). *Att skriva examensarbete inom utbildningsvetenskap*. Lund: Studentlitteratur.
- Taras, Maddalena (2005). Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies*, 53(4), s. 466-478.
- Taras, Maddalena (2009) Summative assessment: the missing link for formative assessment. *Journal of Further and Higher Education*, 33(1), s. 57-69.
- Tholin, Jörgen (2003). *En roligare dans?: svenska skolors första tolkning av innebörden i lokala betygskriterier i tre ämnen för skolår åtta* Borås: Institutionen för pedagogik, Borås: Högskolan i Borås.
- Tholin, Jörgen (2006). *Att kunna klara sig i okänd natur: En studie av betyg och betygskriterier – historiska betingelser och implementering av ett nytt system* Borås: Institutionen för pedagogik, Borås: Högskolan i Borås.
- Tholin, Jörgen (2007). Vilken kunskap räknas?. I A. Pettersson (Red.), *Sporre eller otyg – om bedömning och betyg* (s. 9-20). Stockholm: Lärarförbundets förlag.
- Tholin, Jörgen (2008). Skriftliga omdömen stjäl arbetstid. *Skolvärlden* (19) s. 26.
- Tomlinson, Carol Ann (2005). Grading and Differentiation: Paradox or Good Practice?. *Theory Into Practice*, 44(3), s. 262-269.
- Utbildningsdepartementet (2007). *Regeringens proposition 2006/07:86 Vissa skolfrågor*. Stockholm 22 mars 2007.
- Utbildningsdepartementet (2008a). *Nu införs skriftliga omdömen för alla elever*. Pressmeddelande 2008-06-10. Hämtad 23 februari 2009 från <http://regeringen.se/sb/d/10403/a/106967>
- Utbildningsdepartementet (2008b). *Frågor och svar om skriftliga omdömen*. publicerat 13 juni 2008. Hämtad 23 februari 2009 från <http://regeringen.se/sb/d/10721/item107347>
- Utbildningsdepartementet (2008c). *En individuell utvecklingsplan*. Promemoria 15 januari 2008. Hämtad 23 februari 2009 från <http://regeringen.se/content/1/c6/09/59/33/422eb9d2.pdf>
- Wedman, Ingemar (2003) Från skallmätning till VG – Om skolans betyg. I S. Selander (Red.), *Kobran, nallen och majjen – Tradition och förnyelse i svensk skola och skolforskning* (Forskning i fokus, nr 12) (s. 307-316). Stockholm: Myndigheten för skolutveckling.
- Wehner-Godée, Christina (2005). Att bedöma små barns kunnande. I L. Lindström & V. Lindberg (Red.), *Pedagogisk bedömning – Om att dokumentera, bedöma och utveckla kunskap* (s. 93-109). Stockholm: HLS Förlag.
- Welsh, Megan E. & D'Agostino, Jerome V. (2009) Fostering Consistency Between Standards-Based Grades and Large-Scale Assessment Results. I T.R. Guskey (Red.), *Practical Solutions for Serious Problems in Standards-Based Grading* (s. 75-104). Thousand Oaks: Corwin Press.
- Wigblad, Rune (1997). *Karta över vetenskapliga samband – Orientering i den samhällsvetenskapliga metodjungeln* (2:a uppl.). Lund: Studentlitteratur.
- Wikström, Christina (2006). Profiles of educational assessment systems worldwide – Education and assessment in Sweden. *Assessment in Education: Principles, Policy & Practice*, 13(1), s. 113-128.

- William, Dylan (2006) *Does assessment hinder learning?* Paper presenterat vid ETS Invitational Seminar 11 juli 2006 vid Institute of Civil Engineers, London, UK. Hämtad 3 mars 2009 från <http://www.dylanwilliam.net/>
- Wretman, Steve (2008). Bedömning för att värdera eller stimulera. *Grundskoletidningen*, (3), s. 18-20.
- Ödman, Per-Johan (2007). *Tolkning, förståelse, vetande – Hermeneutik i teori och praktik* (2:a uppl.). Stockholm: Norstedts Akademiska Förlag.