



Vetenskapsrådet

SVENSK FORSKNING OM BEDÖMNING – EN KARTLÄGGNING



SVENSK FORSKNING OM BEDÖMNING – EN KARTLÄGGNING

Eva Forsberg och Viveca Lindberg

SVENSK FORSKNING OM BEDÖMNING – EN KARTLÄGGNING

Eva Forsberg och Viveca Lindberg

Rapporten kan beställas på www.vr.se

VETENSKAPSRÅDET

Box 1035

101 38 Stockholm

© Vetenskapsrådet

ISSN: 1651-7350

ISBN: 978-91-7307-166-6

Rapportnummer 2:2010

Grafisk form: Erik Hagbard Couchér, Vetenskapsrådet

Foto: Stockxchng

Produktion: Matador Kommunikation

Tryck: CM-Gruppen AB, Bromma 2011

FÖRORD

Utbildningsvetenskaplig forskning rymmer en mångfald av forskning om lärande och kunskapsbildning inom utbildning och undervisning. Vetenskapsrådets utbildningsvetenskapliga kommitté arbetar på flera sätt för att sprida kunskap om aktuella forskningsresultat. Ett exempel är den årliga konferensen Resultatdialog, ett annat exempel är genom att sammanställa forskningsresultat inom ett avgränsat område i form av forskningsöversikter.

Den föreliggande forskningsöversikten är den andra i en serie som tas fram i samarbete med Skolverket, som har finansierat översikten. Skolverket ska stödja skolhuvudmän, rektorer och lärare i deras utbildningsverksamhet och bidra till att förbättra deras förutsättningar att arbeta med utveckling av verksamheterna för ökad måluppfyllelse. I detta arbete sprider Skolverket kunskap om forskningsresultat som är av särskilt intresse för utbildning och undervisning.

Det övergripande målet för denna översikt är att presentera och diskutera aktuell svensk forskning om bedömning och att göra det i ljuset av internationell forskning. De slutsatser och reflektioner som presenteras är författarnas egna. Översikten har tagits fram av docent Eva Forsberg, Uppsala universitet, och fil dr Viveca Lindberg, Stockholms universitet. Den har sakgranskats genom peer review.

Stockholm 2010-08-01

Sigbrit Franke

Ordförande i Vetenskapsrådets utbildningsvetenskapliga kommitté

Elisabet Nihlfors

Huvudsekreterare för utbildningsvetenskap vid Vetenskapsrådet

INNEHÅLL

ENGLISH SUMMARY	6
1 BEDÖMNING – INTRODUKTION	8
1.1 Uppdrag och inramning	9
1.2 Utgångspunkter och rapportens struktur	11
2 BEDÖMNING SOM ETT METODOLOGISKT BESTÄMT OBJEKT	13
2.1 Abstractstudierna: avhandlingar och vetenskapliga artiklar	13
2.2 Kompletterande in- och utblickar	16
2.3 Avslutande noteringar	18
3 BEDÖMNINGSPRAKTIKER FRAMVUXNA I SVENSK SKOLA	20
3.1 Mogen för skolan – skolmognadsprövningar	20
3.2 Från husförhör till läxförhör och skrivningar	23
3.3 System för betyg	28
3.4 Sammanhanget för skolan och bedömningen	32
3.4.1 Skolans omstrukturering	33
3.4.2 Globaliseringen och internationella jämförande kunskapsmätningar	35
3.5 Den svenska skolans bedömningskulturer	36
4 BEDÖMNINGSFORSKNING – NÅGRA UTBLICKAR	40
4.1 Bedömningsforskning i internationella forskningshandböcker	40
4.1.1 Undervisning, lärare, lärarutbildning och läroplaner	40
4.1.2 Utvärdering, policy, utveckling och sociologi	43
4.1.3 Språk, estetiska ämnen och informationsteknologi	46
4.2 Internationella organisationer och konferenser	49
4.2.1 AERA: American Educational Research Association	49
4.2.2 EERA: European Educational Research Association	50
4.2.3 EARLI: European Association of Research on Learning and Instruction	51
4.2.4 NFPF/NERA: Nordisk Förening för Pedagogisk Forskning	52
4.2.5 IAEA International Association for Educational Assessment	52
4.2.6 AEA EUROPE: Association for Educational Assessment	53
4.3 Bedömning som epistemiskt objekt	54
4.3.1 Bedömning – ett kunskapsobjekt i omvandling	55
5 SVENSK OCH NORDISK BEDÖMNINGSFORSKNING OCH DESS OMFATTNING	60
5.1 Forskning om bedömning i Sverige	60
5.1.1 Lärosäten engagerade i forskning om bedömning	62
5.1.2 Discipliner involverade i forskning om bedömning	66
5.2 Artiklar om bedömning i nordiska tidskrifter	68
5.3 Norden i internationella sammanhang	70
6 BEDÖMNINGSFORSKNINGENSSTUDIEOBJEKT	73
6.1 En översiktsbild	74
6.1.1 Studieobjekten – beskrivningar och fördelningar	75
6.2 Artiklar i svenska, nordiska och internationella tidskrifter	78

7 SVENSKA AVHANDLINGAR 1990–2009	80
7.1 Lärares didaktiska bedömningspraktik.....	80
7.2 Elevers prestationer och skolors resultat.....	81
7.3 Skolans styrning och lärares arbete.....	82
7.4 Övergångar.....	83
7.5 Organisering av skola, undervisning och elever.....	84
7.6 Bedömningar som fenomen.....	84
7.7 Elevers erfarenheter och upplevelser av bedömning.....	85
7.8 Bedömningsredskapens kvalitet.....	86
7.9 Andra centrala aspekter i avhandlingarna.....	87
8 BEDÖMNINGSFORSKNINGEN – ETT FRAMTIDSPERSPEKTIV	89
8.1 Omfattning och forskningsmiljöer.....	89
8.2 Bedömningsprojektens studieobjekt.....	90
8.2.1 Lärares didaktiska arbete.....	90
8.2.2 Elevers prestationer och skolors resultat.....	91
8.2.3 Skolans styrning och lärares arbete.....	92
8.2.4 Organisering av skola, undervisning och elever.....	93
8.2.5 Övriga kategorier/studieobjekt.....	93
8.3 Avslutandet noteringar.....	95
9 BEDÖMNING SOM OBJEKT FÖR KUNSKAPSBILDNING	96
9.1 Metriska studier – statistiska analysmodeller.....	96
9.1.1 Individuella differenser.....	96
9.1.2 Provkonstruktion.....	97
9.1.3 Övergångar.....	97
9.1.4 Komparativa studier.....	98
9.1.5 Effektstudier.....	99
9.2 Professions- och verksamhetsstudier.....	100
9.2.1 Bedömningsredskap.....	100
9.2.2 Lärande/bedömning.....	101
9.2.3 Lärarprofessionen.....	102
9.3 Systemstudier.....	103
9.3.1 Historisk/sociologisk bedömningsforskning.....	103
9.3.2 Styrnings- och policystudier.....	107
9.4 Ett sammanfattande avstamp.....	109
10 BEDÖMNINGSFORSKNINGEN – NÅGRA UTVECKLINGSLINJER	111
10.1 Svensk bedömningsforskning – en lägesbestämning.....	111
10.2 Svensk bedömningsforskning – ett framtidsperspektiv.....	115
REFERENSER	119
KÄLLOR	136
BILAGA 1: MÖTET MELLAN KUNSKAPSOBJEKT OCH ADMINISTRATIVA SÖKSYSTEM	137

ENGLISH SUMMARY

Swedish research on assessment – a survey

Measuring, recording and judging the value of individual achievements and societal phenomena is nothing new. In one sense, it is a natural part of what it means to be a human among humans, that is, a phenomenon that is part and parcel of culture. The systematic character and scale of evaluation today, however, reflect a more modern, rational approach to this activity. The increasingly frequent use being made of evaluation and assessment is linked to the assumption that individuals and societies are amenable to development. Associated with this is a goals-and-means philosophy, with planning and evaluation as key components. Although – or perhaps precisely because – this rationality has been challenged, much emphasis is placed on the importance of carefully following, monitoring and evaluating activities and their out-comes. Nowadays, moreover, our dependence on knowledge is underscored. Concepts such as a knowledge economy or society, and lifelong and life-wide learning, can be seen as examples of this. In political and economic terms, the result has been a growing interest in the contribution education can make to the development of society.

In 2008 some three million individuals were involved in child day care or in preschool, primary, secondary, adult or higher education in Sweden. In addition, there were the people employed in school and other education and in research, and the 2.1 million who participated in some form of staff training. Education, in other words, is an activity that tangibly affects a great many people. Although education for children and young people is provided in both the public and the private sector, it is funded mainly from the public purse. That means that, economically, education is our biggest policy area, a fact which in itself invites considerable interest in questions concerning its meaning, effectiveness and benefits, for both the individual and society.

From the standpoint of educational science, assessment can be regarded as one of the systems by which education signals what knowledge is important and how knowledge, skills and proficiency can be expressed, discerned and communicated. Together with curriculum and pedagogy, assessment or evaluation is thus one of what the sociologist of education Basil Bernstein (1971/1980) called the “message systems” of education. Grades and diplomas also serve as tools for selection, opening or closing doors to different

educational pathways and careers. Assessment and grading by teachers thus crucially influence students' life chances. Against this background, it is hardly surprising that questions of pupil achievement, school results and the effects of education are high on the agenda of public and educational debate. Here, it will be explored whether this is also the case in research.

1 BEDÖMNING – INTRODUKTION

Att mäta, bedöma och betygsätta enskildas prestationer och samhälleliga företeelser är inte något nytt. Däremot är den systematik och omfattning som karakteriserar dessa företeelser i dag ett mer modernt rationellt uttryck. Det ökade bruket av utvärderingar och bedömningar kan ses som ett uttryck för ett mål – medeltänkande med planering och utvärdering som centrala element och utveckling som huvudsakligt syfte. Även om denna rationalitet har utmanats, eller kanske just därför, framhålls vikten av att noggrant följa, kontrollera och värdera prestationer. I utbildningssammanhang är prestationerna av tradition relaterade till kunskap och vårt beroende av kunskap. Begrepp som kunskapsekonomi och kunskapsamhälle liksom livslångt och livsvitt lärande kan ses som exempel på detta. Politiskt och nationalekonomiskt har dessa faktorer medfört ett ökat intresse för utbildningens bidrag till samhällsutvecklingen. Bedömning har därmed kommit att bli en företeelse som fått en betydligt bredare uppmärksamhet än vad som tidigare var fallet.

År 2008 var cirka tre miljoner individer involverade i förskola, barnomsorg, grund- och gymnasieskola, vuxenutbildning och högre utbildning. Utbildning är med andra ord en verksamhet som rent konkret berör många människor. Även om utbildningen för barn och ungdomar genomförs i både offentlig och privat regi är verksamheten i huvudsak offentligt finansierad. Det innebär att utbildning ekonomiskt sett är vårt största politikområde, något som i sig bäddar för ett stort intresse för frågor om utbildningens mening, effektivitet och avkastning för såväl individen som samhället. I det här sammanhanget har betyg och andra liknande resultatmätt fått ökad uppmärksamhet.

Ur ett pedagogiskt perspektiv kan bedömning betraktas som ett av de system genom vilket utbildningen signalerar vilken kunskap som är viktig, liksom hur kunnande kan uttryckas och hur kunnighet kan urskiljas och kommuniceras. Tillsammans med läro- och kursplaner och undervisningen utgör bedömning och utvärdering därmed vad utbildningssociologen Basil Bernstein (1971/1980) kallade utbildningens "message systems". Betyg, examina och andra resultatmätt används också som redskap för selektion, som öppnar respektive stänger dörrar för olika bildningsgångar och karriärer. Lärares bedömningar och betygsättning har därigenom en avgörande betydelse för elevers livschanser. Mot den här skisserade bakgrunden är det föga förvånande att frågor om elevers prestationer, skolors resultat och utbildningens effekter står högt upp på dagordningen i den

offentliga och utbildningspolitiska debatten. Om så också är fallet inom forskningen och i så fall på vilka sätt är en fråga som här kommer att uppmärksammas.

1.1 Uppdrag och inramning

Avsikten med denna rapport är att analysera kunskapsbildning om den utbildningsvetenskapliga forskningen om bedömning, som är relaterad till grundskolan och gymnasieskolan samt övergången mellan skolformer. I bedömning inkluderar vi praktiker förbundna med bedömning av enskilda elevers prestationer, som de kommer till uttryck i klassrummet och i skolans administrativa system. Det innebär att både internt och externt reglerade bedömningsredskap och -former uppmärksammas.

Vidare fokuserar vi både ett primärt och ett sekundärt bruk av bedömning i forskningen. Det får som konsekvens att översikten behandlar forskning som har bedömning som huvudfokus och forskning som använder resultat från olika bedömningar som nationella prov, betyg och internationella kunskapsmätningar för att analysera olika fenomen förbundna med skola och utbildning. Sammanfattningsvis rymmer översikten därmed både bedömningsforskningens kärna (forskning som traditionellt förknippas med bedömning och betyg) och dess periferi (bruket av bedömningsresultat i forskning). Vi har således valt en bred tolkning av uppdraget. En närmare metodologisk bestämning av bedömning finns i kapitel 2.

Översikten täcker svensk forskning som den utvecklats under tjuugoårsperioden 1990–2009. I syfte att ge perspektiv på den svenska forskningen har vi kompletterat den med några utblickar (jfr kapitlen 3 och 4). Motivet för den valda perioden är såväl nationellt som internationellt förankrat. I Sverige inleds perioden med en kritik av de normrelaterade betygen som under 1990-talet ersätts med det mål- och kunskapsrelaterade betygssystemet. I dag förbereds återigen en revidering av betygssystemet. Under perioden har också deltagandet i internationella komparativa mätningar ökat, liksom bruket av resultaten från dem för sekundäranalyser. Dessutom rymmer perioden en restrukturerings av utbildningssystemet i sin helhet. Gransknings- och redovisningssamhälle, performativ vändning, ”accountability” och kvalitetssäkring är några av de ord som använts för att beskriva förändringen. Dessa företeelser är alla nära förbundna med bedömningsfrågor. Nationellt, men också internationellt, har intresset för och mängden studier av bedömning ökat under perioden.

För översikten har vi, i enlighet med det uppdrag som utformades i samarbete med Vetenskapsrådet, använt abstract eller sammanfattningar.

Valet av abstract motiveras av materialets omfattning i relation till den tid som stod till förfogande för uppdraget. Avhandlingarnas och artiklarnas abstracts kommer att analyseras och kategoriseras med utgångspunkt i följande frågor:

- Vilka forskningsfrågor ställs?
- Vilka forskningsansatser används?
- Vilka huvudsakliga resultat har forskningen genererat?

Basen består av avhandlingar, svensk forskning inom området som den kommer till uttryck i svensk kontext. Vidare ingår svensk forskning som den manifesterar sig i vetenskapliga tidskrifter med peer review: svenska, nordiska och internationella. Slutligen har avhandlingarna och artiklarna kompletterats med en genomgång av forskningsprojekt som under senare år erhållit medel från Utbildningsvetenskapliga kommittén. Därmed kan vi också ge en bild av pågående forskning, resultat som kan förväntas i den nära framtiden.

Med utgångspunkt i uppdraget har vi i första hand sökt forskning genom de svenska forskningsbiblioteken (avhandlingar), internationella databaser (vetenskapliga tidskriftsartiklar) och webbsök. De sistnämnda har använts för att ge en bild av svensk forskning om bedömning ur ett internationellt perspektiv. En sammanfattning av studiens inramning med avgränsningar finns i tabell 1.1.

TABELL 1.1. Studiens inramning.

Publikationstyper	avhandlingar, tidskriftsartiklar, forskningsansökningar
Tidsperiod	1990–2009
Rumslig inramning	Sverige, Sverige i Norden och i det internationella
Innehållslig bestämning	forskning om bedömning eller forskning som brukar prestationsmätt
Skolformer	grund- och gymnasieskolan samt övergångar mellan skolformer
Sökstrategier	svenska forskningsbibliotek, internationella databaser, webbsök

Här uppmärksammas kortfattat de utgångspunkter vi valt för forskningsöversikten och rapportens övergripande struktur.

1.2 Utgångspunkter och rapportens struktur

Vi tar i kapitel 3 avstamp i bedömning som social verksamhet. Det innebär att vi ger en bild av bedömningskulturen i det svenska systemet för utbildning och delvis också dess sammanhang. Ingången är uttryck för uppfattningen att objektet som studeras och kunskapen om detta objekt ömsesidigt strukturerar varandra, även om det inte råder något ett-till-ett-förhållande mellan dem. Med andra ord handlar det om att bedömning som social verksamhet är påverkad av den i samhället inbäddade kunskapen om bedömning och vice versa. Å ena sidan ger kunskap upphov till utvecklandet av nya former för bedömning, å andra sidan utgör nya bedömningspraktiker betingelser för utvecklandet av ny kunskap. Att ta utgångspunkt i en beskrivning som tar fasta på några centrala karakteristika i den svenska bedömningskulturen är således ett sätt att skriva fram några av villkoren för forskningen om bedömning. Beskrivningen kan fungera som en av flera grunder för fenomen och aspekter som vi kan förvänta oss kan eller bör vara föremål för forskning om bedömning.

Vid sidan av den kunskaps sociologiska utgångspunkten är studien inramad av ett läroplansteoretiskt/didaktiskt perspektiv. Det får i detta sammanhang främst tre konsekvenser. För det första betyder det att bedömning ses som en integrerad del av tre internt relaterade signalsystem: läroplan, undervisning och utvärdering. Av särskild betydelse är den didaktiska relationen som lyfter fram förhållandet mellan elev/er, lärare och innehåll, liksom bedömning som en kommunikativ aspekt i denna relation (jfr Blankertz 1987). För det andra betonar läroplansteoretiska perspektiv betydelsen av makt, kontroll och legitimitet i formering, manifestering och reformulering av utbildningen och de olika signalsystemen. För det tredje pekar läroplansteoretiker ut arenor med aktörer som deltar i förhandlingar/kampen om och formandet av utbildningen: utbildningspolitik och administration, utbildningsvetenskaplig forskning och pedagogisk verksamhet (jfr Forsberg red. 2007). Till detta kan fogas dagspress, TV och webben. Hit hör också tidskrifter som dels medierar budskap mellan aktörer på olika arenor, dels medverkar som självständiga aktörer som på ett avgörande sätt kan påverka vilka frågor som kommer upp på dagordningen och hur de kan förstås. För denna rapport har såväl studiens design, val av källmaterial som analys av data formats med utgångspunkt i ett läroplansteoretiskt perspektiv.

Kapitlen 2–4 är ett slags bakgrund för översikten medan översiktens resultat redovisas i kapitlen 5–9.

I kapitel 2 redovisar och diskuterar vi hur bedömning i denna studie fått sin metodologiska bestämning. Därefter tar vi i kapitel 3 upp den svenska skolan och bedömningen. På detta följer kapitel 4 som ger en internationell utblick.

I kapitel 5 beskriver vi omfattningen av den nordiska forskningen om bedömning liksom involverade forskningsmiljöer i den svenska forskningen. Dessutom beskriver vi förekomsten av svensk och till viss del nordisk bedömningsforskning i svenska, nordiska och internationella tidskrifter. I kapitel 6 ligger fokus på kategoriseringen av undersökningarnas studieobjekt. Vidare ges en översiktlig bild för samtliga material. I samband med detta kommenterar vi särskilt tidskriftsartiklarna. I kapitel 7 följer en mer ingående beskrivning av studieobjekten i de svenska avhandlingar som publicerades 1990–2009. I kapitel 8 ger vi med utgångspunkt i projekt som erhållit medel från UVK en bild av pågående projekt. I kapitel 9 riktar vi uppmärksamheten mot bedömningsforskningens skilda inriktningar.

I det avslutande kapitel 10 blickar vi framåt och pekar på några möjliga utvecklingslinjer för den svenska forskningen om bedömning.

2 BEDÖMNING SOM ETT METODOLOGISKT BESTÄMT OBJEKT

Bedömning är, som kommer att framgå med större tydlighet av följande kapitel, ett mångfacetterat kunskapsobjekt. Hur vi metodologiskt närmar oss bedömning får därmed avgörande konsekvenser för vad som kommer att innefattas i bedömning. I detta kapitel redovisar och diskuterar vi hur studien har genomförts och hur bedömning fått sin metodologiska bestämning. Inledningsvis står abstractstudien med avhandlingarna och de vetenskapliga artiklarna i fokus. Därpå följer ett avsnitt som särskilt uppmärksammar mötet mellan bedömning som kunskapsobjekt och de administrativa söksystemen, med sökningarna efter artiklar i databaser som exempel. Avslutningsvis diskuterar vi några kompletterande perspektiv som riktar intresset mot pågående forskning, praktikerfarenheter av bedömning samt bedömningsforskningens internationella sammanhang.

2.1 Abstractstudierna: avhandlingar och vetenskapliga artiklar

I uppdraget ingår alltså att genomföra en abstractstudie av svenska avhandlingar och av artiklar i ett urval vetenskapliga tidskrifter – nationella, nordiska och internationella. Bilden av bedömningsforskningen i Norden, i nordiska och internationell vetenskapliga tidskrifter, används främst som en kontrast till den svenska. Av metodologiska skäl utgör avhandlingarna basen i studien, eftersom deras abstract generellt sett är utförligare. Därför har vi också valt den bredaste sökstrategin för avhandlingarna. Sökningar har genomförts i Libris. Vi har använt olika termer relevanta för bedömning i kombination med avhandling. Det innebär bland annat att både doktors- och licentiatavhandlingar ingår. Vi har också kompletterat med några avhandlingar som vi kände till sedan tidigare och som inte föll ut, men likväl är relevanta.

Hur forskning om bedömning med relevans för utbildning ska avgränsas är inte givet. Det är därför inte heller vilka relevanskriterier som kan eller bör användas när källmaterial ska sökas och väljas ut. För att inte bli i händerna på vardagliga termer och administrativt utvecklade söksystem är det av vikt att man redan från början har en relativt god bild av det kunskaps-

objekt som ska kartläggas. I annat fall riskerar man att missa en del av den forskning som de facto finns inom området, men som inte klassificerats som man förväntat sig. I vissa söksystem används de nyckelord författaren uppgett, medan andra söksystem använder klassificeringar som tilldelas av administratörerna. Man måste således ha en uppfattning om vilka centrala forskare och vilken forskning som bör finnas med i utfallet för att resultatet ska vara acceptabelt. I de fall objektet är mångfacetterat kan man antingen anlägga en bred sökingång som möjliggör ett helhetsperspektiv. Man kan också använda en snävare ingång som betonar någon eller några aspekter.

Vi har generellt sett valt en bred sökstrategi, i synnerhet när det gäller avhandlingarna. Det framgår bland annat av våra söktermer. Parenteserna innehåller ett urval exempel. Söktermerna innefattar övergripande termer (bedömning, utvärdering, psyometri, inspektion) termer relaterade till vad som bedöms (förmåga, bildning, kunskapsbedömning, personlighetsutveckling), former av bedömning (test, prov, kontroll, läxor, formativa/summativa bedömningar, internationella jämförande test och inspektion, kvalitetsredovisning, uppföljning), bedömningens resultat (prestation, skolresultat, godkänd), bedömningens funktion (differentiering, selektion, segregering, urval, utbildningsval, rekrytering, information, motivation, anpassad studiegång, studieavbrott) samt dokumentation och kommunikation av bedömningars resultat (kvartssamtal, utvecklingsamtal, individuella utvecklingsplaner, betyg, examina).

Libris erbjuder, vid sidan av en mer fritt utformad boolesk sökning, fyra fasta vägar för ordbaserade sökningar: fritext, ämnesord, bläddra ämnesvis och index. Fritextsökningen hittar alla poster i ett antal fördefinierade fält, däribland ämnesord och genrebeteckningar. Ämne begränsar sökningen till poster som klassificerats i enlighet med sökorden. Både fritext och ämne kan trunkeas. I Libris finns också en möjlighet att bläddra ämnesvis för att få en översikt över litteratur inom ett område. Dessutom kan man bläddra i alfabetiska listor över olika ämnen. Det kanske kan tyckas att det är självklara saker, men ur vårt perspektiv fick vi två överraskningar. Dels ledde

de olika sökvägarna till skilda resultat¹, dels fick vi pröva rätt många olika sökord för att få upp sådana avhandlingar som vi kände till.

Vi har valt en bred ingång, men det medför inte att allt inkluderas. Den geografiska avgränsningen till Sverige (för avhandlingarna) och Norden (för artiklarna) har tidigare nämnts liksom inriktningen mot grund- och gymnasieskolan samt övergångar mellan skolformer. Om kärnan i vad vi studerar har klargjorts genom sökorden kan gränserna delvis läggas fast genom en negativ bestämning. Vi har inte tagit med studier som explicit behandlar bedömningar med medicinsk, specialpedagogisk eller psykologisk grund, eller forskning som exklusivt avser bedömningar i andra skolformer än de angivna. Inte heller forskning om utvärdering av program, organisation och system har tagits med om den inte kan återföras på enskilda elevers prestationer. Detsamma gäller för de studier som avser övergångar mellan skolformer.

Om vi övergår till artiklarna har vi valt två svenska vetenskapliga tidskrifter. Den ena, *Pedagogisk Forskning i Sverige* som grundades 1996, är den samlande tidskriften för disciplinen pedagogik och utkommer med fyra nummer per år. Den allmändidaktiska tidskriften *Utbildning & Demokrati. Tidskrift för didaktik och utbildningspolitik*, som har funnits sedan 1992, utges tre gånger per år. Den har en institutionell spridning till de flesta svenska högskolor och universitet. I studien ingår också artiklar som beaktar nordisk forskning om bedömning. De två för utbildningsområdet mest centrala tidskrifterna har tagits med. *Nordisk Pedagogik*, som etablerades 1980, publicerar artiklar på de nordiska språken och engelska. *Scandinavian Journal of Educational Research*, som utkommer med sex nummer per år sedan 1957, är den enda engelskspråkiga pedagogiska tidskriften i Norden.²

¹ Ett för vårt syfte intressant sökord är differentiering. Avgränsat till avhandlingar (typ av material) och med trunkering gav frisök 114 träffar och ämnesord 14 träffar. I funktionen bläddra ämnesvis kommer man till differentiering via uppfostran och undervisning, pedagogik och pedagogisk psykologi till differentieringsproblem. Med fortsatt avgränsning till avhandlingar är resultatet 76 träffar. Slutligen ger indexsök följande utfall (även i detta fall avgränsat till avhandlingar) för ord och ordkombinationer som kan vara relevanta: differentiering (0 träffar), differentiering i utbildning/en (8). Här hänvisas man vidare till begåvningsdifferentiering (0), individualiserad undervisning (1) och pedagogisk psykologi (6) liksom social differentiering (1). Utöver dessa finns ett antal träffar på enskilda kombinationer av differentiering. Sammanfattningsvis ger dessa olika sökstrategier denna träffbild för differentiering med avgränsningen avhandlingar: frisök (114), ämnesord (14), bläddra ämnesvis (76) och indexsök (16). Det exempel vi valt här är inte avgränsat till vår tidsperiod, för vårt resonemang spelar det dock ingen roll. Det vi vill peka på är den stora variationen i utfallet från 14–114, och det är dessvärre inte så att de 114 rymmer samtliga i de övriga kategorierna. Att söka relevant material är med andra ord ett grannliga arbete där det inte enbart räcker med att välja rätta sökord. Ett systematiskt tillvägagångssätt är förvisso en nödvändig, men inte säkert en tillräcklig, grund för valda sökningar. Detta är ett exempel på hur vi genom sökprocessen återkommande fått arbeta med att justera sökningen i relation till det material (avhandlingar) vi hittat, men också för att, med hjälp av avhandlingar vi känner till, fått se vilka sökord som ger de avhandlingar som definitivt hör till men inte kommit upp genom de sökord vi valt.

² För ytterligare information om tidskriften se t.ex. Karl-Georg Ahlström (2009).

Sökningarna i de svenska och nordiska tidskrifterna har förutom när det gäller Scandinavian Journal of Educational Research gjorts manuellt och med utgångspunkt i titlar och abstract. Scandinavian Journal of Educational Research söktes via ERIC. Det kommenteras ytterligare i bilaga 1. Som ledstjärna för sökningarna i tidskrifterna har samma sökord använts som i avhandlingsstudien, med en öppenhet för att även annat kan vara relevant. I de nämnda tidskrifterna ingår artiklar på svenska, danska, norska och engelska. Av tabell 2.1 framgår både vilka svenska och nordiska tidskrifter vi använt och vilka tidsperioder som vi täcker in.

TABELL 2.1. Svenska och nordiska vetenskapliga tidskrifter och undersökningsperioder.

Svenska vetenskapliga tidskrifter:	
Pedagogisk Forskning i Sverige	1996–2009
Utbildning & Demokrati	1992–2009
Nordiska/skandinaviska vetenskapliga tidskrifter:	
Nordisk Pedagogik	1990–2009
Scandinavian Journal of Educational Research	1990–2009

2.2 Kompletterande in- och utblickar

I syfte att kontrastera den svenska forskningen om bedömning har vi genomfört vissa kompletterande studier. Den första, som fokuserar pågående forskning, är en framåtsyftande utvidgning. Den andra är en beskrivning av den svenska skolan och bedömning. Den kan närmast karakteriseras som en inblick i vad som pågår i praktiken. Den tredje fokuserar den internationella bedömningsforskningen och representerar en utblick. Såväl blicken in i framtiden som in- och utblicken kommenteras nedan.

Forskning som finns avrapporterad i avhandlingar och tidskrifter är som regel resultat av redan genomförd forskning. Vi har därför kompletterat denna med en analys av forskningsprojekt som fått medel från Utbildningsvetenskapliga kommittén (UVK) under perioden 2006–2009. Därigenom kan vi få en inblick i vad vi kan vänta oss av den svenska bedömningsforskningen inom en nära framtid.

För samma period har vi även använt de kortfattade presentationer av projekt som befinner sig i ett slutskede och som finns i UVK:s Resultat-

dialoger. Å ena sidan kan projektbeskrivningarna användas för att slå fast var den svenska forskningsfronten befinner sig. Å andra sidan kan de ses som uttryck för forskning som saknas och som behöver utvecklas. Notabelt är också att vi genom information på hemsidor kunnat konstatera att andra bidragsgivare till forskning om utbildning endast undantagsvis finansierar forskning om bedömning.

Utifrån det läroplansteoretiska perspektivet redovisar vi i kapitel 3 en beskrivning av den svenska skolan och bedömningen. Genom en rörelse mellan då och nu kan vi få syn på såväl det som är mer grundläggande och beständigt som det som är mer tillfälligt och flyktigt. Därmed kan centrala karakteristika i den svenska bedömningskulturen lyftas fram.

Utöver hur svensk och nordisk forskning framträder på den internationella arenan gör vi i kapitel 4 en utblick genom att översiktligt uppmärksamma hur bedömning tas upp i internationella handböcker. Vi har valt ett brett spektrum forskningshandböcker där bedömningsforskning kan förväntas utgöra antingen huvudfokus eller en central aspekt. Handböckerna avser områden som undervisning, lärare, lärarutbildning och läroplansteoretiska frågor samt utbildningens utvärdering, policy, utveckling och sociologi. Dessutom beaktar vi några specifika områden som språk/engelska, estetiska skolämnen och informationsteknologi.

Vid sidan av detta har vi skannat av förekomsten av bedömningsfrågor i ett antal internationella organisationer som genom sina konferenser fungerar som arenor och mötesplatser för utbildningsvetenskaplig forskning. AERA, EARLI och EERA kan ses som paraplyorganisationer för olika kategorier av utbildningsvetenskaper genom sina avdelningar, nätverk och särskilda intressegrupper (Lindblad, Kyndel & Larson 2004). Uppbyggd på ett likartat sätt är också NFPP, som är en nordisk paraplyorganisation inom utbildningsområdet och även två för bedömning specifikt inriktade organisationer: IAEA och AEA-Europe. De fullständiga benämningarna för organisationernas akronymer finns i tabell 2.2.³

³ Information om organisationerna/konferenserna har sökts på följande webbsidor: www.aera.net, www.eera.ac.uk, www.earli.org, www.nfpf.net, www.iaea.info, www.aea-europe.net.

TABELL 2.2 Internationella paraply- och specialorganisationer inom utbildnings- och bedömningsområdet.

Paraplyorganisationer	
AERA	American Educational Research Association
EARLI	European Association for Research on Learning and Instruction
EERA	European Educational Research Association
NFPF	Nordisk Förening för Pedagogisk Forskning
Specialorganisationer	
IAEA	International Association for Educational Assessment
AEA-Europe	Association for Educational Assessment

Den översiktliga bild som handböckerna och organisationerna/konferenserna kan ge är på intet sätt heltäckande. Men den kan peka ut dels bedömning som en aspekt av utbildning, dels för bedömningsforskningen centrala områden och delvis även förändringar över tid. Därmed kan bilden fungera som en referenspunkt vid diskussionen om den svenska bedömningsforskningen.

2.3 Avslutande noteringar

Studiens övergripande syfte är alltså att ge bild av svensk forskning om bedömning som den framträder, huvudsakligen utifrån en svensk horisont. Den kontrasteras i första hand mot nordisk forskning, men även mot internationell, som den kommer till uttryck i forskningshandböcker och etablerade forskningsorganisationer, allmänpedagogiska eller med specifik inriktning på pedagogisk bedömning. Primärt analyseras abstract eller andra sammanfattande utdrag för den svenska och nordiska forskningens del. Abstracten har delvis olika kvalitet, beroende på vilka anvisningar tidskrifterna ger om deras utformning och vilken variation som accepteras. Tyngdpunkten och utgångspunkten för analyserna utgår därför från avhandlingarna, som kompletteras med artiklarna och UVK-projekten. Utblicken till organisationer med konferenser och till forskningshandböcker ger tillgång till bedömning ur ett internationellt perspektiv. Tabell 2.3 innehåller en översiktlig beskrivning av omfattningen på det material vars abstract och sammanfattningar vi använt för analyser.

TABELL 2.3. Det totala materialets omfattning.

Svenska avhandlingar	95
Vetenskapliga tidskriftsartiklar	388
Artiklar i svenska vetenskapliga tidskrifter	52
Artiklar i nordiska tidskrifter ⁴	109
Svenska och övriga nordiska artiklar i andra internationella tidskrifter	227
UVK-projekt	35
Totalt antal avhandlingar, artiklar och projekt	518

Sammanfattningsvis grundas studiens design såväl i de i kapitel 1 angivna kunskaps sociologiska och läroplansteoretiska/didaktiska utgångspunkterna som i den i detta kapitel angivna metodologiska bestämningen. Det innebär att vi intresserar oss för bedömning både som en social verksamhet och som ett objekt för kunskapsbildning. Vi uppmärksammar också hur både bedömning och forskning om bedömning förändras över tid. Vi rör oss, för att låna ett uttryck från Reinhart Koselleck (2004), mellan erfarenhetsrum och förväntanshorisonter.

⁴ Artiklarna i de nordiska tidskrifterna representeras i tabellen i bilaga 1 av två olika antal, 57 respektive 109, som en konsekvens av två skilda sökstrategier. De 57 artiklarna utgör grunden för siffror som redovisas i relation till materialet i sin helhet i tabell 6.1 respektive 6.3. I övriga fall har vi utgått från de 109 artiklarna.

3 BEDÖMNINGSPRAKTIKER FRAM- VUXNA I SVENSK SKOLA

Här gör vi några nedslag i bedömningspraktiker relaterade till den svenska skolan. Avsikten är att lyfta fram några centrala aspekter av bedömning som en bakgrund till översiktens resultat, inte att ge en heltäckande bild. Skola och bedömning kan förstås i relation till både en samtida och en historisk reproduktionsprocess. Den samtida gäller överförandet av ”kunskaper och färdigheter som är nödvändiga för att återskapa villkoren för produktionen och villkoren för att leva i samhället” (Lundgren 1999 s. 8). Den historiska reproduktionen handlar om traderingen av skolans egen historia, om skolans syfte, mål och innehåll. Även om skolan alltid är stadd i förändring kan vi urskilja avlagringar från dess historia både i utformningen av verksamheten och i vårt sätt att tänka om den. Genom att rikta blicken mot skolans ”institutionella minne” blir det möjligt att peka ut likheter och skillnader i skolans bedömningar av elever och deras prestationer sett över tid.

Vi rör oss mellan då och nu, där folkskolan och tiden för dess etablering är vårt exempel från förr i tiden och nuet primärt utgörs av grundskolan, även om några hänvisningar också är till andra skolformer. Under rubriken *Mogen för skolan – skolmognadsprövningar* intresserar vi oss för själva ”skolporten” och skolmognadsprovet som den möjliga entrébiljetten. I *Från husförhör till läxförhör och skrivningar* uppmärksammar vi interaktionen i klassrummet med fokus på relationen elev-lärare-innehåll liksom de former för bedömning som därmed aktualiseras. Vi betraktar särskilt System för betyg och deras förändring. Därefter behandlar vi *Sammanhanget för skolan och bedömningen*. Avslutningsvis sammanfattar vi och pekar ut några karakteristiska drag i *Den svenska skolans bedömningskultur*.

3.1 Mogen för skolan – skolmognadsprövningar

Institutionaliseringen av utbildningen för de breda folklagen har sina rötter i ett samarbete mellan familjen och kyrkan. Från 1500-talet var det husbonden och prästen som delade ansvaret för undervisningen av barn och husfolk samt kontrollen av deras kunskaper. Kyrkolagen från 1686 stadgade obligatorisk undervisning i läsning och katekesen. Tidigt kom också klockaren att sköta delar av barnundervisningen. I 1842 års folkskolestadga fastslås att undervisningen nu ska bedrivas genom skolgång och att barnen

ska börja senast vid slutet av sitt nionde år. År 1882 bestämdes skolåldern till perioden 7–14 år. Samtidigt slogs fast att skolrådet efter samråd med kyrkostämman kunde skjuta upp skolstarten till barnet fyllde 9 år, om de ansåg det nödvändigt. På vilka grunder ett barn skulle börja skolan vid den ena eller andra tidpunkten angavs emellertid inte.

Med 1962 års skollag knöts skolstarten formellt till skolmognadsbegreppet. Beroende på mognadsgrad kunde barnen börja skolan ett år före eller ett år efter det de fyllt sju år. Redan tidigare hade det aktualiserats genom 1946 års skolkommision och genom olika skolmognadsprov som började användas under 1940-talet. Frits Wigforss konstruerade det första, Kalmarprovet. Linköpingsprovet och Nybörjarprovet är andra exempel (Engström 2003). Enligt skolkommisionen var syftet med skolmognadsprov att avgöra om barnet hade förutsättningar att delta i skolarbetet med framgång, att bestämma i vilken årsklass barnet skulle placeras, att ge vägledning för individanpassning av undervisningen utifrån barnets förkunskaper. På ett överordnat plan handlade det om att genom en tidig differentiering åstadkomma mer homogena undervisningsgrupper (Marklund 1982). Det framhölls också särskilt att icke skolmogna barn skulle delta i förskola. Provet kom att reglera relationen mellan hem och skola, övergången mellan förskola och skola, inträdet till skolan och differentieringen inom skolan.

Någon formell definition av skolmognadsbegreppet finns inte, men av kommissionens skrivningar framgår att de betraktar skolmognad som något som har att göra med fysisk och psykisk förmåga som kan fastställas genom läkarundersökning och ett kortare pedagogiskt prov (Hägglund 1990). Som ett arv från husförhörens tid ansågs den som hade förutsättningar att lära sig att läsa också mogen för skolan. Men närmare bestämmningar fick skolmognadsbegreppet framför allt genom skolmognadsproven och den praktik som utvecklades i anslutning till dessa. Under 1960-talet började proven och deras sorteringsfunktion att ifrågasättas av forskare. I stället för skolmognadsprov rekommenderades en inskolningsperiod. I mitten av 1970-talet föreslog SIA-utredningen (SOU 1974:53) att skolmognadsbegreppet och de 1962 tillkomna särskilda skolmognadsklasserna skulle avskaffas.

De skolmognadsundersökningar som redovisats ger vid handen, att skolmognadsprov som utföres under våren är otillförlitliga mätinstrument. Skolmognadsproven är ej heller förankrade i en utvecklingspsykologisk eller psykologisk-pedagogisk teori. En sådan förankring är också omöjlig, eftersom skolmognadsbegreppet ej kan ges en tillfredsställande definition (SOU 1974:53 s. 452).

Utredningen betonade vidare att i grunden bör diskussionen inte vara en fråga om mognad, utan om metod. Här kan urskiljas en rörelse från en absolut till en relativistisk uppfattning. Det handlar inte längre om huruvida barnet är moget för skolan utan snarare om vad barnet är moget för (Engström 2003). Skolmognadsklasserna avskaffades, men skolmognadsbegreppet behölls även om bruket av skolmognadsprov successivt avtog och den särskiljande praktik som de ingick i ändrade karaktär.

Skolmognadsbegreppets officiella status och skolkommissionens beslut att införa skolmognadsprov skapade ett fält för forskare där de kunde utveckla olika prov. Samtidigt blev det möjligt för politiker att orientera sig åt

det mättekniska hållet i en tid när det beteendevetenskapliga forskarsamhället var upptaget av diskussioner om intelligens- och begåvningsbegrepp, testteknik och -utveckling och experimentell metod (Hägglund 1990 s. 17).

Solveig Hägglunds översikt av studier av skolmognad och skolstart 1955–1985 pekar ut fem delområden som forskarna fokuserat. Skolmognadsprov som prediktions- och prognosinstrument för skolprestation och anpassning är ett av dem. Övriga handlar om åtgärder för skolomogna, samverkan förskola–skola, läsinlärning och kognitiv utveckling samt skolorganisation och inskolning. Samtliga med relevans för forskning om bedömning om än på olika sätt. Som Hägglund påtalar kan vi se hur

frågor om prediktion och prognos av studieframgång och studieförutsättningar har fått stå tillbaka för frågor om elevers upplevelse av skolmiljö och undervisning och organisatoriska aspekter på förskole- och skolmiljö. Vidare har vi sett ett ökat utrymme för barns sociala bakgrund och andra faktorer i deras tidiga utvecklingsvillkor i forskarnas problemformuleringar. Ansatsen i design har förändrats från experimentell och testinriktad till naturliga försök och observation. Skolmognadsproblematik har från att ha varit ett väl avgränsat mätproblem för forskarna, blivit en fråga om socialisationsvillkor i ett vidare perspektiv (Hägglund 1990 s. 53).

Under 1990-talet förskjuts avgörandet om barnens skolstart från skolan till föräldrarna, men alltjämt förväntas grunden vara barnets mognadsnivå (jfr t.ex. Reg. prop. 1990/91:115). Föräldrarna anses vara bättre skickade att avgöra om barnet är moget eller ej. Att de likväl kan behöva stöd i detta är den nu sittande statliga utredningen Flexibel skolstart i grundskolan ett uttryck för som i sina direktiv (Reg. Dir. 2009:98) bland annat ska ”föreslå på vilket sätt föräldrar och elever ska erbjudas skolmognadsprov som ett extra stöd för föräldrars beslut om när deras barn bör börja skolan” (s. 1) och skälet därför är att ”barn mognar i olika takt och att det därför behöver finnas en möj-

lighet att anpassa skolstarten därefter” (s. 2). Vilken effekt skolmognadsprovens återkomst kan få för forskningen om bedömning går inte att avgöra på förhand. Däremot öppnas förnyade möjligheter att utveckla kunskapen om denna form av prov liksom villkoren och konsekvenserna av dem.

3.2 Från husförhör till läxförhör och skrivningar

Under 1600- och 1700-talen lagstodgades de så kallade husförhören. Från 1723 gjordes föräldrar och lärare explicit ansvariga för undervisningen av barn och tjänstefolk. Vid årliga prästvisitationer kontrollerades kunskaperna och resultaten, och prästens omdöme om vars och ens av församlingsmedlemmarnas läsförmåga och religionskunskaper antecknades i kyrkoregistran (jfr Lindmark 2004). Med 1842 års folkskolestadga (SFS 1842:19) överfördes utbildningen till skolan, men både kyrkans och familjens inflytande över undervisningen fanns kvar. Fortsättningsvis låg fokus på läsning och muntliga förhör, skrivandet var förbehållet stadsbor och makthavare. Kyrkans förhörsmetod, ”katekisationen”, med färdiga frågor och svar bestod, det gällde att inte förvanska Guds ord. En särskild typ av muntlig bedömningspraktik etablerades. Kommunikationens funktion i undervisningen var främst att ”informera, påverka och kontrollera” (Adelmann 2002). Prästens husförhör ersattes av skollärares läxförhör och de av skolstyrelsen bestämda offentliga förhören i vilka prästen prövade elevernas kunskaper.

Initialt betonades i folkskolan basfärdigheter och religiöst innehåll samt memorerande av en given text. Den muntliga traditionen återspeglas i undervisningsmetoderna, t.ex. växelundervisningen, men också i bedömningsarbetet: hur eleverna förväntades presentera sitt kunnande och vad som kvalificerade som kunskap. Efter hand och med den i samhället tilltagande skriftspråkligheten förändrades även klassundervisningen (jfr t.ex. Linné 2004) och småningom även formen för prövning. Lärobokens intåg under mitten och slutet av 1800-talet är ett uttryck för denna stora förändring. Andra betydelsefulla faktorer som bidragit till framväxten och utvecklingen av olika bedömningspraktiker är t.ex. de skolämnen som etablerades och utbildningens olika mål, sociala och kognitiva. Sammantaget återspeglas skillnader och likheter mellan olika skolämnen och mål också i vilket innehåll som efterfrågas, hur eleverna förväntas representera sina kunskaper och vad som ligger till grund för bedömning. De olika sätt på vilka kunnigheten kan komma till uttryck får därmed en innehållslig bestämning. Med fokus på skillnader kan några historiska kvardröjande exempel nämnas: essän i svenska, provräkningen i matematik, blindkartan i geografi, 60 meter i idrott och träljustakarna i slöjden. Här bör också

nämnas den pedagogiska textens centrala funktion i skolans verksamhet (Selander 2003), vilket också ända fram till 1991 visat sig i statens intresse för att reglera bruket av läromedel (Johnsson Harrie 2009). Det är alltså flera faktorer som bidragit till vilka bedömningspraktiker som etablerats och upprätthållits, och till hur de förändrats över tid. Tillgången till skriven text, förväntningar på läskunnighet och senare skrivkunnighet bidrar till att förändra undervisningspraktikerna, men också skolämnena och de traditioner som växer fram i relation till dem. Även utvecklingen av läromedel och senare tillgången till ett för lärarna valfritt utbud av läromedel får konsekvenser för undervisningens innehåll och i förlängningen också för vad som blir föremål för bedömning (Parmenius Swärd 2008).

Skolans innehåll är också internt relaterat till interaktionen mellan lärare och elever samt elever emellan och de arbetsformer som används. Med klassundervisningens införande etablerades ett kommunikationsmönster som står sig än i dag. Både internationellt och nationellt bekräftas denna stabilitet. David Tyack och Larry Cuban (1995) framhåller t.ex. att

The basic grammar of schooling, like the shape of classrooms, has remained remarkably stable over the decades. Little has changed in the ways that schools divide time and space, classify students and allocate them to classrooms, splinter knowledge into "subjects", and award grades and "credits" as evidence of learning. (s. 85)

Den lärarcentrerade och läroboksberoende frontalundervisningen med helklassundervisning i kombination med bänkarbete där alla elever arbetar enskilt med samma uppgift är utmärkande för skolans grammatik. Det klassiska kommunikationsmönstret består här av att läraren frågar, eleven svarar och läraren värderar elevens svar, det vill säga exempel på rutinmässigt förekommande bedömningssituationer⁵. Till detta hör också elevernas handuppräkring och lärarens utpekande av vem som får tala (Anward 2003). Det klassiska mönstrets turtagande, initiering, respons och evaluering (IRE), identifierades också tidigt av svenska klassrumsforskare (se t.ex. Lundgren 1972 och Gustafsson 1977). Den så kallade två tredjedelsregeln, med lära-ren som den som tar merparten av talutrymmet i anspråk, har sina rötter i denna tradition. Det gäller också två andra fenomen, dels den så kallade styrgruppen i en klass som fungerar som en kontroll för läraren i fråga om takt, dels det som brukar betecknas som lotsning, det vill säga en metod där lärare stegvis leder elever till rätt svar (jfr Lundgren 1979). Alla dessa

⁵ Paul Black och Dylan Wiliam t.ex. har bedrivit omfattande forsknings- och utvecklingsarbete specifikt relaterat till lärares frågor till elever i klass (t.ex. 2006).

fenomen kan förstås som processer förbundna med det som i dag benämns formativ bedömning i klassrummet.

Även om den angivna beskrivningen av interaktionen i klassrummet har sin motsvarighet i skolans verksamhet kan vi konstatera att beskrivningen passar bättre in på vissa skolämnen än på andra. Dessutom används oftare också andra arbetsformer, i synnerhet från 1970-talet och framåt och särskilt i de lägre årskurserna. Enligt nordiska studier står lärares undervisning framför helklass för cirka 50 procent av tiden. Resten av tiden består av andra interaktionsformer där IRE-modellen inte är dominerande (Sahlström 2008). Både mer kollektiva former där flera elever arbetar tillsammans med en och samma arbetsuppgift och självreglerande former som "eget arbete" och projektarbete förekommer (jfr t.ex. Österlind 2005). De förra används dock betydligt mer sparsamt än de senare, som kontinuerligt har ökat från 1960-talet och framåt (Granström 2003). Nya undervisningspraktiker resulterar i att också bedömningspraktikerna ändrar karaktär. Den ökande andelen eget arbete gör att lärarnas feedback till eleverna i större utsträckning blir skriftlig och att eleverna får del av den vid ett senare tillfälle än då svaret produceras (Eriksson 2010).

Etnografiska och samtalsanalytiskt inriktade studier, som fokuserar elever och elevers handlande i klassrummet, har visat på dels en större variation av lektionstyper, dels andra tolkningar av helklassundervisningen. Det sistnämnda handlar om att det vid sidan av det officiella samtalet pågår ett omfattande samtal elever emellan (Sahlström 2008). Bilden av den passiva mottagande eleven behöver alltså kompletteras med en mer aktiv initiativtagande elev med ett "mikroinflytande" både vid sidan av och i det officiella samtalet. Eleven som kontinuerligt utsatt för omständigheter utanför sin kontroll ifrågasätts också av Eva Forsberg (2000). Olika bilder av interaktionsmönstren i skolan finns också när det gäller könsskillnader. Medan det tidigare framhölls att flickor är tysta i klassrummet och har färre kontakter med lärare än pojkar finns det exempel där flickorna är mer framträdande och utåtriktade, särskilt i högre årskurser och skolformer. Bilden varierar också i relation till ämnen, arbetsformer och situationer liksom elevernas bakgrund. Det finns förskjutningar i könsmönstren, utan att nya interaktionsmönster ersätter de gamla (Öhrn 2002).

I det perspektiv som anlagts här är den vanligaste bedömningsformen i svensk skola det kontinuerligt pågående samtalet i helklassundervisningen. En annan mycket vanlig företeelse i svensk skola är hemarbetet, eller läxan som den mer vardagligt kallas, och det (o)förberedda muntliga eller skriftliga läxförhör som ofta äger rum i anslutning till läxan. Beroende på uppgift kan redovisningen anta även andra former. Det kan vara att visa upp en produkt eller någon specifik färdighet. Notabelt här är dels att dessa former

av bedömning sällan uppmärksammas inom bedömningsforskningen, dels att forskningen om läxor och läxförhör som fenomen i sig är sparsam både nationellt och internationellt (jfr t.ex. Hellsten 1997 och 2000, Lindell 1990, Pettersson & Leo 2005, Österlind 2001). En nyligen utkommen komparativ bok, *International Perspectives on Student Outcomes and Homework*, kan vara ett tecken på att läxor är ett fenomen som håller på att få ökad uppmärksamhet (Deslandes red. 2009).

Andra vanliga former för bedömning som fått betydligt mer uppmärksamhet är prov av skilda slag, och som i skolans värld ibland går under benämningen skrivningar. Något som kan ses som ytterligare ett exempel på skriftspråklighetens genomslag i skolan. Prov, skrivningar, tycks under en längre tid ömsesidigt definiera varandra, åtminstone i många av skolans ämnen. I relation till prov har klassiska testteoretiska frågor rests som hör ihop med psykometrin och dess tillämpning inom skolan, det vill säga edumetrin. Genom förskjutningar i synen på kunskap och lärande, och genom förändringar av skolans verksamhetsformer och bruk av nya tekniker som IKT, växer det fram nya bedömningspraktiker med nya syften vid sidan av de traditionella bedömningsformerna (jfr Roos 2005). Utöver kontrollerande feedback betonas också feed-forward för främjande av lärande (Sadler 1998, Wiggins 1998). En relativt vanlig syn i skolan numera är användningen av olika former av bedömningsmatriser (Lindberg 2009).

Eva Österlind (2005) som särskilt studerat självreglerande arbetsformer, t.ex. "eget arbete", projektarbete och läxor, visar hur elever från skilda sociokulturella bakgrunder förhåller sig på olika sätt till dessa arbetsformer. I relation till bedömning finns här intressanta aspekter som handlar om förskjutningar i ansvar mellan lärare och elever. Eleverna förväntas i större omfattning planera, genomföra-registrera-värdera, sitt arbete. Till sin hjälp har de planerings- eller loggböcker. De kommer in som nya redskap i bedömningen som innefattar både självvärderande och av läraren utvärderande moment, och de brukas för både formativa och summativa syften.

Vår avsikt är inte att fördjupa oss i alla möjliga alternativa bedömningsformer och deras eventuella för- och nackdelar (för exempel se Lindström & Lindberg 2005). Av vikt här är främst att peka på deras förekomst, om än relativt ringa, som en konsekvens av en undervisningspraktik i förändring. Men minst lika viktigt har det varit att visa på några relativt stabila mönster som har betydelse för bedömningen av elevernas prestationer, både sociala och kognitiva. Noteras bör också att målens olika karaktär har betydelse för hur de bedöms. Mål som demokratisk och social kompetens liksom elevs identitetsutveckling visar sig kanske inte bäst eller nödvändigtvis främst i klassrumssituationen, utan lika ofta i andra mindre reglerade situationer som på raster, luncher, håltimmar och aktiviteter utanför skolhuset (jfr t.ex.

Evaldsson 2000). I relation till att skolan ska både fostra och lära ut kunskaper betonas utbildningens moraliska dimension och etik- och maktaspekter lyfts också särskilt fram i lärares bedömningar och betygsättning (jfr Colnerud 1997, Sabbagh 2009). Även bedömningsfrågor i relation till flerspråkighet och kulturell mångfald aktualiseras här (se t.ex. Tornberg red. 2006).

Ytterligare en aspekt behöver beaktas som sammanhänger med interaktionen i klassrummet och skolan. Det gäller bruket av intelligenstest, diagnoser, prov och betyg för att ordna och rangordna barnen utifrån anlag, prestationer och intressen. I förhållande till olika politiska intentioner har över tid olika former för differentiering vuxit fram. I "en skola för alla" uppstår med nödvändighet frågor om hur man ska hantera elevgruppens heterogenitet. Några av lösningarna som förekommit är allmänna och särskilda kurser, nivågruppering, anpassad studiegång, specialundervisning, hjälpklasser och särskilda skolor, liksom särskilda myndigheter för att sköta specialpedagogiska frågor (jfr t.ex. Ahlström, Emanuelsson & Wallin 1986, Wallby, Carlsson & Nyström 2001). Noteras bör också att forskare anlitas flitigt inför 1900-talets stora skolreformer, inte minst i relation till olika former av organisatorisk och pedagogisk differentiering. I dessa sammanhang behandlas bland annat hur elevers individuella differenser, särskilt på begåvningsområdet, kan fastställas genom olika former av bedömningar och hur de kan användas som grund för bland annat nivågruppering och individanpassad undervisning inom klassens eller skolans ram (Dahllöf 1967).

Normalitet och avvikelse är begrepp man kan använda för att beskriva hur så kallade svagbegåvade och talangfulla skiljs från genomsnittseleverna (jfr Isaksson 2009). Något som bland andra Thom Axelsson (2007) visar i sin avhandling *Rätt elev i rätt klass. Skola, begåvning och styrning 1910–1950*. En liknande problematik behandlas av Diana Berthén (2007) i förhållande särskolan. Utan att fastna i beskrivningar om hur utvecklingen sett ut kan vi notera att det finns rörelser mellan såväl segregering och integrering som selektion och elektion (se t.ex. Lund 2006). Ifråga om selektion–elektion gäller det vem eller vilka som, utifrån vilka kunskaper eller erfarenheter, ska avgöra hur eleverna ska delas in i skolformer och grupper. Är det läkarna, psykologerna, lärarna eller eleverna och deras föräldrar (jfr t.ex. Lundahl 2006)? Här aktualiseras hur gränser dras upp, vem som inkluderas eller exkluderas och på vilka grunder, i olika fasta eller mer tillfälliga konstellationer. Mätningar, värderingar, dokumentation, spridning och användning av information om elevers prestationer har varit och är alltjämt avgörande för den skolinterna sorteringsfunktionen. Betygen ingår i detta sammanhang, men har också andra funktioner som vi ska utveckla vidare.

3.3 System för betyg

Redan under husförhörens tid fanns betygen med. Anteckningar från förhören infördes i husförhörlängder med betyg i innanläsning, katekeskunskap, begrepp och förstånd. Såväl betyg i klartext, som anger om något är mer eller mindre bra, som kodade betyg förekommer i längderna och antalet skalsteg varierar. Flyttar vi fram till folkskolan så fick eleverna betyg såväl i ordning och uppförande som i skilda skolämnen. De förra avskaffades när 1969 års läroplan infördes. Det innebär inte att praktiker för att reglera elevers beteenden därmed upphörde. Skolans och lärarens fostrande roll har varit och är ett bekant fenomen (jfr Jackson 1968⁶, Landahl 2006, 2009). Förskjutningar kan iakttas från en yttre disciplinering med skamvrå, kroppsaga och utpekande, till en inre med självdisciplinering där eleven övervakar sig själv genom olika reglerande dokumentations- och selektionsprocesser (jfr Qvarsebo 2006, Samuelsson 2008, Sandell 2007). Om eleverna ska få betyg och hur ofta har återkommande diskuterats och olika skäl har framförts. Betygens kontrollerande, informerande, motiverande och vägledande syfte och deras funktion vid differentieringen inom skolan och vid övergångar mellan skolformer har omväxlande betonats. Vid flera tillfällen har också betygssystemen avlöst varandra.

Med 1897 års folkskolestadga och 1905 års läroverksstadga infördes ett gemensamt och enhetligt betygssystem, det absoluta betygssystemet med bokstavs-beteckningar och sju skalsteg. Detta system ersattes i samband med etableringen av grund- och gymnasieskolan med det grupp- och normrelaterade siffersystemet med fem steg. Detta kom i sin tur att efterträdas av det nuvarande mål- och kunskapsrelaterade betygssystemet som togs i bruk i samband med 1994 års läroplansreform. Ett av skälen för bytet var att den normerande skalan gjorde att betygen sa mer om rangordningen mellan eleverna än vilka kunskaper de hade. I både det absoluta och det kriteriebaserade systemet med Godkänt, Väl Godkänt och Mycket Väl Godkänt finns möjligheten att få Underkänt, dock inte i grundskolan. I det absoluta systemet utgjorde ett underkänt betyg grund för kvarsittning, det vill säga att gå om en årskurs. Det kan ske också i dag, men först efter samråd med elevens föräldrar och en bedömning av elevens samlade studiesituation. Betygssystemet är nu återigen föremål för förändring. Till de centrala förändringarna hör fler skalsteg, nya betygsbeteckningar, betyg från årskurs 6, betygsliknande omdömen för yngre åldrar, utvecklingssamtal med krav på angivande av elevers kunskapsutveckling och nya kursplaner som anger

6 Begreppet den dolda läroplanen aktualiserades i dessa sammanhang.

både centralt innehåll och kriterier för vad eleverna ska kunna. Dessutom vill man återinföra underkänt i grundskolan (jfr Carlgren, Forsberg & Lindberg 2009).

Efterkrigstidens betygssystem vilar på meritokratisk grund. Detta ideal är knutet till idén om att få rätt man på rätt plats⁷ och att tillvarata alla begåvningar.⁸ Det handlar om att eliminera betydelsen av börd och ekonomiska resurser till förmån för skillnader baserade i individernas kognitiva förmåga/begåvning och flit. Extra synligt blir det meritokratiska synsättet i bedömningssystemet och bruket av detta vid urvalsprocesser, när det gäller att skilja ut både de särskilt begåvade och de obegåvade som bör bli föremål för särskilda åtgärder. Det meritokratiska idealet fungerar som en legitimering av de skillnader som uppstår genom de resultat, jämförelser och rangordningar som skolan åstadkommer genom sin bedömningspraktik. Det karakteristiska för både meritokrati och betygssystem kan beskrivas på följande sätt:

Folk skal ha like sjanser til å kunne konkurrere seg til en oppnådd ulikhet. Det er denne særegne kombinasjonen av likhet og konkurranse, ulikhet og legitimitet, som finnes både i karaktersystem og i meritokrati /.../ (Aas 2006 s. 25)

Det gäller med andra ord en kombination av likhets- och konkurrensideal. En kombination som har stått relativt outmanad. Eftersom betygen används som redskap för sortering och har direkt bäring på elevers livschanser är betygens legitimitet av stor vikt. Om nu grunden för betygssystemet inte ifrågasätts i någon större omfattning är det vanligare med frågor om betygen är satta på ett rättvist sätt och om de har lika värde (jfr Wikström 2005a).⁹ Kontroller av hur nationella prov rättas och jämförelser mellan nationella prov och betyg har påvisat vissa systematiska skillnader i lärares bedömningar. Frågor har ställts om såväl lärarnas kompetens att bedöma elevers kunskaper och utveckling som lärarnas förmåga att konstruera prov, anordna situationer och uppgifter för bedömning och sätta välgrundade betyg. Likaså har den ringa förekomsten av dessa moment i lärarutbildningen påtalats och förändringar genomförts.

⁷ Noteras kan i detta sammanhang att urvals- och rekryteringsprocesser både i det militära och i företag och förvaltningar utgjort en bas för kunskapsutvecklingen om bedömning, se t.ex. International Handbook of Selection and Assessment (Anderson & Herriot red. 1997).

⁸ Jfr Torsten Husén och Kjell Härnqvists (2000) studier om begåvningsreserven.

⁹ Att grunderna inte ifrågasätts innebär inte ett likartat förhållningssätt. Såväl lärare och administratörer som forskare har gjort olika uttolkningar både inom respektive kategori och sinsemellan (Selghed 2004, Lindberg 2009).

Förändringarna av betygssystemet har varit relaterade till förskjutningar i synen på kunskap och lärande i läro- och kursplaner. Ett avgörande brott med tidigare traditioner är 1994 års läroplaner för grund- och gymnasieskolan som bland annat kan beskrivas som ett skifte från innehåll som pensum/stoff till förmåga/kompetens. En förskjutning med implikationer för både undervisningens innehåll och bedömningens former. Både brottet med skolans tradition och det icke givna innehållet ökar komplexiteten i bedömningarna och ställer därmed lärarna inför delvis nya uppgifter. Några av de frågor som lärarna nu har att hantera är hur kriterier kan tolkas, vad bedömningar kan avse, hur de kan genomföras och hur kunnskap kan urskiljas inom olika ämnen och områden (jfr Carlgren, Forsberg & Lindberg 2009).

Den omfattande stödverksamhet som centrala myndigheter utvecklat i form av projekt, publikationer, databaser, provbanker, diagnostiska uppgifter och nationella prov kan ses i ljuset av diskussionerna om betygens legitimitet och lärares kompetens. Detsamma gäller införandet av fler och tidigare nationella prov och fler kontroller i form av självvärderingar och den kommunala och nationella nivåns uppföljningar, utvärderingar och inspektioner. Ytterligare andra exempel är kraven på ökad tydlighet och konkretion i styrdokument, som bland annat ligger bakom de kursplaner som nu är under utveckling. Konsekvenserna av detta för både lärarnas yrkesutövning, verksamheten och elevers prestationer har uppmärksamats och diskuterats bland annat i termer av både deprofessionalisering och kontraproduktivitet (jfr Forsberg & Wallin red. 2006).

Andra frågor som aktualiserats är t.ex. vad betygen är ett mått på. Man har konstaterat att de inte bara mäter kunskap utan också har inslag av motivation, intresse och flit. Det är också dokumenterat att socioekonomisk och kulturell bakgrund, kön, etnicitet, föräldrars utbildning, egna och andras förväntningar inverkar på betygen och att dessa i sin tur har betydelse för val av yrke, ekonomi och hälsa (se t.ex. Skolverket 2009, Öhrn 2002). Betyg har ett starkt signalvärde och en förmåga att styra både elevers beteende, undervisningen och utbildningens resultat, ibland på bekostnad av mål i läro- och kursplaner, kognitiva och sociala (Bernstein 1971/1980).

Ytterligare ett uttryck för betygens signalvärde är den så kallade treämnesskolan som lägger stor tyngd på svenska, matematik, engelska, eftersom godkänt i dessa ämnen reglerar inträdet till gymnasieskolan. Fast här är det relationen mellan betyg och behörighetsregler som griper nedåt i systemet. Att styra med svansen är ett uttryck för när betygssystemet tar över på bekostnad av läroplaner och kursplaner och den syn på kunskap och lärande och mål som finns uttryckta i dessa. Med alltför många obesvarade frågor eller otillfredsställande svar på frågor om betygssystem uppstår legitimitetssvikt, något som i sin förlängning kan medföra att systemet eller delar av

det utsätts för ett förändringstryck. Det här gäller inte bara betygen, utan bedömningssystemen i sin helhet. Det är bland annat mot denna bakgrund 1900-talets skiften av betygssystem och den nu förestående förändringen kan förstås. Nämnas kan också att en utredning om möjligheterna för en elev att få sina betyg omprövade pågår (Regeringens Dir. 2009:102).

Förr fick elever betyg både tidigare och oftare. Men i samband med att betygens urvalsfunktion bedömdes minska och deras värde som pedagogiskt redskap ifrågasattes, begränsades och senarelades tiden för betyg. I stället växte andra former för dokumentation fram. I kvartssamtalen, och deras efterföljare utvecklingsamtalen, och vid elevvårdskonferenser används numera individuella utvecklingsplaner och åtgärdsprogram. I dessa ryms både lägesbeskrivningar och planerings- och åtgärdsförslag. Nya former för bruk och spridning av bedömningsresultat kan förstås som en utveckling av nya teknologier som påverkar både innehåll, relationerna mellan dem som medverkar och formen för kommunikation, nya bedömningspraktiker etableras (se t.ex. Andreasson 2007, Asp-Onsjö 2006, Granath 2008, Hofvendahl 2006).

Ovan berördes hur resultat från test, prov och betyg brukas för att differentiera elever internt. Betyg används även vid övergångar mellan olika skolformer. Men det har inte alltid varit så. Tidigare reglerades t.ex. övergången mellan folkskola och realskola vid en inträdesprövning som genomfördes av den mottagande skolan, med viss information från folkskollärarna. På förslag av 1939 års betygssakkunniga, med Frits Wigforss som en av ledamöterna, fick folkskollärarna för första gången standardprov i modersmål och räkning till sitt förfogande. Efter flera års arbete, inklusive normering, började proven användas 1943/44. De användes fram till 1997 när de ersattes av de nationella ämnesproven. Skillnaden mellan dem var att de förra syftade till att göra betyg satta av olika lärare jämförbara, medan de senare primärt ska stödja läraren vid betygsättningen. Förändringen hade sin grund i föreställningen att grundskolans betyg hade eller skulle komma att få en allt mindre betydelse för urvalet till gymnasieskolan (se t.ex. Andersson 1991, Ljung 2000, Lundahl 2009). Vid sidan av detta är avsikten också att de nationella proven ska ge underlag för en analys dels av elevers starka och svaga sidor, dels av i vilken utsträckning kunskapsmålen nås på olika nivåer i utbildningssystemet. I det nationella provsystemet ingår också diagnostiska material för grundskolan, och provmaterial i Skolverkets provbank.

Vid övergången mellan grund- och gymnasieskolan reglerar betygen i dag vilka elever som ska få tillträde till de nationella programmen och vilka som ska hänvisas till de individuella programmen. Inom ramen för den nya gymnasieskolan 2011 kan ytterligare former för differentiering urskiljas: ett extra tionde år för att klara behörighetskraven till nationella programmen och tre olika studievägar: lärlingsprogram, yrkesprogram och högskoleför-

beredande program. De två första leder till yrkesexamen och den senare till högskoleförberedande examen. Kraven på yrkesexamen avser 90 procent G på kurserna i ett nationellt program, lägst G i svenska, engelska, matematik och gymnasiearbetet samt G i de viktigaste yrkeskurserna. För behörighet till högskolan krävs komplettering i svenska och engelska.

Betygens betydelse för urval har förvisso förändrats, men inte eliminerats. Regleringar av behörigheten till gymnasieskolan och den högre utbildningen och ökade möjligheter att välja grund- och gymnasieskola påverkar betygens värde. Förändrade regler för tillträde till gymnasieskolan 2011, inklusive återinförande av examen (högskole- respektive yrkesexamen), och de nya reglerna för tillträde till högskolan innebär samtliga en ökad och dubbel differentiering med utgångspunkt i vilka ämnen eleverna har betyg i och vilka betyg de har. Sammantagna innebär de också både skärpta regler och ökade krav vid både entré och exit (Forsberg 2008). Något som påverkar elevers karriärvägar och bildninggångar om än på skilda sätt i relation till kön, socioekonomisk och etnisk bakgrund (jfr t.ex. Lindblad 2008, Lundahl L. 2009). Betyg och examen återförenas här drygt 40 år efter studentexamens avskaffande och införandet av dess efterträdare slutbetyget (Sjöstedt & Thorén red. 1963).

Vid sidan av betygens här nämnda funktioner för information, kontroll och (s)elektion har betyg också kommit att alltmer användas som mått på skolans resultat och utbildningens utfall. Med betygsdata som grund besvaras frågor om skolans kvalitet och utvecklingstrender. Här uppmärksammar vi några sammanhang som fokuseringen på skolans resultat ingår i.

3.4 Sammanhanget för skolan och bedömningen

Som vi konstaterade inledningsvis kan skolan förstås i relation till både interna och externa processer. Här uppmärksammar vi framför allt några samtida och externa sammanhang och deras betydelse för skolan och då särskilt bedömningen. Med början i en kort tillbakablick på styrning och ledning av folkskolan förflyttas sen fokus till två för skolan och bedömningen centrala sammanhang: skolans omstrukturering och globaliseringen i ljuset av internationella jämförande kunskapsmätningar.

Under folkskolans tid kunde föräldrar om de så önskade välja att antingen undervisa sina barn hemma eller placera dem i ett offentligt läroverk alternativt sätta dem i en enskild skola. I de fall där eleverna undervisades hemma skulle skolläraren pröva om de fick en undervisning som i noggrannhet och omfattning någorlunda motsvarade den som eleverna i folkskolan fick. Styrelsen kunde utfärda varningar till föräldrar som inte såg till att barnen fick

en skolgång i linje med stadgan. Om inte det gav resultat kunde föräldrarna fråntas barnen.

Domkapitlet, kyrkoherden och skolstyrelsen skulle ha uppsikt över ledning och utveckling av verksamheten och se till att folkskolorna användes flitigt. Förteckningar skulle också föras över vilka barn som gick i skolan och varje år skulle skolstyrelsen sända in en berättelse till domkapitlet om förhållandena i skolorna. Domkapitlet skulle i sin tur vart tredje år skicka ett utlåtande till Kunglig Maj:t om folkundervisningens tillstånd inom stiftet. Den styrmodell som framträder i stadgan har drag av både centralisering och decentralisering. Mål och innehåll liksom formerna för kontroll var centralt formulerade. Församlingarna och socknarna i allmänhet förväntades däremot stå för finanserna och ansvara för verksamhetens genomförande och resultat samt avrapportering. Redan från början var folkskolan alltså en gemensam samhällelig angelägenhet och utbildningen var, liksom i dag, ett uppdrag som delades mellan flera aktörer. På plats var också ett system för råd, prövningar, kontroll och inspektion samt sanktioner om resultaten betraktades som oacceptabla.

3.4.1 Skolans omstrukturering

Alltsedan folkskolan infördes och i synnerhet sedan grundskolan etablerades har styrningen av utbildningen diskuterats. Argumenten har gällt om och i vilken omfattning utbildningen ska vara offentligt reglerad och producerad, vad som ska styras centralt och vad som ska överlåtas åt den kommunala och lokala utbildningsledningen. Man har också diskuterat hur ansvar och befogenheter för olika uppgifter ska fördelas mellan å ena sidan politik och administration och å den andra de professionella i skolan. Även relationen mellan hem och skola och mellan skolan och arbetslivet har diskuterats återkommande. Diskussionerna har handlat om vad styrningen ska avse, vilka som ska involveras i styrningen och om tyngdpunkten ska ligga på input och processer och/eller output.

Under de två senaste decennierna har systemet för styrningen av utbildningen genomgått omfattande förändringar. Förskjutningarna från centralisering över decentralisering och mot recentralisering kan ses som en rörelse där ansvar, befogenheter och uppgifter rör sig fram och åter mellan å ena sidan centrum och inter/nationell policy och administration och å andra sidan den kommunala, privata och lokala skolan och de professionella. Styrningens tidigare fokusering på mål, organisation och processer har kompletterats med ett alltmer utbyggt kontrollsystem. Uppföljning, utvärdering och inspektion är nu i förgrunden och mängden kontrollaktiviteter på alla nivåer i systemet har ökat.

Tre övergripande motiv har angetts för omstruktureringen: högre effektivitet, utvidgad professionalism och ökad demokrati. Att involvera många i formande och realisering av policy och praktik kan beskrivas i relation till vad som inom forskningen diskuteras som en förskjutning från "governing" till "governance". Förändringarna avsåg bland annat ökat inflytande för elever och deras föräldrar, möjligheten att välja skola, ökat ansvar för lärare genom fler och mer komplexa arbetsuppgifter och ökat ansvar för den lokala nivån vad gäller utformningen av skolan och dess resultat. Alla nivåer får ansvar för uppföljning och utvärdering av resultat och krav ställs på olika former av dokumentation som kvalitetsredovisningar, nationella uppföljningar, utvärderingar och så småningom inspektionsrapporter (jfr t.ex. Carlgren m.fl. 2006, Forsberg & Lundgren 2004, Nyttell 2006).

Av särskilt intresse här är att omstruktureringen handlar om utkrävande av an/svar genom uppföljning, dokumentation, synliggörande och spridning av resultat på prov och betyg liksom jämförelser mellan skolor. Till bilden hör också marknadiseringen och konkurrensutsättningen av skolan (Lundahl 2002). Betygen blir den hårdvaluta som skolor och kommuner kan använda som konkurrensmedel. I rankningar av skilda slag kan de ge en bild av hur skolorna "ligger till". De som deltar i denna "second-hand"-business med data från elevers prestationer är många. Som exempel kan nämnas media av skilda slag, politiken, utbildningsdepartementet och centrala och lokala myndigheter, Sveriges kommuner och landsting och enskilda skolor (kommunala och fristående), fackliga organisationer, föräldraföreningar och representanter för det privata näringslivet. Till detta kan fogas att forskare och utredare i olika discipliner deltar både i nämnda aktiviteter och i forskningsprojekt i vilka resultat på prov och betyg används och ibland reanalyseras. Perioden utmärks av en strid ström av publikationer, inte minst från de statliga skolmyndigheterna. Skolverket har publicerat flera översikter över vad skolans resultat beror på. Sveriges kommuner och landsting har nyligen släppt öppna jämförelser för gymnasieskolan och utbildningsdepartementet har gett ut en översikt över resultat från andra kunskapsöversikter.

Till denna bild hör också förändringarna av den statliga förvaltningen, med etablerandet av Skolverket som en kunskapsorganisation, Myndigheten för skolutvecklings uppgång och fall samt den 2008 etablerade Statens skolinspektion. Tillsynen av skolan är som framgått inget nytt. Initialt var det en uppgift för kyrkan. Från 1861 till 1958 hade den statliga Folkskoleinspektionen och efter den Folkskoleöverstyrelsen ansvar för att inspektera och rapportera om läget i utbildningsväsendet. Den i sin tur avlöstes av länskolnämnderna som därefter ersattes av Skolverket. Formellt upphörde inspektörsverksamheten när länskolnämnderna lades ner. Men i praktiken

hade inspektörernas uppdrag redan tidigare förskjutits i riktning från kontrollant till vägledare med uppgift att främja skolans utveckling. Under 00-talet ökade dock kravet på inspektion och så småningom kom den från Skolverket fristående inspektionen. En annan manifestering av separationen av bedömningspraktiker på systemnivå är den statliga utredningen om utvärdering och utbildning. Den har nyligen lämnat sitt betänkande Att nå ut och att nå ända fram – hur policyinriktad utvärdering och forskningsresultat inom utbildningsområdet ska tillgodoses (SOU 2009:94). Utredningen föreslår utöver befintliga strukturer att ett Råd för utbildningsutvärdering och analys inrättas, som ska spänna över hela utbildningsområdet och över flera discipliner.

3.4.2 Globaliseringen och internationella jämförande kunskapsmätningar

Även det ökade deltagandet i och bruket av resultaten från internationella jämförande kunskapsmätningar som PISA och TIMSS ingår i bedömningarna som är externa i förhållande till skolan vad gäller initiativ och bruk. Globaliseringen är förvisso inget nytt fenomen inom utbildningen, men policy och praktik formas alltmer i skärningspunkten mellan internationella och nationella arenor. Inom både EU och OECD har det vuxit fram diskurser om kunskapens betydelse för samhälle och ekonomi. Nödvändigheten av livslångt lärande och utvecklandet av nyckelkompetenser betonas särskilt. I anslutning till detta har man tagit fram program för att följa och utveckla de nationella utbildningssystemen. De internationella mätningarna ingår som ett led i dessa program (jfr Forsberg 2009).

I forskningslitteraturen diskuteras utbytet mellan det nationella och det internationella fältet som en fråga om transfer och mer precist som "borrowing" och "lending", det vill säga nationer lånar in och lånar ut idéer om hur utbildningen kan utformas (Ochs & Phillips 2004). Omfattningen av denna påverkan diskuteras, interaktionen är relativt komplex och påverkan är inte enkelriktad. Att det internationella samspelar med och villkorar den nationella politiken står dock klart (jfr t.ex. Daniel Pettersson 2008). Internationella test används som redskap i policydiskussioner för att antingen konfirmera eller reformera utbildningen. Det finns flera tecken på att de internationella mätningarna ökar i betydelse. En mångfald konferenser har anordnats som redovisar resultat eller förordar att data används för reanalyser. Medias rapportering om resultaten är relativt omfattande och testen inordnas i det nationella systemet för uppföljning och utvärdering av resultat. Dessutom slår kunskapssynen med grund i kompetens- och literacybegreppen genom i de nu pågående revideringarna av kursplaner och betygskriterier.

Prov- och betygsdata från olika perioder och upprättande av longitudinella databaser har gjort det möjligt att beskriva trender, förändringar över tid (jfr t.ex. Erikson m.fl. 1999). Jämförelser mellan nationer, mellan kommuner och skolor och över tid används på olika sätt och med skilda motiv. Man jämför för att fastställa skillnader mellan olika kategorier av elever, för att ge ris eller ros till ansvariga för utbildningen, för att skildra utvecklingen av elevernas prestationer i synnerhet i vissa ämnen, som incitament för att dis/legitimera reformering eller status quo av skolan.

Den ökande användningen av kontroll- och utvärderingsinstrument på alla nivåer i systemet bidrar med preciseringar av vad som ska förstås som viktiga kunskaper och hur de kan fastställas. I synnerhet gäller detta test, nationella prov, internationella jämförande kunskapsmätningar och preciserade kriterier för bedömning och betygsättning. Även den ökande kommunikationen om elevers prestationer och skolors resultat bidrar med konkretiseringar, tydliggöranden och synliggöranden av vad skolan och eleverna förväntas ägna sig åt. Uttrycket "the medium is the message" pekar på betydelsen av formen för kommunikationen. Den både möjliggör och begränsar vilket budskap som kan kommuniceras (jfr Forsberg & Lundahl 2006, Krantz 2009). Det här är inget enkelt samband. En reduktion av komplexitet är många gånger förutsättningen för att något över huvud taget ska bli sagt. Mot denna bakgrund kan man t.ex. förstå en del av diskussionen om betyg, betygsliknande omdömen och utvecklingssamtal. Där kan företrädare för de olika formerna för kommunikation om elevernas prestationer samfällt hävda att mer blir sagt med den form de förespråkar.

3.5 Den svenska skolans bedömningskulturer

Med den ena foten i skolans historia och den andra i dagens skola och det sammanhang den är inbäddad i ska vi här avslutningsvis sammanfatta några huvuddrag i den svenska skolans bedömningskultur. En del fenomen är beständiga, några är nya, andra kommer och går och en del försvinner, kanske för gott. En del är sällsynta, medan andra är så vanliga och självklara att de nästan är svåra att få syn på. Med siktet inställt på det närliggande är det lätt att få för sig att bedömningskulturen är stadd i stark förändring. Sett över en längre tid och med ett fågelperspektiv framstår vissa element i bedömningskulturen som mer manifesterade. Bedömningar har återkommande använts vid såväl entrén till skolan, i skolans interna processer och vid exit, antingen vid övergången till högre skolformer eller till arbetslivet. Olika aktörer, representerade i förhållande till skolans interna och externa verksamheter, har också använt bedömningar för olika syften.

Om vi börjar vid skolporten kan vi konstatera både likheter och skillnader. En avgörande skillnad är att merparten av barnen går i förskolan, så inträdet i skolan handlar i dag mer sällan om en övergång från hemmet till skolan. Det innebär att professionella är involverade i både den avlämnande och den mottagande verksamheten. Vi har en professionell grund för att bedöma om en elev ska börja skolan eller ej, även om föräldrarna också är med i bilden. Samtidigt finns signaler på att skolmognadsprov kan bli aktuella igen, men nu som ett stöd för föräldrar att avgöra om deras barn är skolmoget. Något som i sig kan förstås som en rörelse från den barnmogna skolan som förväntas anpassa sig till barnen tillbaka till en skola som barnen ska kunna klara av. Idén om ett skolmoget barn och en mer absolut förståelse av barnets kunskande förefaller med andra ord inte ha lämnat skolans värld. Detsamma gäller föreställningen om att barnets mognad kan fastställas genom prov. Det finns med andra ord tecken på att den tidigare förskjutningen från mätproblem till socialisationsvillkor är på väg att kastas om, sett utifrån inträdet till skolan. I dag är det också möjligt för elever och deras föräldrar att välja skola och här spelar elevers prestationer roll på flera sätt. För det första använder skolor elevernas resultat i sin marknadsföring för att attrahera elever och helst "rätt" elever. För det andra kan en enskild elevs betyg i senare årskurser avgöra var eleven kan få sin utbildning. Det gäller särskilt utbildningar som riktar sig till talangfulla elever och utbildningar på gymnasienivå.

I skolans verksamhet har bedömningar under lång tid använts i såväl formativa och summativa syften som i mer eller mindre formella sammanhang. På ett övergripande plan gäller att formen för bedömning har rört sig från det muntliga mot det skriftliga. Under hela 1900-talet finns en stabilitet i skolans bedömningspraktik, samtidigt som omfattande förändringar genomförs. Det muntliga eller skriftliga läxförhöret, provet och skrivningen är vanliga former för bedömning. Hit hör också bedömningar kopplade till fenomen som handuppräckning, styrgrupp och lotsning. Detsamma gäller lärares kontinuerliga kommunikation med eleverna i enlighet med det väl etablerade mönstret fråga-svar-respons. Samtidigt finns det också skillnader, dels mellan olika skolämnen, dels som en konsekvens av en förändrad undervisningspraktik med bruk också av ny informationsteknologi. Det senare pekar mot en aktivare elev med större ansvar för både planering och genomförande av det egna arbetet inklusive självvärdering av sin insats och prestation. I detta sammanhang har också nya former för dokumentation av elevens prestationer etablerats, t.ex. portfolio och loggböcker. Resultat på diagnoser, test och prov har också en lång historia som redskap för att ordna elever i skilda grupper där de får ta del av delvis olika kurser och innehåll. Bedömningar för att skilja det normala från det avvikande är inget nytt, även om tillvägagångssätten kan variera över tid.

Även om 1900-talet såg flera olika betygsystem komma och gå finns vissa stabila drag. Hit hör bland annat den meritokratiska grunden för betyg. Gemensamt är också förekomsten av system för betyg samt betyg i både grundläggande och högre skolformer. Om än med olika emfas har också betygens informerande, motiverande, vägledande och sorterande funktion lyfts fram. Men skillnader finns i en rad olika avseenden, som när elever får betyg, hur ofta de får betyg, på vad betyg ges etcetera. Med jämna mellanrum diskuteras t.ex. om betyg ska sättas i ordning och uppförande eller ej. Att gå om ett år är i dag mindre vanligt och inte direkt relaterat till formella bedömningar, även om det föreslagna tionde grundskolåret för dem som inte kvalificerar sig för nationella program i gymnasiet kan tolkas i denna riktning.¹⁰

Omstruktureringar av läro- och kursplaner och betygsystem under 1990-talet medförde vissa avgörande förändringar i förutsättningarna för lärarnas bedömningar. Skolans kunskapsinnehåll blev här en öppen fråga. Med skiftet från relativa till mål- och kunskapsrelaterade betyg ökade kraven på både lärare och elever att kunna kommunicera om och med bedömningar. En annan samtida process är en ökad fokusering på den enskilda eleven, som bland annat visar sig i nya former för dokumentation och kommunikation som åtgärdsprogram, individuella utvecklingsplaner och utvecklingssamtal. Tidigt kom standardisering av bedömningar att bli en central fråga kopplad till bedömning, först genom centrala prov och därefter med nationella prov. Omfattningen av dessa, vilka ämnen som ingår, hur ofta de ska genomföras och vilka syften de främst förväntas fullgöra har förvisso skiftat över tid. Men i slutändan har de ansetts nödvändiga, inte minst i ett legitimerings-syfte. Detsamma gällde under lång tid utfärdande av examen som ett bevis på genomgången sekundärutbildning med godkänt resultat, något som i dag är på väg tillbaka.

På organisations- och systemnivå har vi kunnat iaktta förändringar som är relativt drastiska, men också här finns likheter mellan då och nu. Bedömningar och betyg används dels för att ordna elever i skilda grupper, dels för att informera föräldrar om elevers prestationer, dels för att avläsa utbildningssystemets standard och utveckling, dels som redskap för urval och sortering för vidare studier och positioner på arbetsmarknaden. På denna nivå är de som använder bedömningarna många och de representerar flera olika verksamheter. De mål- och kunskapsbaserade betygen med en gräns för godkänt har bidragit till att synliggöra samhällets förväntningar

¹⁰ I en rapport om åldersblandade grupper redovisas att lärare redan i årskurs 2 ger uttryck för att elever som börjat skolan som sexåringar bör gå tio år i skolan (Holmberg 2002).

och krav på skolan och eleverna. Elevers kunskapsbrister som de visar sig i nationella prov och betyg har därmed fått ökad uppmärksamhet. Den avgörande skillnaden jämfört med tidigare gäller framför allt att man använder elevers prestationer som resultatmått för bedömning av organisationen och systemets kvalitet. Fenomenet är inte nytt, men det är omfattningen och systematiken liksom mångfalden involverade intressenter, inte minst på den internationella arenan. Förändringarna sammanhänger med en legitimitetssvikt i systemet, som i sin tur bidrog till en förändring av styrningen som innebar en förskjutning från input och processer till output. Något som i sin tur förändrat den centrala förvaltningens karaktär och uppdrag. När utfall står i förgrunden, som nu, ökar trycket på kontroll, inspektioner och bedömningar i hela utbildningssystemet. I samband med detta har också samarbetet mellan den politiska, den administrativa och den praktiskt pedagogiska nivån omförhandlats. Uttryck för detta finns inom både den evidensbaserade rörelsen och den omfattande stödverksamheten som utvecklats på central nivå. Aktiviteter som också förbinder forskningen med både den statliga förvaltningen och den pedagogiska praktiken. Här möts frågor om makt, etik och legitimitet. Frågor som inte är nya i förhållande till skolan och dess intressenter.

Skolan är, som påpekats i kapitel 1, en av samhällets grundläggande institutioner med många intressenter. Den är mycket omfattande och har ett komplext ut/bildningsuppdrag som delvis är motsägelsefullt sett utifrån skolans målskrivningar och skolans funktion. Som politikområde tar skolan dessutom en stor del av skattemedlen i anspråk. Det här tillsammans med skolans betydelse för elevers bildningsgångar och livschanser gör beskrivningar och analyser av utbildningens utfall och kommunikationen om detta till ett centralt område för studier med bedömning som objekt.

4 BEDÖMNINGSFORSKNING – NÅGRA UTBLICKAR

Här gör vi några korta utblickar på internationell forskning om bedömning i syfte att skapa en utgångspunkt för den i kapitel 10 avslutande diskussionen om svensk forskning om bedömning. Vi presenterar internationell forskning som den framträder utifrån nedslag i ett urval internationella forskningshandböcker och några för utbildningsområdet centrala internationella organisationer. Avslutningsvis sammanfattar vi huvuddragen i den internationella bedömningsforskningen.

4.1 Bedömningsforskning i internationella forskningshandböcker

Bedömning utgör en aspekt av flera skilda pedagogiskt didaktiska fenomen. Därför har vi inkluderat internationella handböcker från olika områden: undervisning, lärare, lärarutbildning och läroplansteoretiska frågor samt utbildningens utvärdering, policy, utveckling och sociologi. Dessutom tittar vi på några ämnesinnehållsliga områden som språk/engelska, estetiska skolämnen och informationsteknologi.

4.1.1 Undervisning, lärare, lärarutbildning och läroplaner

En första bild av forskning om bedömning kan ges utifrån den amerikanska forskningsorganisationen AERA:s Handbook of Research on Teaching, en handbok som har ambitionen att belysa samtida forskningstrender, perspektiv, frågor och metoder. Genom att den utkommit i fyra upplagor kan de olika upplagorna bidra med att belysa förändringar i forskningen även i relation till bedömning. Trots att organisationen är amerikansk inbjuds internationellt erkända forskare att bidra med artiklar. Tabell 4.1 visar de fyra utgåvorna, redaktörerna för respektive utgåva och året för publicering.

TABELL 4.1 AERAs Handbook of Research on Teaching.

Titel	Utgåva	Redaktör	Årtal
Handbook of Research on Teaching	1. ed.	Gage	1963
Second Handbook of Research on Teaching	2. ed.	Travers	1973
Handbook of Research on Teaching	3. ed.	Wittrock	1986
Handbook of Research on Teaching	4. ed.	Richardson	2001

Den första utgåvan från 1963 består av fyra delar och är på totalt 1 172 sidor. Fyra kapitel med direkt relevans för bedömning återfinns i del II *Methodologies in Research on Teaching*. I kapitlen behandlas mättekniska frågor i relation till elevers deltagande, beteende och prestationer samt förändringar av dessa. Observationer, metoder för skattningar och test och mätningar av både kognitiva och ickekognitiva variabler beaktas. I förgrunden står hur forskare kan mäta och bedöma elevers prestationer och hur forskarna kan utveckla sina instrument för bedömning.

I 1973 års utgåva som består av 1 322 sidor finns några kapitel av intresse i två av de fyra delarna: *Methods and Techniques of Research and Development and Research on Special Problems of Teaching*. Precis som i den första utgåvan behandlas mätfrågor, men nu ingår även kvalitativa tekniker för insamling av data. Dessutom diskuteras lärarens effektivitet utifrån kriterier som mäter elevers lärande. Under speciella problem redovisas också forskning om undervisning för så kallade mentalt efterblivna, emotionellt störda och talangfulla elever. Här handlar det om definitioner och att inordna elever i grupper, men bedömningar som grund för differentieringen betonas inte uttryckligen.

Den tredje utgåvan från 1986 på 987 sidor består av fem delar. Av dessa är ingen uttryckligen inriktad på bedömning, men en del, *Adapting teaching to differences among learners*, består av fyra kapitel. Ett med samma namn, och dessutom tre som handlar om talangfulla elever, tvåspråkiga elever och specialundervisning. I en annan del finns också ett kapitel som beaktar skolans effekter i ett samhällsligt perspektiv. I kapitlet *Teacher Behavior and Student Achievement* på knappt 50 sidor knyts elevers lärare mer direkt till undervisningen och lärarens strategier och handlingar. Därutöver finns också ett kapitel som tar upp Measurement of Teaching, som kan ses som en efterträdare till test och prov.

Den senaste utgåvan från 2001 innehåller betydligt fler delar (åtta) än föregångarna. Den består av 1 278 sidor. Ingen del handlar enbart om

bedömning. I *Methodology* är t.ex. test och prov inte föremål för ett eget kapitel. I den åttonde delen Instruction finns det enda kapitlet, *The Role of Classroom Assessment in Teaching and Learning*, som entydigt signalerar att det handlar om bedömning. Här är det bedömning för att främja lärande som är centralt och bedömning diskuteras i relation till olika teorier för lärande.

Två utgåvor av *International Encyclopedia of Teaching and Teacher Education* från 1987 (Dunkin ed.) och 1995 (Anderson ed.) kan komplettera bilden något. Även den här handboken representerar en internationell bredd av forskare och forskning av relevans för lärare, undervisning och lärarutbildning, vilket är skälet till att den tagits med. De båda utgåvorna skiljer sig från varandra genom att den första, på 878 sidor, är fokuserad på forskningstraditioner om bedömning som använder strukturerade observationer och frågeformulär liksom traditionella papper-penna-mätningar av elevers förmågor och prestationer. I den andra utgåvan på 684 sidor har dessa metoder kompletterats, ibland nästan helt ersatts av etnografiska fältstudier, semistrukturerade intervjuer och narrativ. Bedömning har i den senare utgåvan också en mer framträdande roll. Här finns ett avsnitt på cirka 50 sidor som särskilt behandlar bedömning. I *Assessing and Evaluating* redovisas bland annat forskning om lärares frågor till och värderingar av elevers svar, diagnoser, test och mätningar, bedömningar i klassrummet, dokumentation, etik och betyg. Det framhålls dessutom att bedömningsmetoder skiljer sig åt inom olika områden och i förhållanden till skilda syften. Därför inkluderas frågor om bedömning även i avsnitt som handlar om t.ex. minneskunskaper, problemlösning, kreativitet, kritiskt läsande och metakognitiva aspekter.

Handbook of Research on Curriculum, publicerad 1992 (Jackson ed.), har valts för att den har fokus på läroplansforskning, ett område som är internt relaterat till bedömning. Handboken består av drygt 1 000 sidor. Även i denna ingår kapitel om bedömning. I ett kapitel om utvärdering av läroplaner och bedömning definieras inledningsvis skillnaden mellan assessment och evaluation. Assessment ses som bedömningar som tar fasta på enskilda elevers prestationer medan evaluation knyts till utvärderingen av en specifik läroplan eller ett program. Samtidigt framhålls hur de alltmer sammanfaller när aggregerade data för individer används tillsammans med andra indikatorer för utvärdering av utbildningssystem. Utvärderingens nära koppling till teströrelsen betonas. Sex olika motiv för bruk av test lyfts fram. Test har omväxlande använts för att fatta beslut om förändringar av klassificering av elever, skolorganisationen, utbildningsprogram, tid för olika skolämnen, metoder för undervisning och strategier för kontroll. Testen har också använts för att utvärdera hur effektiva lärarna är. Under 1960- och

1970-talen växer ett intresse för mål, input, processer och både avsiktliga och icke avsiktliga utfall fram och därmed sker också en förskjutning mot kvalitativa ansatser. Med 1980-talet vänds uppmärksamheten åter mot output och indikatorer och standardiserade kvantitativa metoder hamnar på nytt i förgrunden.

I bedömningar av enskilda elever identifieras i ett längre historiskt perspektiv en rörelse från praktiska och muntliga examinationer, med fokus på kvalitativ ranking, mot uppsatser och skriftliga prov som betonar kvantitativa resultat som provpoäng och senare betyg. Även betydelsen av ny teknik lyfts fram. I ett läroplansteoretiskt perspektiv skiljer man på den avsedda och den av elever erfarna och uppnådda läroplanen. Bedömningar av enskilda elevers prestationer ses som en intern aspekt av läroplanen och en att lärarnas viktigaste och mest komplexa uppgifter. Det är emellertid en uppgift som författarna menar behöver uppmärksammas i större omfattning. Inte minst med tanke på den maktutövning som bedömningen är ett uttryck för.

Despite the centrality of assessment in teaching we know relatively little about the actual practice of teachers in schools. /.../

A preoccupation with the technical refinement /.../ has meant that other major issues in as-sessment have received relatively little attention – issues such as the role of assessment in the work of the school: the effects of using standardized tests; an examination of teachers' informal methods of assessment: and the assessment of equally important, if more elusive, noncognitive variables. (Madaus & Kellaghan 1992, s. 126–127)

I det andra kapitlet som är relevant för bedömning beaktas differentiering och dess konsekvenser för elevers möjligheter, prestationer och meningsskapande.

4.1.2 Utvärdering, policy, utveckling och sociologi

De handböcker som inriktar sig på utvärdering, policy, utveckling och sociologi är alla publicerade på 00-talet och omfattar cirka 1 000 sidor, fränsett handboken i sociologi som är på knappt 600 sidor. De har valts med tanke på att utvärdering kan ses som ett överordnat begrepp i vilket bedömning inkluderas (jfr åtskillnaden mellan evaluation och assessment på föregående sida), att bedömning är en av de centrala policyfrågorna inom utbildningen, att utvecklingsfrågor relaterade till utbildning ofta är relaterade till bedömning och utvärdering samt att bedömning är en fråga med koppling till flera utbildningssociologiska strukturbegrepp (genus, klass, etnicitet) och ses som ett av signalsystemen.

TABELL 4.2 Handböcker med fokus på utbildningens policy, utvärdering, utveckling och sociologi.

Titel och redaktör	Förlag	År
INTERNATIONAL HANDBOOK OF EDUCATIONAL EVALUATION Kellaghan & Stufflebeam	Kluwer	2003
INTERNATIONAL HANDBOOK OF EDUCATIONAL POLICY Bascia, Cumming, Datnow, Leithwood & Livingstone	SpringerLink	2005
INTERNATIONAL HANDBOOK OF SCHOOL EFFECTIVENESS AND IMPROVEMENT Townsend m.fl	SpringerLink	2007
HANDBOOK OF THE SOCIOLOGY OF EDUCATION Hallinan	SpringerLink	2006

Handboken om utvärdering består av två delar som tar upp perspektiv på utvärdering och utvärderingspraktiker. Perspektiven presenteras under de fem huvudområdena teori, metod, användning, profession och social kontext. Utvärderingspraktiker diskuteras i relation till fem typiska objekt för utvärdering: elever, personal, program, skolor och utbildningssystem. Alla dessa sektioner har på olika sätt relevans för bedömning. Men av särskilt intresse för bedömning med grund i elevers prestationer är sektionen *New and Old in Student Evaluation*. I den finns fyra kapitel som redovisar forskning om psykometriska principer vid bedömningar av elevers förmågor och prestationer, bedömningar i klassrummet, alternativa former för bedömning och externa examinationer. Dessutom finns en sektion som analyserar hur utvärdering griper in i och formar relationerna mellan olika nivåer i ett utbildningssystem. I detta sammanhang redovisas forskning om nationella system för kontroll och utvärdering samt internationella jämförande kunskapsmätningar.

I handboken om utbildningspolicy finns tre för våra syften relevanta kapitel. Ett handlar om internationella jämförande test som incitament i storskaliga utbildningsreformer med grund i så kallad evidensbaserad policy och praktik. I ett annat kapitel fokuseras statliga interventioner som metod för överförande av ansvar. Tre sätt att undvika ickeproduktiva effekter av "accountabilitypolicy" lyfts fram. De handlar om att utveckla redskap för bedömning som kan fånga alla utbildningens mål (sociala, emotionella och kognitiva), att använda data av hög kvalitet som underlag för beslut och

att utveckla nya attityder gentemot misslyckanden. Attityder som betraktar misslyckanden som en väsentlig aspekt av lärande. I ett kapitel om bedömning betonas vikten av att skilja mellan olika syften och att anpassa formerna och innehållet i bedömningen därefter. Författarna till kapitlet menar också att det inte är fråga om några enkla val. Olika typer av bedömningar har sin grund i "different underlying conceptions of knowledge and of learning" (Gipps & Cumming 2005 s. 695). De fortsätter med att konstatera att det mångtydiga fenomenet "assessment" kan beskrivas som ett kontinuum med psykometriska och testmättningsmodeller i den ena änden och tolkande konstruktivistiska modeller för bedömning i den andra inklusive en ny variant av den senare, bedömning *för* lärande, i vilken bedömning är en inneboende aspekt av lärandeprocessen.

I den 2007 publicerade *International Handbook of School Effectiveness and Improvement* betonas relationen mellan forskning, policy och praktik. Forskningen om framgångsrika skolor innefattar data om utfallet för enskilda elever. Trots detta är bedömningens innehåll eller mätningen av det inte en huvudfråga. Men i ett kapitel framhålls vikten av att problematisera det som händer mellan input och output, den så kallade svarta lådan. Den under 1980-talet dominerande femfaktormodellen för framgångsrika skolor betonade skolans ledarskap, fokusering på undervisning, klimat som främjar lärande, lärare med höga förväntningar och regelbunden uppföljning och utvärdering av elevers prestationer. Modellen har sedan dess utvecklats, men fokuseringen på skolnivån, och i mindre utsträckning klassrummet och läraren och eleverna, är densamma. Vad denna forskning hitintills visat är emellertid att det främst är eleven, läraren och klassrummet som är avgörande för elevens lärande. Att utveckla kunskap om denna påverkan ses som en viktig uppgift för den framtida forskningen.

Handboken från 2006 som redovisar utbildningssociologisk forskning är indelad i sex sektioner och 24 kapitel. I sektionen teoretiska och metodologiska orienteringar handlar ett kapitel om skolans effekter. Dessutom finns ett kapitel som behandlar klassrumsprocesser och elevers prestationer. Sektionen *The Study of School Outcomes* rymmer flera kapitel som är relevanta för bedömning. Några handlar om relationen mellan skola och arbetsliv. Ett kapitel tar upp differentieringen av elever och innehåll samt social stratifiering. I ett annat kapitel diskuteras skolans betydelse för elevers livschanser, bland annat med referens till elevers betyg och examina. Genom denna forskning har kritiska frågor om jämlikhet och demokrati riktats mot systemen för utbildning sett utifrån elevers olika bakgrund och erfarenheter. Inom utbildnings-sociologin används testresultat, betyg och examina för att analysera relationen mellan skolan och andra samhälleliga institutioner. Den empiriska traditionen är stark och bruket av longitudinella databaser

vanligt. Mot den bakgrunden framhålls betydelsen av att utveckla valida och reliabla data som tar hänsyn till utbildningens alla mål.

4.1.3 Språk, estetiska ämnen och informationsteknologi

Slutligen har vi också några exempel som gör det möjligt att säga något utifrån ett urval specifika frågor. Hit hör bedömning i relation till olika skolämnen. Här har vi valt å ena sidan språk/engelska, som exempel på ett kärnämne, och å andra sidan art/estetiska ämnen som till sin karaktär har vissa särdrag som skiljer det från mer traditionella ”teoretiska” ämnen. Dessutom har vi som ett svar på förändringar i samhället och inom utbildningen tagit med en handbok om informationsteknologi i undervisningen. De forskningshandböcker vi valt framgår av tabell 4.3.

TABELL 4.3 Handböcker inom specifika områden

Titel och redaktör	Förlag	År
INTERNATIONAL HANDBOOK OF ENGLISH LANGUAGE TEACHING Cummins & Davison	Springer	2007
INTERNATIONAL HANDBOOK OF RESEARCH IN ARTS EDUCATION Bresler	Springer	2007
INTERNATIONAL HANDBOOK OF INFORMATION TECHNOLOGY IN PRIMARY AND SECONDARY EDUCATION Voogt & Knezek	SpringerLink	2008

I boken om undervisning av engelska på 1 138 sidor ingår en sektion om bedömning: *Assessment and Evaluation in ELT: Shifting Paradigms and Practices*. Sektionen rymmer kapitel som tar upp standardbaserade ansatser och rörelser, ”high-stakes testing” och bedömning, nya bedömningsformer vid rekrytering till högre utbildning, bedömningspraktikers påverkan på undervisningen, möjligheter och svårigheter i den klassrumsbaserade bedömningen, bedömning och maktutövning samt hur skilda sätt att definiera språk och lärande av språk har implikationer för bedömning. Tre breda teman och riktningar inom forskningen diskuteras. Först och främst att det behövs mer forskning om olika nyckelbegrepp som ”standards, ethics, trustworthiness, and fairness, the relationship between testing, assessment, and evaluation, and the interaction between validity and reliability” (Davison &

Cummins 2007, s. 420). För det andra påpekas att det krävs mer detaljerade klassrumsstudier av alternativa bedömningar och deras effekter på elevers lärande. Dessutom bör man genomföra fler longitudinella studier om relationen mellan formativ bedömning, återkoppling och lärande samt studier av bedömningspraktiker som portfolio, självvärdering, kamratbedömning och interaktiva och dynamiska bedömningar. För det tredje behövs forskning som relaterar förändringar på systemnivå till utveckling av bedömningspraktiker i klassrummet och elevers lärande. Dessutom bör man uppmärksamma konflikter mellan motstridiga bedömningskulturer med ursprung i olika ideologier och grundantaganden.

Handboken om undervisning i estetiska skolämnen som är på 1 516 sidor innehåller sektionen *Assessment and Education* med åtta kapitel på drygt 100 sidor. Här redovisas forskning om bedömning i relation till olika kunskapsområden som dans, musik, litteratur, drama och bild. Dessutom diskuteras bedömningsforskningen mer övergripande. Regina Murphy och Magne Espeland (2007) lyfter fram den dubbla betydelsen av assessment/evaluation "to assist as well as to value something" (s. 338). De betonar också att lärare inom estetiska ämnen ofta inte använder explicita kriterier. Ibland skapar egna kriterier alternativt förlitar sig på ett gemensamt ethos eller standard inom ett specifikt fält. De grundar sina bedömningar på ett professionellt utvecklat seende och vad de vet om själva processen. Ett dilemma i undervisningen inom dessa ämnen är

how to capture non-verbal activity – mime, music, visual image, movement, and all its subtleties – through verbal expression. Despite the advances in authentic assessment and performance assessment, the unique ways of knowing are subsumed into descriptors which cannot always capture the spiritual, ephemeral or visceral qualities of an artwork that is listened to, observed or sensed in a single moment. (s. 339)

Elliot Eisner (2007) diskuterar några uppfattningar om bedömning som han finner problematiska, bland annat föreställningen att man måste använda mätningar och test.

It is almost as though measurement and evaluation are thought to be the same. This is hardly the case. One can measure without evaluating, and one can evaluate without measuring. This is not an argument about measuring, it simply is a recognition that evaluation requires making judgments about the value of what one is paying attention to. Measurement has to do with the quantitative description of a set of conditions. Most of our evaluation practices both in school and in daily life outside of school have little or nothing to do with measurement. The assumption that it does should be parked where it belongs. (s. 423–424)

Vad det gäller är snarare att utveckla en förfinad känslighet att observera "what is subtle but significant, and the ability to provide reasons for one's appraisal of their worth" (s. 426). I ett kapitel som knyter an till denna tanke förordar Chris Higgins (2007) att den som ska bedöma värdet av något måste

use their imagination to move past the cardboard versions of things. The question for educational evaluation is not which method to choose or how to employ it, but how to notice the aspects of schools hidden by our stereotypes of schools, the qualities of teachers and learners obscured by our cynicism or sentimentalism, the dimensions of classrooms that are hiding in plain view. (s. 393)

Handboken om informationsteknologi i undervisningen är på 1 178 sidor. Bedömning behandlas främst i kapitlet *Changing Assessment Practices and the Role of IT*. Ola Erstad (2008) knyter ihop olika traditioner om lärande och kunskap, behavioristiska, konstruktionistiska och sociokulturella, med skilda former för bedömning. Han noterar också att forskningen om bedömning och IT är begränsad. Det är först under 1990-talet som det växer fram som forskningsfält. Informationsteknologin har inte medfört att äldre former för bedömning ersatt nya, i stället existerar de sida vid sida, eftersom skolor måste anpassa sig till "national standards and systems over which they have no control" (s. 184). Men han menar att nya teknologier har skapat ett intresse för att mäta det icke mätbara och

IT represents some new possibilities for developing assessment practices, especially formative assessment, and how the complexity of these tools can be used to assess higher order thinking skills, such as problem solving, that are difficult to assess by paper and pencil. (s. 189)

Samtidigt anser Erstad att många initiativ inte har handlat om att förändra bedömningspraktiker, utan främst om att utveckla traditionella summativa former för bedömning, även om exempel på motsatsen finns. Men sammantaget finns det ett stort behov av utvecklingsstudier med fokus på IT och bedömning. Vid sidan av dessa frågor beaktas också hur man gör bedömningar och utvärderingar genom storskaliga oftast kvantitativa nationella eller internationella mätningar. Här påpekas att IT:s inflytande på elevers lärande beror på både vad som mäts och hur detta mäts. I samband med detta påpekas att det behövs nya och olika indikatorer för skilda syften. Att studera påverkan på elevers lärande är komplicerat. Vikten av att använda och utveckla såväl experimentella, multimetodansatser och design research som metoder anpassade till specifika IT-tillämpningar framhålls.

Handböckerna har dock sin begränsning, vanligen är produktionstiden lång. Som exempel kan nämnas att arbetet med den fjärde utgåvan av *Hand-*

book on Research on Teaching (Richardson, red. 2001) hade satts igång redan 1997. Det innebär att även om de olika årgångarna speglar en förändring så är det som framstår som nytt i den sista delen redan etablerat då boken väl kommer ut. Lorrie Shepards (2001) artikel i den, om kopplingen mellan teorier om lärande och bedömning, har haft avgörande betydelse för möjligheten att ställa nya typer av forskningsfrågor. Men den tar inte upp det lärandeperspektiv som under 00-talet kommit att få stor uppmärksamhet, det sociokulturella. Det gör i stället Caroline Gipps (2002), som visar vilka utmaningar pedagogisk bedömning ställs inför i relation till sociokulturella teorier om lärande. En schweizisk forskare, mindre känd i engelskspråkiga sammanhang, som även hon pekar på hur sociokulturella perspektiv kan förändra vår förståelse av bedömning och kunskap, är Linda Allal (jfr Allal & Pelgrims Ducrey 2000). På motsvarande sätt har det arbete som Paul Black och Dylan Wiliam (1998, 2006, 2009) bedrivit i förhållande till lärares klassrumsbedömning, frågor om formativ bedömning och feedback (jfr även Hattie & Timperley 2007), haft avgörande betydelse för internationell forskning. På senare tid har också kritik riktats mot polariseringen formativ–summativ (t.ex. Newton 2007). Även metodutvecklingen, intresset för videobandad klassrumsforskning, har bidragit till att synliggöra bedömningen i klassrumskommunikationen och dess betydelse för vad som för eleverna framstår som viktigt att kunna (Torrance & Pryor 1998, 2001, Torrance 2007).

4.2 Internationella organisationer och konferenser

Ett stort antal internationella organisationer fungerar som mötesplatser för utbildningsvetenskaplig forskning. Som framgår av kapitel 2 har vi valt att fokusera de fyra paraplyorganisationerna AERA, EARLI, EERA och NFPF samt de två specifikt inriktade organisationerna IAEA och AEA-Europe. I det följande framgår hur bedömningsforskning uppmärksammas i dessa sammanhang.

4.2.1 AERA: American Educational Research Association

AERA är en organisation med medlemmar från många länder. Den ger ut flera vetenskapliga publikationer och anordnar årligen en stor konferens. AERA består av tolv divisioner och över hundra särskilda intressegrupper, SIG. Två av sektionerna kan genom sina kortfattade beskrivningar direkt anses som relevanta för forskning om bedömningar av elevers prestationer.

Hit hör *Measurement and Research Methodology* med sitt intresse för psykometri, mätningar och förbättrade statistiska procedurer samt *Research, Evaluation, and Assessment in Schools* med fokus på tillämpad forskning, programutvärdering, bedömning av elevers prestationer och accountability. Dessutom finns ett antal SIG som fokuserar på bedömning. De är inriktade mot klassrumsbedömningar, kognition och bedömning, storskaliga mätningar, portfolio, test- och mätinstruments validitet samt utvärdering. Motiven för varför dessa frågor är intressanta är både inom- och utomvetenskapliga. Flera SIG ger också uttryck för att bedömning bör förstås i relation till såväl policy och pedagogisk verksamhet som forskning. Särskilt framhålls också ambitioner som handlar om att främja teoretiska, metodologiska och empiriska bidrag. På AERA:s hemsida finns konferensernas teman för 1991–2009 angivna. Även om bedömning inte varit framträdande finns kopplingar till frågor om bedömning, särskilt under 00-talet då teman om utbildningens kvalitet och accountability ingått.

4.2.2 EERA: European Educational Research Association

EERA grundades 1994 för att utveckla samarbete mellan forskare och organisationer inom utbildningsvetenskap i Europa. Organisationen har också en egen vetenskaplig tidskrift. I dag finns 27 nätverk. Ett av dessa, *Assessment, Evaluation, Testing and Measurement*, som etablerades 1997 beskriver målet för nätverket på följande sätt:

The network aims to promote information exchange and active collaboration among European researchers in the field of assessment and evaluation in all levels and all sectors of education and training, to further develop assessment techniques and approaches, and to encourage the intelligent interpretation of assessment results. In doing so, the network wants to cover not only research on the development and utilization of assessments and evaluations, and the secondary analyses of data of (international) assessment studies, but explicitly also research aimed at fundamental issues and theoretical developments in assessment, evaluation, testing and measurement. (www.eera.ac.uk)

Bedömning och utvärdering beskrivs också som en inom utbildning tilltagande aktivitet på alla nivåer och över hela världen. Vikten av att studera dessa fenomen understryks också av att resultaten används i många olika sammanhang: som vägledning för utformning av undervisningen, vid selektion av studenter och sökande till olika arbeten och för utvärdering av utbildningen på organisations- och systemnivå. Mot den bakgrunden fokuserar nätverket på forskning om "the quality of assessment in education and the uses made of assessment results in teaching, learning, evalu-

ation and policy making” (www.eera.ac.uk). Synen på bedömning framgår också av de nyckelord som används för att beskriva nätverket: assessment, evaluation, testing, measurement, system monitoring, assessment quality och impact on policy and practice.

Den redd som kan urskiljas manifesterade sig 2008 i ett namnbyte för nätverket som före 2008 hette *Student Assessment*. Skälet för bytet var att många papers handlade om bedömning, men inte nödvändigtvis enbart av elevers prestationer. De kunde också handla om utvärdering av program och kontrollfunktioner av skilda slag. Initialt var nätverket inriktat mot validitet, reliabilitet och bruket av nya former för bedömning och etablerade teknikens tillämpning i nya sammanhang för att mäta elevers intellektuella förmågor, färdigheter, attityder och personlighetsdrag. Vid den årliga konferensen, *ECER: European Conference of Educational Research*, 2009 genomförde nätverket fyra symposier och många papersessioner. Jämfört med 2008 hade mängden papers fördubblats. Ökningen kan delvis men inte helt förklaras som en generell effekt av ökat deltagande i konferensen. Tilläggas kan också att ingen av ECER-konferenserna 2006–2009 har haft bedömning som övergripande tema. Det gäller också konferensen 2010.

4.2.3 EARLI: European Association of Research on Learning and Instruction

Syftet med EARLI är att främja teoretisk och empirisk forskning med särskilt fokus på lärande och undervisning samt utbyte av kunskap mellan forskare. I likhet med de andra paraplyorganisationerna ordnar EARLI konferenser och publicerar vetenskapliga texter. Konferensen hålls vartannat år. Under perioden 2003–2009 handlade ingen om bedömning. Inom EARLI finns sexton SIG. En av dem, *Assessment and Evaluation*, avser bedömning. I den betonas nödvändigheten av att uppmärksamma behovet av pedagogiska modeller för bedömning liksom digitala bedömningar som kan öppna för elever att i högre grad själva kontrollera sitt lärande. Därtill finns det ett intresse för att utveckla bedömning och utvärdering teoretiskt, betydelsen av återkoppling, livslångt lärande och bedömning, e-bedömning och kollektiva former för bedömning.

En annan SIG, som behandlar *Educational Effectiveness*, har som syfte att främja utvecklingen av ansatser som är longitudinella, storskaliga och inkluderar flera nivåer som studenter, lärare, skol- och systemnivå. Även kvasiexperimentella studier betraktas som väsentliga. Ambitionen är att förstå de multipla processer som formar utbildningens utfall och dess variation, stabilitet och förändring.

4.2.4 NFPF/NERA: Nordisk Förening för Pedagogisk Forskning

NFPF, som har funnits i drygt trettio år, har sin förankring i de nordiska länderna. Föreningen främjar utvecklingen av pedagogisk forskning i Norden, bland annat genom att anordna årliga konferenser och genom spridning av den vetenskapliga publikationen *Nordisk Pedagogik* och den med föreningen associerade *Scandinavian Journal of Educational Research*. NFPF har i dag 25 nätverk. Av dem är inget direkt relaterat till bedömning, vilket delvis kan förklaras av grunden för kategorisering av nätverken som mer sällan fokuserar specifika kunskapsobjekt. Informationen om konferensernas tema gäller endast sex. Inget av dem handlar om bedömning.

4.2.5 IAEA International Association for Educational Assessment

IAEA grundades 1975 med det uttalade syftet att stödja utbildningsorganisationer i utvecklingen och tillämpningen av pedagogiska bedömnings-tekniker för att främja utbildningens kvalitet. IAEA:s främsta mål är:

- to improve communication among organizations involved in educational assessment by sharing professional expertise through conferences and publications, and by providing a framework within which cooperative research, training, and projects involving educational assessment can be undertaken.
- to make expertise in assessment techniques more readily available for the solution of educational problems.
- to cooperate with other agencies having complementary interests.
- to engage in other activities for the improvement of assessment techniques and their appropriate use by educational agencies around the world. (www.iaea.info)

Här framgår närheten mellan policy, praktik och forskning som också kommer till uttryck i titeln på organisationens tidskrift *Assessment in Education: principles, policy & practice*. I tabell 4.4 kan man se vilka olika teman IAEA:s årliga konferenser 2002–2009 har haft. I tabellen ingår också temat för konferensen 2010. Sammantagna pekar dessa teman på bedömning som ett objekt i omvandling, som är knutet till många verksamheter, flera nivåer och som kan brukas för olika syften.

TABELL 4.4 Tema för IAEAs årliga konferenser 2002–2010

År	Konferens
2010	ASSESSMENT FOR THE FUTURE GENERATIONS
2009	ASSESSMENT FOR A CREATIVE WORLD
2008	RE-INTERPRETING ASSESSMENT: SOCIETY, MEASUREMENT AND MEANING
2007	MUTUAL DEPENDENCE OF NATIONAL ESTIMATION SYSTEMS AND EDUCATION STANDARDS
2006	ASSESSMENT IN AN ERA OF RAPID CHANGE: INNOVATIONS AND BEST PRACTICES
2005	ASSESSMENT AND THE FUTURE OF SCHOOLING AND LEARNING
2004	ASSESSMENT IN THE SERVICE OF LEARNING'
2003	SOCIETIES' GOALS AND ASSESSMENT
2002	REFORMING EDUCATION ASSESSMENT TO MEET CHANGING NEEDS

Slutligen kan vi notera, med hjälp av det medlemsregister som finns på hemsidan och listan för deltagare i konferensen 2007, att det svenska deltagandet förefaller ytterst begränsat. Det gäller enbart konferensen och inte med nödvändighet dess tidskrift.

4.2.6 AEA EUROPE: Association for Educational Assessment

AEA Europe är en organisation vars mål är att fungera som en plattform för reflektion över utvecklingen av bedömning inom utbildningssystemen i Europa. Bedömning inkluderar här akademiska, professionella och arbetslivssammanhang samt processer och produkter. Syftet är att ta fram ett ramverk som kan utgöra grunden för kooperativa forskningsprojekt. Genom organisationen är avsikten också att delta i aktiviteter som kan främja utvecklingen av och kunskapen om bedömningar och hur man använder dem. Den första konferensen hölls 2000. De årliga konferenserna har sedan dess handlat om att förbättra bedömningar, att utröna vunna erfarenheter, bedömningar i relation till olika värden som förändrade behov, jämlikhet och demokrati eller ansvar, kvalitet och standard. Det framgår närmare av tabell 4.5.

TABELL 4.5 AEA-Europe – teman för konferenserna 2000–2009

År	Konferens
2009	INNOVATION IN ASSESSMENT TO MEET CHANGING NEEDS
2008	ACHIEVING QUALITY IN ASSESSMENT: VALIDITY AND STANDARDS
2007	ASSESSMENT FOR EDUCATIONAL QUALITY
2006	ASSESSMENT AND EQUITY
2005	ASSESSMENT AND ACCOUNTABILITY
2004	ASSESSMENT AND CULTURE: TRADITION OR INNOVATION?
2003	ASSESSMENT CHALLENGES FOR DEMOCRATIC SOCIETIES
2002	ASSESSMENT IN EUROPE: WORKING TOGETHER TOWARDS UNDERSTANDING
2001	ASSESSMENT IN EUROPE: ACHIEVEMENTS AND EXPERIENCES
2000	IMPROVING ASSESSMENT IN EUROPE

Tilläggs kan också att Christina Wikström vid Institutionen för beteendevetenskapliga mätningar i Umeå är styrelseledamot och generalsekreterare i AEA-Europe. Vi kan dessutom notera att endast ett fåtal personer från Sverige deltog i den årliga konferensen 2009.

4.3 Bedömning som epistemiskt objekt

Vårt intresse här gäller i första hand bedömning som ett objekt för vetenskaplig kunskap, inte bedömning som socialt fenomen, även om de är förbundna med varandra. I den vetenskapliga verksamheten är det objekten som är i fokus för vår verksamhet. I vanemässiga praktiker är objekten däremot frusna, enhetliga och inneslutna i vår vardagliga förståelse (Knorr Cetina 2001). De skiljer sig därmed från vetenskapens kunskapsobjekt.

Objects of knowledge are characteristically open question-generating and complex. They are processes and projections rather than definitive things. Observations and inquiry reveals them by increasing rather than reducing their complexity (s. 181).

Kunskapsobjekt, eller epistemiska objekt som Karin Knorr Cetina också kallar dem, existerar alltså samtidigt i en rad olika variationer och kan aldrig slutgiltigt fastställas. Som en konsekvens av forskarnas systematiska studier är epistemiska objekt i ständig rörelse.

Om vi överför detta resonemang till bedömning kan vi identifiera en rörelse från en situation där kunskapen om bedömning är mer eller mindre oreflekterad till en situation där kunskapen och metoderna för att utveckla den är mer distinkta och differentierade. Som en konsekvens har vi i dag en mer komplex förståelse av bedömningar och de olika praktiker de är inbäddade i. I vilken omfattning det här är en giltig beskrivning för olika aspekter av bedömning är emellertid en empirisk fråga. Genom att sammanfatta de redovisade bilderna av forskning om bedömning är avsikten att kvalificera bedömning som epistemiskt objekt. Här diskuterar vi bedömning som ett kunskapsobjekt i omvandling.

4.3.1 Bedömning – ett kunskapsobjekt i omvandling

Den internationella forskningen om bedömning kan beskrivas i relation till ett antal distinktioner. En första som uppmärksammas av många gäller termerna *assessment* och *evaluation*. De flesta knyter *assessment* till utvärderingen av enskilda elevers prestationer, medan *evaluation* relateras till värderingen av skolan som organisation och system. Alltmer förstås de dock som uttryck för, om inte en och samma sak, så ändå fenomen som är nära förbundna med varandra genom både innehåll och konsekvenser. Utvärderingen av individer, organisationer och system är allt oftare baserade på individdata. Samtliga av dessa utvärderingar kan handla om att både assistera/främja utveckling och att värdera utfall.

En annan distinktion tar fasta på om forskningen, i förhållande till skolan, studerar interna eller externa bedömningar och bedömningspraktiker. De interna bedömningarna är antingen knutna till undervisningen, och genomförs av lärare, eller test och undersökningar av skilda slag, som utförs av specialpedagoger, skolpsykologer och skolläkare. Interna bedömningar kan också vara initierade av lärare och elever eller av externa aktörer och ingå i lokala eller inter/nationella system för kontroll, uppföljning och utvärdering. Forskningen har också analyserat externa aktörers bruk av bedömningar, främst som en aspekt av förändringar av styrningen av utbildning, det vill säga som en del av kravet på *accountability*, standards och kvalitet.

Medan vissa forskare främst är inriktade på att utveckla kunskap om bedömningar i de pedagogiska verksamheterna är andra mer upptagna av användningen av bedömningar inom policy och administration. Några ägnar sig särskilt åt forskningens bidrag till och funktion inom olika

bedömningspraktiker. Andra uppmärksammar relationen mellan samtliga tre verksamheter policy, pedagogisk praktik och forskning. Inom den redskapsfokuserade forskningen har vi både dem som utvecklar instrument för klassrumsbedömningar och dem som främst utvecklar redskap för utvärderings- och/eller forskningsändamål. Både kvantitativa och kvalitativa tekniker för insamling och analys studeras och utvecklas.

Forskningen förbinder bedömning med flera olika nivåer. Bedömning beskrivs, analyseras och förklaras med hänvisning till såväl individen, gruppen och organisationen som systemet. Ibland är det lärares och elevers erfarenheter, uppfattningar om och inställning till bedömning som står i förgrunden, men oftast handlar det om elevers prestationer. Mestadels fokuseras elevers kognitiva resultat och mer sällan hur de utvecklas socialt och emotionellt. En stor del av forskningen redovisar hur resultaten varierar i förhållande till kön, socialgrupp och etnicitet.

Skolan som organisation uppmärksammas i första hand på två skilda sätt. Forskningen om framgångsrika skolor har förklarat elevers resultat med hänvisning till organisationens ledarskap, klimat etcetera. Annan organisationsinriktad forskning har studerat hur bedömningar används för att kategorisera och ordna elever, innehåll och studiegångar. Bedömningar används för att placera elever i grupper som svarar mot prestationer och/eller intressen. Något som i förlängningen får konsekvenser för vilka läroplaner elever blir erbjudna. Forskningen om system ses i relation till både inter/nationella jämförande kunskapsmätningar och utbildningens värden (demokrati, jämlikhet, etc.), styrning och utveckling. Forskningen på denna nivå handlar om att fastställa läget, jämföra olika utbildningssystem med varandra och urskilja trender. Därtill analyseras avsedda och icke avsedda konsekvenser av den ökade användningen av bedömningar för undervisningen och utbildningen. I samband med detta ställs också frågor om funktion, makt och legitimitet.

Vid sidan av forskning med rötter i psykometrin med fokus på utvecklingen av test och statistiska analyser förekommer också klassrumsforskning som är inriktad på både traditionella och alternativa former för bedömning. Inom den senare studeras och betonas utvecklingen av kunskap och multimodala redskap för att mäta olika aspekter av kunnande, exempelvis verbala/icke-verbala förmågor, praktiska/teoretiska kunskaper, processer/produkter, kognitiva/sociala/emotionella förmågor. Teoretiskt förbinds inom denna riktning teorier om lärande och kunskap med olika syften (summativa, formativa) och former för bedömning (prov, portfolio, självvärderingar) och innehåll (skilda skolämnen och olika förmågor). Även frågor om seende, etik och professionellt omdöme beaktas inom den alternativa bedömningsforskningen.

Som framgått har mängden nivåer, verksamheter, aktiviteter, handlingar, redskap och aktörer som studeras mångfaldigats. Dessutom ställs allt fler och mer differentierade forskningsfrågor, som beaktar skillnader både i fråga om vad som bedöms och sammanhanget för bedömning. Vad som innefattas i forskningen om bedömning har utvidgats kontinuerligt, men framför allt på sistone. Antalet traditioner har mångfaldigats och det gäller också mängden metodansatser. Forskningen innefattar numera psykometriska, didaktiska/läroplansteoretiska, organisationsteoretiska och sociologiska traditioner. Variationen av metodansatser sträcker sig från kvantitativa longitudinella storskaliga mätningar med statistiska analyser över strukturerade observationer till kvalitativa ansatser med tolkande och kritiska analyser. En rörelse kan iakttas som sträcker sig från kvantitativa till kvalitativa metoder med ett på senare tid förnyat intresse för kvantitativa tekniker för insamling och analys. Frågor som validitet och reliabilitet ställs på delvis nya sätt där den ena eller andra ges företräde, beroende på tradition.

Vi kan också notera ett ökat samarbete mellan olika aktörer. Antingen genom att resultat framtagna inom en verksamhet används av andra eller för att olika aktörer samarbetar om att utveckla kunskap om bedömning. Som exempel på det förra kan evidensbaserad praktik och policy nämnas och i det senare fallet aktions- och deltagarorienterad forskning. Mot bakgrund av den här skisserade bilden framträder bedömning som ett mångfacetterat kunskapsobjekt. Det kan också betraktas som ett gränsobjekt, det vill säga ett objekt som kan delas av flera intressenter och som är så plastiskt att det kan preciseras och operationaliseras i relation till specifika syften, kunskapsintressen, forskningsfrågor och metodansatser (jfr Star & Griesemer 1989).

I det här sammanhanget är det intressant att notera den begränsade uppmärksamhet som frågan om bedömning får i en av de få handböcker, om inte den enda internationella, som särskilt fokuserar elever. *International Handbook of Student Experience in Elementary and Secondary School* (Thiessen & Cook-Sather red. 2007) är på 872 sidor och innehåller tre sektioner om hur elever deltar och förstår livet i klassrum och skola, elevers identitet och utvecklingen av denna samt involverandet av elever för utveckling av verksamheten. Bedömning står inte i förgrunden varken för någon av sektionerna eller för något av de 33 kapitlen. Inte heller förekommer termer som assessment, evaluation, measurement eller test i ämnesindex. Det innebär inte att bedömning helt lyser med sin frånvaro. Som exempel kan nämnas *Students Perspective on Good Teaching: Implications for Adult Reform Behavior* (Wilson & Corbett 2007) som bland annat tar upp problemet med att använda standardiserade test som grund för skolutveckling.

The problem with using standardized tests scores as the bases for spurring adult action is that producing good results can be accomplished in ways that do not necessarily represent good instruction. Numerous educational experts have pointed this out. Popham (2004), for example, has argued that standardized tests (at least the ones currently in place around in our nation's schools) are "instructionally insensitive" since both good and bad instruction can lead to increased test score. Likewise, Braun and Mislevy (2005) make the point that policy makers are developing accountability systems on an "intuitive test theory" that may look commonsensical but is in reality full of faulty assumptions. Standardized tests no doubt leverage action, but the actions may not be ones that are educationally sound (s. 307–308).

Utifrån analyser av intervjuer med elever lyfter de fram tre frågor som bör vara vägledande för god undervisning, varav två är direkt knutna till bedömning.

- Does every child complete every assignment at an acceptable level of quality?
- Is every assignment worth doing?
- Does every adult know the name of every child in the building, and is every child known well by at least several adults in the building? (s. 308).

Även om också andra exempel finns på hur bedömning är en aspekt av att vara elev står det klart att bedömning sett utifrån elevers erfarenheter och attityder behöver studeras mer ingående. Mot bakgrund av genomgången i detta kapitel kan vi också lyfta fram ett antal områden som forskarna pekar ut som centrala att studera framöver. Nödvändigheten av att forska om olika nyckelbegrepp framhålls. Bedömningsforskningen har i stor utsträckning studerat vissa skolämnen medan andra inte uppmärksamats alls eller endast i begränsad mening. Dessutom har betydelsen av olika sammanhang för bedömning inte beaktats i någon större omfattning. Många forskare betonar också särskilt vikten av mer forskning om hur kunskap, lärande och bedömning är internt relaterade till varandra oavsett om det är bedömning av lärande, bedömning som lärande eller bedömning för lärande som är syftet med bedömningen. Några framhåller också vikten av att öka kunskapen om alternativa former för bedömning.

Betydelsen av informationsteknologi lyfts också fram som något som behöver studeras ytterligare. Både som redskap inom såväl storskaliga kvantitativa mätningar som utveckling av bedömningspraktiker i klassrummet. Ytterligare ett område som lyfts fram som centralt att studera vidare är hur relationen mellan klassrumsbedömningar, standarder och systemkontroller kan förstås. Genomgående betonas också att forskningen bör bidra med

kunskap och utveckla redskap som kan användas av andra, t.ex. administratörer och lärare. Att forskning om bedömning är ett angeläget kunskapsområde beror bland annat på att det är både inom- och utomvetenskapligt motiverat och att frågor om etik, makt och legitimitet står på spel. Det är frågor vars relation till bedömning, enligt forskarna, behöver beskrivas och analyseras ytterligare. Sammanfattningsvis kan vi notera att bedömning framträder som ett alltmer angeläget område för forskning inom såväl pedagogik som andra discipliner.

5 SVENSK OCH NORDISK BEDÖMNINGSFORSKNING OCH DESS OMFATTNING

Här beskriver vi svensk och till viss del också nordisk bedömningsforskning som den framträder utifrån svenska avhandlingar och artiklar i svenska vetenskapliga tidskrifter samt svenska artiklar och artiklar från andra nordiska länder i såväl nordiska vetenskapliga tidskrifter som andra internationella tidskrifter. Därutöver uppmärksammar vi också inslaget av forskningsprojekt som beviljats medel från Utbildningsvetenskapliga kommittén (UVK) och som kan knytas till forskning om bedömning. Som framgår av kapitel 2 varierar de undersökta tidsperioderna i första hand som en konsekvens av när tidskrifter etablerades, men också beroende av syfte. Det sistnämnda avser forskningsprojekten som tagits med för att möjliggöra ett framtidsperspektiv och som därför omfattar perioden 2006–2009. Inledningsvis beskrivs förekomsten av forskning publicerad i Sverige och därefter nordisk och internationell forskning. I samband med det första uppmärksammar vi också vilka forskningsmiljöer som är engagerade i forskning om bedömning.

5.1 Forskning om bedömning i Sverige

Omfattningen och spridningen över tid för svenska avhandlingar och vetenskapliga artiklar publicerade i *Pedagogisk Forskning i Sverige* (PFS) och *Utbildning & Demokrati* (U&D) framgår av tabell 5.1.

TABELL 5.1 Omfattning och spridning av forskning om bedömning över tid och i absoluta tal i svenska avhandlingar och artiklar publicerade i två svenska vetenskapliga tidskrifter. Absoluta tal.

År	Avhandlingar 1990–2009	PFS 1996–2009	U & D 1992–2009	UVK-projekt 2006–2009
2009	8	1	0	7
2008	13	2	1	8

2007	20	1	0	12
2006	16	4	3	8
2005	4	2	0	-
2004	3	3	0	-
2003	4	1	4	-
2002	1	8	0	-
2001	2	4	0	-
2000	3	2	0	-
1999	2	3	0	-
1998	4	1	0	-
1997	1	4	0	-
1996	2	5	0	-
1995	3	-	0	-
1994	1	-	0	-
1993	2	-	3	-
1992	1	-	0	-
1991	1	-	-	-
1990	4	-	-	-
Σ	95	41	11	35

Under perioden 1990–2009 publicerades totalt 95 svenska avhandlingar med relevans för bedömning, varav 89 doktorsavhandlingar och sex licentiatavhandlingar. En snabb överblick visar att spridningen under den undersökta perioden är allt annat än jämn. Under det första decenniet (1990–1999) utkommer endast 21 avhandlingar jämfört med det senare (2000–2009) som omfattar 74 avhandlingar. Men även inom den senare perioden finns stora skillnader. 2000–2005 publiceras 17 avhandlingar. Den markanta ökningen sker 2006. Det totala antalet avhandlingar under perioden 2006–2009 är 57. Det innebär att 60 procent av samtliga avhandlingar om bedömning är publicerade under de fyra senaste åren.

Om vi förflyttar oss till tidskrifterna blir bilden en annan. Spridningen är här jämnare eller annorlunda fördelad. Med *Pedagogisk Forskning i Sverige* som exempel kan vi notera att det finns en variation mellan en till åtta artiklar per år och att det är ungefär lika många år med en till två artiklar som med tre, fyra eller fem artiklar. Skillnaderna mellan de fjorton åren förefaller mer tillfällig och kanske slumpartad, frånsatt 2002. Med anledning av professor emeritus Kjell Härnqvists åttioårsdag utgavs temanumret *Individuella förutsättningar för utbildning* med flera artiklar relevanta för forskning om bedömning. Totalt publicerades under perioden 1996–2009 41 artiklar, ett relativt stort antal.

Utbildning & Demokrati hade endast ett fåtal artiklar under perioden 1992–2009. Tidningen utger emellanåt olika temanummer, och även om inget explicit behandlar bedömning finns ett nummer 2006 med flera artiklar som fokuserar bedömning. Även därutöver finns det utspritt över åren exempel på artiklar om bedömning, sammanlagt elva artiklar om bedömning.

Vid sidan av omfattning och spridning har vi också granskat fördelningen mellan manliga och kvinnliga författare. För avhandlingarnas del dominerar de kvinnliga författarna med 56 avhandlingar, 60 procent av samtliga avhandlingar. Till de 41 artiklarna i *Pedagogisk Forskning i Sverige* finns det sammanlagt 50 författare och de elva artiklarna i *Utbildning & Demokrati* har 13 författare. I PFS är två tredjedelar av författarna män. Tio artiklar är samproducerade, men endast en artikel har både en man och en kvinna som författare. U&D har en jämnare könsfördelning, sju kvinnor och sex män står enskilt eller tillsammans som författare till artiklarna om bedömning.

5.1.1 Lärosäten engagerade i forskning om bedömning

Med utgångspunkt i den forskning som är producerad i Sverige och som ingår i studien kan man i någon mån beskriva miljöerna som bedömningsforskningen växer fram i under den studerade perioden. Vi har identifierat vid vilka lärosäten avhandlingarna, artiklarna publicerades i PFS och U&D och de av UVK finansierade projekten har publicerats.

Hur de fyra publikationskällorna fördelar sig över olika lärosäten framgår av tabell 5.2. Redan en snabb överblick visar på en tydlig dominans för vissa lärosäten. De tre lärosäten som har publicerat det största antalet avhandlingar är samtliga universitet, men även vid Malmö högskola och den tidigare Lärarhögskolan i Stockholm (ingår numera i Stockholms universitet) har det lagts fram flera avhandlingar.

TABELL 5.2 Svenska publikationer och ansökningar: fördelning mellan lärosäte. Absoluta tal

Lärosäte	Avhandlingar	PFS	U & D	UVK-projekt	Totalt
UMEÅ UNIVERSITET	16	6	-	7	29
GÖTEBORGS UNIVERSITET	15	23	1	8	47
UPPSALA UNIVERSITET	15	5	3	5	28
STOCKHOLMS UNIVERSITET ¹¹	13	3	-	5	20
LINKÖPINGS UNIVERSITET	12	1	2	1	16
MALMÖ HÖGSKOLA	9	1	1	2	13
LULEÅ TEKNISKA UNIVERSITET	4	-	-	1	5
VÄXJÖ UNIVERSITET	4	-	3	1	8
HÖGSKOLAN I BORÅS	2	-	-	-	2
HÖGSKOLAN I JÖNKÖPING	1	-	-	-	1
KARLSTAD UNIVERSITET	1	1	-	-	2
KAROLINSKA INSTITUTET	1	-	-	-	1
LUNDS UNIVERSITET	1	-	-	-	1
ÖREBRO UNIVERSITET	1	-	1	1	3
MÄLARDALENS HÖGSKOLA	-	-	-	2	2
MITTUNIVERSITETET	-	-	-	1	1
HÖGSKOLAN I KALMAR	-	1	-	-	1
HÖGSKOLAN I SKÖVDE	-	-	-	1	1
TOTALT	95	41	11	35	181

Av de sex lärosäten som har flest publicerade avhandlingar inom området ingår fyra i den av Vetenskapsrådet, under perioden 2007–2010, finansierade *Nationella forskarskolan i pedagogisk bedömning* (Pettersson 2006). Dessa läro-

¹¹ Stockholms universitet inkluderar också texter publicerade inom den tidigare Lärarhögskolan i Stockholm.

säten är Umeå universitet, Göteborgs universitet, Stockholms universitet (tidigare Lärarhögskolan i Stockholm) och Malmö högskola. Dessutom har 62 procent av UVK-projekten om bedömning fördelats till dessa lärosäten som också är relativt flitiga skribenter i PFS. Deras starka representation kan förutom delaktigheten i forskarskolan förstås mot bakgrund av att lärosätena i Umeå, Göteborg och Stockholm sedan flera år tillbaka har haft olika myndighetsuppdrag.

Bland uppdragen kan nämnas konstruktionen av nationella prov (tidigare centralprov) och medverkan i de internationella jämförande kunskapsmätningarna. Även uppgiften att utveckla provbanker, bland annat i naturvetenskapliga ämnen, hem- och konsumentkunskap med flera skolämnen för grund- och gymnasieskolan, bidrar till att upprätthålla forskningsmiljöer där frågor om kunskap, bedömning, prov och betyg återkommande ingår i arbetsuppgifterna. Umeå universitet har dessutom ansvar för Högskoleprovet och Stockholms universitet har tillsammans med Växjö haft ansvaret för att utveckla provbanken i yrkesämnena. Även Uppsala universitet som har publicerat flest antal avhandlingar under perioden har ansvar för nationella prov (svenska). Vid sidan av nämnda högskolor har också Linköping ett relativt stort antal avhandlingar, flera dock i periferin i förhållande till bedömningsforskningens kärna. Sett utifrån tabellen kan också framhållas den mer spridda bilden som framträder i U&D, där Örebro universitet (tidskriftens hemvist), Uppsala universitet och Växjö universitet tillhör de lärosäten som har flest artiklar om bedömning.

I Göteborg har Institutionen för pedagogik och didaktik en dominerande ställning. Tretton av de femton avhandlingarna är publicerade där. Institutionen dominerar den nationella avhandlingstatistiken inom bedömningsforskningen. De två andra avhandlingarna är utgivna av Institutionen för idéhistoria och vetenskapsteori respektive Sociologiska institutionen.

Vid Stockholms universitet och Lärarhögskolan i Stockholm är åtta av de tretton avhandlingarna producerade vid antingen någon av Lärarhögskolans institutioner eller Institutionen för pedagogik alternativt Institutionen för internationell pedagogik. De tre avhandlingar som publicerats vid den senare är speciella genom att fokus ligger antingen på fenomen i något annat land än Sverige (t.ex. Benincasa 1995) eller på komparativa studier (t.ex. Engström 1993, Zhao 1993). Vid Centrum för tvåspråkighetsforskning i Stockholm har en avhandling producerats med fokus på svenska som andraspråk (Ganuza 2008). Av övriga avhandlingar från Stockholm är tre publicerade vid Institutionen för ekonomi och en vid Institutet för social forskning (SOFI).

Åtta av avhandlingarna från Umeå universitet är publicerade vid Pedagogiska institutionen eller vid två av de institutioner som under 00-talet

brutits ut ur denna, Institutionen för beteendevetenskapliga mätningar respektive Institutionen för barn- och ungdomspedagogik, specialpedagogik och vägledning. Övriga avhandlingar fördelar sig mellan flera olika institutioner. Samtliga avhandlingar från Malmö är producerade vid lärarutbildningen. För Uppsalas del gäller att sju av avhandlingarna framlagts vid Pedagogiska institutionen, tre vid Institutionen för nordiska språk, medan enstaka är från Institutionen för psykologi, Institutionen för nationalekonomi respektive Institutionen för kulturgeografi. Därtill är två utgivna av forskare som arbetar på Institutet för arbetsmarknadspolitisk utvärdering som har en ekonomisk inriktning.

Vid Linköpings universitet har institutionerna omorganiserats och Institutionen för pedagogik och psykologi (IPP) heter numera Institutionen för beteendevetenskap och lärande (IBL). Sammanlagt har åtta av de tolv avhandlingar som tagits med i materialet producerats inom IPP/IBL. De tre återstående avhandlingarna är framlagda vid Institutionen för samhälls- och välfärdsstudier¹² (Gruber 2007, Åström 2008), Institutionen för språk och kommunikation (Hofvendahl 2006)¹³ och Institutionen för Tema Barn (Axelsson 2007).

Göteborgs universitet är det lärosäte som under perioden 2006–2009 har flest beviljade UVK-ansökningar. Av de åtta projekten har sju gått till Institutionen för pedagogik och didaktik. Medel för fem projekt och två forskarassistenttjänster har gått till Umeå universitet. Här finns ingen enskild dominerande institution utan projektmedlen är spridda till olika institutioner. Sammanlagt fyra ansökningar vid Stockholms universitet har erhållit medel från UVK. De är fördelade på tre institutioner och ett institut. Ansökningarna avser tre projekt och en forskarskola. Uppsala universitet har beviljats tre ansökningar (två projekt och en forskarassistenttjänst). Till dem har vi också fogat två projekt från *Institutet för arbetsmarknadspolitisk utvärdering* (IFAU). Även högskolorna i Mälardalen och Malmö har beviljats medel för vardera två projekt. Flera lärosäten har ett projekt var. Det gäller Linköpings universitet, Luleå tekniska högskola, Mittuniversitetet, Högskolan i Skövde, Växjö universitet och Örebro universitet. I det här sammanhanget är det viktigt att notera en begränsning i redovisningen som bygger på huvudsökandes hemvist. I många fall samarbetar flera lärosäten, särskilt universitet och högskolor och därmed ökar också inslaget särskilt av högskolor som oftare är medsökande än huvudsökande.

¹² Med anknytning till lärarutbildningen vid Campus Norrköping.

¹³ Inom ramen för Arbetslivsinstitutets projekt Skolliv.

Då vi ser till fördelningen mellan pedagogikrelaterade institutioner och övriga, framgår att en tredjedel av dem som beviljats medel är institutioner som representerar andra samhällsvetenskapliga ämnen, men också humanistiska och naturvetenskapliga ämnen (kulturgeografi, tema etnicitet, IFAU, matematik/matematik och statistik, matematik, teknik och naturvetenskap, SOFI, idéhistoria och vetenskapsteori, statsvetenskap/humaniora och statsvetenskap).

5.1.2 Discipliner involverade i forskning om bedömning

För att bli mer specifika och för att urskilja i vilka disciplinära sammanhang forskning om bedömning produceras har vi kartlagt avhandlingsförfattarnas institutionstillhörighet. Den kan fungera som en god indikator på den disciplinära tillhörigheten, även om det inte är ett perfekt mått. Även om pedagogiska och didaktiska institutioner dominerar den bedömningsrelaterade forskningen är de inte ensamma aktörer på scenen. Vid sidan av att intresset för forskning om bedömning har ökat har också antalet discipliner som bedriver forskning relaterad till bedömning blivit fler.

Av tabell 5.3 framgår fördelningen av avhandlingar på institutioner och under perioden 1990–2009. Inom det pedagogiska fältet finns en rad olika institutionsnamn som i tabellen går under beteckningen pedagogikrelaterade institutioner.

TABELL 5.3 Fördelning av avhandlingar över tid och i relation till institutioner

1990–1999	
Pedagogikrelaterade institutioner	20
Andra institutioner	-
2000–2004	
Pedagogikrelaterade institutioner	12
Andra institutioner	2
Nordiska språk	1
Institutet för social forskning (SOFI)	1
Ekonomiska institutionen	

2005–2009	
Pedagogikrelaterade institutioner	40
Andra institutioner	21
Institutionen för matematik och statistik	2
Nordiska språk (eller liknande)	3
Nationalekonomi/ekonomi	6
Institutionen för statsvetenskap	1
Institutionen för socialt arbete	1
Sociologiska institutionen	2
Institutionen för idéhistoria och vetenskapsteori	1
Humanistiska institutionen	2
Institutionen för psykologi	1
Institutionen för kulturgeografi	1
Karolinska institutet/Odontologi	1

Under perioden 1990–2009 produceras totalt 95 avhandlingar, varav 72 på olika pedagogikrelaterade institutioner och 23 på institutioner som representerar andra discipliner. Pedagogikämnet dominerar scenen under 1990-talet och det tidiga 00-talet. Tydligt är att inslaget av avhandlingar från andra institutioner än pedagogikrelaterade ökar markant efter 2005 då 21 av dem publicerades. Dessförinnan är utompedagogiska avhandlingar relaterade till bedömning en marginell företeelse. Vi kan se ett ökat intresse framför allt från språkvetare och olika samhällsvetenskapliga discipliner som national/ekonomi, sociologi, statsvetenskap och kulturgeografi. Men enstaka exempel finns också på avhandlingar relaterade till bedömning inom humaniora och naturvetenskapliga discipliner, t.ex. idé- och lärdoms historia respektive matematik och statistik.

Utan att fördjupa oss i avhandlingarnas innehåll kan vi notera ytterligare några iakttagelser. I flera fall finns överlappningar. Det innebär att avhandlingar med likartat fokus återfinns inom olika discipliner. Avhandlingar som producerats på institutioner för nordiska språk eller liknande handlar om svenskämnets didaktik (t.ex. Ask 2007, Eklund Heinonen 2009, Lundström 2007, Norlund 2009, Palmér 2008). På liknande sätt framläggs vid Institutio-

nen för matematik och statistik i Umeå matematikdidaktiska avhandlingar (Boesen 2006, Taflin 2007). Avhandlingar med dessa inriktningar återfinns också vid flera pedagogikrelaterade institutioner. Flera exempel finns inom både svenska (Bergman 2007, Norlund 2009, Parmenius Swärd 2008) och matematik (Bjerneby Häll 2006). Den här typen av avhandlingar har med andra ord inte en självklar institutionsplacering och disciplinär tillhörighet.

Sett utifrån pedagogikämnet är intresset för utbildning inte förvånande då disciplinen har en mångfacetterad botten med rötter i såväl filosofi, psykologi och sociologi som historia (Härnqvist 1997). Ämnets identitet öppnar med andra ord för många intressenter (jfr Dahllöf 1990). Även det generellt ökade intresset för kunskap, lärande och utbildningens betydelse för samhället i stort och för enskilda individer kan ses som motiv för ett tilltagande engagemang i utbildningsfrågor. Dessutom har ju pedagogiken genom forskning riktad mot skolan en direkt koppling också till andra discipliner genom deras förbindelse med olika skolämnen. Till detta kan vi också foga den utmaning som framväxten av ett utbildningsvetenskapligt fält inneburit och som medfört att allt fler discipliner numera har möjlighet att bedriva forskning med skolan som studieobjekt (Fransson & Lundgren 2003, Lindberg 2002). Även omorganiseringen med nya gränsdragningar mellan institutioner och discipliner bidrar till ökad rörlighet. Sammanfattningsvis bidrar samtliga dessa faktorer på skilda sätt till att allt fler forskare vid svenska lärosäten inom olika discipliner är engagerade i frågor om bedömning.

5.2 Artiklar om bedömning i nordiska tidskrifter

Ett första steg på vägen mot att belysa hur svensk forskning relaterad till bedömning kommuniceras och sprids utåt i vetenskapliga sammanhang har varit att gå igenom vad som publicerats i de två nordiska tidskrifterna *Nordisk Pedagogik* och *Scandinavian Journal of Educational Research*.

Antalet artiklar i *Scandinavian Journal of Educational Research* (76) är mer än dubbelt så stort som antalet i *Nordisk Pedagogik* (33). Av de senare är knappt hälften författade av forskare eller doktorander vid svenska lärosäten. Skillnaden i antal kan delvis förklaras med att *Scandinavian Journal of Educational Research* utkommer med sex nummer per år medan *Nordisk Pedagogik* bara utkommer med fyra. Sammanlagt har svenska forskare skrivit 47 av de totalt 109 artiklarna om bedömning, medan forskare vid danska, finska, isländska och norska lärosäten producerat sammanlagt 62 artiklar. Sverige står med andra ord för drygt två femtedelar (43 procent) av samtliga artiklar om bedömning publicerade i dessa tidskrifter.

TABELL 5.4 Forskning sett utifrån artiklar i två nordiska tidskrifter 1990–2009. Absoluta tal

År	Nordisk Pedagogik	Scandinavian Journal of Educational Research
2009	2	4
2008	5	9
2007	7	5
2006	0	9
2005	1	6
2004	0	10
2003	0	1
2002	1	3
2001	0	3
2000	0	1
1999	0	2
1998	4	2
1997	0	0
1996	1	2
1995	1	4
1994	0	1
1993	2	6
1992	6	3
1991	3	4
1990	0	1
Σ	33	76

Om vi riktar uppmärksamheten mot fördelningen av artiklar över tid kan vi se vissa olikheter. Artiklarna i *Nordisk Pedagogik* fördelar sig t.ex. jämnt mellan perioderna 1990–1999 och 2000–2009 med 17 respektive 16 publicerade artiklar. Noteras bör dock att merparten av artiklarna återfinns i början av

1990-talet och i slutet av 2000-talets första decennium. Det förefaller som bedömning tidigare var en fråga som fick en viss uppmärksamhet, men sedan under en period är relativt frånvarande för att därefter återigen träda in på scenen. Det tidigare intresset hade som framgår ovan inte sin motsvarighet i avhandlingarna. Kanske kan förekomsten av artiklar under denna period förstås som ett resultat av studier genomförda under en tidigare period. Det kan med andra ord handla om en viss förskjutning i tiden mellan nationell och nordisk publicering.

För *Scandinavian Journal of Educational Research* ser däremot spridningen annorlunda ut. Under perioden 1990–2000 publiceras 21 artiklar om bedömning, det vill säga mindre än en tredjedel av samtliga. I likhet med *Nordisk Pedagogik* finns en tyngdpunkt både i det tidiga 1990-talet och mot den andra hälften av första decenniet på 2000-talet. En markant ökning sker 2004 och därefter är den genomsnittliga nivån något högre än tidigare. Tre möjliga förklaringar kan vara dels arten av avhandlingar som publiceras, dels relationen till internationella jämförande test, dels ett ökat tryck på forskare att publicera artiklar i internationella tidskrifter. Inom disciplinen pedagogik kompletteras de tidigare helt dominerande monografiavhandlingarna med sammanläggningsavhandlingar. Ökningen i antalet artiklar per år blir märkbar två år innan antalet avhandlingar ökar. Då artikelproduktionen i svenska, nordiska och internationella tidskrifter relateras till avhandlingar och författare framgår det att flera av artiklarna ingår i sammanläggningsavhandlingarna (jfr Berggren 2007, Klapp Lekholm 2008, Nyström 2003 och 2004a, Näsström 2008). Genomslaget för internationella kunskapsmätningar såväl i politiken och det offentliga samtalet som i forskningen har tidigare betonats. Här kan vi identifiera en ökning särskilt från och med 2003 (jfr Eklöf 2007, Fredriksson m.fl. 2009, Myrberg & Rosén 2006). Slutligen är det också möjligt att systemen för kvalitet och effektivitet som utvecklats vid högskolor och universitet och som premierar internationell publicering fått en viss effekt. Även om likartade processer i andra länder kan bidra till en viss trängsel och ökade svårigheter med att få artiklar publicerade.

5.3 Norden i internationella sammanhang

Sökningar efter svenska och nordiska artiklar om bedömning i tidskrifter med peer review har gjorts i tre databaser: ERIC, EBSCO och IBSS. Sökningar har också genomförts i databasen som är direkt kopplad till *Scandinavian Journal of Educational Research* och manuellt i *Nordisk Pedagogik*. Avgränsningarna avser perioden 1991–2008. Vi har tagit med artiklar som antingen är skrivna av svenska eller andra nordiska forskare eller som fokuse-

rar bedömningar i de nordiska ländernas utbildningssystem. Sökningar och avgränsningar redovisas och diskuteras utförligare i kapitel 2. Som ett resultat av sökningarna i de internationella databaserna har vi identifierat 279 artiklar. I tabell 5.5 presenteras fördelningen över tid och mellan länderna.

TABELL 5.5 Artiklar om bedömning i internationella tidskrifter 1991–2008. Absoluta tal

År	Sverige	Finland	Norge	Danmark	Island	Totalt
2008	28	13	8	7	1	57
2007	16	12	7	5	1	41
2006	24	16	8	4	2	54
2005	10	4	9	3	0	26
2004	10	13	5	5	0	33
2003	3	4	2	3	0	12
2002	2	5	1	3	1	12
2001	2	1	1	0	0	4
2000	1	0	1	0	0	2
1999	2	0	1	0	0	3
1998	3	1	0	0	0	4
1997	1	0	1	0	0	2
1996	2	2	1	0	0	5
1995	2	0	2	0	0	4
1994	3	0	0	0	0	3
1993	6	1	2	0	0	9
1992	3	0	0	0	0	3
1991	0	3	1	0	0	4
1990	1	0	0	0	0	1
Σ	119	75	50	30	5	279

Den tidigare identifierade skillnaden mellan de nordiska länderna återfinns också här på den internationella arenan. Av de 279 artiklarna avser 43 procent Sverige. Därpå följer i fallande ordning Finland (27 procent) Norge (18 procent), Danmark (11 procent) och till sist Island med endast fem artiklar. I linje med spridningen av avhandlingarna och till viss del också tidskrifterna har vi också här ett begränsat antal artiklar (14 procent) under 1990-talet. Intresset för forskning om bedömning växer kontinuerligt under 00-talet, i synnerhet under den senare delen. Åren 2000–2005 publiceras knappt en tredjedel (32 procent) av artiklarna. Den stora majoriteten (54 procent) finns i tidskrifter som utkom 2006–2008. Bilden är likartad i samtliga nordiska länder, även om Danmark och Island, som har få artiklar, saknar artiklar under 1990-talet och början av 2000-talet.

I likhet med vad vi tidigare framhöll om den ökande förekomsten av artiklar under slutet av perioden kan vi även här notera relationen till sammanläggningsavhandlingar. Detsamma kan sägas om de internationella jämförande mätningarna som har bidragit till att fler artiklar relaterade till bedömning produceras. Det gäller inte minst Finland där flera artiklar refererar till PISA om än på skilda sätt. Några använder dem för att förklara vilka faktorer som bidragit till Finlands framgångar i PISA (t.ex. Kivirauma & Ruoho 2007). Andra problematiserar både resultat och förklaringar (t.ex. Simola 2005). Dessutom finns det forskare som tar sin utgångspunkt i PISA för att uttala sig om något annat, t.ex. vilka frågor begåvade skolbarn i olika åldrar formulerar med fokus på framtiden (t.ex. Tirri & Nokelainen 2006).

Med de svenska artiklarna som exempel kan vi också konstatera att en övervägande majoritet av författarna endast har skrivit en artikel. Det är med andra ord inte många svenska forskare som återkommer ofta i den internationella debatten. Enskilt eller tillsammans med andra har 14 forskare publicerat två artiklar. Sex andra forskare har skrivit vardera tre artiklar och ytterligare tre har deltagit med fyra artiklar var. Sammanfattningsvis innebär det att 74, 17, 7 respektive 2 procent av författarna producerat en, två, tre respektive fyra artiklar om bedömning i de internationella tidskrifterna. Vi kan också notera att det är vanligt att artiklarna har två eller flera författare (63 procent) och att en majoritet (58 procent) av författarna är manliga.

6 BEDÖMNINGSFORSKNINGENS STUDIEOBJEKT

Att bestämma vad forskningen handlar om kan göras med delvis olika utgångspunkter. Här fokuserar vi bedömningsforskningens innehåll definierat som dess studieobjekt. Närmare bestämt handlar det om vad forskarna uppmärksammar i sina studier, objekten de studerar för att bli klokare på bedömning, vårt kunskapsobjekt (jfr Törnebohm 1971 och Forsberg 2000). Med analogin karta och territorium avser "kunskapsobjektet den karta utifrån vilken ett visst territorium studeras" och studieobjektet "det territorium som studeras" (Fransson & Lundgren 2003 s. 84). Med detta som utgångspunkt riktar vi här således intresset mot de skilda studieobjekt som forskarna fokuserar i avhandlingar och artiklar. Det handlar om ett antal områden som "är avgränsade på olika sätt, ofta utifrån administrativa indelningar som skolform, administration, rekrytering" (Lindblad, Kyndel & Larson 2004). De har sammanställts i form av listor med sammanfattande begrepp, som sedan jämförts sinsemellan med avseende på likheter och skillnader. Resultatet av jämförelsen är de åtta huvudsakliga kategorierna av studieobjekt som presenteras översiktligt i tabell 6.2. I kapitel 7 konkretiseras kategorierna utifrån avhandlingarna och i kapitel 8 utifrån UVK-projekten. I kapitel 9 för vi analysen ett steg vidare genom att identifiera olika aspekter av kunskapsobjektet bedömning.

Men det finns anledning att än en gång göra läsaren uppmärksam på den relativt generösa avgränsning vi använder för att identifiera bedömningsforskningen. Inkluderad är både forskning som studerar bedömning som fenomen och forskning som använder resultat från bedömningar för att studera andra fenomen. I det senare fallet är det inte bedömning som utgör forskningens kunskapsobjekt. I det perspektivet utgör den inte bedömningsforskningens kärna, utan snarare dess periferi. Vi har dock valt att ta med denna forskning på grund av den avgörande betydelse resultatmåttan har för den kunskap som utvecklas inom traditioner som nyttjar betyg och provresultat som mått på framgång eller misslyckande. Ett annat argument för att ta med den typ av forskning som använder sig av resultatmått hämtas från ett av den klassiska bedömningsforskningens begrepp nämligen validitet. Samuel Messick (1989) har i en ofta citerad artikel introducerat ett vidgat validitetsbegrepp, där han argumenterat för att validitet måste ses i relation till två dimensioner. Han pekar å ena sidan ut utgångspunkterna för validering av provets utformning, dess kvalitet i förhållande till vad som ska

bedömas (provets inferenser respektive konsekvenser), och provets resultat och funktioner (hur provresultaten tolkas respektive används) å den andra. Det sistnämnda handlar om hur våra uppfattningar och värderingar påverkas av hur resultaten tolkas och används. En del av den forskning vi tagit med har finansierats med medel från UVK, sannolikt bland annat för att den använder betygsresultat (men inte provresultaten) som utgångspunkt för studierna. Vi menar att just att ta med dessa studier, som vi bedömt som perifera men ändå relevanta, reser frågor om vilka de möjliga konsekvenserna är av de bedömningar som görs.

Analyserna utgår från de respektive materialens abstract eller sammanfattningar. Som framgår i kapitel 2 skiljer sig dessa åt kvalitetsmässigt. Tyngdpunkten är därför förlagd till avhandlingarnas abstract, medan övriga används för att komplettera och kontrastera dessa. Det innebär konkret att analysernas resultat, skilda kategorier av studieobjekt, har sin grund i avhandlingarna och att artiklarna och projekten sorterats med utgångspunkt i dessa.

Inledningsvis ger vi en bild av spridningen i förhållande till grundskolan och gymnasieskolan samt övergångar till och från dessa. Därefter redovisar vi kategorierna med bedömningsforskningens studieobjekt. Efter en kortfattad presentation av kategorierna presenterar vi en översiktsbild av hur fördelningen ser ut för materialet i sin helhet. På detta följer en kortare utveckling om de svenska och nordiska artiklarna i de internationella tidskrifterna. Avhandlingarna och UVK-projekten redovisas mer ingående i kapitlen 7 och 8.

6.1 En översiktsbild

Om vi börjar med en granskning av vilka skolformer och vilka övergångar som uppmärksammas inom bedömningsforskningen kan vi se stora skillnader både i en generell mening och mellan publikationsmaterialen. Generellt gäller att bedömningsforskningens fokus ligger på grundskolan. Få studier gäller övergångar, särskilt till och från grundskolan. Det är inte heller särskilt vanligt med artiklar som studerar utbildningssystemet, åtminstone inte med referens till bedömning eller elevers prestationer och skolors resultat. Därefter föreligger skillnader beroende på om vi är inriktade på avhandlingarna, artiklarna eller projekten. Men variationer finns också mellan artiklar publicerade i svenska, nordiska eller övriga internationella tidskrifter. Fördelningen framgår av tabell 6.1. I de fall en avhandling, en artikel eller ett projekt handlar om flera skolformer alternativt både skolformer och övergångar har det utgjort grund för flera markeringar. Det totala antalet markeringar överstiger därför antalet studerade objekt (abstract/sammanfattningar) i respektive material.

TABELL 6.1 Bedömningsrelaterad forskning per skolform. Procent.

	Avhandlingar N=102	Artiklar i ... tidskrifter			UVK-projekt N=43	Totalt N=466
		Svenska N=55	Nordiska N=74	Internationella N=235		
FÖRSKOLAN/ GRUNDSKOLAN	1	0	1	2	7	3
GRUNDSKOLAN	54	42	70	57	56	57
GYMNASIESKOLAN	20	6	3	15	14	13
GRUNDSKOLAN/ GYMNASIESKOLAN	3	10	4	2	7	4
GYMNASIESKOLAN/ HÖGSKOLAN/ARBETE	12	30	7	9	7	11
UTBILDNINGSSYSTEM/ ÖVERGRIPANDE	10	12	15	15	9	12
Σ %	100	100	100	100	100	100

Vissa material har relativt få markeringar och det krävs därför en viss försiktighet vid tolkningen av data. Det gäller främst antalet markeringar för artiklar i svenska tidskrifter och UVK-projekt som totalt har 55 respektive 43 markeringar. Små skillnader kan framträda som mer omfattande och vi riktar därför intresset mot de stora variationerna. Bilden är likartad för avhandlingarna, de nordiska artiklarna i de internationella tidskrifterna och UVK-projekten. De har främst fokus på grundskolan, men de har också en något större uppmärksamhet på gymnasieskolan. I de svenska och nordiska tidskrifterna lyser däremot gymnasieskolan nästan helt med sin frånvaro. Övergångar framför allt mellan gymnasieskolan och högskolan är mer frekventa i svenska tidskrifter än i övriga publikationer. Tilläggas kan också att dominansen för studier av grundskolan är särskilt tydlig i de nordiska tidskrifterna.

6.1.1 Studieobjekten – beskrivningar och fördelningar

Forskningen om bedömning eller forskning med referens till bedömningsresultat har kategoriserats i åtta grupper. Såväl antalet som valet av kategorier kan diskuteras. Det handlar inte om några absoluta kategorier, gränserna är delvis flytande och överlappningar förekommer. Åtta olika studieobjekt har identifierats. De beskrivs kortfattat i tabell 6.2.

TABELL 6.2 Beskrivningar av studieobjekten

<p>LÄRARES DIDAKTISKA BEDÖMNINGSPRAKTIK: studier som placerats i denna kategori utmärks av att de studerar lärares didaktiska överväganden relaterat till bedömning. Det kan då handla både om vad som bedöms, hur det bedöms och uppfattningar och inställningar till bedömningar.</p>
<p>ELEVERS PRESTATIONER OCH SKOLORS RESULTAT: i dessa avhandlingar fokuserar eleverens resultat och/eller skolors resultat. De används som ett led i att fastställa antingen vad mätningarna avser och vad de påverkas av eller hur resultaten påverkar eleverna i ett framtidsperspektiv.</p>
<p>SKOLANS STYRNING OCH LÄRARENS ARBETE: i studierna som ingår här uppmärksammas den statliga styrningen och hur den inramar lärares arbete och med vilka konsekvenser. Ibland handlar det om ett styrinstrument som är direkt knutet till bedömning som betygskriterier. Men ofta är det fråga om mer generella fenomen med konsekvenser för bedömning.</p>
<p>ÖVERGÅNGAR: studier av urval och rekryteringsprocesser samt de bedömningsinstrument som brukas i samband med dessa t.ex. behörighetsregler, betyg, högskoleprov, särskilda urvalsprov. Hit hör också studier av elever som agenter i val av skola, program, etcetera.</p>
<p>ORGANISERING AV SKOLA, UNDERVISNING OCH ELEVER: här finns studier som beskriver processer och dokument som används för att urskilja vilka elever som ska ingå i vilka grupper och hur detta samordnas med organiseringen av skolans innehåll och utformning.</p>
<p>BEDÖMNINGAR SOM FENOMEN: studier som förlagts till denna kategori har ett övergripande fokus på bedömningar och intresserar sig för dem som pedagogiska redskap i skolan, som instrument för styrning av utbildningen, som ett instrument för forskningen att kartlägga skolan och/eller som ett sätt att reglera relationen mellan såväl skolan och samhället som individen och samhället.</p>
<p>ELEVERS ERFARENHETER OCH UPPLEVELSER AV BEDÖMNING: i denna grupp ingår avhandlingar som studerar eleverens erfarenheter och upplevelser av olika former av bedömningar såväl i en mer generell mening som i relation till specifika bedömningsredskap inom olika ämnen.</p>
<p>BEDÖMNINGSPRAKTIKENS KVALITET: avhandlingar som placerats i denna kategori fokuserar i vilken omfattning bedömningar är tillförlitliga och till exempel prov förmår mäta det de gör anspråk på att mäta. Det kan avse både forskarnas och administrationens mätinstrument och lärarnas redskap för bedömning.</p>

Med dessa kategorier som grund för bedömningsforskningens studieobjekt har vi analyserat abstract och sammanfattningar från samtliga material. De analyserade texterna är inte internt enhetliga. Det gäller särskilt sammanläggningsavhandlingarna, men ibland också monografierna och artiklarna liksom projekten som kan bestå av flera delstudier. Inordnandet av respektive sammanfattning/abstract har gjorts med utgångspunkt i var tyngdpunkten ligger, men med viss hänsyn till den interna variationen. Det innebär att några objekt kategoriserats i flera kategorier. Precis som i tabell 6.1 finns det således en skillnad vad avser det totala antalet markeringar för respektive material jämfört med antalet studerade objekt. Fördelningen i procent framgår av tabell 6.3.

TABELL 6.3 Fördelningen av avhandlingar, artiklar och projekt över studieobjekt. Procent

	Avhandlingar N=109	Artiklar i ... tidskrifter			UVK-projekt N=40	Totalt N=519
		Svenska N=73	Nordiska N=60	Internationella N=237		
LÄRARES DIDAKTISKA ARBETE	23	11	9	11	15	15
ELEVERS OCH SKOLORS RESULTAT	18	21	46	47	28	35
SKOLANS STYRNING/ LÄRARES ARBETE	12	7	5	6	12	8
ÖVERGÅNGAR	12	15	5	4	5	8
ORGANISERING	12	14	10	10	12	11
BEDÖMNING SOM FENOMEN	11	6	7	9	10	8
ELEVERS ERFARENHETER/ UPPLEVELSER	9	7	11	8	8	8
BEDÖMNINGSRÄDSKAPENS KVALITET	3	19	7	5	10	7
Σ %	100	100	100	100	100	100

En blick på helheten visar att en majoritet av artiklarna handlar om elevers prestationer och skolors resultat. Därutöver är fördelningen mellan de skilda studieobjekten relativt jämn även om det är en något större andel artiklar som behandlar lärares didaktiska arbete. Det finns variationer mellan texter publicerade i olika tidskrifter, men likheterna är stora. I samtliga fall utom i de svenska tidskrifterna gäller att de två redan nämnda kategorierna dominerar. Artiklarna i de svenska tidskrifterna har, jämfört med övriga, en större andel artiklar som behandlar bedömningsredskapens kvalitet. Förutom i avhandlingarna är elevers prestationer och skolors resultat den största kategorin. Avhandlingarna liksom artiklarna i de svenska tidskrifterna utmärks av en jämnare fördelning mellan studieobjekten. Även UVK-projekten ligger åt det hållet, trots en tyngdpunkt i kategorin elevers prestationer och skolors resultat. Artiklarna i både de nordiska och de internationella tidskrifterna har, jämfört med övriga, den starkaste kopplingen till elevers prestationer och skolors resultat. Nästan hälften av artiklarna hör till denna kategori. Vi redogör här kortfattat för innehållet i artiklarna i tidskrifterna.

6.2 Artiklar i svenska, nordiska och internationella tidskrifter

Inledningsvis kan vi konstatera att många forskare publicerar enstaka artiklar och att variationen av tidskrifter med artiklar är stor. Flera svenska artiklar i tidskrifterna, särskilt efter 2004, hör ihop med sammanläggnings-avhandlingar. En betydande del av dem fokuserar elevers prestationer och skolors resultat med utgångspunkt i pedagogiska/psykologiska perspektiv. Till detta perspektiv kan vi även knyta många artiklar som handlar om bedömningsredskapens kvalitet. Det kan t.ex. handla dels om att söka förklaringar till vilka faktorer som bidrar till provresultat och/eller betyg och i vilken omfattning (t.ex. Klapp Lekholm & Cliffordson 2009), dels om att separera individuella prestationer från skolprestationer (Åberg-Bengtsson & Erickson 2006).

Antalet artiklar som behandlar bedömning som fenomen är få. Men i ett exempel jämförs de svenska husförhörslängderna med ungerska kyrkoregister (Zipernovsky 2003). Ett annat exempel går in på formativ bedömning, en fråga som fått särskild uppmärksamhet internationellt sedan slutet av 1990-talet. Bertil Roos och David Hamilton (2005) beskriver den formativa bedömningens historiska och teoretiska rötter som kan spåras tillbaka till 1930- och 1940-talen. Till artiklar som behandlar bedömning i förhållande till lärares didaktiska arbete hör en studie om vilka lärar- och klassrumsfaktorer som bidrar till framgångsrik läsinlärning (Damber 2009). En annan fokuserar på lärares föreställningar om kursplanens mål i idrott och hälsa och vilka förmågor lärarna värderar när de talar om betygskriterier (Redelius, Fagrell & Larsson 2009). Innehållsfrågan med avseende på icke-kognitiva aspekter behandlas i Per Gerrevalls (2003) artikel *Bedömning av demokratisk kompetens – en pedagogisk utmaning*. Även Birgitta Garne (2003) har fokus på demokrati men i relation till svenskämnet i artikeln *Skolan, proven och demokratin*. De svenska artiklarna som berör temat skolans styrning-lärares arbete har tyngdpunkten på styrningsaspekten medan lärares arbete behandlas perifert eller implicit (t.ex. Osborn 2006). Ett undantag från detta är Inger Erikssons (2009) arbete relaterat till det svenska Timplaneprojektet.

Beträffande artiklar i tidskrifter relaterade till de övriga nordiska länderna har vi redan konstaterat att de isländska artiklarna är synnerligen få. För Danmarks del är de flesta artiklarna relaterade till resultaten från de internationella kunskapsmätningarna (t.ex. Rangvid 2007) eller till dessa i en mer generell mening (Egelund 2008). Utöver dem finns det ytterligare några artiklar inom den psykometriska/statistiska traditionen. Artiklarna är

snarlika motsvarande artiklar i Sverige: resultatmått används som en variabel i sambandsanalyser. Även i Danmark finns det ekonomer som bedriver studier med grund i resultatmått på elevers prestationer (t.ex. Browning & Heinesen 2007). Det innebär dock inte att artiklarna är särskilt många. Övriga danska artiklar har inget samlat fokus, utan framstår närmast som enstaka artiklar.

Bland de finska artiklarna är det ganska många som relateras till eller motiveras av de finska PISA-resultaten. Några gör reanalyser i syfte att förklara resultaten (t.ex. Linnakylä 2004, Rautalin & Alasuutari 2007, Välijärvi 2003). Andra problematiserar dem i en kritisk belysning (t.ex. Simola 2005). Även forskare från andra länder använder de finska PISA-resultaten i komparativa studier (t.ex. Ammermueller 2007). Några artiklar är också relaterade till TIMSS. Även för Finlands del gäller att flera artiklar är riktade åt många olika håll. Några artiklar har fokus på bedömning som fenomen, t.ex. Hannu Simola (2002) som analyserar ordningsbetygen i den finska skolan ur ett historiskt perspektiv.

Endast ett fåtal artiklar handlar om elevers uppfattningar om, upplevelser av och inställning till bedömning och betyg. I en artikel om övergången till grundskolan i Finland undersöks elevers attityder till, motivation för och upplevelser av (prov)uppgifter (Aunola, Leskinen & Nurmi 2006). En annan ingång till elevers upplevelser representeras av en finsk studie (Kärkkäinen, Rätty & Kasanen 2008) som studerar hur elevers erfarenheter av och prestationer i olika skolämnen bidrar till att forma deras uppfattningar om sig själva. Ofta relateras elevernas attityder till deras prestationer. Ett exempel på detta är en norsk enkätstudie som beskriver sambandet mellan elevers betyg och sociala passivitet i relation till skolämnen och lärare (Paulsen & Bru 2008). I övrigt bekräftar artiklarna om Norge trenden i de andra nordiska länderna. Även här finns PISA-artiklar (t.ex. Hivistendahl & Roe 2004, Schleicherf 2006). Mönstret är detsamma frånsatt att andelen psykometriska artiklar är mindre. Om vi bortser från antalet artiklar finns många likheter mellan de nordiska länderna, det gäller både när bedömningsforskningen tar fart på den nordiska och internationella arenan liksom forskningens innehållsliga fokus.

7 SVENSKA AVHANDLINGAR 1990–2009

Här redovisar vi hur svenska avhandlingar under perioden 1990–2009 kan beskrivas med hjälp av de i kapitel 6 presenterade kategorierna för forskningens studieobjekt. I kartläggningen ingår 95 avhandlingar som antingen direkt fokuserar bedömningar av skilda slag eller använder t.ex. nationella prov eller betyg för att uttala sig om skolans standard, utveckling eller något liknande. Exempelen är valda så att de ska peka ut både det typiska och variationen inom respektive kategori.

7.1 Lärares didaktiska bedömningspraktik

Den största andelen av avhandlingarna, knappt en fjärdedel, finns inom denna kategori. Gemensamt är en fokusering på lärares didaktiska överväganden inom olika bedömningspraktiker. Allmäntdidaktiska och ämnesdidaktiska avhandlingar utgör underkategorier och den senare dominerar klart. Som exempel kan nämnas *Gymnasieskolans svenskämnen. En studie av svenskundervisningen i fyra gymnasieklasser* (Bergman 2007). Den handlar om hur svenskämnet tar form i fyra gymnasieklasser i flera olika program. Undervisningen i klasserna följs under två år och kompletteras med intervjuer av lärare och elever. Vad som framstår som viktigt och vad som bedöms i de fyra klasserna utgör en aspekt av hennes resultat. I *Godkänd i svenska? Bedömning och analys av gymnasieelevers texter* (Östlund-Stjärnegårdh 2002) är bedömning mer i förgrunden. Skillnaderna mellan godkända och icke godkända elevtexter analyseras liksom lärares och externa bedömares omdömen. Dessutom ingår en enkätstudie som fokuserar lärares motiveringar till gränsdragningen mellan godkänt och icke godkänt betyg. Maria Andrée (2007) följer i *Den levda läroplanen. En studie av naturorienterande undervisningspraktiker i grundskolan* NO-undervisningen i två klasser i grundskolan under ett år. Utifrån observationer, ljudinspelade lektioner och intervjuer med lärare och elevgrupper beskriver hon två undervisningspraktiker: den kriteriestyrda, där eleverna deltar i produktionen och reproduktionen av rätta svar som grund för godkänt betyg, och laborationspraktiken som handlar om att utveckla begrepp och begreppsliga relationer. De allmäntdidaktiska avhandlingarna studerar fenomen som förekommer i flera ämnen. *I ett fall, Milda makter! Utvecklingssamtal och loggböcker som*

disciplinerings tekniker (Granath 2008), fokuseras hur lärares bedömningar kommuniceras till elever och föräldrar genom utvecklingssamtal.

Avhandlingarna från tidigt 1990-tal skiljer sig från dem som publicerades under 2000-talet genom att bedömningsfrågan är mindre framträdande. I stället framstår frågor om val av innehåll, material eller metoder som mer centrala. Det är vanligare att tala i termer av kunskapsutveckling och lärande än explicit om bedömning. Som exempel kan nämnas *Matematisk operativitet: En analys av relationen mellan form och innehåll i skolmatematik*. (Bergsten 1990) och *Logoprogrammering på mellanstadiet: en studie av fördelar och nackdelar med användning av Logo i matematikundervisningen under årskurserna 5 och 6 i grundskolan* (Hedréen 1990). Med en snävare urvalsgrund för avhandlingarna hade dessa inte kommit med i vårt material. Men eftersom de belyser viktiga aspekter av lärares bedömningsarbete har vi tagit med dem.

7.2 Elevers prestationer och skolors resultat

I knappt en femtedel av avhandlingarna behandlas elevernas prestationer och samband mellan dessa och andra variabler. Det gemensamma är att de använder ett, ibland flera resultatmått: betygens meritvärden, resultaten från nationella prov eller från internationella mätningar. Studierna har sin hemvist främst i mättraditioner relaterade till såväl pedagogik/psykologi och ekonomi som sociologi. I flera studier finns ett intresse för sambandet mellan betyg/prestationer och sociala faktorer som klass, genus och etnicitet. Inom pedagogik/psykologi uppmärksammas både kognitiva och icke-kognitiva faktorer. I *Grades and Grade Assignment, Effects of Student and School Characteristics*. Kombinerar Alli Klapp Lekholm (2008) icke-kognitiva faktorer som intresse, motivation och föräldrastöd med olika skolfaktorer för att analysera vad som bidrar till elevers resultat. Här kan också nämnas Christina Wikström (2005b) som i sin avhandling *Criterion-referenced Measurement for Educational Evaluation and Election* undersöker skillnader i resultat från skolor med olika huvudmän. En avhandling inom den ekonomiska traditionen är Anna Nilssons (2005) *Indirect Effects of Unemployment and Low Earnings: Crime and Children's School Performance* (sammanläggningsavhandling). I en av delstudierna analyseras sambandet mellan föräldrars arbetslöshet och elevers betygsresultat.

Tre avhandlingar skiljer sig något från övriga i den här kategorin. Den ena är *Segregation och karriärposition: En studie av bostadsomgivningens betydelse för utbildning, sysselsättning och inkomst bland yngre i stockholmsregionen* (Sundlöf 2008) som representerar en för forskning relaterad till bedömning ovanlig disciplin, kulturgeografi. Elevernas resultat relateras här till

bostadsområdet och dess betydelse. En annan är Staffan Stenhags (2007) licentiatavhandling, *Vad säger matematikbetyget? En kvantitativ studie av 2079 elevers betyg i årskurs nio*. Studien handlar om vilka argument som används för att motivera matematikämnets plats i utbildning och hur de har förändrats över tid. Vidare prövar han hållbarheten i dagens argument genom en statistisk analys: kan grundskolans matematikbetyg användas för att predicera studieresultat i andra teoretiska ämnen. Ytterligare en avhandling "Att göra sina uppgifter, vara tyst och lämna in i tid". Om elevansvar i det högmoderna samhället (Söderström 2006) kan ses som ett särfall genom att den använder sig av en kvalitativ ansats. Studien analyserar ansvarets, en icke-kognitiv faktor, betydelse för godkänt betyg.

7.3 Skolans styrning och lärares arbete

Det som speciellt karakteriserar avhandlingarna i gruppen är relationen mellan styrningen av skolan och lärares arbete. Flera av avhandlingarna i gruppen har således varken bedömning eller betyg som ett ingångsintresse, men genom att bedömning är en aspekt av lärarbetet bidrar även dessa studier till att belysa olika aspekter av bedömning. Oavsett om studierna avser lärares dilemman, etnicitet, lärares kunskapsbehov eller formulering av lokala betygskriterier relateras de till styrningen av skolan.

Exempel på avhandlingar inom den här gruppen är En rolig dans? Svenska skolors första tolkning av innebörden i lokala betygskriterier i tre ämnen för skolår åtta och Att kunna klara sig i ökad natur: En studie av betyg och betygskriterier – historiska betingelser och implementering av ett nytt system (Tholin 2003, 2006).¹⁴ I dessa analyseras lärarnas tolkning av styrdokument för bedömning genom skolornas lokala betygskriterier i kemi, engelska samt idrott och hälsa på 93 skolor (2003) och 85 skolor (2006) och förändringarna av kriterierna över en femårsperiod. I Ann-Sofi Wedins (2007) avhandling *Lärares arbete och kunskapsbildning. Utmaningar och inviter i den vardagliga praktiken* studeras lärares dagliga arbete, vad de behöver kunna och hur de lär sig. Av den empiri som produceras genom samtal med två lärare och fältstudier under grundskolans senare år framgår att bedömning och betygsättning är ett återkommande tema som lärarna uppmärksammar stort.

I några avhandlingar finns en förskjutning från lärares arbete till mer organisatoriska aspekter relaterade till skolan. De kan därför ses som en

¹⁴ Tholin bidrar med två avhandlingar i och med att han 2003 avlade licentiatexamen och disputerade 2006.

undergrupp. Hit hör t.ex. Linda Rönnbergs (2007a) studie *Tid för reformering. Försöksverksamheten med slopad timplan i grundskolan* inom Timplanprojektet. Den syftade till att analysera hanteringen, från initiering till genomförande och konsekvenser, av försöksverksamheten med slopad timplan i grundskolan. Inte heller här står frågor om bedömning och betyg i centrum, men de aktualiseras likväl som en konsekvens av empirin. Drygt en tiondel av avhandlingarna beaktar frågor om skolans styrning och lärares arbete.

7.4 Övergångar

Avhandlingarna i denna kategori är inriktade på bedömningar i samband med övergången mellan olika skolformer, men också mellan skolan och arbetslivet. Det kan å ena sidan avse behörighetsregler, det vill säga vilka som får tillträde till högskolan. Å andra sidan kan det handla om vilka elever och grupper som väljer högskolan eller vilken högskola de väljer. Avhandlingarna beaktar också grunderna för urvalet av sökande (högskoleprov, gymnasiebetyg, särskilt urvalsprov). Frågorna studeras mest frekvent i relation till övergången mellan gymnasieskolan och högskolan, men det finns också studier som uppmärksammar övergången mellan grundskolan och gymnasieskolan. I *Marknad och medborgare: elevers valhandlingar i gymnasieutbildningens integrations- och differentieringsprocesser* (Lund 2006) studeras elevers val av gymnasieskola och -program samt hur de hanterar den programkultur de möter under den första tiden i gymnasiet. Detta diskuteras i termer av integrerings- och differentieringsprocesser inom gymnasieskolan. Där ingår även elevers val av t.ex. vilket betyg de vill arbeta för i olika kurser.

Utveckling och utvärdering av ett särskilt urvalsprov för tandläkarutbildningen studeras av Karin Röding (2005) *University Admission Based on Tests and Interviews: Implementation and Assessment*. Det handlar om metodutveckling och frågan är hur studenter som antagits genom urvalsprov klarar studierna i jämförelse med dem som antagits på traditionellt sätt (betyg eller högskoleprov). Empirin är relaterad till tre antagnings-tillfällen och uppföljningen består av såväl resultat på examinationer av olika slag som omdömen från lärarna. En annan studie av ett urvalsprov är relaterad till Grekland. Luciana Benincasas (1995) avhandling *A Journey, a Struggle, a Ritual: Higher Education and the Entrance Examinations in a Greek Province Town* kan ses som en kontrast till de svenska avhandlingarna. Den handlar om den kulturella innebörden av att delta i det grekiska antagningsprovet till högskolan, som det upplevs av deltagare och deras föräldrar i en grekisk landsortsstad. Studien bygger på observationer och djupintervjuer.

I avhandlingen *Varför inte högskola? En longitudinell studie av olika faktors betydelse för studiebegåvade ungdomars utbildningskarriär* (Hammarström 1996) studeras ungdomar som (inte) har börjat studera på högskolan. Centralt är motiven för personer med höga gymnasiebetyg som väljer att påbörja eller avstå från högre utbildning. Studiens empiriska material bygger på en enkätstudie till 2 000 personer i tjugofemårsåldern. Drygt en tiondel av avhandlingarna beaktar övergångar av något slag.

7.5 Organisering av skola, undervisning och elever

Med resultaten från bedömningar rangordnas och sorteras elever i olika grupper som tilldelas/erbjuds mer eller mindre o/lika läroplaner. Majoriteten av den dryga tiondelen avhandlingar som placerats i denna kategori är publicerade efter 2006. Flera uppmärksammar makt- och genusaspekter. Thom Axelsson (2007) studerar ordnandet av elever i ett historiskt perspektiv i avhandlingen *Rätt elev i rätt klass. Skola, begåvning och styrning 1910–1950*. Dokumentationens funktion i dessa differentieringsprocesser lyfts fram i flera avhandlingar. I *Dokumenterat utanförskap – Om skolbarn som inte når målen* följer Gunilla Ingestad (2006) 77 elever i en kommun. I studien används en kombination av kvalitativa och kvantitativa data: intervjuer med elever och skolpersonal, dokumentanalyser och prov- och betygsresultat från olika bedömningstillfällen under skoltiden. I *Fostran till andrarang: en studie av dominansprocessen vid skolstart och via vägar genom utbildningssystemet ur ett kvinnovetenskapligt perspektiv* (Käller 1990) genomförs en observationsstudie i syfte att urskilja olika dominansprocesser i klassrumsinteraktionen.

7.6 Bedömningar som fenomen

Avhandlingarna som ryms inom denna grupp förenas av sitt intresse för kunskapsbedömningar eller betyg som fenomen. Ofta används ett historiskt perspektiv. I en delstudie i den tidiga avhandlingen *Relativa betyg: Några empiriska studier och en teoretisk genomgång i ett historiskt perspektiv* studerar Håkan Andersson (1991) betygets historiska bakgrund och framför allt den utveckling som ledde fram till att de relativa betygen infördes i svensk skola. Ett annat exempel är Wolmar Christensson (1995) som har en kunskapsfilosofisk ingång i *Subjektiv bedömning som besluts- och handlingsunderlag*. Den implicita kunskapssynen, det vill säga idén om att kunna åstadkomma objektiva kunskapsbedömningar, i det moderna projektet

problematiseras. Genom att relatera subjektivitetsbegreppet till profession och praktikgrundad kunskap är avsikten att öppna för andra sätt att förstå det som i den positivistiska traditionen tolkas som subjektivism.

Ytterligare tre exempel med en mer kunskapssociologisk och/eller idéhistorisk ingång kan nämnas. I licentiatavhandlingen *Kunskapsbedömning – hur, vad och varför* ger Helena Korp (2003) både en historisk beskrivning av bedömning i svensk skola och en översikt av forskningsläget inom den anglosachsiska världen. Hon relaterar förändringar i bedömningsforskningen till tre teoretiska perspektiv på lärande med skilda kunskapsteoretiska antaganden. Hon pekar på behovet av att skilja mellan bedömning på individ- och systemnivå. I *Viljan att veta vad andra vet. Kunskapsbedömning i tidigmodern, modern och senmodern skola* (Lundahl 2006) analyseras bedömningar ur ett kunskapssociologiskt och läroplansteoretiskt perspektiv. Genom textanalyser av historiska material som är relaterade till politik, forskning och pedagogisk praktik analyseras hur de samspelar i formeringen, manifesteringen och reformuleringen av skolans bedömningspraktiker. Bedömning uppmärksammas därmed som såväl ett pedagogiskt redskap som ett administrativt instrument. Avhandlingen *Likhet uten solidaritet? Idéhistoriske studier av karakterer i utdanning og meritokrati* (Aas 2006) är idéhistorisk. Den handlar dels om bruket av betygssystem i skola och utbildning, dels om meritokrati och meritokratiska normer. Den grundläggande idén bakom meritokrati, som systemet för betyg är inbäddat i, problematiseras. En annan avhandling i samma kategori är *Psykoteknik: kulturell fabricering av personlig identitet* där Rikard Eriksson (1999) fokuserar psykoteknik som fenomen i ett sociokulturellt perspektiv. Psykoteknik studeras med avseende särskilt på hur psykoteknik bidrar till att normer och värderingar upprättas som griper in i människors liv och livsstil och i deras relation till samhället utifrån kunskapsteori, biografi, etnografi och historia. Ungefär en tiondel av avhandlingarna återfinns inom denna kategori.

7.7 Elevers erfarenheter och upplevelser av bedömning

Elevers erfarenheter och upplevelser av bedömningar behandlas i knappt en tiondel av avhandlingarna. Såväl elevgrupperna som sammanhangen (skolform, skolämne, typ av bedömningsituation och bedömningens funktion) varierar mellan avhandlingarna. Likaså kan studierna beröra bedömning som generellt fenomen i skolan eller handla om specifika redskap för bedömning. Ett exempel på en avhandling av mer generell karaktär är *Sko-*

lan som barnarbete och utvecklingsprojekt: en studie av hur grundskoleelevers arbetsmiljö skapas–förändras–förblir som den är av Jan-Olof Hellsten (2000). Elevernas arbetsmiljö i grundskolan studeras genom deltagande observation. En aspekt av elevernas skolvardag är läxförhören – en traditionell klassrumssituation där lärare ofta bedömer elevers kunskaper muntligt och eleven vanligtvis får omedelbar respons.

I en annan avhandling, *Living Away from Blessings. School Failure as Lived Experience* av Carina Henriksson (2004), studeras elevers upplevelser av misslyckanden i skolan. Elever från gymnasieskolans individuella program och elever på ungdomsvårdsskolor har kontaktats. De har sedan skrivit brev om situationer i skolan där de upplevt att de misslyckats. Breven kompletterades med intervjuer. En aspekt av resultatet är att eleverna sällan nämner kognitiva misslyckanden, i stället handlar en stor del av det empiriska materialet om misslyckanden relaterade till elevernas beteende. En annan avhandling, *Upplevd kompetens: en fenomenologisk studie av ungdomars upplevelser av sin egen kompetens i skolarbetet* av Maj-Lis Hörnqvist (1999), visar att inte heller kompetens är något ungdomar enbart eller ens i första hand relaterar till skolämnen utan även i detta fall har de bredare referensramar.

Jessika Karlsson (2008) undersöker i *A Novel Approach to Teaching Emotional Expression in Music Performance* om feedback kan bidra till att utveckla det känslomässiga uttrycket i elevers och musikers musicerande. I studien prövas ett datorprogram på en grupp, en annan grupp får feedback av en lärare. Sedan jämför hon hur eleverna/musikerna upplever den feedback de får av datorprogrammet respektive av läraren.

7.8 Bedömningsredskapens kvalitet

Ytterst få avhandlingar har särskilt fokus på bedömningsredskapens kvalitet, endast tre procent. Centralt är här frågor om tillförlitlighet i termer av validitet och reliabilitet. Innebörden i begreppen har förändrats över tid. I två avhandlingar behandlas olika aspekter av validitetsbegreppet. Simon Wolming (2000) problematiserar i *Validering av urval* om prognostisk validitet i förhållande såväl till urvalet till högskolan som till hur prestationer i högskolan mäts. Han pekar på alternativa sätt att hantera validitetskravet i relation till urvalsprovet. Han visar att prestationerna i högskolan påverkas av en mängd andra faktorer än dem provet avser att mäta, vilket ställer krav på andra (eller kompletterande) ansatser. I Peter Nyströms (2004b) avhandling *Rätt mätt på prov. Om validering av bedömningar i skolan* kopplas validitetsfrågan till gymnasieskolan och bedömningar av elevers matematikkunskaper. I och med att eleverna differentieras på olika sätt (yttre differentiering genom elevernas

val av program) och inre differentiering genom nivågruppering påverkas förutsättningarna för bedömningarna. Det ger utslag i ett reliabilitetsproblem som en aspekt av validitet. Även validiteten i bedömningen av det nationella kursprovet i matematik för gymnasieskolan problematiseras. Före att lösa problemet föreslås en annan beräkningsgrund för provresultatet.

Från och med 1990-talet har även begreppet alignment (samstämmighet) fått en central betydelse som ett kvalitetsbegrepp i samband med kunskapsbedömningar. Den samstämmighet som avses är relationen mellan utbildnings-/kursplanemål och de kunskaper som prövas. Gunilla Näsström (2008) börjar i *Measurement of Alignment between Standards and Assessment* med att på teoretisk grund välja ett redskap för att mäta graden av samstämmighet mellan mål och prov. Hon prövar sedan redskapet i några olika sammanhang och drar slutsatser om dess användbarhet.

7.9 Andra centrala aspekter i avhandlingarna

Den valda kategoriseringen av studieobjekten lyfter fram vissa aspekter medan andra hamnar i bakgrunden. Det finns t.ex. ett antal avhandlingar som undersöker aspekter relaterade till internationella fenomen. Antingen uppmärksammas förhållanden i andra länder (Benincasa 1995, Zhao 1993) eller så används data från internationella jämförande studier (Eklöf 2006, Engström 1993, Myrberg & Rosén 2006, Yang 2003, Åström 2008) alternativt finns det ett särskilt fokus på internationella arenor, t.ex. OECD (Pettersson 2008).

En annan aspekt som också kan betonas är alternativa former för bedömning. Till en sådan kategori skulle vi ha fört Ann Dragemark Oscarssons (2009) studie *Self-Assessment of Writing in Learning English as a Foreign Language. A Study at the Upper Secondary School Level*. I avhandlingen behandlas elevers lärande och elevers självbedömningar av texter skrivna på engelska. Eleverna förväntas å ena sidan lära sig att bedöma egna texter i syfte att utveckla dem och å andra sidan utveckla sin förmåga att skriva specifika texter. Maria Hurtig (2007) beskriver i *Jag vågar visa att jag kan: Om meningsskapande med digitala portföljer* elevers och lärares arbete med den bedömningspraktik som är knuten till digitala portföljer. Fokus ligger på vad eleverna erbjuds och vad som görs tillgängligt för dem i dessa situationer. I avhandlingen *Om livets kontinuitet. Undervisning och lärande om växters och djurs livscyklar – en fallstudie i årskurs 5* följer Eva Nyberg (2009) en lärares arbete med två klasser. En aspekt av lärarens klassrumsarbete som får särskild uppmärksamhet i avhandlingen är lärarens formativa bedömning. Även Jessika Karlssons (2008) tidigare nämnda avhandling *A Novel Approach to Teaching Emotional Expression in Music Performance* om feedback hör till

den här gruppen. Utöver dessa finns också två avhandlingar som studerar utvecklingssamtal. Vid sidan av Gunilla Granaths (2008) avhandling som även beaktar loggböcker ingår här Johan Hofvendahls (2006) studie *Riskabla samtal – en analys av potentiella faror i skolans kvarts- och utvecklingssamtal*.

Den tredje aspekten vi vill förtydliga är att det finns ett relativt stort antal ämnesdidaktiska avhandlingar, framför allt inom svenska (Ask 2007, Bergman 2007, Eklund Heinonen 2009, Folkesson 1998, Lundström 2007, Norlund 2009, Palmér 2008, Parmenius Swärd 2008, Östlund-Stjärnegårdh 2002) och matematik (Bergsten 1990, Bjerneby Häll 2006, Boesen 2006, Hedrén 1990, Pettersson 1990, Stenhag 2007, Taflin 2007). När det gäller andra skolämnen är några ämnen representerade av enstaka avhandlingar. Det gäller t.ex. musik (Karlsson 2008), naturkunskap (Andrée 2007, Nyberg 2009) och idrott och hälsa (Ericsson 2003).

Andra indelningar skulle ha kunnat baseras på om det är individen, organisationen eller samhället som står i förgrunden, alternativt om det är mikro, meso och/eller makro som behandlas. En annan möjlighet hade varit att utgå från vilken typ av verksamhet och aktörer det är som uppmärksammas, t.ex. policy, administration, forskning, fackliga organisationer och/eller pedagogisk praktik. På ett övergripande plan kan vi konstatera att dessa nivåer, verksamheter och aktörer finns företrädade om än i olika omfattning. Det gör att det är svårt att ge en enhetlig bild av forskningen inom området, för det mesta framstår avhandlingarna som fallstudier av olika fenomen, skolämnen och skolformer. Den pedagogiska praktiken är däremot inte särskilt väl representerad i artiklarna även om det finns enstaka exempel på det. De övriga nivåerna representeras i både avhandlingar och artiklar.

8 BEDÖMNINGSFORSKNINGEN – ETT FRAMTIDSPERSPEKTIV

Att sja om framtiden är inte lätt, men att blicka in i den nära framtiden genom redan initierade aktiviteter är lättare. Vi använder projekt som erhållit medel av Utbildningsvetenskapliga kommittén (UVK) för att få en indikator på vad vi kan förvänta oss av bedömningsforskningen inom de närmast kommande åren. Vi har tagit del av material från UVK som täcker perioden 2006–2009, men på skilda sätt. Tyngdpunkten ligger på projekt som ansökt om och erhållit medel under perioden. Som ett komplement har vi också gått igenom de av UVK årligen utgivna *Resultatdialogerna* (2006, 2007, 2008, 2009). De innehåller primärt projekt som avslutas under det år de finns med i *Resultatdialogen*. Därigenom täcker vi en något längre period. Här finns projekt som kan ha fått medel 2003 till projekt som kommer att avslutas först 2012.

8.1 Omfattning och forskningsmiljöer

Mindre än 15 procent av forskarna som söker medel från UVK får ett positivt svar. Konkurrensen och kravnivån är med andra ord att betrakta som relativt hög. Vilka som erhåller medel och på vilka grunder kan naturligtvis diskuteras. Oavsett utfallet av en sådan diskussion kan projekten som erhållit medel ses som ett tecken på forskning som, enligt företrädare inom det utbildningsvetenskapliga fältet, saknas eller behöver utvecklas. Samtidigt har rimligtvis också forskare som fått medel ansetts ha förutsättningar att genomföra projekten på ett framgångsrikt sätt. De traditioner och de miljöer som forskarna är förankrade i har betraktats som relevanta och fruktbara för forskningsområdet ifråga. Sammantaget kan projekten därmed fungera som indikatorer på både forskningsfronten och kunskapsluckor som behöver fyllas. Projekten gör det också möjligt att urskilja delar av vad vi kan förvänta oss av forskningen inom en snar framtid.

Av de beviljade UVK-ansökningarna med relevans för forskning om bedömning avser tre forskarsassistenttjänster, en forskarskola och 31 forskningsprojekt. Sammanlagt tilldelades alltså 35 ansökningar medel under perioden 2006–2009. Fördelningen är relativt jämn sett över åren, även om 2007 rymmer fler projekt. Om vi jämför med projekten i *Resultatdialogerna* kan noteras att dessa är genomgående några färre per år och med en ökning från 2009 som innehåller åtta projekt av relevans jämfört med fyra till fem

ären dessförinnan. Som redovisas i översikten i kapitel 6 avser en övervägande majoritet av projekten Grundskolan och i övrigt är de relativt jämnt fördelade mellan gymnasieskolan och övergångar mellan olika skolformer. Mönstret känns igen också i övriga studerade publikationer liksom i de projekt som finns avrapporterade i Resultatdialogerna. Noteras kan dock att dominansen här är mindre och att fler studier griper över fler skolformer och att andelen som intresserar sig för gymnasiet och systemstudier är något större.

Jämfört med den bild vi redovisar i kapitel 5 om vilka forskningsmiljöer, eller närmare bestämt discipliner, som är involverade i bedömningsforskningen visar UVK-projekten på ett starkare genomslag från andra discipliner än pedagogik. Av de 35 projekten är andra discipliner involverade enskilt eller tillsammans med pedagogik i 20 av dem. Variationen av discipliner är dessutom stor: special/pedagogik, flera ämnesdidaktiska (fysik, kemi, matematik, språk), samhällsvetenskap (sociologi, kulturgeografi, vetenskapsteori, national/ekonomi, ekonometri), naturvetenskap (biokemi, matematik, statistik, astronomi) språkvetenskap (allmän språkvetenskap, lingvistik, nordiska språk, finska) och beteendevetenskap (psykologi). Den ökade variationen framträder också i relation till de tidigare projekten i resultatdialogerna.

8.2 Bedömningsprojektens studieobjekt

Materialets fördelning över de skilda studieobjekten framgår av kapitel 6, där vi noterade en viss skillnad mellan avhandlingarna och UVK-projekten. Här redovisar vi kortfattat exempel på projekt inom de respektive kategorierna. Några av projekten är sådana att de kan placeras i flera kategorier. De första kategorierna följer samma presentationsordning som tidigare. Men på grund av att projekten är färre till antalet har vi bedömt det som meningsfullt att endast beskriva fyra av kategorierna utförligare, medan övriga presenteras under rubriken övriga kategorier/studieobjekt.

8.2.1 Lärares didaktiska arbete

Sju projekt går att knyta till kategorin *Lärares didaktiska arbete*. Två av dem har en variationsteoretisk grund och är relaterade till "learning studies": *Hur lärares yrkesutövning påverkas av att de deltar i gemensamt bygge av yrkets vetenskapliga grund* (Runesson 2009)¹⁵ och *Effekter av likheter och skillnader*

¹⁵ I det följande anges huvudsökande för projekten, men nästan samtliga har en eller flera medsökande forskare och i flera fall från olika lärosäten.

mellan exempel valda för att främja lärande (Marton 2007). Medan det första projektet undersöker relationen mellan lärares professionella kunskap och elevers lärande är syftet med det senare mer uttalat att vidareutveckla variationsteorin. Fyra projekt har en ämnesdidaktisk inriktning. I *Nationella matematikprov som katalysator för implementering av utbildningsreformer* intresserar sig Johan Lithner (2007) för de nationella provens inverkan på lärares och läromedelsförfattares uppfattningar om och hur de genomför sina arbetsuppgifter. Det andra ämnesdidaktiska projektet *Uppkomst av skillnader i prestationsförmåga i matematik – en empirisk studie av klassrumspraktik* är komparativt och jämför Sverige, Tyskland och Canada med avseende på vilka skillnader i klassrumspraktiken som bidrar till att skillnader i prestationsförmåga i matematik uppkommer (Jablonka 2007). Projektet *Kemitexter som redskap för naturvetenskapligt lärande: en komparativ studie av undervisningspraktiker i svenska och finlandssvenska klassrum* (Eriksson m.fl. 2006) tar sin utgångspunkt i de statistiskt signifikanta skillnaderna i PISA-resultaten mellan svenska och finlandssvenska skolor. I projektet jämförs kemiundervisningens innehåll som den ter sig i klassrumspraktikerna, i proven och i lärarnas bedömningsarbete. Vid sidan av dessa finns också projektet *Kan mera naturvetenskap i förskollärarytbildningen påverka dagisbarnens attityder, skolresultat och studieinriktning?* (Mellander 2009). Det är en longitudinell studie av effekterna av en naturvetenskapligt inriktad förskollärarytbildning. Den är inriktad på de barn som har haft dessa förskollärare och barnens därpå följande prestationer och utbildningsval. I resultatdialogerna var det en större andel av samtliga projekt som tillhörde denna kategori och det var också här de flesta projekten kunde återfinnas.

8.2.2 Elevers prestationer och skolors resultat

Den största andelen projekt (elva) återfinns i kategorin *Elevers prestationer och skolors resultat*. Här finns projekt som använder longitudinella databaser, har sin grund i differenser och som genomför statistiska analyser av skilda slag. Ett av dessa är *Betygens jämförbarhet, prognosförmåga och effekter på lärande* (Cliffordson 2008) som studerar dels betygens jämförbarhet över tid mellan lärare och skolor, dels betygens påverkan på elevers lärande och motivation, dels betydelsen av kognitiva och icke-kognitiva faktorer för betygens prediktiva värde. Bland studierna, om än mer sällsynt, finns också projekt som är inriktade mot skolans sociala mål, t.ex. *Skolornas institutionella karaktär och elevernas medborgarkompetens: en jämförelse av olika kommunala och fristående skolor över tid och rum* (Amnå 2009).

Flera projekt i den här kategorin analyserar effekterna av skolval. Kajsa Yang Hansens (2007) projekt *Skolsegregation i den svenska skolan – orsakar*

och effekter undersöker t.ex. förändringar i prestationer under perioden 1991–2006 i förhållande till möjligheten att välja skola och skillnader mellan elever och skolor vad avser socioekonomiska förhållanden och etnicitet. På liknande sätt, men med fokus på skolornas geografiska placering, reses också frågor om boendesegregation och integrerade skolor i projektet *Den svenska skolans nya geografi: rörlighet, attityder och resultat* (Malmberg m.fl. 2009).

Till denna kategori kan också föras några studier som fokuserar internationella jämförande mätningar. Projektet *Analys och modellering av svenska elevers prestationer i TIMSS och PISA i ett internationellt perspektiv* (Wiberg 2008) undersöker test och analysmetoder som används i de komparativa mätningarna. Avsikten var också att utveckla en multinivåmodell för att skapa ett mer informativt och effektivt mått på elevers prestationer. Anders Jakobssons (2008) projekt *Storskaliga studier och elevers prestationer i naturvetenskap i ett longitudinellt perspektiv* är i sin tur inriktat mot att analysera skillnader mellan svenska och danska trender och tendenser för elevers prestationer i PISA och TIMSS. Även här studeras testens validitet. Sett utifrån Resultatdialogerna är kategorin prestationer och resultat mindre framträdande.

8.2.3 Skolans styrning och lärares arbete

I gruppen ingår fem projekt som alla har en direkt referens till skolans styrning och samtidigt uppmärksammar lärares verksamhet. Två studerar skolinspektionen, dels ett forskarassistentprojekt (Rönneberg 2007b), dels ett komparativt projekt *Inspektion som styrning. Skolinspektion och utbildningsstyrning i Sverige, England och Skottland* (Segerholm 2009). I det senare är intresset inriktat på såväl det internationella och det nationella som policy och skolpraktik. Även projektet *Internationella kunskapsmätningar, nationell bedömningskultur och reformhistoria* (Forsberg m.fl. 2006) är komparativt med fokus på det inter/nationella och bedömningskulturen i mötet mellan policy, forskning och pedagogisk praktik i Sverige och Tyskland. Samtliga dessa studier har också ett fokus på bedömning som fenomen. Till den här kategorin har vi också fogat Sverker Lindblad m.fl. (2008) som i *Bildningsgångar och skolresultat i det senmoderna samhället* studerar innebörden av resultatstyrning för skolors verksamhet och för hur elever med olika kulturell bakgrund formas i relation till sina skolprestationer och skolors resultat. Ytterligare ett projekt som kan nämnas är Ann-Christine Vallberg Roth m.fl. (2007) som i *Reglerad barndom och lärarprofession – möten mellan förskola och mångfaldens medborgare genom ett filter av dokumentation* fokuserar de individuella utvecklingsplanerna (IUP), specifikt med avseende på relationen mellan lärare, föräldrar och barn. Detta studeras mot bakgrund av de politiska beslut som gäller IUP och implementeringen av dem i

skolan. Inga större skillnader i representativitet finns för denna kategori när vi jämför ansökningarna och resultatdialogerna.

8.2.4 Organisering av skola, undervisning och elever

De fem projekten i gruppen är knutna till specialpedagogiska frågeställningar och till elevers utbildningskarriärer. Ett av de förstnämnda är Mara Westling Allodis (2008) projekt *Delaktighet, socialt klimat och lärande i skolan. En intervention på organisationsnivå i skolans lärandemiljöer för en utveckling av skolans förmåga till förändring och anpassning till elevernas behov*. I ett annat projekt *Specialpedagogiska stödåtgärder i grundskolan: omfattning, former och effekter* analyserar Joanna Giota (2007) exkluderande och inkluderande processer i skolan.

Projekten om utbildningskarriärer är här förbundna med frågor om mångkulturalitet, etnicitet respektive språklig kompetens. De ställs i förhållande till organiseringen av eleverna och undervisningens innehåll och utformning. Erik Olsson (2006) studerar i *Transnationella utbildningskarriärer. Studieval, möjlighetshorisont och sociala nätverk för unga med utländsk anknytning* skilda elevgrupper från olika sociokulturella sammanhang och med olika skolerfarenheter. I en annan ansökan, *Utbildningsval, sociala karriärer, språkliga kompetenser och identiteter* (Lainio 2006), är det centrala betydelsen av språklig kompetens i modersmålet (finska/bosniska/svenska) för elevernas möjligheter att lära sig svenska och för deras sociala, utbildnings- och yrkeskarriärer. I resultatdialogerna var andelen projekt i denna kategori ungefär lika stor.

8.2.5 Övriga kategorier/studieobjekt

Forskarassistentprojektet *Utbildning och intelligens* (Lundahl 2007) kan främst knytas till kategorin *Bedömning som fenomen*. I projektet analyseras den psykometriska kunskapsproduktionen 1880–1960 som den kommer till uttryck i forskningsmaterial och lärarfackliga tidskrifter och kursprogram. Ytterligare en studie har bäring på bedömning som fenomen. Projektet *Evidensbaserad pedagogik: Epistemologiska förutsättningar för evidensrörelsens expansion in i utbildningsvetenskap* (Bohlin 2007) analyserar både expansionen av evidensrörelsen till nya områden och kontroversen om evidensbaserad utbildning, med referens till den ökande rutiniseringen och formaliseringen av epistemiska procedurer. Även några av projekten (Forsberg m.fl. 2006 och Segerholm 2009) kategoriserade under styrning kan föras till denna kategori. Båda projekten är särskilt intresserade av bedömning som ett samhällsligt fenomen på både policy- och verksamhetsnivå.

Även i kategorin *Elevers upplevelser och erfarenheter av bedömning* finns ett antal ansökningar och några kompletterar avhandlingar. I studien *Vad gör testandet med eleverna? En studie kring nationella prov i matematik* (Sjöberg 2008) studeras hur traditionella matematikprov och nationella prov i matematik påverkar elevers matematikinläring. Särskilt uppmärksammas vad detta betyder för elever som inte uppnår godkända resultat. I ett annat projekt *Interaktioner mellan den lärande och lärandesituationen: effekter på affektiva upplevelser och lärandeutfall* undersöker Mikael Winberg (2007) hur elever med olika förkunskaper och kunskapsyn upplever olika typer av lärande- och provsituationer, vilka situationskaraktäristika som är mest betydelsefulla för olika typer av elever och hur upplevelserna är relaterade till kvaliteten på det lärda. Även Hanna Eklöfs (2006) forskarsistentprojekt *Elevers uppfattning av och inställning till prov och tester i skolan* har förts till denna kategori. Eklöf utvecklar den i avhandlingen behandlade frågan om elevers testmotivation i samband med internationella kunskapsmätningar. I projektet undersöks genom intervjuer och enkäter elevers inställning till prov under sin skoltid och dess relation till elevers prestationer.

Några projekt som inordnats i kategorin elevers prestationer och resultat har också relevans för *Bedömningsredskapens kvalitet*. Det gäller såväl Christina Cliffordson (2008), Marie Wiberg (2008) och Anders Jakobsson (2008) som fokuserar olika provs validitet. Utöver dessa kan projektet *Testet under lupp – Konstruktion av läsförståelse i internationella läsundersökningar* (Liberg 2007) sorteras in här. Med textanalytiska och statistiska metoder undersöks särskilt vad det är elever kan som enligt bedömningsmallen anses inte förstå. Mer övergripande fokuseras de kulturella drag som gestaltas genom texterna, texternas uppbyggnad och innehållsliga struktur samt hur det samvarierar med bedömningar av elevsvar. Elever med olika språklig och sociokulturell bakgrund samt kön jämförs också.

Bland UVK-projekten finns vid sidan av en ansökan som avser den tidigare nämnda forskarskolan i pedagogisk bedömning (Pettersson 2006) också några projekt som studerar antingen övergångar och bedömningsfunktion i dessa eller kunskapsbedömningar som fenomen. Det finns också en ansökan som avser medel för den tidigare nämnda forskarskolan i bedömning. Projektet om *Övergångar* fokuserar andra övergångar än avhandlingarna. Den ena studien *Övergångar i förskolan som kritiska händelser i barns liv, särskilt med avseende på delaktighet och marginalisering* (Garpelin 2008) avser att ge en bred belysning av övergångar: vad karakteriserar dem, vilken betydelse har de för barnen specifikt med avseende på segregering/inkludering, vilka personalens intentioner är samt hur övergångar i förskolan skiljer sig från andra övergångar, t.ex. i skolan. Ett annat projekt, *Osäkra övergångar. Unga utan fullständig gymnasieutbildning: vägarna och åtgärderna i longitudinellt*

perspektiv, är relaterat till elever som avslutar gymnasieskolan utan fullständiga betyg (Lundahl m.fl. 2009). Projektet belyser såväl individrelaterade aspekter (ungdomars upplevelser av de åtgärder som erbjuds) som socialt och regionalt relaterade aspekter (bl.a. orsaker till variationen i mängden studieavbrott mellan kommuner). Skillnaderna mellan resultatdialogerna och ansökningarna är inte stora, även om vissa variationer förekommer för de kategorier som redovisas under denna rubrik.

8.3 Avslutandet noteringar

Vid en jämförelse mellan avhandlingar och UVK-projekten (ansökningarna och resultatdialogerna) kan noteras att avhandlingarna och resultatdialogerna ligger närmare varandra. När det gäller ett så begränsat urval finns det naturligtvis anledning att vara försiktig med slutsatser, men det förefaller som det skett en förskjutning som gör att vi framöver kan vänta oss mer forskning med fokus på elevers prestationer och resultat och mindre på lärares didaktiska arbete, även om det inte helt har försvunnit ur bilden. I övrigt är variationerna relativt ringa. Nämnas kan också att några av projekten är utvecklingar av avhandlingsprojekt.

I resultatdialogerna från 2006–2009 finns det fyra till fem projekt med fokus på bedömningsforskning, frånsett 2009. Den senaste ökningen kan ses i relation till att projekten i ansökningarna för samma period var sju till åtta, frånsett 2007 då tolv projekt identifierades. Framför oss ligger således avrapporteringar från en växande andel projekt under den närmaste tiden. Notabelt är också att projekten är relativt komplexa med studier som studerar flera olika aspekter av bedömning. Många gånger använder de också ett flertal metoder för både insamling och analys av data. Till detta kan slutligen fogas det ökade intresset för forskning om bedömning från flera olika discipliner, varav många icke pedagogirelaterade sådana. Flera projekt bedrivs också i samarbete mellan pedagogik/didaktik och andra discipliner. Vi ser alltså fram emot en situation där frågor om bedömning kommer att belysas med en större variation av begrepp och metodansatser liksom att teorier utvecklas som kan bidra med förklaringar av skilda slag.

9 BEDÖMNING SOM OBJEKT FÖR KUNSKAPSBILDNING

Med utgångspunkt i materialet om avhandlingarna, artiklarna och UVK-projekten riktar vi här intresset mot bedömningsforskningens kunskapsobjekt. Med den breda ingång som vi använt för att identifiera relevant forskning om bedömning är det rimligt att förvänta sig ett utfall som rymmer en viss variation. Vi kommer att diskutera denna som uttryck för tre inriktningar inom bedömningsforskningen: metriska studier, professions- och verksamhetsstudier samt systemstudier. De presenteras med exempel från det svenska materialet som helhet.

9.1 Metriska studier – statistiska analysmodeller

Det gemensamma för studierna är att de använder sig av statistiska metoder och analyser för att producera data och resultat. Flera studier har ett syfte bortom de empiriska resultaten. De ska t.ex. bidra till utvecklingen av de statistiska metoderna och modellerna eller kombinera registerdata från olika longitudinella databaser på nya sätt. På ett övergripande plan kan resultaten relateras till frågor om individuella differenser, provkonstruktion, övergångar, komparativa studier och effektstudier. De discipliner som företräds inom den här inriktningen är främst psykologi/pedagogik, ekonomi/nationalekonomi och sociologi. Det förenande i de metriska studierna är inte kunskapsobjekten utan metodansatsen.

9.1.1 Individuella differenser

Studier som bidrar med kunskap om individuella differenser hör traditionellt hemma i gränslandet mellan pedagogik och psykologi. Mått på intelligens (IQ) och/eller på specifika begåvningsfaktorer relateras till frågor om skolframgång eller till utbildningens längd och inriktning (t.ex. Cliffordson & Gustafsson 2008). Studierna bidrar med metodologiskt nya sätt att förklara samband mellan de faktorer som studeras eller genom att föra in nya mått av annat slag, t.ex. registerdata. Resultat och/eller framgång kopplas även till frågor om kön. Monica Roséns (1998) avhandling *Gender Differences in Patterns of Knowledge* är ett exempel på det senare. Avhandlingen är uppbyggd runt tre delstudier med grundskoleelevers resultat från kognitiva test, standardprov och en IEA-

studie. Skillnader i resultat mellan pojkar och flickor beskrivs som ett bidrag till feministisk forskning om genuskillnader i kunskapsmönster.

Utöver elevers kognitiva prestationer i samband med nationella prov eller internationella kunskapsmätningar uppmärksammas även så kallade icke-kognitiva aspekter, t.ex. motivation, engagemang och intresse, och deras betydelse för de kognitiva resultaten. I en enkätstudie om relationen mellan elevers motivation för provet och deras provresultat visar Hanna Eklöf (2007) att svenska elever är motiverade för att skriva internationella prov. Motivationen förefaller spela mindre roll för flickorna. Även mätinstrumentet (enkäten) är en del av resultatet. Vidare tar en del av forskningen om de nationella proven sin utgångspunkt i provresultaten. I en delstudie i en sammanläggningsavhandling i nationalekonomi, *Empirical Studies of Public Policies within the Primary School and the Sickness Insurance* (Lindahl 2008), studeras skillnader mellan elevers provresultat och betyg. Resultaten visar statistiskt signifikanta skillnader på två punkter. Flickor får oftare generösare betyg än pojkar. Det gäller också elever med utländsk bakgrund jämfört med svenska elever. Ett mer utförligt bidrag till den svenska forskningen om individuella differenser finns i *Pedagogisk Forskning i Sverige* (2002 nr 3).

9.1.2 Provkonstruktion

Några, dock inte så många, studier är inriktade på provkonstruktion, ett av psykometrins klassiska områden. I vårt material är de relaterade till högskoleprovet och använder statistiska metoder t.ex. för uppgiftsanalys. I artikeln *Dimensions of performance in the interpretation of diagrams, tables and maps. Some gender differences in the Swedish scholastic aptitude test* (Åberg-Bengtsson 1999) analyseras uppgifter för att identifiera latent dimensioner av en specifik uppgiftsgrupp (tabeller, kartor, diagram) i högskoleprovet. En sådan latent dimension är genus. Resultatet är betydelsefullt med avseende på framtida provkonstruktion. Med övergången från det normrelaterade till det mål- och kunskapsrelaterade betygssystemet minskade behovet av studier av psykometrisk provkonstruktion. En alternativ forskning om provkonstruktion är inte framträdande i materialet.

9.1.3 Övergångar

Till de studier som utvecklar kunskap om övergångar hör t.ex. rekryteringsstudier och valideringsstudier. I materialet finns det studier som jämför prognosvärdet för gymnasiebetygen och högskoleprovet med avseende på framgång i högskolan. Resultaten pekar på att betygen har ett något bättre prognosvärde (Cliffordson 2004). Men det finns skillnader mellan

olika typer av utbildningar. Hittills är det huvudsakligen professionsutbildningar som studerats (Svensson 2004). Ett empiriskt bidrag till frågan om antagning till högskolan på "särskild grund" visar att det antagningsprov som utvecklats och prövats vid Karolinska institutets tandläkarutbildning gav positivt resultat (Röding 2005). Tre omgångar studenter som antogs på basis av provet visade sig klara såväl teoretiska som praktiska prövningar bra eller bättre än studenter som antagits på basis av gymnasiebetyg eller högskoleprovet.

Antagning till högskolan handlar också om vilka som väljer högskola och på vilka grunder. En generell och klassisk bild är att valet av högskolestudier är relaterat till social klass och genus. Bland de studier som ingår i vårt material finns sådana som bidrar med empiriskt material som kan nyansera dessa generella sociala kategorier. Margareta Hammarström (1996) visar t.ex. i sin avhandling *Varför inte högskola? En longitudinell studie av olika faktorer betydelse för studiebegåvade ungdomars utbildningskarriär* att unga människors bedömning av ekonomiska faktorer (studiestödets storlek i förhållande föräldrarnas möjlighet att ge bidrag) och konjunkturläget (möjligheterna att arbeta vid sidan om studierna respektive möjligheten till ett arbete direkt efter gymnasieskolan) är avgörande för om de som avslutat gymnasiet väljer att studera eller inte. I studien består urvalet av unga människor med höga betyg från gymnasieskolan. I avhandlingen *Sökandemönster till högskolan: en fallstudie vid Luleå tekniska universitet* (Perzon 1998) kompletteras Hammarströms resultat med att geografiskt avstånd snarare än betyg är avgörande för val av högskola. Även om merparten av studierna handlar om övergången från gymnasieskolan till högskolan, förekommer det också studier om övergången mellan grund- och gymnasieskolan. En studie av rekryteringen till gymnasieskolans program visade att det fortfarande 1998 fanns starka samband mellan elevernas socialgrupp och deras val av gymnasieprogram (Svensson 2001).

9.1.4 Komparativa studier

Från tidigt 1990-tal och framåt finns exempel på artiklar som tar sin utgångspunkt i resultat från internationella test och komparativa studier och där statistiska metoder används för att förklara antingen det huvudsakliga resultatet eller delresultat. Lena Adolfsson och Widar Henriksson (1999) studerar svenska elevers resultat på TIMSS och söker alternativa förklaringar till deras resultat. Resultatets giltighet sätts under lupp. Författarna menar att resultatet inte kan tolkas utan hänsyn till läromedlens innehåll och det antal år eleverna har gått i skola eftersom åldern för skolstart varierar mellan deltagande länder. Resultaten används också för att jämföra länder, vilket framgår av en komparativ studie med PISA-resultaten för Sverige och Schweiz, två

länder vilkas utbildningssystem internationellt sett anses ha hög kvalitet på utbildningen (Fredriksson, Holzer, McCluskey-Cavin & Taube 2009). Vissa aspekter av huvudresultaten granskas. Man konstaterar bland annat att schweiziska elever klarar sig bättre i matematik än svenska elever och att flickor i båda länderna läser bättre än pojkarna. Skillnaden mellan könen är större i Sverige än i Schweiz och pojkar klarar matematik bättre än flickor.

9.1.5 Effektstudier

Inom effektstudierna används i huvudsak samma eller likartade metoder oavsett forskarnas disciplinära hemvist. Effektstudierna reser allt oftare frågor utifrån ett intersektionellt perspektiv, som i varierande omfattning tar hänsyn till elevers kön och sociala, ekonomiska, kulturella och geografiska bakgrund. I flera fall beaktas också skolans inramning, resurser, humankapital, organisering och placering samt elevers tidiga skolkarriärer och effekterna av dessa för motivation, intresse och engagemang. Dessa faktorerers betydelse för val av utbildning och utbildningsframgång liksom utbildningens bidrag till position i arbetslivet, inkomst och hälsa studeras. Ofta diskuteras resultaten i relation till olika välfärdsfrågor. Det är inte heller ovanligt att effektstudier särskilt utvecklar kunskap om utfallet av specifika reformer, t.ex. införandet av fristående skolor och betydelsen av konkurrensutsättningen av skolan. En utförlig och aktuell genomgång av vad som påverkar skolans resultat och konsekvenser av olika resultat finns i Skolverkets rapport (2009) *Vad påverkar resultaten i svensk grundskola: kunskapsöversikt om betydelsen av olika faktorer*. Där sätts resultaten i relation till olika samhällsfaktorer, skolreformer, skilda resurser och skolans inre arbete. Resultaten diskuteras i termer av segregering, decentralisering, differentiering och individualisering.

Grunden för effektstudier är olika former av individdata från olika longitudinella databaser som aggregeras och som möjliggör kunskapsbildning på både organisations- och systemnivå. Det finns ett fåtal studier (Ahlström 2009), som vi här inte går närmare in på, som har skolans organisation och utveckling som objekt för sina studier. De är relaterade till frågan om den framgångsrika skolan eller skolledaren. Här är organisationsteoretiska perspektiv vanliga och metodansatsen mer varierad även om kvantitativa inslag är en bas.

Perspektiv som aktualiseras i effektstudier är knutna till såväl sociologisk, national/ekonomisk och kulturgeografisk som pedagogisk teoribildning. Ett exempel på en studie är Jan O. Jönssons och Carina Moods (2008) om sociala kontrastmekanismer och om huruvida de påverkar elevers anspråksnivå om de har högpresterande kamrater. Datamaterialet består av två elevkohorter

(160 417 elever och 829 skolor) i grundskolans skolor 9, med matchad information om socialt ursprung och prestation. Med statistisk kontroll för skolor, elev- och familjekarakteristika finner de att tendensen att göra ett ambitiöst val ifråga om gymnasieprogram är lägre för elever med högpresterande skolkamrater. Då en medelelev flyttar från en medelskola till en lågpresterande ökar sannolikheten för att eleven kommer att sträva högre.

9.2 Professions- och verksamhetsstudier

Under den här rubriken ryms studier relaterade till lärares bedömningspraktiker, samtalsforskningen och professionsforskningen. Generellt kan studierna inom gruppen beskrivas i termer av om de studerar lärares bedömningspraktiker direkt eller indirekt. Medan de förstnämnda har fokus på bedömning har de sistnämnda fokus på något annat, men bidrar genom resultaten eller empirin till bilden av lärares bedömningsarbete.

9.2.1 Bedömningsredskap

I några verksamhetsstudier uppmärksammas explicit lärares bedömningsredskap. Ett exempel på detta är Gunilla Granaths (2008) avhandling som fokuserar dels utvecklingssamtalet, en politiskt påförd bedömningspraktik, dels ett av lärare infört planeringsredskap (loggboken). Med utgångspunkt i ett maktperspektiv är analysen inriktad på hur utvecklingssamtalet och loggboken tar form i samspelet mellan lärare och elever och, när det gäller utvecklingssamtalet, även föräldrarna. Resultaten pekar på konsekvenserna för eleverna. Loggböcker och utvecklingssamtal är viktiga redskap för att påverka eleverna och deras förhållningssätt till sig själva, sitt skolarbete, sina liv och sin framtid. Redskapen formar elevernas syn på sig själva och bidrar till att skapa normativa bilder av hur de bör vara (t.ex. att alltid visa sig kompetent). Bilder som eleverna jämför med föreställningarna om sig själva. Även etnografiska studier av klassrumsarbetet förekommer. I en avhandling (Hurtig, 2007) finns t.ex. en studie av klassrumsarbetet med specifikt fokus på digitala portföljer. Avhandlingen ramar in av sociokulturell och multimodal teori. Det empiriska bidraget visar på ömsesidigheten i lärare-elevrelationen i klassrummet. I utformningen av de digitala portföljerna utvecklade både elever och lärare nya begrepp och rutiner.

Vi har redan i flera sammanhang nämnt andra exempel på studier som infört nya redskap i praktiken: individuella utvecklingsplaner (IUP), utvecklingssamtal, kvalitetsredovisningar, nationella prov, betyg, provbanker och portföljer. I de flesta fall finns det enstaka studier om respektive redskap,

endast IUP och åtgärdsplaner återkommer. Här pekar resultaten åt olika håll. Studier som fokuserar IUP, utvecklingssamtal och åtgärdsprogram visar att lärarna glider mellan bedömning av kunskaper och elevernas beteende, att bedömningarna formuleras relativt olika om de avser pojkar eller flickor och att det är förhållandevis vanligt att åtgärdsprogram inte ger önskad effekt. Kort sagt, skolan misslyckas med att ge det extra stöd som behövs för de elever som behöver det. Andra redskap, t.ex. portfolio, självbedömning och formativ bedömning förefaller däremot ge mer positiva resultat.

9.2.2 Lärande/bedömning

Inom forskning relaterad till lärande och bedömning finns såväl konstruktionistiska och sociokulturella/verksamhetsteoretiska som variationsteoretiska studier i materialet. Ett exempel relaterat till ämnesdidaktik och kulturell reproduktion i gymnasieskolan finns i Lotta Bergmans (2007) avhandling om hur svenskämnet tar form i fyra gymnasieklasser. Särkilt uppmärksammas vad som framstår som viktigt och vad som är föremål för lärarens bedömning i de fyra klasserna. Avhandlingens empiriska bidrag kan sammanfattas i tre huvudresultat. För det första har svenskundervisningen i samtliga klasser en tendens att dras mot formell färdighetsträning. För det andra är lärarnas kursplanearbete ensidigt inriktat på lägre betygs-kriterier och mål att uppnå. För det tredje utnyttjas inte det handlingsutrymme som styrdokumentet medger, bland annat i valet av innehåll för olika elevgrupper. Avancerade eller reducerade varianter av samma innehåll erbjuds till elever på olika program. Bergman menar att det hänger samman med en stark föreställning om att det inte går att bedöma elevernas kunskaper om det inte är baserat på samma innehåll. Svenskämnets tradition återspeglas i klassrumsarbetet som lärarna mött det under sin utbildning, där språk och litteratur hålls åtskilda från varandra.

Maria Andrée (2007) fokuserar vad elever erbjuds att lära sig i grundskolans NO-undervisning. Hon identifierar två praktiker med sinsemellan olika föreställningar om vad som ska åstadkommas. Det präglar klassrumsarbetet och ger olika förutsättningar för vad som blir möjligt att lära. Bedömning i karaktärsämnen inom gymnasieskolans yrkesförberedande program företräds av en avhandling, *Därför fick jag bara Godkänt: Bedömning i karaktärsämnen på HR-programmet* (Tsagalidis 2008). Den är skriven inom ett socialkonstruktionistiskt perspektiv. Utifrån video/intervjuer (stimulated elicitation) urskiljs sju nyckelkvalifikationer och fem kategorier av specifika yrkeskunskaper som lärarna fokuserar vid bedömning i karaktärsämnen. Nyckelkvalifikationer värderas högre än specifika yrkeskunskaper, som utgör förutsättningen för godkänt betyg. Lärarnas värderingar reflekterar

deras yrkeskulturella kunskap och erfarenhet. Det gör att elever utan bransch erfarenhet har svårt att nå högre betygsnivåer.

Den lärandeforskning som utgår från variationsteorin syns inte i avhandlingarna, vilket kan tyckas lite märkligt eftersom en central del av learning studies handlar om att utforma för- och eftertest som ska visa vad eleverna redan kan och vad de lärt sig. Däremot har två UVK-projekt som tar variationsteorin som sin utgångspunkt (Marton 2007, Runesson 2009) beviljats medel. Samtalsforskningen är inte särskilt väl företrädd i vårt material. Ulla Mauritzon och Roger Säljö (2001) har vi nämnt tidigare. Resultatet av den studien visar att barns förmåga att förstå vuxnas frågor, som de är avsedda att förstås, i stor utsträckning är beroende av om de delat samma inramning som den vuxne som ställer frågan.

9.2.3 Lärarprofessionen

En del studier är inriktade mot att utveckla kunskap om lärarprofessionen mer eller mindre uttalat i relation till bedömning. Den förändrade styrningen av skolan och dess konsekvenser för yrket har rest en rad frågor om lärares autonomi, ansvar, etik, uppgifter, kompetens och kunskapsbildning (bland annat i relation till praktknära forskning). Studierna visar sammantaget, via empirin mer än i abstracten, att bedömning och betygssättning är frågor som lärarna i hög grad är upptagna av. Avhandlingen *Lärares dilemman* (Räihä 2008) pekar ut bedömning och betygssättning som en av de arbetsuppgifter lärare upplever som motstridig. En annan avhandling *Lärares arbete och kunskapsbildning. Utmaningar och inviter i den vardagliga praktiken* (Wedin 2007) lyfter fram bedömning och betygssättning som ett av de områden lärarna upplever behov av att lära sig mer om. En tredje avhandling, *Lärares lärande om elever – en sociologisk studie om yrkespraktik* (Ranagården 2009), visar att en stor del av lärares lärande sker i olika slags möten med just eleverna, men även i samtal med kollegor. Studierna representerar humaniora, pedagogik och sociologi, men ofta används flera perspektiv, språk teori, organisationsteori, och/eller begrepp. Metodologiskt används intervjuer eller observationer. Platsen för observationer avgränsas inte enbart till klassrummet som i flera av studierna i avsnittet 9.2.2 utan även lärarrummet blir en plats där andra aspekter av lärararbetet studeras.

Bland artiklarna kan nämnas Gunnel Colneruds (1997) som jämför artiklar som representerar två traditioner om den goda läraren, den moderna kognitivistiska diskursen och teacher-thinking-diskursen. Hon ser lärares bedömning, värdering och kritik som en aspekt av det dagliga arbetet och betygssättning, den öppna maktutövningen, betraktas som en central aspekt av lärararbetet. Samtidigt betraktar Colnerud bedömning som en av lärar-

praktikens motstridiga arbetsuppgifter. Resultatet av hennes innehållsliga jämförelse visar att båda traditionerna förefaller förtränga dessa arbetsuppgifter. Det pekar på behovet av att utveckla en vokabulär som byggs ”på de konfliktfyllda utmaningar som lärarnas praktik innehåller” (a.a. s. 103).

9.3 Systemstudier

Forskningen som kan inordnas under inriktningen systemstudier är både teoretiskt och metodologiskt olikartad. Utbildningssystemet i sin helhet och/eller delarna som ingår i detta samt förutsättningar och villkor för dessa studeras. Även konsekvenser av olika utformningar av systemet och förändringar av det är här centralt. Oavsett fokus är avsikten inom denna forskning att utveckla kunskap på systemnivå. Det kommer till uttryck i de teoretiska grundvalarna liksom i vad studierna gör anspråk på att förklara och förstå. Vi ska här ta upp exempel på den historiskt/sociologiska bedömningsforskningen samt styrnings- och policystudier. Till systemstudierna kan också fogas de nämnda effektstudierna. Även om dessa studier väl platser här redovisas de i samband med övriga metriska studier.

9.3.1 Historisk/sociologisk bedömningsforskning

Med historisk/sociologisk forskning avser vi studier som i någon mån är idé/historiska och som samtidigt är kunskaps/sociologiskt inriktade. Denna forskning är nationellt relativt sparsam, särskilt vad gäller historiska studier som täcker perioden före 1900. Under 00-talet har emellertid ett flertal studier som mer eller mindre starkt betonar det historiska eller det sociologiska lagts fram. Inom den historiskt inriktade svenska forskningen om kunskapsbedömningar har tre traditioner identifierats (Lundahl 2006). Det handlar om historiesociologiska studier i vilka bedömningar, betyg och/eller examen relateras till samhällets sociala reproduktion. Det handlar också om studier om psykometrins historia och framväxt samt en reform-historisk tradition som antingen fokuserar systemet för betyg specifikt eller i policyhistoriska studier relaterar bedömningar av elevers kunskaper till olika reforminitiativ. Lundahls exempel är två studier som ligger före den period vi undersöker och en som ligger precis i början av perioden. För att peka på nytillskottet inom dessa traditioner har vi som exempel valt tre relativt nya avhandlingar. Till dem har vi fogat ytterligare två avhandlingar som får exemplifiera två delvis nya inriktningar. Den ena kombinerar ett kunskapsociologiskt och läroplansteoretiskt perspektiv. Den andra är en idéhistorisk avhandling.

Helena Korp har två avhandlingar som båda kan förstås som exempel på systemstudier, dels licentiatavhandlingen *Kunskapsbedömning – hur, vad och varför* från 2003, dels doktorsavhandlingen *Lika chanser i gymnasiet? En studie om betyg, nationella prov och social reproduktion* (2006). Det är den sistnämnda vi här tar som ett exempel på en historiesociologisk studie och då med tyngdpunkt på det sociologiska. Dess syfte är att belysa de nationella provens användning och komplikationer som stöd för en likvärdig betygsättning i olika gymnasie miljöer ur ett reproduktions- och konfliktteoretiskt perspektiv som integrerar ras, kön och etnicitet. I avhandlingen finns både en historisk beskrivning av den svenska gymnasieskolan och en samtida analys av gymnasieskolans system för bedömning. Den historiska analysen, baserad på tidigare forskning, utgör en grund för analysen av samtidsstudien där lärare, elever och rektorer intervjuas om nationella prov (NP) och betygsättning. Resultaten visar en stor variation mellan skolor och program vad avser undervisningens inriktning när det gäller förberedelser, genomförande och bedömning av NP och deras användning i betygsättning. Såväl inramningen som klassifikationen av ämnen varierar mellan teoretiska och yrkesprogram. Den erbjudna och den erhållna läroplanen i svenska var influerad av gymnasieskolans tidigare utformning när programmen hade sin hemvist i olika typer av skolor. Å ena sidan den klassiska gymnasieskolan med betoning på textanalys, akademiskt skrivande och retoriska färdigheter, å andra sidan yrkesskolan med betoning på grundläggande läs- och skrivfärdigheter. På samma sätt varierade lärares bedömningar mellan programmen på ett sådant sätt att de inte kan betraktas som likvärdiga. Då teoretiska program rekryterar elever främst från medelklassen och yrkesprogrammen från arbetarklassen bidrar dessa differentierade pedagogiska praktiker till den sociala reproduktionen.

Rikard Erikssons (1999) avhandling *Psykoteknik. Kulturell fabricering av personlig identitet* behandlar en för skolan relevant aspekt av bedömning, den psykologiska testningens betydelse för normbildningen i samhället. Avhandlingen är med andra ord ett nutida exempel på en psykometrisk studie. Samtidigt tillför den med sitt kunskapssociologiska perspektiv, och den etnografiska analysen av psykotekniken som kulturellt fenomen, nya sätt att skriva och förstå historien om den psykologiska testningen och dess betydelse för det moderna samhällets utveckling. Empiriskt analyseras en bred variation av texter, test, offentliga utredningar, artiklar i tidskrifter, foton, ljudupptagningar och även testföremål knutna till Psykotekniska institutet (PTI) som etablerades 1943. Undersökningsperioden täcker primärt åren 1943–1969, men den griper både bakåt och framåt i tiden. Med utgångspunkt i analyser av bland annat en stor mängd utlåtanden beskrivs hur den naturvetenskapliga tankestil som utmärkte institutets testverksamhet fabricerar personliga identiteter. Proven skulle genomföras snabbt

och den testade förväntades komma fram till rätt svar på rätt sätt. Vägen fram till resultatet var minst lika viktig. Det handlar om att följa en i förväg utstakad linje eller ordning. Fram träder ett etos med normativa levnadsregler för hur den moderna människan (inte) bör vara. Sund, välvårdad och ordentlig är några positiva kännetecken och bräcklig, slapp, ombytlig och oengagerad några av de negativa. På ett övergripande plan beskrivs hur psykoteknikens världsbild och etos formar ett kulturellt mönster i den svenska kulturen. Ett mönster där psykologerna/själens ingenjörer medverkar i både meningsskapande och sortering t.ex. som ett led i yrkesvägledning.

Ett exempel på den reformhistoriska traditionen utgör Jörgen Tholins (2006) avhandling *Att klara sig i okänd natur. En studie av betyg och betygskriterier – historiska betingelser och implementering av ett nytt system*. Studien omfattar dels en historisk bakgrund till det svenska betygssystemet, dels en analys av lärares förståelse av den 1995 introducerade nya läroplanen i grundskolan och det nya betygssystemet där lärare till en början själva förväntades utveckla kriterier för två av de tre betygen. Den historiska delstudien har sin grund i olika officiella dokument, både statliga utredningar och dokument från centrala myndigheter. Den omfattar perioden från folkskolans etablering 1842 fram till tiden för studiens genomförande. I den andra delstudien används olika lokala dokument, planer och kriterier för bedömning i engelska, kemi och idrott och hälsa. Dokumenten insamlades vid två olika tidpunkter, 1995 och 2005. Med referens till den historiska studien konstaterar författaren att bedömningen liksom utbildningen i övrigt reglerades alltmer från 1842 fram till 1970-talet då en förskjutning från centralisering till decentralisering påbörjas. Särskilt noteras konsekvenserna av detta på myndighetsnivå där Skolverket initialt inte gav svar på frågor om hur den lokala skolverksamheten skulle utformas. Något som successivt förändras under 2000-talet, bland annat genom etableringen av Myndigheten för skolutveckling 2003. Myndigheterna kom därefter att publicera ett stort antal texter och råd om bedömning och betygssättning. Problem med implementering identifieras som gäller relationen mellan betyg och kunskapssyn, betygssättning när mål saknas samt bristande information till lärare. Den skolrelaterade studien analyseras med distinktionen mellan formulering och realisering som utgångspunkt och med intresse för frågor om makt och socialisation. Författarens slutsats är att ovan nämnda problem i ljuset av den svenska skolans bedömningshistoria ledde till att lärare skrev sina egna kriterier och därmed fick större makt över betygen än någonsin förut. Elevers och deras föräldrars inflytande var alltså ytterst begränsat. Genom det empiriska bidraget kan en förändring mellan 1995 och 2005 iakttas. Under denna period rikriktas i hög grad kriterierna för bedömning.

Till de sociologiska, psykometriska och reformhistoriska perspektiven kan vi med utgångspunkt i vårt material också foga dels en kombination av en kunskaps sociologisk och en läroplanshistorisk tradition, dels en idéhistorisk. Exemplet inom den förstnämnda utgörs i vårt material av Christian Lundahls (2006) avhandling *Viljan att veta vad andra vet. Kunskapsbedömning i tidigmodern, modern och senmodern skola*. Ambitionen är att i relation till tidigare traditioner förstå bedömningar primärt ur ett pedagogiskt perspektiv och med fokus på både interna och externa bedömningspraktiker och hur dessa formats, manifesterats och reformulerats. Studierna som ingår i avhandlingen är empiriskt grundade i historiska och samtida textmaterial som utbildningsvetenskaplig forskning, offentliga utredningar, tidskrifter som riktar sig till lärare och kursmaterial för fortbildning av lärare. Valet av material är ett uttryck för det läroplansteoretiska perspektivet i vilket forskningen, politiken/administrationen och den pedagogiska praktiken betraktas som samproducenter av utbildningen och dess villkor. Avhandlingen i första hand teoretisk, även om också tidigare icke studerade aspekter av kunskapspraktiker uppmärksammas, t.ex. bedömningar sett utifrån tidiga skolordningar eller utifrån olika lärarfackliga tidskrifter. Kunskapsbedömningar beskrivs och analyseras i förhållande till skolans interna och externa praktiker som både producerar och reproducerar vetande. Pedagogiska och administrativa bedömningar förstås som ömsesidigt strukturerande. Särskilt uppmärksammas hur de senare fungerar styrande och påverkar lärares autonomi och läroplansvalida bedömningar.

Gro Hanne Aas (2006) avhandling *Likhet uten solidaritet? Idehistoriske studier av karakterer i utdanning og meritokrati* är, som framgår av titeln, ett idéhistoriskt exempel. I centrum står betygssystemet, men i förhållande till tidigare studier i en bredare kontext än utbildning. Betygssystemet relateras till meritokratiska idéer och normer. Avhandlingens övergripande fråga handlar om varför bruket av betygssystem blivit så självklara, så naturaliserade. Inkluderad i avhandlingen är en tidigare publicerad studie Utdanningsentreprenørene. *Idéhistoriske studier av karaktergivning hos jesuitter og andre ordensmenn* (Aas 1997/2006). Intresset var här inriktat på motiveringar för att införa betyg. Analysen visar att jesuiterna införde betyg primärt utifrån administrativa och inte pedagogiska hänsyn. Avhandlingen består av fyra delar som bidrar till huvudfrågan med textanalyser av olika material (offentliga utredningar, meritokratirelevanta källor av olika slag) och delfrågor. I den inledande delen fokuseras meritokrati utifrån analyser av Michael Young (1961) och boken *The Rise of the Meritocracy* och dess sammanhang. I en annan del analyseras normalfördelning som betygspolitik. Aas studerar här det norska utbildningspolitiska klimatet på 1930-talet och

re/konstruerar fyra diskurser som normalfördelningen förstås inom: enhetspolitik, reformpedagogik, vetenskapliggörande och befolkningspolitik. Därefter tar Aas upp den norska skolmannen Ribsskog och hans betydelse inom pedagogikfältet på 1930-talet. Därigenom kunde förståelsen av norsk reformpedagogik som både rörelse och hållningar lyftas fram. Den fjärde delen är en analys av 1970-talets betygskritik i norsk utbildningspolitik. Här konstruerar Aas betygssystemet som en social teknologi med många olika funktioner för flera skilda intressenter. Sammantaget visar Aas hur både betygssystemet och de meritokratiska normerna kombinerar ideal av jämlikhet och konkurrens på ett särskilt sätt. I utbildning uppstår en legitim och rättvis ojämlikhet genom lika möjligheter att konkurrera om den. Men med referens till Michael Youngs bok om meritokrati framhåller hon att ett samhälle baserat på principen om jämlika chanser till meriter har små chanser att överleva som kultur och livsstil.

9.3.2 Styrnings- och policystudier

Om vi nu vänder oss till systemstudier som har fokus på styrning och/eller policy kan vi för det första notera att vi även här har att göra med studier av skilda slag, såväl teoretiskt som metodologiskt. Det är inte ovanligt att svenska avhandlingar har referenser till statliga intentioner, ofta för att legitimera att ett specifikt område studeras. Det handlar om studier som analyserar material, kommunikation, aktiviteter och händelser med referens till de utbildningspolitiska rummen och de beslut som fattas där. Hit hör också studier som uppmärksammar centrala myndigheters verksamheter. Studieobjekten kan här variera från mikro- till makronivå och från det lokala och nationella till det internationella alternativt eller vara en kombination av dessa. Av de studier vi nämnt är det flera som också specifikt utvecklar kunskap om styrningen av skolan. Likväl har vi valt att addera ytterligare tre exempel som delvis skiljer sig åt vad gäller disciplinär hemvist, teoretiska utgångspunkter, studieobjekt och metodologiska ansatser. Två UVK-projekt och en avhandling används som exempel.

Det ena UVK-projektet gäller en forskarassistenttjänst. I *Den statliga skolinspektionen: En centraliserad resultatstyrning möter en decentraliserad skola* riktar Linda Rönnberg (2007b) intresset mot den samtida centraliseringen och decentraliseringen. Syftet med projektet är att beskriva, analysera och kritiskt granska spänningen mellan dessa. Med skolinspektionen som exempel studeras utifrån en policyteoretisk tradition problemsituationen, programmet, genomförandet och konsekvenserna. Hon utför såväl dokumentstudier (offentligt tryck) som intervjuer (politiker och tjänste-

män på nationell och kommunal nivå samt skolledare och lärare). Ytterst handlar projektet om utbildningssystemets politiska villkor och styrbarhet. Avsikten är också att bidra empiriskt med kunskap inom ett område som dels är en gammal företeelse i en ny kontext, dels är föga studerad i Sverige. Projektet pågår fram till 2011.

Det andra UVK-projektet, *Bildningsgångar och skolresultat i det senmoderna samhället* (Lindblad m.fl. 2008), har som syfte att analysera innebörden av resultatstyrning för skolors verksamhet och hur elever med olika social och kulturell bakgrund formar och formas utifrån skolprestationer och skolors resultat. Det är med andra ord relationen mellan resultat och verksamhet i den performativa vändningen som projektet vill utveckla kunskap om. Mot denna bakgrund krävs en begreppsapparat som kan hantera både pedagogiska system och processer och innebörden av politiska och sociala kontexter. Snarare än att utgå från ett implementeringstänkande är det den självreglerande verksamhet som observerar förändringar i omvärlden, och eventuellt agerar utifrån denna, som utgör den teoretiska utgångspunkten. Det innebär också att utbildningssystemet är självreglerande visavi det politiska och ekonomiska systemet. Centralt för förandet av elevers bildningsgång är frågor om kunskap, differentiering och kategorisering. För att fånga in detta studeras verksamheten i skolan, de levda läroplanerna. För att kunna studera dessa frågor i termer av både stabilitet och förändring genomför projektet dels en longitudinell studie med upprepade mätningar, dels en etnografisk skolstudie. I den senare analyseras dokument, klassrumsobservationer och intervjuer. Å ena sidan studeras hur skolor/klasser hanterar skolresultat som organiserade hågkomster, å andra sidan undersöks hur skolresultaten hanteras av personer, vad som noteras, hur de hanteras och hur personerna agerar utifrån dessa. Studien är omfattande och förväntas avrapporteras 2011.

Vårt sista exempel är Daniel Petterssons (2008) avhandling *Internationell kunskapsbedömning som inslag i nationell styrning av skolan*. Avsikten med studien är att utveckla kunskap om hur internationella kunskapsmätningar formar en politisk styrning av skolan i ett samspel mellan internationella och nationella faktorer. Formeringen studeras utifrån hur det internationella landar, diskuteras och medieras i den nationella utbildningspolitiska praktiken. Avhandlingen är förankrad i en sociologisk konfliktteoretisk tradition där makt, förhandlingar och ideologisk hegemoni är några av studiens analytiska begrepp. Avhandlingen rymmer en historisk översikt över IEA och OECD, med hänvisning till internationella kunskapsmätningar. Dessutom ingår en studie om de internationella mätningarna i den nationella kontext som studerar hur politiker, administratörer och aktörer inom media agerar när mätningar som PISA och TIMSS landar i det natio-

nella sammanhanget. En rad olika dokument från riksdag och myndigheter liksom pressmeddelande och artiklar studeras. Det svenska belyses också särskilt i relation till ett urval andra länder i samband med lanseringen av PISA 2000 i media. Med textanalyser exemplifieras hur ramar förändras, varför de förändras och vilka determinanterna är. Studierna visar att det internationella kommer till uttryck i det nationella på olika sätt. För det första bidrar det internationella till att omformulera relationer genom en kamp om vad som ska anses som utbildningspolitiska problem och legitima sätt att hantera dessa. Nationella aktörer omformulerar för det andra den internationella ideologin genom att omvandla externa prioriteringar till interna. För det tredje kan den internationella ideologin förstås som en isomorfisk ideologi som måste medieras. Författaren lyfter fram tre resultat. Internationella kunskapsmätningar används som en arena för kommunikation i vilken argument kan lyftas fram och diskussioner föras. Mätningarna används också som referenspunkter för beskrivningar och utgångspunkter för en vision om den framtida utbildningen. Slutligen används de för att legitimera reformering eller status quo. Sammanfattningsvis konstateras att de nationella utbildningspolitiska aktörerna har en avgörande roll för att omformulera och mediera den internationella utbildningsideologin.

9.4 Ett sammanfattande avstamp

Vi har ovan identifierat tre huvudområden med undergrupper: metriska studier, professions- och verksamhetsstudier samt systemstudier. Utifrån dessa huvudområden vill vi avslutningsvis lyfta fram några punkter. För det första ger avhandlingarna sammantaget ett kalejdoskopiskt intryck, de framstår som separata bitar, med få inbördes relationer. Få av dem bidrar till en sammanhållen kunskapsuppbyggnad. Både studieobjekten och kunskapsobjekten är mångfaldiga och perspektiven är flera och varierande. De olika metriska studierna, om individuella differenser, provkonstruktion, övergångar, komparationer och effektstudier, är dock ett undantag. Både var för sig och i relation till varandra framstår de som relativt tydligt avgränsade från de övriga avhandlingarna. Bristen på gemensamt fokus och sammanhållning är densamma i relation till UVK-projekten och artiklarna. Inom de två inriktningarna professions- och verksamhetsstudier samt systemstudier kommer detta bland annat till uttryck genom att de flesta studierna brukar perspektiv, begrepp och metoder som ännu inte är etablerade inom bedömningsforskningen. En beskrivning av bedömning som objekt för kunskapsbildning i Sverige är därför att det är ett både mångfacetterat och

fragmentariskt objekt där några delar framstår som tydligare än de andra, men där det är svårt att precisera beskrivningen av vad de är exempel på. En annan beskrivning är att det är ett objekt under utveckling, att det är ett epistemiskt objekt under re/formering, och att de aspekter som först blivit tydliga har fått konkurrens på många fronter samtidigt. Här finns goda möjligheter till vidare utveckling av forskningen om bedömning.

10 BEDÖMNINGSFORSKNINGEN – NÅGRA UTVECKLINGSLINJER

Här gör vi en återkoppling till bedömning som social verksamhet (jfr kapitel 3), utblicken till den internationella forskningen om bedömning (jfr kapitel 4) och redovisningen av den svenska bedömningsforskningen (jfr kapitlen 5–9). Med dessa som referenspunkter riktar vi uppmärksamheten mot framtiden och pekar ut vad vi menar kan vara fruktbara sätt att vidareutveckla forskningen om bedömning.

10.1 Svensk bedömningsforskning

– en lägesbestämning

I kapitel 3 står bedömning som social verksamhet i centrum som den framträder med det material och de avgränsningar vi gjort. Den bild som tonar fram är skriven med utgångspunkt i ett läroplansteoretiskt/didaktiskt perspektiv. Där använder vi studier om bedömning, men också forskning som inte nödvändigtvis brukar förbindas med bedömning. Avsikten var bland annat att visa eventuella kunskapsluckor i den forskning som tagits med i översikten. Beskrivningen fokuserar på centrala karakteristika som tar fasta på både stabila mönster och mer temporära inslag. Bedömningar såväl vid entrén till skolan och i skolans inre processer som vid exit uppmärksammades. Detsamma gäller informella kontinuerligt pågående bedömningar och mer formella bedömningar som prov, betyg och examina. Dessutom beaktades betydelsen av skillnader mellan olika innehåll och utformning liksom vikten av olika kommunikationsformer. Därtill fokuserades särskilt bedömningsars olika syften och deras funktion som redskap för att ordna eleverna både i och efter skolan. Det svenska systemet för bedömning och dess relation till utbildningssystemet i sin helhet och dess sammanhang med globaliseringen och omstruktureringen av systemet uppmärksammades liksom frågor om legitimitet, makt och etik.

Om vi vänder blicken mot den internationella bedömningsforskningen, genom översikt bilderna av innehållet i handböcker och virtuella besök i ett urval paraplyorganisationer och några för bedömning specifika organisationer, framstår bedömning som ett kunskapsobjekt i omvandling. De senaste åren har medfört en utvidgning av vad som inkluderas och assessment och

evaluation förstås alltmer som avhängiga av varandra och som aspekter av en och samma sak. Fram träder ett objekt som inkluderar såväl psykologiska, pedagogiska och ekonomiska som sociologiska perspektiv och som förbinder fenomen på mikro- och makronivå.

Psykometrin med dess intresse för begåvning, test och provkonstruktion är en central inriktning med fokus på statistiska analyser och modellutveckling. Vid sidan av denna finns inom klassrumsforskningen studier av både traditionella och alternativa former för bedömning. Inom de senare framhålls vikten av att utveckla multimodala redskap för lärares bedömningar av elevers kunskande. Här betonas också det professionella omdömet och seendet. Både läroplansteoretiska och didaktiska perspektiv används inom klassrumsforskningen. Särskilt betonas här också vikten av att analysera hur teorier om lärande, kunskap och bedömning är relaterade till varandra. En annan inriktning som lyfts fram är system- och policystudier med fokus på styrning, demokratiska värden, utveckling och skolan som organisation.

Utifrån den bild av svensk bedömningsforskning som tonar fram i kapitlen 5–9 kan vi för det första notera att det är ett fält under tillväxt. Antalet studier har ökat markant under den studerade perioden oavsett om vi beaktar avhandlingarna eller artiklarna. Ökningen gäller särskilt 00-talet. Sammantaget publicerades 95 avhandlingar, 63 artiklar i svenska vetenskapliga tidskrifter och 119 i internationella tidskrifter. Dessutom erhöll 35 bedömningsrelaterade projekt medel från UVK under perioden 2006–2009. Genom att studera vilka forskningsmiljöer som bidrar med kunskap inom bedömningsforskningen kunde vi konstatera att intresset i dag kommer från många olika discipliner, även om pedagogikrelaterade institutioner alltjämt dominerar. Dessutom kunde vi slå fast att några miljöer var särskilt framträdande och att de utmärktes av att de hade både statliga myndighetsuppdrag (främst provkonstruktion) och ingick i den nationella forskarskolan i bedömning.

Vi har också konstaterat att svenska forskare är frekventa skribenter i de nordiska och internationella tidskrifterna jämfört med övriga nordiska länder. I de internationella tidskrifterna utgör de svenska artiklarna nästan hälften av alla nordiska som publicerats. Det innebär att Danmark, Finland, Island och Norge gemensamt producerat drygt hälften av artiklarna. Sett utifrån avhandlingarna är författarna till övervägande delen kvinnor. Men förhållandet är omvänt när vi betraktar det svenska inspelet på den internationella arenan där männen är i lika klar majoritet. Tilläggas kan att andelen sammanläggningsavhandlingar ökat, med påföljden att delstudier ofta återfinns i nordiska eller andra internationella tidskrifter. Artiklarna är publicerade i närmare 80 olika tidskrifter. De flesta forskarna har enstaka artiklar, även om ett tjugotal har två och ett fåtal flera. Många skriver till-

sammans med andra när de publicerar sig internationellt. En fråga man kan ställa sig är: vad innebär det att många svenska forskare representeras av en enda artikel? Samtidigt ser vi att flera av dem som skrivit avhandlingar av relevans för fältet inte alls producerat artiklar, varken nationellt, nordiskt eller internationellt.

Om vi riktar uppmärksamheten mot studieobjekten kan vi slå fast att den största andelen studier fokuserar elevers prestationer och skolors resultat följt av bedömning i relation till lärares didaktiska arbete, som emellertid är i majoritet i avhandlingarna. Resultatstudierna dominerar artiklarna, i synnerhet de nordiska och internationella. Det finns också en tydlig övervikt av publikationer som behandlar bedömning i grundskolan, men även gymnasieskolan och övergångar från dem uppmärksammas. Systemövergripande studier förekommer också. Flera studieobjekt studeras, dock mer sparsamt, t.ex. ordnandet av elever i skolor (differentiering), övergångar, bedömning som fenomen, elevers erfarenheter och upplevelser av bedömning samt bedömningsredskapens kvalitet.

Den forskning som tar sin utgångspunkt i elevers prestationer uttryckta i samlad meritvärde, nationella prov och IEA-studier är det mest omfattande forskningsområdet med flest antal studier. Betygen i enskilda ämnen får mindre uppmärksamhet och då enbart i matematik, svenska och engelska, medan betygen i övriga ämnen får föga eller ingen uppmärksamhet. Ett annat område som det finns flera studier om och där man kan tala om en mer kumulativ kunskapsproduktion är studier relaterade till övergångar från gymnasieskolan till högskolan, medan övergångar från gymnasieskolan till arbetslivet samt till och från grundskolan är sparsamt representerade. Utöver dessa finns det bland avhandlingarna en numerärt stor del som ger antingen empiriska eller teoretiska tillskott till bedömningsforskningen, men som representeras av såväl varierande studie- och kunskapsobjekt som perspektiv. Andelen sådana studier förekommer mer sällan i artiklarna, men bredden kvarstår. Som helhet betraktade är det bidrag de ger därför främst av fallstudiekaraktär. Undantaget utgörs av ämnesdidaktiska studier i svenska och matematik, där det redan går att tala om en mer samlad kunskap. Även det kunskaps sociologiska och historiska området har utvecklats under senare år.

Om vi nu betraktar denna beskrivning av bedömningsforskningen i relation till bedömning som social praktik och den internationella bedömningsforskningen kan vi för det första konstatera att den svenska forskningen i linje med den internationella vidgats i dag omfattar den fler perspektiv, objekt och metoder och producerar därmed också en mer varierad bild av bedömningen i skola och undervisning. Samtidigt kan vi notera att det över huvud taget är sparsamt med studier som fokuserar de informella bedöm-

ningarna både som ett studieobjekt i sig och vad gäller deras relation till de mer formella bedömningarna. Det förefaller som om de två positioner som är inriktade på att utveckla kunskap om prov av traditionellt slag respektive mer alternativa former båda har fokus bortom det mer vardagliga. I det ena fallet är det fokus på det befintliga och mer formella och i det andra är det behovet av att ersätta det befintliga med något annat som betonas. Oavsett position tycks de informella bedömningsformerna falla mellan stolarna.

En annan aspekt som uppmärksammas i olika sammanhang men som hittills inte fått sin motsvarighet i forskningen är nödvändigheten av ämnesdidaktisk bedömningsforskning. Som framgått är det i första hand svenska och matematik som uppmärksammas och, som en konsekvens av de internationella jämförande studierna, i någon mån även naturvetenskapliga ämnen. Enstaka exempel finns också på några andra ämnen, men många skolämnen är föremål för forskning i mycket liten utsträckning. Det ökade inslaget av forskare från olika discipliner bäddar för möjligheten till studier av fler skolämnen. Exempel på detta finns också i de pågående UVK-projekten. Den allmänpedagogiska och samtalsanalytiska forskningen har pekat på förändrade interaktions- och kommunikationsmönster i skolan. Förekomster av flera olika mönster öppnar för komparativa studier där innehåll och utformning av undervisning kan studeras med avseende på hur olika varianter och kombinationer samspelar med skilda former för bedömning och konsekvenserna av detta. I det här sammanhanget kan också lärares behov av redskap för pedagogisk bedömning lyftas fram. Sådana studier förekommer, men de är totalt sett få. Det behöver undersökas på skilda sätt med hänsyn tagen till olika ämnen och arbetsformer och i relation till elevers olikheter.

Vidare behöver lärares perspektiv på och erfarenheter av bedömningsarbetet synliggöras i högre grad än tidigare. Detta gäller exempelvis hur lärare resonerar om och motiverar sina bedömningar, rationaliteten för lärarnas bedömningsarbete. Vidare är det angeläget att kartlägga vilka olika redskap för bedömning som är i bruk på skolorna i olika ämnen och hur de används. Internationell forskning har visat på behovet av att utveckla redskap för bedömning, vilket kan ses som ett tydligt uppdrag för praxisutvecklande forskning. Trots att det finns ca 165 000 lärare i Sverige har intresset för deras bedömningsarbete och erfarenheter hitintills varit marginellt. Detsamma gäller i ännu högre utsträckning eleverna. Översikten visar att studier där elevers upplevelser och erfarenheter, eller konsekvenserna, av bedömning kommer till uttryck är fortfarande starkt underrepresenterade. Likaså saknas nästan helt studier där elever inte ses som (strukturellt determinerade) objekt för bedömning utan även som aktörer i bedömnings- och betygssättnings-sammanhang.

Bedömning som en central aspekt av elevers skolgång och dess relation till frågor om makt och moral i relation till eleverna hör också till områden som behöver studeras. Området är beforskat och teorier och metoder utvecklade för att studera t.ex. normalitet och avvikelse. Vad som också behövs är studier av hur dessa frågor kommer till uttryck vid sidan av problematiken om in- och exkludering. Det finns i dag ett antal studier som utvecklat kunskap om systemaspekter som accountability, performativitet etcetera. Det finns också några pågående och nyligen avslutade UVK-projekt som uppmärksammar relationen mellan mikro, meso och makro. Men i tider av omförhandling och förändring av betygssystemet finns det anledning att särskilt uppmärksamma dessa frågor och att göra det med en varierad uppsättning metodansatser. I det här sammanhanget kan också framhållas vikten av att fortsatt analysera hur politik, administration, forskning och pedagogik ömsesidigt påverkar varandra och på vilket sätt i relation till olika bedömningspraktiker. Hur träder man in i varandras verksamheter och på vilket sätt påverkar det skolan och bedömningarna.

I den internationella forskningen utpekas ett antal områden i behov av forskning. Det är viktigt att utveckla kunskap om nyckelbegrepp, bedömningskontext, bedömning och IKT, kunskap/lärande/bedömning och relationen mellan klassrumsbedömningar, standarder och systemkontroller. Från diskussionen den förda kan vi också känna igen en uppmaning om att fokusera frågor om legitimitet, makt och etik. Särskilt betonas också att forskningen bör bidra med kunskap och metoder/redskap som kan brukas inom policy/administration och i lärares yrkesverksamhet. Flera saker som vi redan nämnt återkommer här. Nytt är intresset för att studera relationen mellan IKT och bedömningar på såväl klass- och skol- som systemnivå. Denna typ av studier är ytterst begränsade i vårt material och inte i något fall huvudfokus. Med referens till den internationella forskningen kan vi alltså peka ut områden som endast i ringa omfattning har studerats i Sverige. Här finns med andra ord utrymme för framtida insatser.

10.2 Svensk bedömningsforskning

– ett framtidsperspektiv

Vi har gjort en sammanfattande bestämning av läget för den svenska bedömningsforskningen som den framträder i relation till valda avgränsningar och tar den som avstamp för att diskutera några möjliga vägar framåt. Vi kontrasterar den mot hur den svenska bedömningsforskningen har beskrivits

tidigare. Under slutet av 1990-talet utvärderades pedagogikämnet i Sverige (Rosengren & Öhngren red. 1997). I samband med utvärderingen konstaterades i ett kapitel om lärande och individdifferenser att

There has been a substantial volume of work on testing and assessment in recent years in Sweden. Some research and development work has continued in the psychometric tradition of standardized testing, while other projects have involved both conventional and 'authentic' tests for use in national testing of various school subjects. ... There are substantial difficulties in combining a programme of research with the demands of producing a test for national use: the opportunities for using novel testing techniques is inevitable somewhat constrained. (Entwistle 1997 s. 87–88)

Vid sidan av konstaterandet att det finns en omfattande forskning om test och bedömningar framhålls, som också vi gör i kapitel 7, att de institutioner som är aktiva här är de som har myndighetsuppdrag av olika slag. Utvärderaren utvecklar detta ytterligare.

In Sweden as in other countries however, test development is driven to a large extent by the demands of government and the education community at large. In this way, what can be developed is to a significant extent constrained by policy decisions, and a considerable amount of time and energy is spent on producing materials in a fairly routine way. (s. 89)

Sen dess har miljöerna framgångsrikt erövrat medel från UVK för både projekt och den nationella forskarskolan i pedagogisk bedömning. Samtidigt har de erhållit kanske ännu fler statliga uppdrag, som en konsekvens av det ökade trycket på att producera information om utbildningssystemets standard och utveckling. Mot den bakgrunden är situationen i dag förbättrad. Det är också i dessa miljöer som traditionen av ett nära samarbete med skolans lärare och elever vid utprovning och utveckling av test upprätthålls. Här finns med andra ord också goda grunder för att utveckla praktisknära forskningsprojekt, projekt i skärningspunkten mellan forskning, politik och praktik. Sådana exempel finns i vårt material men förekommer sparsamt.

I utvärderingen framhålls ytterligare två problem med bedömningsforskningen som primärt är avgränsad till provkonstruktions- och testtraditionen. Ett handlar om frånvaron av fokus på ickekognitiva faktorer. Av vårt material framgår att en utveckling har skett även i detta avseende. Både avhandlingar och artiklar har numera producerats inom området, om än inte i någon större omfattning. Dessutom påtalas att en närmande mellan forskning om lärande i klassrummet och forskning om individdifferenser är önskvärt.

The methodological and theoretical separation between research on individual differences and on learning in classrooms remains strong, with little evidence of researchers seeking to this divide by looking at individual differences in learning.

Det här kan kopplas ihop med det tidigare påtalade behovet av studier som förbinder lärande–kunskap–bedömning med varandra. Det är frågor som har börjat studeras, men här återstår mycket att göra. I en översikt i en bilaga till *Kunskapsbedömning – hur, vad och varför* konstaterade Helena Korp (2003)

att det i Sverige har publicerats anmärkningsvärt lite forskning om kunskapsbedömning. I synnerhet tycks det vara ont om forskning av "grundforskningskaraktär". Vid en snabb genomgång synes den vetenskapliga publikationen om kunskapsbedömning till den allra största delen vara inriktad mot pragmatiska frågeställningar, medan mera teoretiska frågor om exempelvis bedömningens sociala och ideologiska funktioner och dess pedagogiska konsekvenser är mera sällsynta som utgångspunkt för forskning och diskussion. (s. 170)

Forskningen av grundforskningskaraktär är alltjämt ringa, men en avgörande utveckling har skett. Det har medfört både ett empiriskt nytillskott och en teoretisk utveckling av hur bedömning som fenomen kan förstås i både ett pedagogiskt och ett administrativt perspektiv. Vidare har statistiska modeller förfinats, nya kategorier fogats till tidigare existerande, bland annat genom nya discipliners intresse för bedömning etcetera. Vi har identifierat flera nya avhandlingar och några UVK-projekt som utgör bidrag här, dels inom redan etablerade områden som förstärkts, dels inom tidigare relativt utforskade områden. Denna utveckling är en början och behöver förstärkas ytterligare. Av särskilt intresse är empiriskt grundade teoretiseringar.

Alla dessa förändringar till trots kan vi fortfarande se att det Viveca Lindberg (2005) tidigare konstaterat om svensk bedömningsforskning alltjämt gäller, de psykometriska mätningarna dominerar starkt och övrig forskning är relativt fragmentariserad. Men det här är en sanning med modifikation. Den psykometriskt inriktade forskningen motsvarar i grova tal cirka en fjärdedel av forskningen om bedömning. Det som gör att den framstår som dominant är snarare frånvaron av enhetlighet i övrig forskning. Vi har tidigare kategoriserat bedömningsforskningen i tre övergripande områden som i sin tur består av flera undergrupper. Några av dem är relativt nya och även om det kan finnas ganska många studier inom en riktning är de inte teoretiskt eller metodologiskt sammanhållna.

Att forskning inom vissa områden inte bedrivs i Sverige är inte självklart ett problem eller något som betyder att vi ska satsa på just det området.

Forskning är till sitt väsen primärt ett internationellt fenomen och i den mån kunskap finns om ett fenomen och kontexten inte sätter hinder i vägen kan den nyttiggöras i olika sammanhang. Det finns dock några förhållanden där det är rimligt att satsa på en utveckling på nationell nivå. En aspekt handlar om möjligheten att utveckla spetskompetens. Det kan då gälla att tillvarata redan befintlig kompetens och utveckla den vidare. En annan situation då det finns anledning att satsa på forskning är då kunskap saknas på både den nationella och den internationella arenan. En tredje aspekt handlar om de fall då den nationella och sociokulturella kontexten antingen genom komparativa studier kan främja forskningen i ett internationellt perspektiv eller kan belysa frågor av betydelse för aktörer och verksamheter nationellt. Det sistnämnda är särskilt angeläget när utbildningssystemet genomgått omfattande förändringar, vilket nu är fallet. Det finns också ett ytterligare skäl för att studera bedömning nationellt. Som ett gränsobjekt mellan olika sfärer (hem, skola, arbete), mellan skilda aktörer (elever, lärare, administratörer, politiker, media, forskare) och med många olika funktioner är det kontextuellt sensitivt och ett fenomen som har många lokala och nationella intressenter, för vilka frågan är av stor betydelse.

Vi pekade inledningsvis på bedömning som en integrerad aspekt av läroplanen och som en fråga som skär genom många områden. Detta är både en möjlighet och en svårighet. Det förra pekar på sannolikheten att vi kan förvänta oss ytterligare tillväxt av forskningen om bedömning, eftersom den berör så många fält och olika kunskaps- och studieobjekt. Samtidigt kan just detta bidra till att området inte får den särskilda uppmärksamhet det förtjänar och som behövs mot bakgrund av dess betydelse. Det kan därför finnas anledning att fundera över att utveckla ett forskningsprogram som kan innesluta olika inriktningar och bidra till utveckling av vart och ett av dem, men också relationerna dem emellan. En avgörande slutsats av översikten är att det inom vissa områden endast finns ytterst få studier. Men i många fall finns det forskning sin inte förstås eller förbinds med bedömning. Vi ser flera forskningstraditioner som kan bidra med kunskap om bedömning: den psykometriska/edumetriska, den läroplansteoretiska, den ämnesdidaktiska, klassrumsforskningen, barn- och ungdomsforskningen, den specialpedagogiska forskningen, den kunskaps sociologiska, den samtalsanalytiska, den historiska och den samhällsvetenskapliga forskningen i en mer generell mening. Dessa forskningstraditioner behöver relateras till varandra för att ytterligare förstärka bedömning som ett objekt för kunskapsbildning, såväl teoretiskt och metodologiskt som empiriskt.

REFERENSER

- Aas, Gro Hanne (2006). *Likhet uten solidaritet? Idéhistoriske studier av karakterer i utdanning og meritokrati*. Göteborgs universitet, Institutionen för idéhistoria och vetenskapsteori.
- Aas, Gro Hanne (2006). *Utdanningsentreprenørene. Idéhistoriske studier av karaktergivning hos jesuitter og andre ordensmenn*. Göteborgs universitet, Institutionen för idéhistoria och vetenskapsteori (original från 1997).
- Adelmann, Kent (2002). *Att lyssna till röster. Ett vidgat lyssnandebegrepp i ett didaktiskt perspektiv*. Lunds universitet.
- Adolfsson, Lena & Henriksson, Widar (1999). Different methods – different results. How the Swedish results in mathematics vary depending on methodological approach. *Educational Research and Evaluation, Vol. 5*(2) pp. 127–13.
- Ahlström, Karl-Georg (2009). En kanal ut i världen. Svensk pedagogik i Scandinavian Journal of Educational Research. I L. Wikander, C. Gustafsson, U. Riis & L. Larson, (red.) *Pedagogik som examensämne 100 år*. Uppsala universitet, Pedagogiska institutionen.
- Ahlström, Karl-Georg, Emanuelsson, Ingemar, & Wallin, Erik (1986). *Skolans krav – Elevernas behov*. Lund: Studentlitteratur.
- Allal, Linda & Pelgrims Ducrey, Greta (2000). *Assessment of – or in – the zone of proximal development. Learning and Instruction, Vol 10*(2), pp. 137–152
- Ammermueller, Andreas (2007). PISA. What Makes the Difference? Explaining the Gap in Test Scores Between Finland and Germany. *Empirical Economics, Vol. 33*(2), pp. 263–288.
- Amnå, Erik, Englund, Tomas, Eriksson, Cecilia, Ljunggren, Carsten, Ekman, Joakim & Uenmar Öst, Ingrid (2009). *Skolornas institutionella karaktär och elevernas medborgarkompetens. En jämförelse av olika kommunala och fristående skolor över tid och rum*. Örebro universitet, Akademin för humaniora/statsvetenskap. Forskningsprojekt finansierat av Vetenskapsrådet nr. 2009-4402.
- Anderson, Lorin W. (ed.) (1995). *International Encyclopedia of Teaching and Teacher Education*. 2. ed. Oxford: Pergamon.
- Anderson, Neil & Herriot, Peter (eds) (1997). *International Handbook of Selection and Assessment*. Chichester: John Wiley.
- Andersson, Håkan (1991). *Relativa betyg. Några empiriska studier och en teoretisk genomgång i ett historiskt perspektiv*. Umeå universitet, Pedagogiska institutionen.
- Andreasson, Ingela (2007). *Eleveplanen som text – Om identitet, genus, makt och styrning i skolans elevdokumentation*. Göteborgs universitet, Institutionen för pedagogik och didaktik, No. 259.
- Andrée, Maria (2007). *Den levda läroplanen. En studie av naturorienterande undervisningspraktiker i grundskolan*. Lärarhögskolan i Stockholm, Institutionen för samhälle, kultur och lärande.
- Anward, Jan (2003). Språkutveckling i skolan: praktik, teoretisk praxis och praktisk teori. I I. Carlgren (red.) *Forskning av denna världen: praxisnära forskning inom utbildningsvetenskap*. Stockholm: Vetenskapsrådet. (pp. 9–14.)

- Ask, Sofia (2007). *Vägar till ett akademiskt skriftspråk*. Växjö universitet, Institutionen för humaniora.
- Asp-Onsjö, Lisa (2006). *Åtgärdsprogram – dokument eller verktyg? En fallstudie i en kommun*. Göteborgs universitet, Institutionen för pedagogik och didaktik, No. 248.
- Aunola, Kaisa, Leskinen, Esko & Nurmi, Jari-Erik (2006). Developmental Dynamics between Mathematical Performance, Task Motivation, and Teachers' Goals during the Transition to Primary School. *British Journal of Educational Psychology*, Vol. 76(1), pp. 21–40.
- Axelsson, Thom (2007). *Rätt elev i rätt klass. Skola, begåvning och styrning 1910–1950*. Linköpings universitet, Institutionen för tema barn, No. 379.
- Bascia, Nina, Cumming, Alistair, Datnow, Amanda, Leithwood, Kenneth & Livingstone, David (Eds) (2005). *International Handbook of Educational Policy*. Dordrecht: Springer
- Benincasa, Luciana (1995). *A Journey, a Struggle, a Ritual. Higher Education and the Entrance Examinations in a Greek Province Town*. Studies in Comparative and International Education, 41. Stockholm University, Institute of International Education.
- Berggren, Caroline (2007). Broadening Recruitment to Higher Education Through the Admission System. Gender and Class Perspectives. *Studies in Higher Education*, Vol. 32(1), pp. 97–116.
- Bergman, Lotta (2007). *Gymnasieskolans svenskämnen. En studie av svenskundervisningen i fyra gymnasieklasser*. Malmö högskola, Lärarutbildningen.
- Bergsten, Christer (1990). *Matematisk operativitet. En analys av relationen mellan form och innehåll i skolmatematik*. Linköpings universitet, Institutionen för pedagogik och psykologi.
- Bernstein, Basil (ed.) (1971/1980). *Class, Codes and Control. Vol. 1, Theoretical Studies Towards a Sociology of Language*. London: Routledge.
- Berthén, Diana (2007). *Förberedelse för särskildhet: särskolans pedagogiska arbete i ett verksamhets-teoretiskt perspektiv*. Karlstads universitet, Avdelningen för pedagogik.
- Bjerneby Häll, Maria (2006). *Allt har förändrats och allt är sig likt. En longitudinell studie av argument för grundskolans matematikundervisning*. Linköpings universitet, Institutionen för beteendevetenskap.
- Black, Paul & Wiliam, Dylan (2009). *Developing the theory of formative assessment. Educational Assessment Evaluation and Accountability. Vol 21(1)*, pp. 5–31.
- Black, Paul & Wiliam, Dylan (2006). Assessment for learning in the classroom. I J. Gardner (ed.), *Assessment and learning* (pp. 9–26). London, Thousand Oaks & New Delhi: SAGE Publications.
- Black, Paul & Wiliam, Dylan (1998). *Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, Vol. 5 (1)*, pp. 7–73.
- Blankertz, Herwig (1987). *Didaktikens teorier och modeller*. Stockholm: HLS Förlag.
- Boesen, Jesper (2006). *Assessing Mathematical Creativity. Comparing National and Teacher-made Tests, Explaining Differences and Examining Impact*. Umeå University, Department of Mathematics and Mathematical Statistics, No. 34.

- Bohlin, Ingemar (2007). *Evidensbaserad pedagogik. Epistemologiska förutsättningar för evidensrörelsens expansion in i utbildningsvetenskap*. Göteborgs universitet, Institutionen för idéhistoria och vetenskapsteori. Forskningsprojekt finansierat av Vetenskapsrådet, nr 2007-3665.
- Bresler, Liora (Ed.) (2007). *International Handbook of Research in Arts Education*. Dordrecht: Springer.
- Browning, Martin & Heinesen, Eskil (2007). Class Size, Teacher Hours and Educational Attainment. *Scandinavian Journal of Economics*, Vol. 109(2), pp. 415-438.
- Carlgren, Ingrid, Forsberg, Eva & Lindberg, Viveca (2009). *Perspektiv på den svenska skolans kunskapsdiskussion*. SKIP-rapport nr 7. Stockholm, Stockholms universitets förlag.
- Carlgren, Ingrid, Klette, Kirsti, Myrdal, Sigurjón, Schnack, Karsten & Simola, Hannu (2006). Changes in Nordic Teaching Practices. From Individualised Teaching to the Teaching of Individuals. *Scandinavian Journal of Educational Research*, Vol. 50(3), pp. 301-326.
- Christensson, Wolmar (1995). *Subjektiv bedömning som besluts- och handlingsunderlag*. Göteborg Studies in Educational Sciences, No. 101.
- Cliffordson, Christina (2008). *Betygens jämförbarhet, prognosförmåga och effekter på lärande*. Forskningsprojekt finansierat av Vetenskapsrådet, Utbildningsvetenskapliga kommittén, No 2008-4994.
- Cliffordson, Christina (2004). De målrelaterade gymnasiebetygens prognosförmåga. *Pedagogisk forskning i Sverige*, vol 9 (2) pp. 129-140.
- Cliffordson, Christina & Gustafsson, Jan-Eric (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence*, Vol. 36(2), pp. 143-152.
- Colnerud, Gunnel (1997). En omöjlig jämförelse. *Pedagogisk Forskning i Sverige*, Vol. 2(2), pp. 95-104.
- Cummins, Jim & Davison, Chris (Eds) (2007). *International Handbook of English Language Teaching*. New York: Springer.
- Dahllöf, Urban (1990). Det svenska pedagogikämnets identitet. I K. Fridjonsdottir (red.) *Svenska samhällsvetenskaper*. Stockholm: Carlssons bokförlag.
- Dahllöf, Urban (1967). *Skoldifferentiering och undervisningsförlopp*. Stockholm: Almqvist & Wiksell.
- Damber, Ulla (2009). Using Inclusion, High Demands and High Expectations to Resist the Deficit Syndrome. A Study of Eight Grade Three Classes Overachieving in Reading. *Literacy*, Vol. 43(1), pp. 43-49.
- Davison, Chris & Cummins, Jim (2007). Assessment and Evaluation in ELT. Shifting Paradigms and Practices. I J. Cummins & C. Davison (Eds). *International Handbook of English Language Teaching*. New York: Springer. (pp. 415-420.)
- Deslandes, Rollande (ed.) (2009). *International Perspectives on Student Outcomes and Homework. Family_School_Community Partnerships*. London: Routledge.

- Dragemark Oscarssons, Anne (2009). *Self-Assessment of Writing in Learning English as a Foreign Language. A Study at the Upper Secondary School Level*. Göteborgs universitet, Institutionen för pedagogik och didaktik, No. 277.
- Dunkin, Michael J. (ed.) (1987). *The International Encyclopedia of Teaching and Teacher Education*. 1. ed. Oxford: Pergamon.
- Eisner, Elliot (2007). Interlude. Assessment and Evaluation in Education and the Arts. I L. Bresler (ed.) *International Handbook of Research in Arts Education*. Dordrecht: Springer.
- Egelund, Niels (2008). The Value of International Comparative Studies of Achievement – A Danish Perspective. *Assessment in Education. Principles, Policy and Practice, Volume 15*(3), pp. 245–251.
- Eklund Heinonen, Maria (2009). *Processbarhet på prov. Bedömning av muntlig språkfärdighet hos vuxna andraspråksinlärare*. Uppsala universitet, Institutionen för nordiska språk, No. 78.
- Eklöf, Hanna (2007). Test-taking Motivation and Mathematics Performance in TIMSS 2003. *International Journal of Testing, Vol. 7*(3), pp. 311–326.
- Eklöf, Hanna (2006). *Motivational Beliefs in the TIMSS 2003 Context. Theory, Measurement and Relation to Test Performance*. Umeå universitet, Institutionen för beteendevetenskapliga mätningar. Anställning som forskarassistent finansierad av Vetenskapsrådet nr 2006-2605.
- Engström, Arne (2003). *Specialpedagogiska frågeställningar i matematik. En introduktion*. Ny, omarb. uppl. Örebro universitet.
- Engström, Jan-Åke (1993). *Science Achievement and Student Interest. Determinants of Success in Science Among Swedish Compulsory School Students*. Stockholm University, Institute of International Education.
- Entwistle, Noel (1997). Learning and individual differences. I K-E. Rosengren & B. Öhngren (eds). *An Evaluation of Swedish Research in Education*. (pp. 83-98.) Uppsala: Swedish Science Press.
- Ericsson, Ingegerd (2003). *Motorik, koncentrationsförmåga och skolprestationer: en interventionsstudie i skolor 1–3*. Malmö Studies in Educational Sciences, 6. Malmö högskola, Lärarutbildningen.
- Erikson, Robert, Stenlås, Niklas & Forskningsrådsnämnden (1999). *Longitudinella studier och samhällsforskningens infrastruktur: en bok om forskningens informationsförsörjning, individbaserade databaser och kunskapsuppbyggnad för samhällsarbete*. Stockholm: Forskningsrådsnämnden (FRN).
- Eriksson, Inger (2010). "Skriftspråkliga redskap i självreglerande undervisningspraktiker" I I. Eriksson, V. Lindberg & E. Österlind (red.). *Uppdrag undervisning – kunskap och lärande*. Lund: Studentlitteratur.
- Eriksson, Inger (2009). Re-interpreting Teaching. A Divided task in Self-Regulated Teaching Practices. *Scandinavian Journal for Educational Research*. 53(1), pp. 53–70.
- Eriksson, Inger, Berg, Astrid, Danielsson, Kristina, Ekvall, Ulla Lindberg, Viveca, Löfgren, Ragnhild & Ståhle, Ylva (2006). *Kemitexter som redskap för naturvetenskapligt lärande: en komparativ studie av undervisningspraktiker i svenska och finlandssvenska klassrum*. Forskningsprojekt finansierat av Vetenskapsrådet, Utbildningsvetenskapliga kommittén 2006-2689.

- Eriksson, Rikard (1999). *Psykoteknik: kulturell fabricering av personlig identitet*. Linköpings universitet, Linköping Studies in Arts and Science, No. 200.
- Erstad, Ola (2008). Changing Assessment Practices and the Role of ICT. I J. Voogt & G. Knezek (eds). *International Handbook of Information Technology in Primary and Secondary Education*. (pp. 181–194.) New York: Springer.
- Evaldsson, Ann-Carita (2000). "För jag är bäst på svenska!". Sociala och språkliga ordningar i en mångkulturell skola. *Utbildning & Demokrati*, Vol. 9(2). pp. 29–48.
- Folkesson, Ann-Marie (1998). *Muntlig framställning i årskurs 5. Utveckling av kriterier för bedömning*. Stockholm: Almqvist & Wiksell International.
- Forsberg, Eva (2009). Kompetensrelaterade kunskapsdiskurser – i inter/nationell policy om utbildning. I I. Carlgren, E. Forsberg & V. Lindberg. *Perspektiv på den svenska skolans kunskapsdiskussion*. Stockholms universitetsförlag & Centrum för studier av skolans kunskapsinnehåll. (pp. 40–77)
- Forsberg, Eva (2008). Framtidsvägen – en huvudled eller en skiljeväg? *Utbildning & demokrati*, Vol. 17(1), pp. 75–98.
- Forsberg, Eva (2000). *Elevinflytandets många ansikten*. Uppsala Studies in Education. Uppsala universitet.
- Forsberg, Eva (red.) (2007). *Curriculum Theory Revisited*. STEP, Uppsala University, Department of Education.
- Forsberg, Eva, Lundahl Christian, Román Henrik & Waldow Florian (2006). *Internationella kunskapsmätningar, nationell bedömningskultur och reformhistoria*. Forskningsprojekt finansierat av Vetenskapsrådet, Utbildningsvetenskapliga kommittén, nr 2006-2447.
- Forsberg, Eva & Lundahl, Christian (2006). Kunskapsbedömningar som styrmedia. *Utbildning & Demokrati*, Vol. 15(3), pp. 7–29.
- Forsberg, Eva & Lundgren, Ulf P. (2004). Sweden: A Welfare State in Transition. *Change and Tradition in Global Education Reform*. I I. C. Rotberg (ed.). Change and Tradition in Global Education Reform. Oxford: Scarecrow Education.
- Forsberg, Eva & Wallin, Erik (red.) (2006). *Skolans kontrollregim – ett kontraproduktivt system för styrning? En antologi*. Stockholm: HLS förlag.
- Fransson, Karin & Lundgren, Ulf P. (2003). *Utbildningsvetenskap – ett begrepp och dess sammanhang*. Rapport 2003:1. Stockholm: Vetenskapsrådet.
- Fredriksson, Ulf, Holzer, Thomas, McCluskey-Cavin, Huguette & Taube, Karin (2009). Strengths and Weaknesses in the Swedish and Swiss Education Systems. A Comparative Analysis based on PISA Data. *European Educational Research Journal*, Vol. 1(1), pp. 54–68.
- Gage, Nathaniel L. (Ed.) (1963). *Handbook of Research on Teaching. A Project of the American Educational Research Association*, Department of the National Education Association. Chicago: Rand McNally.
- Ganuza, Natalia (2008). *Syntactic Variation in the Swedish Adolescents in Multilingual Urban Settings. Subject-verb Order in Declaratives, Questions and Subordinate Clauses*. Stockholms universitet, Centrum för tvåspråkighetsforskning.

- Garme, Birgitta (2003). Skolan, proven och demokratin. *Utbildning & demokrati*, Vol. 12(3), pp. 41–66.
- Garpelin, Anders (2008). *Övergångar i förskolan som kritiska händelser i barns liv, särskilt med avseende på delaktighet och marginalisering*. Mälardalens högskola, Akademin för utbildning, kultur och kommunikation. Forskningsprojekt finansierat av Vetenskapsrådet nr 2008-4823.
- Gerrevall, Per (2003). Bedömning av demokratisk kompetens – En pedagogisk utmaning. *Utbildning & demokrati*, Vol. 12(2), pp. 105–117.
- Giota, Joanna, Emanuelsson, Ingemar, Taube, Karin & Asp Onsjö, Lisa (2007). *Specialpedagogiska stödåtgärder i grundskolan: omfattning, former och effekter*. Göteborgs universitet, Institutionen för pedagogik och didaktik. Forskningsprojekt finansierat av Vetenskapsrådet nr 2007-3341.
- Gipps, Caroline & Cumming, Joy (2005). Assessing Literacies. I N. Bascia, A. Cumming, A. Datnow, K. Leithwood & D. Livingstone (eds). *International Handbook of Educational Policy*. (pp. 695–713.) New York: Springer.
- Gipps, Caroline (2002). Sociocultural perspectives on assessment. I G. Wells & G. Claxton (eds.), *Learning for life in the 21st century: Sociocultural perspectives on the future of education*. (pp. 73–83). Malden: Blackwell Publishing.
- Granath, Gunilla (2008). *Milda makter! Utvecklingssamtal och loggböcker som disciplinerings-tekniker*. Göteborg Studies in Educational Sciences, No. 263.
- Granström, Kjell (2003). Förändring av roller och arbetsrelationer. I G. Berg & H-Å. Scherp (red.) *Skolutvecklingens många ansikten*. (pp. 179-298.) Stockholm, Myndigheten för skolutveckling.
- Gruber, Sabine (2007). *Skolan gör skillnad. Etnicitet och institutionell praktik*. Linköping Studies in Arts and Science, 387. Linköpings universitet, Campus Norrköping, Institutionen för samhälls- och välfärdsstudier.
- Gustafsson, Christina (1977). *Classroom Interaction. A Study of Pedagogical Roles in the Teaching Process*. Lund: LiberLäromedel/Gleerup.
- Hallinan, Maureen T. (ed.) (2006). *Handbook of the Sociology of Education*. New York: Springer.
- Hammarström, Margareta (1996). *Varför inte högskola? En longitudinell studie av olika faktors betydelse för studiebegåvade ungdomars utbildningskarriär*. Göteborg Studies in Educational Sciences, No. 107.
- Hattie, John & Timperley, Helen (2007). The power of feedback. *Review of Educational Research*, Vol. 77(1) pp. 81–112.
- Hedré, Rolf (1990). *Logoprogrammering på mellanstadiet: en studie av fördelar och nackdelar med användning av Logo i matematikundervisningen under årskurserna 5 och 6 i grundskolan*. Linköpings universitet.
- Hellsten, Jan-Olof (2000). *Skolan som barnarbete och utvecklingsprojekt: en studie av hur grundskoleelevers arbetsmiljö skapas-förändras-förblir som den är*. Uppsala studies in education, 86. Uppsala universitet, Pedagogiska institutionen.
- Hellsten, Jan-Olof (1997). *Kampen mot kaos: studie av hur en grundskola utformar elevernas arbetsmiljö*. Uppsala universitet, Pedagogiska institutionen.

- Henriksson, Carina (2004). *Living Away from Blessings. School Failure as Lived Experience*. Växjö University Press.
- Higgins, Chris (2007). The reflections on a line from Dewey. I L. Bresler (ed.), *International Handbook of Research in Arts Education* (pp. 389–394). Dordrecht: Springer.
- Hofvendahl, Johan (2006). *Riskabla samtal – en analys av potentiella faror i skolans kvarts- och utvecklingssamtal*. Linköpings universitet, Institutionen för Språk och Kultur & Arbetslivsinstitutet.
- Holmberg, Lena (2002). *Elever i svårigheter. Delrapport 1. Elever i en åldersblandad skolform. 10/2002*. Malmö högskola, Lärarutbildningen.
- Hurtig, Maria (2007). *Jag vågar visa att jag kan. Om meningsskapande med digitala portföljer*. Luleå tekniska universitet, Institutionen för pedagogik och lärande.
- Husén, Torsten & Härnqvist, Kjell (2000). Begåvningsreserven. *En återblick på ett halvsekels forskning och debatt*. Årsböcker i svensk undervisningshistoria, nr. 193. Uppsala: Föreningen för svensk undervisningshistoria.
- Hvistendahl, Rita & Roe, Astrid (2004). The Literacy Achievement of Norwegian Minority Students. *Scandinavian Journal of Educational Research*, Vol. 48(3), pp. 307–324.
- Hägglund, Solveig (1990). *Skolmognad och skolstartsproblem i svensk grundskoleforskning*. Stockholm: Skolöverstyrelsen.
- Härnqvist, Kjell (1997). Educational Research in Sweden. Infrastructure and Orientation. I K.-E. Rosengren & B. Öhngren (Eds). *An Evaluation of Swedish Research in Education*. Uppsala: Swedish Science Press.
- Hörnqvist, Maj-Lis (1999). *Upplevd kompetens: en fenomenologisk studie av ungdomars upplevelser av sin egen kompetens i skolarbetet*. Luleå tekniska universitet, 1999:22.
- Ingestad, Gunilla (2006). *Dokumenterat utanförskap – Om skolbarn som inte når målen*. Lunds universitet, Sociologiska institutionen.
- Isaksson, Joakim (2009). *Spänningen mellan normalitet och avvikelse. Om skolans insatser för elever i behov av särskilt stöd*. Umeå universitet, Institutionen för socialt arbete.
- Jablonka, Eva (2007). *Uppkomst av skillnader i prestationsförmåga i matematik – en empirisk studie av klassrumspraktik*. Luleå tekniska högskola, Institutionen för matematik. Forskningsprojekt finansierat av Vetenskapsrådet, Utbildningsvetenskapliga kommittén, nr 2007-3285.
- Jackson, Philip W. (ed.) (1992). *Handbook of Research on Curriculum. A Project of the American Educational Research Association*. New York: Macmillan Pub. Co.
- Jackson, Philip W. (1968). *Life in classrooms*. New York: Holt, Rinehart & Winston.
- Jakobsson, Anders, Egelund, Niels, Karlsson, Karl-Göran, Davidsson, Eva & Sörensen, Helene (2008). *Storskaliga studier och elevers prestationer i naturvetenskap i ett longitudinellt perspektiv*. Malmö högskola, Lärarutbildningen. Forskningsprojekt finansierat av Vetenskapsrådet nr 2008-4717.
- Johnsson Harrie, Anna (2009). *Staten och läromedlen: en studie av den svenska statliga förhandsgranskningen av läromedel 1938–1991*. Linköpings universitet.

- Jonsson, Jan O. & Mood, Carina (2008). Choice by Contrast in Swedish Schools: How Peers' Achievement Affects Educational Choice. *Social Forces*, Vol. 87(21), pp. 741–765.
- Karlsson, Jessika (2008). *A Novel Approach to Teaching Emotional Expression in Music Performance*. Uppsala universitet, Institutionen för psykologi.
- Kellaghan, Thomas, Stufflebeam, Daniel L. & Wingate, Lori A. (eds) (2003). *International Handbook of Educational Evaluation*. Dordrecht: Kluwer Academic Publishers.
- Kivirauma, Joel & Ruoho, Kari (2007). Excellence through Special Education? Lessons from the Finnish School Reform. *International Review of Education*, Vol. 53(3), pp. 283–302.
- Clapp Lekholm, Alli (2008). *Grades and Grade Assignment, Effects of Student and School Characteristics*. Göteborg Studies in Educational Sciences, No. 269.
- Clapp Lekholm, Alli & Cliffordson, Christina (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, Vol. 15(1), pp. 1–23.
- Knorr Cetina, Karin (2001). Objectual Practice. I T.R. Schatzki, K. Knorr Cetina & E. von Savigny (Eds). *The Practice Turn in Contemporary Theory*. London: Routledge.
- Korp, Helena (2006). *Lika chanser i gymnasiet? en studie om betyg, nationella prov och social reproduktion*. Malmö studies in educational sciences. Lund: Lunds universitet.
- Korp, Helena (2003). *Kunskapsbedömning – hur, vad och varför*. Stockholm: Myndigheten för skolutveckling.
- Koselleck, Reinhart (2004). *Erfarenhet, tid och historia: om historiska tiders semantik*. Göteborg: Daidalos.
- Krantz, Joakim (2009). *Styrning och mening – anspråk på professionellt handlande i lärarutbildning och skola*. Växjö universitet, Pedagogik, No. 181.
- Källér, Kathrine (1990). *Fostran till andrarang: en studie av dominansprocessen vid skolstart och via vägar genom utbildningssystemet ur ett kvinnovetenskapligt perspektiv*. Stockholm: Almqvist & Wiksell.
- Kärkkäinen, Riitta, Rätty, Hannu & Kasanen, Kati (2008). Children's notions of the malleability of their academic competencies. *Social Psychology of Education*, Vol. 11(4), pp. 445–458.
- Lainio, Jarmo, Sundgren, Eva, Tryggvason, Marja-Terttu, Ylikiiskilä, Antti & Kangassalo, Raija (2006). *Utbildningsval, sociala karriärer, språkliga kompetenser och identiteter*. Mälardalens högskola, Institutionen för humaniora. Forskningsprojekt finansierat av Vetenskapsrådet nr 2006-2418.
- Landahl, Joakim (2009). *Den läsande läraren: pedagogiska skrifter som bildnings- och moderniseringsprojekt 1898–1984*. Stockholm: Stiftelsen SAF i samarbete med Lärarförbundet.
- Landahl, Joakim (2006). *Auktoritet och ansvar: lärares fostrans- och omsorgsarbete i historisk belysning*. Stockholms universitet.
- Liberg, Caroline (2007). *Testet under lupp – konstruktion av läsförståelse i internationella läsuundersökningar*. Forskningsprojekt finansierat av Vetenskapsrådet nr 2007-4562.
- Lindahl, Erica (2008). *Empirical Studies of Public Policies within the Primary School and the Sickness Insurance*. Uppsala University, Department of Economics.

- Lindberg, Leif (2002). Is "Pedagogik" as an Academic Discipline in Sweden just a Phenomenon for the Twentieth Century? The Effects of Recent Educational Reform. *European Educational Research Journal*, Vol. 1(1), pp. 65–81.
- Lindberg, Viveca (2009). Kunskapsuppfattningar och inter/nationella redskap för bedömning. I I. Carlgren, E. Forsberg & V. Lindberg, *Perspektiv på den svenska skolans kunskapsdiskussion*. (pp. 78–119.) Stockholms universitetsförlag & Centrum för studier av skolans kunskapsinnehåll.
- Lindberg, Viveca (2005). Svensk forskning om bedömning och betyg 1995–2005. *Studies in Educational Policy and Educational Philosophy. E-tidskrift* 2005:01.
- Lindblad, Sverker (2008). *Bildningsgångar och skolresultat i det senmoderna samhället. Forskningsprojekt finansierat av Vetenskapsrådet*, Utbildningsvetenskapliga kommittén, nr 2008-4999.
- Lindblad, Sverker, Kyndel, Dag & Larson, Lena (2004). *Internationella arenor för utbildningsvetenskaper: om forskarorganisationer, vetenskapliga tidskrifter och överstatliga organisationer*. Stockholm: Vetenskapsrådet.
- Lindell, Ebbe (1990). *Läxor: hemarbetets utformning och effekter*. Stockholm: Skolöverstyrelsen.
- Lindström, Lars & Lindberg, Viveca (red.) (2005). *Pedagogisk bedömning: att dokumentera, bedöma och utveckla kunskap*. Stockholm: HLS förlag.
- Lindmark, Daniel (2004). *Reading, Writing, and Schooling. Swedish Practices of Education and Literacy, 1650–1880*. Umeå universitet, Institutionen för litteraturvetenskap och nordiska språk.
- Linnakylä, Pirjo (2004). Finnish Education – High Quality and Equity. *Education Review*, Vol. 17(2), pp. 35–41.
- Linné, Agneta (2004). Lektionen som pedagogisk text. I T. Englund (red.) *Skillnad och konsekvens: mötet lärare–studerande och undervisning som meningserbjudande*. (pp. 37–64.) Lund: Studentlitteratur.
- Lithner, Johan, Bergqvist, Ewa, Boesen, Jesper, Nyström, Peter & Palm, Torulf (2007). *Nationella matematikprov som katalysator för implementering av utbildningsreformer*. Umeå universitet, Institutionen för matematik och matematisk statistik.
- Ljung, Bengt-Olov (2000). *Standardproven – 53 år i skolans tjänst*. Lärarhögskolan i Stockholm, Institutionen för undervisningsprocesser, kommunikation och lärande. PRIM-gruppen.
- Lund, Stefan (2006). *Marknad och medborgare: elevers valhandlingar i gymnasieutbildningens integrations- och differentieringsprocesser*. Växjö University Press, No. 82.
- Lundahl, Christian (2009). *Varför nationella prov? Framväxt, dilemman, möjligheter*. Lund: Studentlitteratur.
- Lundahl, Christian (2007). *Utbildning och intelligens. Relationen mellan psykometriska undersökningar och svenskt utbildningsväsende 1880–1960*. Forskningsassistenttjänst finansierad av Vetenskapsrådet, Utbildningsvetenskapliga kommittén 2007-4048.
- Lundahl, Christian (2006). *Viljan att veta vad andra vet. Kunskapsbedömning i tidigmodern, modern och senmodern skola*. Uppsala universitet, Pedagogiska institutionen.

- Lundahl, Lisbeth (2009). *Osäkra övergångar. Unga utan fullständig gymnasieutbildning: vägarna och åtgärderna i longitudinellt perspektiv*. Forskningsprojekt finansierat av Vetenskapsrådet, Utbildningsvetenskapliga kommittén, No. 2009-5964.
- Lundahl, Lisbeth (2002). Sweden: Decentralisation, Deregulation, Quasi-Markets – And Then What? *Journal of Education Policy*, Vol. 17(4), pp. 687–697.
- Lundgren, Ulf P. (1999). Skolan och tusenårsskiftet. www.skeptron.uu.se/broadly/sec/p-lundgren-990930.
- Lundgren, Ulf P. (1979). *Att organisera omvärlden. En introduktion till läroplansteori*. Stockholm: LiberFörlag.
- Lundgren, Ulf P. (1972). *Frame Factors and the Teaching Process. A Contribution to Curriculum Theory and Theory on Teaching*. Göteborg University.
- Lundström, Stefan (2007). *Textens väg. Om förutsättningar för texturval i gymnasieskolans svenskundervisning*. Umeå universitet, Institutionen för litteraturvetenskap och nordiska språk.
- Madaus, George F. & Kellaghan, Thomas (1992). Curriculum Evaluation & Assessment. I P. W. Jackson (Ed.) *Handbook of Research on Curriculum. A Project of the American Educational Research Association*. New York: Macmillan Pub. Co.
- Malmberg, Erik, Andersson, Eva, Östh, John, Niedomysl, Thomas, & Bergsten, Zara (2009). *Den svenska skolans nya geografi: rörlighet, attityder och resultat*. Stockholms universitet, Kultur-geografiska institutionen. Forskningsprojekt finansierat av Vetenskapsrådet nr 2009-6299.
- Marklund, Sixten (1982). *Skolsverige 1950–1975. D. 2, Försöksverksamheten*. Stockholm: Liber Utbildningsförlag.
- Marton, Ference (2007). *Effekter av likheter och skillnader mellan exempel valda för att främja lärande*. Göteborgs universitet, Institutionen för pedagogik och didaktik. Forskningsprojekt finansierat av Vetenskapsrådet, Utbildningsvetenskapliga kommittén, nr 2007-3198.
- Mauritzon, Ulla & Säljö, Roger (2001). Adult Questions and Children's Responses: Coordination of Perspectives in Studies of Children's Theories of other Minds. *Scandinavian Journal of Educational Research*, Vol. 45(3), pp. 213–231.
- Mellander, Erik (2009). *Kan mera naturvetenskap i lärarutbildningen påverka dagisbarnens attityder, skolresultat och studieinriktning?* Forskningsprojekt finansierat av Vetenskapsrådet, Utbildningsvetenskapliga kommittén, No. 2009-6045.
- Messick, Samuel (1989). Validity. I R. L. Linn (Ed.), *Educational measurement*, Vol. 3 (pp. 13–103). New York: American Council on Education.
- Murphy, Regina & Espeland, Magne (2007). Prelude. Making Connections in Assessment and Evaluation in Arts Education. I L. Bresler (ed.). *International Handbook of Research in Arts Education*. Dordrecht: Springer.
- Myrberg, Eva & Rosén, Monica (2006). Reading Achievement and Social Selection in Independent Schools in Sweden. Results from IEA PIRLS 2001. *Scandinavian Journal of Educational Research*, Vol. 50(2), pp. 185–205.
- Newton, Paul, E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, Vol. 14(2), pp. 149–170.

- Nilsson, Anna (2005). *Indirect Effects of Unemployment and Low Earnings. Crime and Children's School Performance*. Stockholm University, Department of Economics.
- Norlund, Anita (2009). *Kritisk sakprosaläsning i gymnasieskolan. Didaktiska perspektiv på läroböcker, lärare och nationella prov*. Göteborg Studies in Educational Sciences, Göteborgs universitet, No. 273.
- Nyberg, Eva (2009). *Om livets kontinuitet. Undervisning och lärande om växters och djurs livscyklar – en fallstudie i årskurs 5*. Göteborgs universitet, Institutionen för pedagogik och didaktik.
- Nyström, Peter (2004a). Reliability of Educational Assessments – The Case of Classification Accuracy. *Scandinavian Journal of Educational Research*, 48(4), pp. 427–440.
- Nyström, Peter (2004b). *Rätt mätt på prov. Om validering av bedömningar i skolan*. Umeå universitet, Pedagogiska institutionen, No. 71.
- Nyström, Peter (2003). Lika barn leka bäst? En gymnasielärardiskurs om nivågruppering i matematik. *Pedagogisk Forskning i Sverige, Vol. 8(4)*, pp. 225–245.
- Nytell, Hans (2006). *Från kvalitetsidé till kvalitetsregim. Om statlig styrning av skolan*. Uppsala universitet, Pedagogiska institutionen.
- Näsström, Gunilla (2008). *Measurement of Alignment between Standards and Assessment*. Umeå universitet, Institutionen för beteendevetenskapliga mätningar.
- Ochs, Kimberly & Phillips, David (2004). Processes of Educational Borrowing in Historical Context. I D. Phillips & K. Ochs (eds). *Educational Policy Borrowing. Historical Perspectives*. Oxford Studies in Comparative Education.
- Olsson, Erik, Osman, Ali, & Larsson, Staffan (2006). *Transnationella utbildningskarriärer – Studieval, möjlighetshorisont och sociala nätverk för unga med utländsk anknytning*. Linköpings universitet, Tema etnicitet. Forskningsprojekt finansierat av Vetenskapsrådet nr 2006-2443.
- Osborn, Marilyn (2006). Changing the Context of Teachers' Work and Professional Development. A European Perspective. *International Journal of Educational Research, Vol. 45(4-5)*, pp. 242–253.
- Palmér, Anne (2008). *Samspel och solostämmor. Om muntlig kommunikation i gymnasieskolan*. Uppsala universitet, Institutionen för nordiska språk, No. 74.
- Parmenius Swärd, Suzanne (2008). *Skrivande som handling och möte – gymnasieelever om skrivuppgifter, tidsvillkor och bedömning i svenskämnet*. Malmö högskola, Lärarutbildningen.
- Paulsen, Erik & Bru, Edvin (2008). Social Passivity and Grades Achieved among Adolescents in Junior High School. *School Psychology International, Vol. 29(2)*, pp. 248–26.
- Pedagogisk forskning i Sverige* (2002). Tema: Individuella förutsättningar för utbildning, Vol. 7(3).
- Perzon, Håkan (1998). *Sökandemönster till högskolan: en fallstudie vid Luleå tekniska universitet*. Luleå tekniska universitet: Industriell ekonomi och samhällsvetenskap/Industriell marknadsföring, 1998:45
- Pettersson, Astrid (1990). *Att utvecklas i matematik: en studie av elever med olika prestationsutveckling*. Stockholm: Almqvist & Wiksell International.

- Pettersson, Astrid, Gustafsson, Jan-Eric, Lindberg, Viveca, Lindström, Lars, Nyström, Peter & Svingby, Gunilla (2006). *Nationella forskarskolan i pedagogisk bedömning*. Forskarskola finansierad av Vetenskapsrådet nr 2006-2534.
- Pettersson, Daniel (2008). *Internationell kunskapsbedömning som inslag i nationell styrning av skolan*. Uppsala universitet, Pedagogiska institutionen.
- Pettersson, Daniel & Leo, Ulf. (2005). Läxor – en oreglerad bedömningspraktik. *Studies in Educational Policy and Educational Philosophy. E-tidskrift* 2005:2. www.upi.artisan.se
- Qvarsebo, Jonas (2006). *Skolbarnets fostran: enhetsskolan, agan och politiken om barnet 1946–1962*. Linköping: Linköpings universitet.
- Ramstedt, Kristian (1996). *Elektriska flickor och mekaniska pojkar. Om gruppskillnader på prov – en metodutveckling och en studie av skillnader mellan flickor och pojkar på centrala prov i fysik*. Umeå universitet, Pedagogiska institutionen, No. 39.
- Ranagården, Lisbeth (2009). *Lärares lärande om elever – en sociologisk studie om yrkespraktik*. Göteborgs universitet, Sociologiska institutionen, No. 39.
- Rangvid, Beatice (2007). School Composition Effects in Denmark. Quantile Regression Evidence from PISA 2000. *Empirical Economics, Vol. 33*(2), pp. 359–388.
- Rautalin, Marjaana & Alasuutari, Pertti (2007). The Curse of Success. The Impact of the OECD's Programme for International Student Assessment on the Discourses of the Teaching Profession in Finland. *European Educational Research Journal, Vol. 6*(4), p. 3.
- Redelius, Karin, Fagrell, Birgitta & Larsson, Håkan (2009). Symbolic Capital in Physical Education and Health. To Be, To Do or To Know? That Is the Gendered Question. *Sport, Education and Society, Vol. 14*(2), pp. 245–260.
- Reg. Dir. 2009:102. Omprövning av betyg inom skolväsendet. Stockholm: Utbildningsdepartementet.
- Reg. Dir. 2009:08. Flexibel skolstart i grundskolan. Stockholm: Utbildningsdepartementet.
- Reg. Prop. 1990/91:115. Om vissa skolfrågor m.m. Stockholm: Utbildningsdepartementet.
- Resultatdialog 2006: forskning inom utbildningsvetenskap. (2006). Stockholm: Vetenskapsrådet.
- Resultatdialog 2007: forskning inom utbildningsvetenskap. (2007). Stockholm: Vetenskapsrådet.
- Resultatdialog 2008: forskning inom utbildningsvetenskap. (2008). Stockholm: Vetenskapsrådet.
- Resultatdialog 2009: aktuell forskning om lärande. (2009). Stockholm: Vetenskapsrådet
- Richardson, Virginia (ed.) (2001). *Handbook of Research on Teaching*. 4. ed. Washington, D.C.: American Educational Research Association.
- Roos, Bertil (2005). *ICT and formative assessment in the learning society*. Umeå: Umeå universitet.
- Roos, Bertil & Hamilton, David (2005). Formative Assessment: A Cybernetic Viewpoint. *Assessment in Education Principles Policy and Practice, Vol. 12*(1), pp. 7–20.
- Rosén, Monica (1998). *Gender Differences in Patterns of Knowledge*. Göteborg Studies in Educational Sciences, No. 124.
- Rosengren, Karl-Erik & Öhngren, Bo (Eds) (1997). *An Evaluation of Swedish Research in Education*. Uppsala: Swedish Science Press.

- Runesson, Ulla, Marton, Ference, Häggström, Johan, Vikström, Anna & Nilsson, Pernilla (2009). *Lärares gemensamma kunskapsproduktion (LGK). Hur lärares yrkesutövning påverkas av att de deltar i gemensamt bygge av yrkets vetenskapliga grund*. Göteborgs universitet, Institutionen för pedagogik och didaktik.
- Räihä, Helge (2008). *Lärares dilemman*. Studier från Örebro i svenska språket 4. Örebro universitet, Humanistiska institutionen.
- Räty, Hannu, Kasanen, Kati & Kärkkäinen, Riitta (2006). School Subjects as Social Categorisations. *Social Psychology of Education, Vol. 9(1)*, pp. 5–25.
- Röding, Karin (2005). *University Admission based on Tests and Interviews. Implementation and Assessment*. Stockholm: Karolinska institutet, 05:123.
- Rönnerberg, Linda (2007a). *Tid för reformering. Försöksverksamheten med slopad timplan i grundskolan*. Umeå universitet, Statsvetenskapliga institutionen.
- Rönnerberg, Linda (2007b). *Den statliga skolinspektionen: En centraliserad resultatstyrning möter en decentraliserad skola*. Anställning som forskarassistent finansierad av Vetenskapsrådet, Utbildningsvetenskapliga kommittén, nr 2007-4031.
- Sabbagh, Clara (2009). Ethics and Teaching. I L. J. Saha & A. G. Dworkin (eds). *International Handbook of Research on Teachers and Teaching*. (pp. 683–693.) New York: Springer.
- Sadler, D. Royce (1998). Formative Assessment. Revisiting the Territory. *Assessment in Education, Vol. 5(1)*, pp. 77–84.
- Sahlström, Fritjof (2008). *Från lärare till elever, från undervisning till lärande: utvecklingslinjer i svensk, nordisk och internationell klassrumsforskning*. Stockholm: Vetenskapsrådet.
- Samuelsson, Marcus (2008). *Störande elever korrigerande lärare: om regler, förväntningar och lärares åtgärder mot störande flickor och pojkar i klassrummet*. Linköping: Linköpings universitet.
- Sandell, Anna (2007). *Utbildningssegregation och självsortering: om gymnasieval, genus och lokala praktiker*. Lund: Lunds universitet.
- Schleicherf, Andreas (2006). Where Immigrant Students Succeed. A Comparative Review of Performance and Engagement in PISA 2003. *Intercultural Education, Vol. 17(5)*, pp. 507–516.
- Segerholm Christina, Forsberg Eva, Lindgren Joakim, Nilsson Ingrid & Rönnerberg Linda (2009). *Inspektion som styrning. Skolinspektion och utbildningsstyrning i Sverige, England och Skottland*. Mittuniversitetet, Institutionen för utbildningsvetenskap. Forskningsprojekt finansierat av Vetenskapsrådet nr 2009-5770.
- Selander, Staffan (2003). Pedagogiska texter: socialt minne och multimodala artefakter för kommunikation. I E. Forsberg (red.) *Skolan och tusenårsskiftet: en vänbok till Ulf P. Lundgren*. STEP, Uppsala University, Department of Teacher Education.
- Selghed, Bengt (2004). Ännu icke godkänt. *Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Malmö högskola, Lärarutbildningen.
- SFS 1842:19. Kungl. Maj:ts Nådige Stadga angående folkundervisning i Riket.
- Shepard, Lorrie (2001). The role of assessment in a learning culture. I V. Richardson (ed.), *Handbook on research on teaching*. (pp. 1066–1101.) 4. ed. Washington, D.C.: American Educational Research Association.

- Simola, Hannu (2005). The Finnish Miracle of PISA. Historical and Sociological Remarks on Teaching and Teacher Education. *Comparative Education*, Vol. 41(4), pp. 455–470.
- Simola, Hannu (2002). From Exclusion to Self-selection. Examination of Behaviour in Finnish Primary and Comprehensive Schooling from the 1860s to the 1990s. *History of Education*, Vol. 31(3), pp. 207–226.
- Sjöberg, Gunnar, Nyroos, Mikaela & Nyström, Eva (2008). *Vad gör testandet med eleverna?* Umeå universitet, Institutionen för matematik, teknik och naturvetenskap. Forskningsprojekt finansierat av Vetenskapsrådet nr 2008-4646.
- Sjöstedt, Carl Erik & Thorén, Birger (red.) (1963). *Studentexamen 100 år*. Stockholm: Natur & Kultur.
- Skolverket (2009). *Vad påverkar resultaten i svensk grundskola: kunskapsöversikt om betydelsen av olika faktorer*. Stockholm: Skolverket.
- SOU 2009:94. *Att nå ut och att nå ända fram – hur policyinriktad utvärdering och forskningsresultat inom utbildningsområdet ska tillgodoses*. Stockholm: Fritzes.
- SOU 1974:53. *Utredningen om skolans inre arbete. Skolans arbetsmiljö: betänkande*. Stockholm: Allmänna förlaget.
- Star, Susan & Griesemer, James R. (1989). Institutional Ecology, "Translations" and Boundary Objects. Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, Vol. 19, pp. 387–420.
- Stenhag, Staffan (2007). *Vad säger matematikbetyget? En kvantitativ studie av 2 079 elevers betyg i årskurs nio*. Uppsala University, Department of Mathematics, Report 2007:20.
- Sund, Krister (2007). *Teachers, Family and Friends. Essays in Economics of Education*. Stockholms universitet, Ekonomiska institutionen.
- Sundlöf, Petra (2008). *Segregation och karriärposition. En studie av bostadsomgivningens betydelse för utbildning, sysselsättning och inkomst bland yngre i stockholmsregionen*. Uppsala universitet, Kulturgeografiska institutionen.
- Svensson, Allan (2004). Gymnasiebetyg eller högskoleprov urvalsinstrument? Fallet civilingenjörsutbildningarna. *Pedagogisk Forskning i Sverige*, Vol. 9(1), pp. 15–36.
- Svensson, Allan (2001). Består den sociala snedrekryteringen? Elevernas val av gymnasieprogram hösten 1998. *Pedagogisk Forskning i Sverige*, Vol. 6(3), pp. 161–172.
- Söderström, Åsa (2006). *"Att göra sina uppgifter, vara tyst och lämna in i tid". Om elevansvar i det högmoderna samhället*. Karlstads universitet, Estetisk-filosofiska fakulteten.
- Taflin, Eva (2007). *Matematikproblem i skolan – För att skapa tillfällen till lärande*. Umeå universitet, Institutionen för matematik och statistik.
- Thiessen, Dennis & Cook-Sather, Alison (2007). *International Handbook of Student Experience in Elementary and Secondary School*. Dordrecht: Springer.
- Tholin, Jörgen (2006). *Att kunna klara sig i okänd natur. En studie av betyg och betygskriterier – historiska betingelser och implementering av ett nytt system*. Högskolan Borås: Institutionen för pedagogik och didaktik.

- Tholin, Jörgen (2003). *En roligare dans? Svenska skolors första tolkning av innebörden i lokala betygsriterier i tre ämnen för skolår åtta*. Högskolan Borås: Institutionen för pedagogik och didaktik.
- Tirri, Kirsi & Nokelainen, Petri (2006). Gifted Students and the Future. *KEDI Journal of Educational Policy*, Vol. 3(2), pp. 55–66.
- Tornberg, Ulrika (red.) (2006). *Mångkulturella aspekter på språkundervisningens kommunikativa praktiker: en konferensrapport*. Örebro universitet, Pedagogiska institutionen.
- Torrance, Harry (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*. Special Issue: *Assessment in post-secondary education and training*, Vol. 14(3), pp. 281–294.
- Torrance, Harry & Pryor, John (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, Vol. 27(5), pp. 615–631.
- Torrance, Harry & Pryor, John (1998). *Investigating formative assessment: teaching, learning and assessment in the classroom*. Buckingham: Open University Press.
- Townsend, Tony (ed.) (2007). *International Handbook of School Effectiveness and Improvement*. Dordrecht: Springer.
- Travers, Robert M. W. (ed.) (1973). *Second Handbook of Research on Teaching. A Project of the American Educational Research Association*. 2. pr. Chicago: Rand McNally.
- Tsagalidis, Helena (2008). *Därför fick jag bara Godkänt... Bedömning i karaktärsämnen på HR-programmet*. Doktorsavhandlingar från Pedagogiska institutionen nr 149. Stockholms universitet.
- Tyack, David & Cuban, Larry (1995). *Tinkering Toward Utopia. Century of Public School Reform*. Cambridge, Mass.: Harvard University Press.
- Törnebohm, Håkan (1971). *Reflections on Scientific Research*. Göteborgs universitet, Institutionen för vetenskapsteori.
- Vallberg Roth, Ann-Christine, Tallberg Broman, Ingegerd & Tvingstedt, Anna-Lena (2007). *Reglerad barndom och lärarprofession – möten mellan förskola och mångfaldens medborgare genom ett filter av dokumentation* (planeringsbidrag nr 2007-654). Malmö högskola, Lärarutbildningen.
- Wallby, Karin, Carlsson, Synnöve & Nyström, Peter (2001). *Elevgrupperingar – en kunskapsöversikt med fokus på matematikundervisning*. Stockholm: Skolverket.
- Wedin, Ann-Sofie (2007). *Lärares arbete och kunskapsbildning. Utmaningar och inviter i den vardagliga praktiken*. Linköpings universitet, Institutionen för beteendevetenskap och lärande.
- Westling Allodi, Mara Vittoria, Fischbein, Siv, Helldin, Rolf & Rydelius, Per-Anders (2008). *Delaktighet, socialt klimat och lärande i skolan. En intervention på organisationsnivå i skolans lärandemiljöer för en utveckling av skolans förmåga till förändring och anpassning till elevernas behov*. Stockholms universitet, Specialpedagogiska institutionen.

- Wiberg, Marie & Nyström, Peter (2008). *Analys och modellering av svenska elevers prestationer i TIMSS och PISA i ett internationellt perspektiv*. Umeå universitet, Institutionen för beteendevetenskapliga mätningar. Forskningsprojekt finansierat av Vetenskapsrådet nr 2008-4027.
- Wiggins, Grant (1998). *Educative Assessment*. San Francisco: Jossey-Bass.
- Wikström, Christina (2005a). Grade Stability in a Criterion-referenced Grading System. The Swedish Example, *Assessment in Education*, Vol. 12(2), pp. 125–144.
- Wikström, Christina (2005b). *Criterion-referenced Measurement for Educational Evaluation and Election*. Umeå University, Department of Educational Measurement, No. 1.
- Wilson, Bruce & Corbet, Dick (2007). Students Perspective on Good Teaching: Implications for Adult Reform Behavior. I D. Thiessen & A. Cook-Sather. *International Handbook of Student Experience in Elementary and Secondary School*. (pp. 283–311.) Dordrecht: Springer.
- Winberg, Mikael (2007). *Interaktionen mellan den lärande och lärandesituationen: effekter på affektiva upplevelser och lärandeutfall*. Forskningsprojekt finansierat av Vetenskapsrådet nr 2007-3216.
- Wittrock, Merlin C. (ed.) (1986). *Handbook of Research on Teaching. A Project of the American Educational Research Association*. 3. ed. New York: Macmillan.
- Wolming, Simon (2000). *Validering av urval*. Umeå universitet, Pedagogiska institutionen, No. 54.
- Voogt, Joke & Knezek, Gerald (Eds) (2008). *International Handbook of Information. Technology in Primary and Secondary Education*. New York: Springer.
- Väljjärvi, Jouni (2003). The System and How Does It Work. Some Curricular and Pedagogical Characteristics of the Finnish Comprehensive School. *Education Journal*, Vol. 31(2), pp. 31–55.
- Yang Hansen, Kajsa (2007). *Causes and effects of school segregation in Swedish schools*. Anställning som forskarassistent finansierad av Vetenskapsrådet nr 2007-4054.
- Yang, Yang (2003). *Measuring Socioeconomic Status and its Effects at Individual and Collective Levels. A Cross-country Comparison*. Göteborg University, Department of Education, No. 193.
- Young, Michael (1961). *The Rise of the Meritocracy, 1870–2033: An Essay on Education and Equality*. Baltimore, MD: Penguin Books.
- Zhao, Shangwu (1993). *Chinese Science Education. A Comparative Study of Achievement in Secondary Schools Related to Student, Home and School Factors*. Stockholm University, Institute of International Education, No. 26.
- Zipernovszky, Hanna (2003). Are the Swedish Parish Examination Records a Unique Phenomenon? *Interchange*, Vol. 34(2), pp. 179–189.
- Åberg-Bengtsson, Lisbeth (1999). Dimensions of performance in the interpretation of diagrams, tables and maps. Some gender differences in the Swedish scholastic aptitude test. *Journal of Research in Science Teaching*, Vol. 36(5), pp. 565–582.
- Åberg-Bengtsson, Lisbeth & Erickson, Gudrun (2006). Dimensions of National Test Performance. A Two-level Approach. *Educational Research & Evaluation*, Vol. 12(5), pp. 469–488.
- Åström, Maria (2008). *Defining Integrated Science Education and Putting It to Test*. Linköping University, Studies in Science and Technology Education, No. 26.

- Öhrn, Elisabet (2002). *Könsmönster i förändring? En kunskapsöversikt om unga i skolan*. Stockholm: Skolverket.
- Österlind, Eva (red.) (2005). *Eget arbete – en kameleont i klassrummet: perspektiv på ett arbetssätt från förskola till gymnasium*. Lund: Studentlitteratur.
- Österlind, Eva (2001). *Elevers förhållningssätt till läxor*. Högskolan i Falun.
- Östlund-Stjärnegårdh, Eva (2002). *Godkänd i svenska? Bedömning och analys av gymnasieelevers texter*. Uppsala universitet, Institutionen för nordiska språk.

KÄLLOR

Svenska och Nordiska tidskrifter

Nordisk Pedagogik

Pedagogisk Forskning i Sverige

Scandinavian Journal of Educational Research

Utbildning & Demokrati. Tidskrift för didaktik och utbildningspolitik

Utbildningsorganisationernas/konferensernas hemsidor

AEA-Europe: Association for Educational Assessment. www.aea-europe.net

AERA: American Educational Research Association. www.aera.net

EARLI: European Association for Research on Learning and Instruction.
www.earli.org

EERA: European Educational Research Association. www.eera.ac.uk

IAEA: International Association for Educational Assessment. www.iaea.info

NFPF: Nordisk Förening för Pedagogisk Forskning. www.nfpf.net

Databaser (för sökning av artiklar i internationella vetenskapliga tidskrifter)

ERIC: Education Resource Information Center

csaweb107v.csa.com.ezproxy.its.uu.se/ids70/quick_search.php?SID=hg656afpri796sbo3uo20a7gb5

EBSCO: Education Research Complete

web.ebscohost.com.ezproxy.its.uu.se/ehost/search?vid=i&hid=113&sid=98b63e4a-32de-4d44-972d-55e190b6bdc6%40sessionmgr114

IBSS: International Bibliography of the Social Sciences

csaweb107v.csa.com.ezproxy.its.uu.se/ids70/quick_search.php?SID=e771afr17m50duk8llrsrouu2

Databaser (för sökning av svenska avhandlingar)

LIBRIS: www.libris.kb.se

Databas (för sökning av pågående forskningsprojekt)

Utbildningsvetenskapliga kommitténs (UVK:s) projektdatabas:

vrproj.vr.se/default.asp?funk=s

Rankingslista över tidskrifter

Top Education Journals by Impact Factor – ISI Web of Science (Social Sciences) 1981–2007.

http://sciencewatch.com/dr/sci/09/jan18-09_2/

BILAGA 1: MÖTET MELLAN KUNSKAPS- OBJEKT OCH ADMINISTRATIVA SÖK- SYSTEM

Svenska och nordiska artiklar som ingår i internationella vetenskapliga tidskrifter har införskaffats genom sökningar för perioden 1990–2008 i referensdatabaserna ERIC,¹⁶ EBSCO¹⁷ och IBSS.¹⁸ Det är förvisso så att databaserna sträcker sig en bra bit bakåt i tiden. Men det innebär inte med nödvändighet att de tidskrifter man är intresserad av ingår för hela den period undersökningen avser. Vi har undersökt detta närmare i relation till bedömning inom utbildningsområdet. För att fastställa att relevanta tidskrifter om utbildning tagits med utgick vi från Top Education Journals by Impact Factor – ISI Web of Science (Social Sciences). För att undvika korttidseffekter av tidskriftsinkluderingen använde vi den lista som är baserad på vikt under perioden 1981–2007. Väl etablerade och för våra syften relevanta tidskrifter täcks av databaserna. Det gäller både högt rankade tidskrifter inom utbildning och centrala tidskrifter för forskning om bedömning.¹⁹ Beträffande *Assessment in Education: Principles, Policy and Practice* ingår artiklar från 1994 och vi har därför sökt separat på tidskriftens hemsida i årgångarna 1990–1993.

Konstruktionen av en sökstrategi för databaserna kräver insikter om både kunskapsobjektet och de administrativa söksystemen. Det är i mötet mellan dessa som bedömning åtminstone till en del får sin metodologiska bestäm-

¹⁶ ERIC – Education Resource Information Center – är världens största referensdatabas inom utbildning och angränsande ämnesområden t.ex. psykologi. Databasen innehåller referenser till cirka 900 internationella pedagogiska tidskrifter samt böcker, avhandlingar, forskningsrapporter, konferensrapporter och andra dokument. Täcker material från 1966 och framåt.

¹⁷ EBSCO – Education Research Complete ger tillgång till fulltextartiklar och referat från engelskspråkiga tidskrifter som täcker alla utbildningsområden. Databasen innehåller abstract från mer än 2 100 tidskrifter och fulltext från mer än 1 200 tidskrifter.

¹⁸ IBSS – The International Bibliography of the Social Sciences (IBSS) är en webbaserad resurs för samhällsvetenskaplig och tvärvetenskaplig forskning. Sökningarna omfattar titlar och abstract.

¹⁹ Som exempel kan nämnas: *Review of Educational Research*, *Reading Research Quarterly*, *American Educational Research Journal*, *Sociology of Education*, *Review of Research in Education* och *Harvard Educational Review* bland de mer generellt inriktade tidskrifterna. Några exempel från forskning om bedömning är: *Educational Evaluation and Policy Analysis*, *Evaluation and Program Planning*, *Applied psychological measurement*, *Applied Measurement in Education*, *Evaluation Review*, *Educational and psychological measurement*, *Educational measurement: issues and practice* och *Journal of educational measurement*. Vid sidan av dessa ingår som framgår av noterna 1–3 också en mycket stor mängd övriga tidskrifter.

ning. Med utgångspunkt i bedömningsforskningen kan konstateras att tre termer är av särskild relevans för våra sökningar: measurement, assessment och evaluation. De kan på en övergripande nivå förbindas med psykometriska studier, klassrumsbedömningar och utvärdering av skolan som organisation och system. Databassökningars tillförlitlighet kan emellertid förstärkas genom informationen som finns i databasernas respektive thesaurus (ordlistor). Med measurement som exempel kan vi konstatera att i ERIC ingår forskning med fokus på large scale assessment or testing, achievement rating, cognitive measurement, merit rating, predictive measurement och scoring. I EBSCO ingår studier som tar fasta på mätning som en generell process och även termen measuring instruments är medtagen, dock inte testing. I IBSS som har en bredare samhällsvetenskaplig bas ingår ett större antal snävare termer som sträcker sig från anthropometrics över craniometry och personality assessment till welfare measurement för att bara nämna några.

Om vi går över till assessment finns inte ordet i någon av databasernas ordlistor, som i stället hänvisar till termen evaluation. I ERIC ingår i evaluation forskning som handlar om "appraising or judging persons, organizations, or things in relation to stated objectives, standards, or criteria". Mer specifika termer som medföljer är t.ex. alternative assessment, curriculum evaluation, educational assessment, psychological evaluation, writing evaluation, self and peer evaluation, student evaluation (utvärdering av elevers prestationer), formative evaluation, portfolio assessment, summative evaluation, informal evaluation och program evaluation. Här finns också information om närliggande termer som kan vara av intresse, t.ex. measurment, testing, achievement, credentials, differences och standards. EBSCO ger ett liknande utfall. Evaluation avser här "general works on the process of appraisal, assessment, and study of an object, person, or specific condition to determine its qualities and value." En sökning på evaluation omfattar också snävare termer, t.ex. educational evaluation, informal evaluation, needs assessment, outcome assessment, personality assessment, disability evaluation, self-evaluation, teacher effectiveness. Relaterade termer som inte ingår är bland annat examinationer, standards och rating of students. I IBSS, slutligen, ges ingen specificering av evaluation förutom att termen inkluderar assessment liksom snävare termer som self-evaluation, self-assessment och program evaluation. Sammanfattningsvis valde vi initialt att använda evaluation och measurement som nyckeltermer.

På liknande sätt gick vi också igenom ett antal andra för forskning om bedömning centrala begrepp som achievement, performance, ability, psychometrics, tests, examination, outcomes, marking, gradings, attainment, ability grouping, tracking, accountability, effective schools, school success, school failure, drop-outs etcetera. Den breda ingången svarar mot avsikten att fånga in

traditioner som både studerar bedömning som fenomen och använder elevers prestationer och resultat som grund för analyser av skol- och systemeffekter. Utifrån en genomgång av sökorden i databasernas thesaurus skapade vi hierarkiska listor. Med utgångspunkt i dessa har vi upprättat vad som närmast kan karakteriseras som sökträd. Vi har identifierat roten (den mest basala termen) och får därmed med såväl stammen och grenarna som lövverket, det vill säga alltmer specifika termer. Med detta tillvägagångssätt har vi undvikit både överlappningar och felaktiga slutsatser, t.ex. att forskning inom ett område saknas när det i stället är så att den inordnats under en annan sökterm. Strategin leder också till att vi för att fånga in likartad forskning i de olika databaserna måste använda delvis olika sökbilder och söktermer.

Vi ska här ta upp ytterligare ett exempel på hur de administrativa systemen kan sätta hinder i vägen för att få fram relevant material. Tidskriften *Scandinavian Journal of Educational Research* ingår i databasen ERIC och därför valde vi att söka den vägen. Utfallet framstod emellertid som begränsat och vi sökte därför också direkt i tidskriftens egen databas och utfallet blev i stort sett detsamma med något undantag. Arbetet fortskred, men i slutskedet bestämde vi oss likväl för att göra ytterligare en sökning. Vi gjorde tre sökningar på *assess**, *eval** och *test** i titel och abstract och inte som tidigare på ämnes- eller nyckelord. Utfallet blev 114, 48 respektive 60 träffar. Av dessa var det 54 som var relevanta men som inte fanns med i vårt material. De fördelar sig på de nordiska länderna i enlighet med tabell 2.2.

TABELL X:1 Antal träffar vid repeterad sökning

	Assess*	Eval*	Test*	Σ
SVERIGE	2	4	11	17
DANMARK	0	0	0	0
FINLAND	9	4	3	16
ISLAND	1	1	0	2
NORGE	10	5	4	19
Σ	22	14	18	54

Fördelningen över femårsperioder var: 1990–1994 (15), 1995–1999 (10), 2000–2004 (12) och 2005–2009 (17). Jämfört med den första bilden är tyngdpunkten alltså jämt mot den senaste perioden. Men den stora skillnaden ligger i tillkomsten av artiklar under de tidigare perioderna. Resultatet skilde sig

rejält från det tidigare, i stället för 24 artiklar fick vi nu 76 relevanta artiklar. Sammantaget är det framför allt psykometriska studier från Sverige, Finland och Norge som tillkommer. Det som är anmärkningsvärt är att de flesta svenska artiklarna faller ut på sökordet test, medan norska och finska artiklar faller ut på assessment.

De svenska artiklarna bidrar med ytterligare information om de internationella kunskapsmätningarna och om högskoleprovet, framför allt är flera av dem publicerade under tidigt 1990-tal, vilket hjälper till att förstå den forskning som i dag produceras inom detta område. Sett i förhållande till svenska lärosäten är det artiklar från Umeå, den nuvarande institutionen för beteendevetenskapliga mätningar, tidigare en del av den pedagogiska institutionen, som bidrar med fler artiklar. Vidare ser vi ett samarbete mellan Umeå och Göteborg. Ett exempel på en artikel som representerar en annan typ av forskningsintresse handlar om lärares bedömningar i klassrummet och fokuserar hur elever förstår lärares frågor (Mauritzon & Säljö 2001). Liksom resultaten i vår primära sökning representerar dessa varierande kunskapsintressen. Även bland de finska artiklarna finns en ökning av antalet artiklar relaterade till IEA-studier. Här kan man se att Finland redan tidigt klarade sig bra, vilket verkar bidra till att artiklar produceras. I artiklarna ger forskarna, liksom under 00-talet förklaringar till framgångarna. Även didaktiska studier finns representerade i det nytillkomna materialet. De norska artiklarna tar också avstamp i IEA-studierna. I några fall relateras resultaten till hur undervisningen går till?

Redovisningen som gäller Norden specifikt är justerad i relation till de artiklar som tillkom. I de fall redovisningen görs i förhållande till övriga internationella artiklar är analyserna baserade på det tidigare utfallet. Det gäller redovisningarna i tabellerna 6.1 och 6.3. Men de huvudtrender som redovisas där påverkas inte på något avgörande sätt av detta. Sammanfattningsvis menar vi att det i princip alltid finns skäl att vara återhållsam med de anspråk man reser i samband med kunskaps- och forskningsöversikter. Det gäller såväl forskningens omfattning och fokus som dess variation.

För att avsluta redovisningen av sökningarna kan vi notera att vi också, fast först i ett andra steg och manuellt, gjort avgränsningar som refererar till geografi och skolform. Det förra betyder fokus på forskning om bedömning i svenska eller övriga nordiska utbildningssystem alternativt svensk eller nordisk forskning om bedömning. Det senare innebär att vi tagit med artiklar som handlar om grund- och/eller gymnasieskolan samt övergångar mellan skolformer Rekryteringsstudier som handlar om antagningsprov och urval till högre studier ingår t.ex. Detsamma gäller forskning om elevkarriärer och bildningsgångar för vilka elevers prestationer, provresultat och betyg har en avgörande betydelse. Materialet har utifrån dessa utgångs-

punkter reducerats efter genomläsningar av abstract. Därtill har dubletter och uppenbart icke relevanta artiklar rensats ut. Gränsfall har identifierats som vi diskuterat gemensamt och därefter tagit med eller uteslutit.

Vetenskapsrådets utbildningsvetenskapliga kommitté arbetar för att sprida kunskap om forskningsresultat inom området lärande, kunskapsbildning, utbildning och undervisning. Skolverket verkar för att lärare, rektor och andra som arbetar inom skolan ska kunna ha användning av forskningsresultaten i sitt dagliga värv. Tillsammans tar nu Vetenskapsrådet och Skolverket fram ett antal forskningsöversikter, detta är den andra i serien. Den presenterar och diskuterar aktuell svensk forskning om bedömning. Översikten har tagits fram av docent Eva Forsberg, Uppsala universitet, och fil dr Viveca Lindberg, Stockholms universitet.



Vetenskapsrådet

Skolverket



Västra Järnvägsgatan 3 | Box 1035 | 103 78 Stockholm | Tel 08-546 44 000 | vetenskapsradet@vr.se | www.vr.se

Vetenskapsrådet är en statlig myndighet som ger stöd till grundläggande forskning av högsta vetenskapliga kvalitet inom alla vetenskapsområden. Utöver forskningsfinansiering arbetar myndigheten med strategi och analys samt forskningskommunikation. Målet är att Sverige ska vara en ledande forskningsnation.

ISSN 1651-7350

ISBN 978-91-7307-166-6
