

Pedagogisk utvärdering som styrning

– En historia från präster till PISA

OVE KARLSSON VESTMAN OCH INGER M. ANDERSSON



Forskning i fokus, nr. 35



MYNDIGHETEN FÖR
SKOLUTVECKLING

SERIEN FORSKNING I FOKUS VID MYNDIGHETEN FÖR SKOLUTVECKLING

Är en skriftserie som etablerats för att möjliggöra utgivning av material som producerats med stöd eller på uppdrag av Myndigheten för skolutveckling.

Det gemensamma för skrifterna är att Myndigheten för skolutveckling gjort bedömningen att materialet är av intresse för t.ex. pedagogiskt yrkesverksamma, beslutsfattare och forskare.

Författarna svarar själva för innehållet och de ställningstaganden som görs.

Publikationerna finns också på hemsidan www.skolutveckling.se

Från och med 2003-03-01 övergick serien genom en omorganisation från Skolverket till Myndigheten för skolutveckling.

BESTÄLLNINGSADRESS:
LIBER DISTRIBUTION
162 47 STOCKHOLM

TEL: 08-690 95 76
FAX: 08-690 95 50

E-POSTADRESS: skolutveckling@liber.se
www.skolutveckling.se

Best.nr. U07:176

©OVE KARLSSON VESTMAN OCH INGER M. ANDERSSON SAMT
MYNDIGHETEN FÖR SKOLUTVECKLING

Att mångfaldiga innehållet i denna bok, helt eller delvis, utan medgivande av Myndigheten för skolutveckling är förbjudet enligt lagen (1960:729) om upphovsrätt.

FORSKNING I FOKUS, NR 35

**Pedagogisk utvärdering som styrning
– En historia från präster till PISA**

ISBN 978-91-85589-34-0
ISSN 1651-3460

OMSLAGSBILD: ©JOHNÉR BILDBYRÅ AB
TRYCK: LENANDERS GRAFISKA AB, KALMAR 2007 · 25768

Svanenmärkt trycksak  Licensnummer 341 145

Pedagogisk utvärdering som styrning

- En historia från präster till PISA

OVE KARLSSON VESTMAN OCH
INGER M. ANDERSSON

MYNDIGHETEN FÖR SKOLUTVECKLING

INNEHÅLL

FÖRORD.....	8
SAMMANFATTNING.....	11
SUMMARY	16
INTRODUKTION TILL UTVÄRDERING.....	21
1.1 Utvärderingens komplexitet.....	22
1.2 Pedagogisk utvärdering	23
1.3 En historisk dimension.....	24
1.4 Utvärdering som styrning och kunskapsområde.....	26
2 TIDIGA FORMER AV UTVÄRDERING	28
2.1 En vid tolkning av utvärderingsbegreppet	28
2.2 Statens behov av styrning och kontroll.....	29
2.2.1 Läskunnighet i kyrkans och statens intressen ...	30
2.2.2 Vad bedömdes och hur?.....	31
2.2.3 Arbetslivet ställer nya kunskapskrav	33
2.3 Folkskolan inrättas	34
2.3.1 Skolinspektörer som granskare av skolan	34
2.3.2 Inspektörsrapporterna standardiseras	36
2.3.3 Folkskollärarna kritiserar inspektionen.....	37
3 UNDERVISNING SOM VETENSKAP	40
3.1 Professorer i pedagogik inrättas.....	40
3.1.1 Forskning och utvärdering får lärarnas stöd ...	41
3.1.2 Intelligensmätning och test	43
3.1.3 Två olika referenspunkter	44
3.1.4 Begreppet "evaluering" etableras.....	45
3.2 Rationella undervisningsmetoder	48
3.2.1 Differentieringsfrågan	48
3.2.2 Individualisering inom klassens ram	50
3.2.3 Utvärdering – ett verktyg för flera	51
3.3 Nya perspektiv på utvärdering.....	52
3.3.1 Influenser från USA.....	53
3.3.2 Att bedöma värde och nytta.....	54
3.3.3 Koder, ramfaktorer och processer.....	56
3.3.4 Teoriinriktad utvärdering.....	59

4	UTVÄRDERING OCH SKOLANS STYRNING	61
4.1	Minskad central styrning	61
4.1.1	Lokalt utvecklingsarbete – motorn i reformarbetet	64
4.1.2	Från reform till programutvärdering.	65
4.1.3	Försök med teoriinriktad utvärdering	66
4.1.4	Decentralisering, målstyrning och utvärdering. .	68
4.2	Utvärdering på nationell nivå	70
4.2.1	Fem professorer kritiserar utvärderingen	71
4.2.2	Det goda exemplets strategi.	73
4.2.3	Central tillsyn och lokal skolutveckling separeras	74
4.3	Utvärdering på lokal nivå.	77
4.3.1	Krav på utvärderarkompetens	78
4.3.2	Utvärdering för reflektion och lärande	80
4.3.3	Kvalitet och marknadstänkande.	82
4.4	Utvärdering på internationell nivå	86
4.4.1	Kritisk granskning av TIMSS och PISA	88
4.4.2	En nyliberal marknadsfilosofi	89
4.4.3	Kombination av extern och intern utvärdering. .	90
5	FORSKNING OM UTVÄRDERING	94
5.1	Metod och material	94
5.2	1970-tal – Forskning om utvärdering formas . . .	96
5.2.1	Utvärdering som systematisk verksamhet	97
5.2.2	Politik och påverkan från USA	99
5.2.3	Målen och referensramens betydelse.	101
5.3	1980-tal – Rationalitet och nytta ifrågasatt.	103
5.3.1	Utvärderingens värdemässiga aspekter	104
5.3.2	Rättvisa, kunskap och användning.	105
5.4	1990-tal – Intressenter, styrning och makt	106
5.4.1	Rationalitet och institutionalisering	108
5.4.2	Externt och internt perspektiv.	108
5.4.3	Styrning och decentralisering	109
5.4.4	Kriterier och intressenter.	111
5.4.5	Feministiskt perspektiv och maktfrågan	113
5.5	2000-tal – Forskning med pluralism	113
5.5.1	Mätbarhet av miljö och energi.	115
5.5.2	Konkurrens och ekonomisk jämförbarhet.	116
5.5.3	Nya vetenskapliga inriktningar	118

5.5.4	Intressentperspektiv och politisk styrning.	121
5.5.5	Granskning av kunskap och kvalitet	122
5.6	Trender inom utvärderingsforskning.	124
6	SAMMANFATTNING OCH KOMMENTARER	128
6.1	Från präster till skolinspektörer.	128
6.1.1	En skola på vetenskaplig och rationell grund . . .	129
6.1.2	Centralstyrning och utvärdering stärks	130
6.2	Från regelstyrning till målstyrning	132
6.2.1	Utvärdering för utveckling.	132
6.2.2	Indirekt centralstyrning med fokus på resultat . .	133
6.2.3	Kvalitetsredovisning ersätter utvärdering	134
6.3	Forskning och debatt	137
6.3.1	Ramar och processer i fokus	137
6.3.2	Politik, värderingar och legitimering.	139
6.3.3	Standardisering och ranking i marknadens tjänst	140
6.4	Riktningar för framtida utvärdering	141
6.4.1	Kvalitetsgranskning som informellt styrmedel . .	141
6.4.2	Internationalisering och standardisering.	142
6.4.3	Behov av en kunskapsgenererande utvärdering .	143
	REFERENSLITTERATUR	145
	AVHANDLINGAR I UTVÄRDERING 1970-2006 .	152
	SERIERNA MONOGRAFIER OCH FORSKNING I FOKUS	160
	Monografiserien	160
	Forskning i Fokus	163

Förord

I denna kunskapsöversikt dokumenteras hur företeelsen utvärdering vuxit fram inom utbildningsområdet. Författarna Ove Karlsson Vestman och Inger M. Andersson – och som medarbetare i kapitel 5 även Anette Månsson – sticker inte under stol med att de ser utvärdering som en normativ analys. Det innebär, att sammanhanget därmed blir avgörande för den utvärdering som utformas i olika tider. Utvärdering är praktisk och pragmatisk och grundar sig på både vetenskap och politik.

Utvecklingen i Sverige är ett tydligt exempel på detta! Staten behövde kontrollera genom präster och inspektörer. Staten behövde differentiera genom intelligenstest och pedagogiska mätningar. Staten behövde styra genom kunskapsmätningar. 1960-talet innebar en viktig förändring och utveckling. Forskning betonade, att även själva undervisningsprocessen och dess förutsättningar måste studeras. Detta för att man skulle kunna förstå och eventuellt förklara skillnader mellan skolor och/eller länder. Genom reformeringen av det politiska styrsystemet under 1980-talet blev utvärdering ett starkt styrinstrument med möjlighet till insyn i den pedagogiska praktiken.

När det gäller den forskningsinriktade verksamheten menar författarna, att utvärderingsforskning ofta handlat om tillämpningar inom olika ämnesdimensioner och forskningsområden. Det handlar om att diskutera och förhålla sig till principer för en vetenskaplig utvärdering. Utvärderarens teoretiska utgångspunkter och egna värderingar verkar vara avgörande även i en tolkande, hermeneutisk ansats.

Författarna lyfter också fram och diskuterar tidens trend att återigen prioritera det mätbara och kvantifierbara. Men nu är det inte system utan elever som står i fokus. Med viss oro pekar de på att dagens komplexa problem tycks hanteras med nygamla metoder eller metoder, som lånas från andra discipliner. Egentligen skulle dessa frågor kräva helt nya utvärderingsmodeller.

Detta är en intressant och lärorik läsning, som stämmer till eftertanke och reflektion. Många har nytta av att läsa och lära av denna bok. Utvärdering kräver grundlig reflektion samt kunskap om olika problemområden. Myndigheten för skolutveckling hoppas, att innehållet ska bidra till både fördjupade diskussioner om utvärdering och mer forskning om utvärdering och därmed en ökad kvalitet i de utvärderingar som genomförs på olika nivåer i utbildningssystemet. Författarna svarar själva för innehållet och de ställningstaganden som görs i boken.

Annika Andræ Thelin
UNDERVISNINGSRÅD

Sammanfattning

Utvärdering är en företeelse som från senare delen av 1900-talet har fått en allt vidare utbredning. Uttrycket utvärdering används idag tämligen frekvent av såväl beslutsfattare som befolkning. Det kan samtidigt ge upphov till oklarheter när man brukar samma ord eller uttryck men menar eller definierar det på skilda sätt. Vi ser utvärdering som en normativ analys av halten av en åtgärd, dess resultat och vägarna dit. Det finns olika idéer om hur denna normativa analys ska gå till och någon enhetlig, för alla tider, given metod eller modell för utvärdering finns inte. Det är vår övertygelse att sammanhanget (kontexten) är avgörande för vilken utvärdering som utformas och hur den tar gestalt i olika tider genom historien. Här spelar såväl vetenskap som politik en avgörande roll. Detta faktum utgör bakgrunden till det arbete vi redovisar i denna bok.

Vårt uppdrag handlade inledningsvis om att göra en kunskapsöversikt med fokus på utvärdering. Med den expansion inom utvärderingsfältet som vi bevittnar, och den mängd publikationer som återfinns idag, blev det svårt att göra en översikt av utvärdering som specifikt kunskapsområde. Utvärdering tar gestalt på skilda sätt inom olika verksamhetsområden och den fyller olika syften och funktioner i olika tider. För att förstå den utvärdering som vi möter idag har det därför varit nödvändigt att känna dess historia. Det vi vill förmedla i denna bok är därför hur utvärdering inom i första hand skolan har vuxit fram och hur dess funktion och utformning förändrats fram till idag. Vår berättelse grundar sig på den kunskap vi byggt upp utifrån tidigare forskning och relevanta källor. Vår egen mångåriga erfarenhet av utvärdering ligger också till grund för den bild vi tecknar i följande text. För att få en helhetsbild innefattar vår berättelse flera dimensioner, alla redovisade utifrån en kronologisk struktur.

Den historiska dimensionen, som är övergripande i vår framställning, för oss tillbaka till reformationen och nationalstatens

tillkomst. Staten behövde kontrollera att lagar och förordningar efterföljdes, och de protestantiska prästerna blev vid denna tid de som utförde kontrollen. I kyrkolagen 1686 påbjöds att alla i hushållsståndet skulle lära sig att läsa de heliga skrifterna, och i en kunglig resolution 1723 föreskrevs att föräldrarna skulle ansvara för den grundläggande undervisningen i läsning och kristendom. Genom stiftens nådde kyrkan och staten ut till hela den svenska befolkningen och genom prästernas återkommande förhör och systematiska anteckningar om medlemmarna i varje hushåll, kom prästerna att tidigt lägga grunden till en svensk befolkningsstatistik. Prästerna kom också att vara i majoritet i den statliga inspektörsinstitution som infördes i början av 1860-talet. Inspektörernas uppgift var att kontrollera att varje socken och stadsförsamling i landet följde 1842 års skollag och fullgjorde den statliga förordningen om allmän undervisningsplikt.

Med upplysningen och den naturvetenskapliga forskningens framväxt från 1700-talet fick en vetenskaplig dimension allt större genomslag. Nya forskningsområden som psykologi och pedagogik infördes som universitetsämnen i början av 1900-talet. Krav på ordning och systematik kännetecknade den naturvetenskapliga forskningen och dessa krav blev också vägledande principer för de nya disciplinerna. Med vetenskapligt utprövade test och standardiserade prov kunde individens intelligens, anlag och kunskaper mätas, jämföras och sorteras i olika nivåer eller grupperingar. En sådan systematisk bedömning gav också underlag för att göra prognoser och placera den enskilde individen på rätt plats både i skolan och i arbetslivet. Om en organiserad utvärdering i det agrara samhället hade funktionen att kontrollera att hushållen och kommunerna fullgjorde sina lagstadgade uppgifter att fostra och att undervisa den uppväxande generationen, så tilldelades utvärdering i form av intelligens-test och pedagogiska mätningar funktionen att differentiera de blivande medborgarna till olika uppgifter i det framväxande industrisamhället.

Vetenskaplig ledning fortsatte att gälla som vägledande princip i det moderna svenska samhälle som tog form efter andra världskriget. Rationalitet och effektivitet blev normgivande ideal för all samhällsplanering och alla verksamhetsområden, såväl offent-

liga som privata. Med vetenskaplig ledning och expertis skulle de mest optimala lösningarna sökas och det goda och jämlika samhället skapas för alla landets medborgare. Utvärdering blev i detta sammanhang ett objektiva och neutralt verktyg för systematisk bedömning av resultaten av genomförda reformprogram och insatser. Förebilder för en sådan bedömning (evaluation) hämtades från USA och fick inte minst genomslag inom pedagogisk utvärdering.

Reformeringen av det svenska skolsystemet inleddes med ett omfattande kommittéarbete och en försöksverksamhet med enhetsskola under perioden 1950–1962. Den nioåriga grundskolan för alla beslutades av riksdagen och enhetlighet och likvärdighet blev riktningssivande för den fortsatta skolutvecklingen. Liksom bostadsplaneringen skulle skolplaneringen följa en enhetlig norm eller standard för alla i hela landet. Vad som skulle undervisas och hur undervisningen skulle bedrivas formulerades och utprövades på central nivå för att sedan realiseras på lokal nivå. Den nationella läroplanen var styrdokument för verksamheten i varje enskild skola och Skolöverstyrelsen ansvarade för den nationella utvärderingen. Kunskapsmätningar var alltjämt den metod som dominerade, även om utvärdering nu skulle gälla genomförda utbildningsinsatser och undervisningsprogram.

Under 1960-talet kom kunskapsmätningar som metod för systemutvärdering att allt mer ifrågasättas av de pedagogiska forskarna i såväl USA som Sverige. För att förstå och kunna förklara skillnader i skolornas uppmätta resultat och göra jämförelser mellan olika länder, var det nödvändigt att också söka information om själva undervisningsprocessen och dess förutsättningar. De kvantitativa kunskapsmätningarna gav inte tillräcklig information om den pedagogiska praktiken, och datainsamlingen måste därför kompletteras med hjälp av kvalitativa metoder. Centralt utprövade undervisningsprogram och metoder hade inte heller visat sig vara de optimala pedagogiska lösningarna. Utifrån den nationella läroplanens mål måste lärarna själva utforma undervisningen och utvärdera vilka arbetsformer och arbetssätt som fungerade i klassen. Såväl den nationella utvärde-

ringen som den lokala skulle ha en kunskapsbyggande funktion och ge underlag för en fortgående skolutveckling.

I början av 1980-talet inleddes en reformering av det politiska styrsystemet och här urskiljer vi ytterligare en dimension i vår berättelse om utvärdering. Reformeringen av det svenska förvaltningssystemet innebar en överföring av befogenheter och ansvar från centrala organ till lokala. Med reformen följde samtidigt ett behov av att granska vad som faktiskt pågick i praktiken, och utvärdering blev i detta sammanhang ett starkt instrument för styrning och insyn i den pedagogiska praktiken. Förskjutningen mot ett ökat lokalt inflytande och ansvar följdes samtidigt av en central styrning i form av särskilda statliga satsningar på tidsbegränsade projekt. Med centralt formulerade betygskriterier och allt fler inslag av nationella och internationella prov förskjuts styrning ytterligare mot central nivå. Under den senaste tioårsperioden har också ett marknadstänkande fått ett ökat genomslag och kvalitet har blivit en tvingande regim även inom den offentliga sektorn.

Av den forskningsinriktade dimension som vi redovisar framgår det tydligt att utvärderingsforskning i huvudsak handlat om tillämpningar inom olika ämnesdiscipliner och forskningsområden. Den större andelen av de avhandlingsarbeten som vi har granskat rör särskilda insatser eller program inom olika verksamhetsområden eller praktiker som exempelvis byggteknik, sjukvård, socialtjänst, utbildning och undervisning. En mindre andel av utvärderingsforskningen handlar om att diskutera och förhålla sig till principer för en vetenskaplig utvärdering – allt från principen om objektivitet och neutralitet till principerna om demokratisk utvärdering och mångfald av intresser. Utvärderarens teoretiska utgångspunkter och egna värderingar verkar avgörande hur han/hon betraktar sitt utvärderingsobjekt. Även här har skilda principer diskuterats bland utvärderingsforskare och det återspeglas även i det avhandlingsmaterial som vi har analyserat. Frågan om utvärderingens relativitet har allt mer kommit i fokus. Det är vanligt att man som forskare tar avstånd från den traditionella synen på utvärdering som objektiv och rationell och mer argumenterar för en hermeneutisk, tolkande ansats vid

utvärdering. Utvärdering som politiskt verktyg lyfts också fram och paras med diskussionen om utvärderingsforskning som en icke-neutral verksamhet.

Samtidigt finns en trend i dagens globala samhälle, att återigen prioritera det mätbara och kvantifierbara. Påtalade problem, exempelvis i vår miljö, som med det snaraste måste åtgärdas ställer krav på nya utvärderingsmodeller som kan brukas för utvärdering av komplexa insatser för hållbar utveckling. Det europeiska samarbetet, som förutsätter ökad rörlighet och växande konkurrens på marknaden, ställer krav på tillförlitliga och effektiva mätinstrument för komparativa undersökningar. I skolan blir prov och betyg åter den optimala metoden, men riktas nu mot elever istället för system. Under det senaste årtiondet har utvärdering i form av kvalitetsgranskning och kvalitetsredovisning fått en ytterst stark ställning som informativt styrmedel. Inom skolan har också nygamla utvärderingsmodeller, som inspektion, väckts till liv och såväl nationella som internationella prov och betyg har fått ökad betydelse som indikatorer på kvalitet. "Evidens-baserade" eller bevisbara åtgärder är också en aktuell trend, företrädesvis inom medicin och sjukvård, men har även fått gehör inom socialtjänst. Något som också håller på att växa sig starkt är utvärdering som en del i en övernationell styrning, där inte minst utbildning är ett centralt problemområde.

Summary

In the last decade evaluation has been disseminated to ever more sectors of society. This expansion is mirrored in the fact that the term evaluation is frequently used by public decision makers as well as by receivers of public sector services, however circumscribed in many different ways – a source of potential misunderstandings. In this book evaluation is defined as a normative analysis of an intervention, its results and the causal links between the two. There are different conceptions of how such a normative analysis should be conducted and in our view there can be no eternal model; contextual factors, mainly policies and science, determine which models are used. That is the background of our book.

Our initial task was to present an overview of the current conceptions of evaluation. However, the tremendous expansion of evaluation made it impossible to map evaluation as a knowledge area; the publications on evaluation are innumerable and its functions vary with the sector in which it is conducted, let alone over time. Thereby the current features of evaluation cannot be explained without knowledge of their history. In this book we describe the development of evaluation in the educational sector; the alterations of its functions and the introduction of new evaluation models. Our own experiences and research together with other relevant sources constitute the foundation of this multi-dimensional and chronological description.

Our recapitulation of history starts with the Reformation and the births of the centralised states. The state needed surveillance of whether laws and regulations were satisfied and the Protestant priests came to fill the function of controllers/evaluators. The Swedish Church act of 1686 stated that all members of every household should learn to read the holy texts and the Royal resolution of 1723 stated that parents were responsible for teaching their children to read and study basic ecclesiology. The parishes became the link as well between the Church and the people as also between the state and the people. The notes on the knowledge's of every person in the parishes that were

made by parish priests during control rounds in the households are the origin of Statistics Sweden. In the 1860s a state inspector institution of the schools was founded and it was constituted to a majority by priests. The task of the inspectors was to find out whether the rural districts and the cities satisfied the School act of 1842 and the regulation on general duty of teaching duty.

The Enlightenment and the development of science in the 18th century disseminated scientific methods into the evaluation practices. At the beginning of the 20th century new disciplines were introduced at the universities, for example psychology and pedagogy. The emphasis on systematic enquiry was features of these new disciplines. Scientifically developed standardised tests were used for estimating intelligence, knowledge and skills, whereby individuals could be compared and categorised. Prognoses were made to supply individual pupils with appropriate education and to direct them toward working places for which they were suited. Thereby evaluation was given a new function; in rural Sweden it had been used to control that the districts and households fulfilled their duties concerning educating the members of the household and the youngest generation, as by contrast in early industrial Sweden evaluation in pedagogy was used to allocate sufficient resources and to direct appropriate tasks to every citizen.

After world war two science continued to guide social change as rationality, effectiveness and efficiency became guiding principles in all sectors, public and private. Expertise was used for finding the optimal ways of creating the good, equalitarian society. Evaluation was in this context considered an objective tool for systematic enquiry on the results of reform programmes. Evaluations conducted in the USA were used as shining examples, not the least in pedagogical evaluation.

A period of major reforms of the Swedish school system began with a parliamentary committee investigation. From 1950 to 1962 the Swedish school system was centralised. The parliament decided on a nine year preliminary school, whereby standardised national norms were made a necessary prerequisite for

further development of the school system. This mirrored the development in other sectors, for example the planning of housing. The central level was to conduct quasi-experiments and to decide what should be taught by which methods. The national curriculum for education in Swedish schools was binding instructions for every school and the conducted national evaluations of the educational sector. Knowledge assessments were still the dominant method, although evaluations of the implementations of educational programmes should be done.

During the 1960s the quantitative knowledge assessments as a method for evaluation of systems got ever more questioned by American as well as Swedish researchers on pedagogy. Supplementary qualitative information on the educational process and its prerequisites was now considered necessary for explaining differences between results of different schools or countries. Furthermore, the centrally developed programmes and methods had not proved satisfactory and therefore teachers were now given the tasks to specify how the education should be conducted and to evaluate the implementation with the goals of the national guidelines as benchmarks. The view gained ground that as well local as also central evaluations should generate knowledge on which the school system could be further developed.

At the beginning of the 1980s the methods of governance were modified. Competences were re-allocated from central to local levels of government. This justified the introduction of evaluation as an instrument for central control of the pedagogical practice. Central control was also strengthened by allocation of resources to time-limited projects, by formulating grading criteria centrally and by an increase in the number of national and international tests. During the last decade market oriented adages and quality assessments have become ever more important in the public sector. Today the general school inspections is the dominant commission for the National board of Schools and different kind of tests and measurement of knowledge have been more and more important.

Our research focused description of evaluation in the educational sector clearly shows that evaluation research has dealt mainly with applications of evaluation within different disciplines and research areas. The majority of the dissertations under study discuss particular projects or programmes in different sectors, or they evaluate practices in for example construction, public health, social work or education. Evaluation research is mainly about discussing and relating to principles of a scientifically conducted evaluation that is principles on the broad spectrum from objectivity and neutrality to democratic evaluation and diversity of interest groups. The evaluator's own theoretical starting points and values seem decisive for his/her conception of the evaluation object. Also concerning this theme different principles have been discussed among evaluation researchers, which is mirrored in the dissertations that we have analysed. The question of the relativity and subjectivity of evaluation is ever more focused upon. Today it is usual that evaluation researchers dissociate from the traditional view on evaluation as objective and rational and argue for a hermeneutic, interpreting approach in conducting evaluations. Evaluation is further described as a not all neutral tool for governance.

At the same time the trend to strive for the measurable and the quantifiable has been revitalised; to measure and compare are explicit political demands in today's society. Observed environmental problems, which call for immediate measures, highlight the need for new evaluation models to be used in evaluations of complex interventions for sustainable development. The European Union, which assumes increasing mobility and competition on markets, requires reliable and effective tools for comparative studies. Exams and grades are again becoming the optimal methods, this time however used for controlling pupils rather than systems. During the last decade evaluation in the form of quality assessment has gained a strong position as an informative tool of governance. The older models have not vanished though. For example inspections have been brought to new life as a method. At the same time evaluation gets ever more prominent as a part of supranational governance concerning, among other policy areas, education.

I Introduktion till utvärdering

Vid utvärdering brukar blicken riktas mot olika företeelser i samhället för att granska och bedöma effekterna av vad som pågår eller har genomförts i en given verksamhet.¹ Det är däremot inte lika vanligt att intresset riktas mot utvärdering som verksamhet, och särskilt då utvärdering som styrning, vilket är syftet med denna bok.

Att rikta blicken mot utvärdering kan motiveras med att det är en verksamhet som under den senaste femtioårsperioden invaderat allt fler områden i samhället och i många fall närmast blivit något av ritual. På alla nivåer i en organisation, såväl offentlig som privat, skall utfallet av verksamheten kontinuerligt utvärderas. Även i den privata sfären förväntas vi utvärdera. Det kan gälla den produkt vi inhandlat eller den service vi fått under hotellvistelsen, på charterresan, av banken och som användare av olika tjänster och program. Men vi möter utvärderingar inte bara som kunder och brukare. Utvärdering har även genomförts av stora världshändelser som Gulf-kriget, och en särskild utvärdering av den svenska regeringens hantering av Tsunamikatastrofen genomfördes under 2005 (SOU 2005:104). En utvärdering har också efterlysts av utfallet av folkomröstningen om EU-konstitutionen i Frankrike och Holland våren 2005.

Utvärdering förefaller således vara en mycket populär och efterfrågad verksamhet. Man kan fråga varför? Ett svar kan vara att den ger en möjlighet att lära oss och dra nytta av våra misstag, eller i vart fall visa att vi vill kunna göra detta. Utvärdering signalerar med andra ord läraktighet, effektivitet och seriositet. En utvärdering kan också visa om det blev som vi kom överens om och bekräftar på så sätt konsensus och förmedlar därtill en känsla av trygghet – att verksamheten är under kontroll.

¹ Se t.ex. titlar som *Det värderande ögat* av Björndal 2005 och *Den värderande blicken* av Eliasson, m.fl. 1990.

Om och hur utvärdering kan uppfylla dessa förväntningar hoppas vi kunna belysa med följande presentation.

1.1 Utvärderingens komplexitet

Vårt arbete med att färdigställa en kunskapsöversikt inriktad på utvärdering har pågått länge, och vägen fram till det resultat som här redovisas har varit kantad med rader av uppehållsstationer för att tänka igenom och försöka förstå det objekt vi undersöker. Vi hade visserligen god kännedom om utvärderingsforskning genom mångåriga studier och som forskare i pedagogik med inriktning på utvärderingsområdet. Och vi har själva också genomfört en rad omfattande utvärderingar på uppdrag av olika statliga myndigheter, i internationella sammanhang och åt en rad kommuner. (Se t.ex. Karlsson 1996, 1998, 2004; Karlsson & Kullberg 1998; Karlsson & Kuivanen 2000; Karlsson, Andersson & Lundin 2002). Men att gå utanför vår egen kunskapsbas och generera en mer generell kunskap om pedagogisk utvärdering förutsatte både ett kritiskt förhållningssätt till utvärdering som företeelse och en distansering från den utvärderingsverksamhet som vi själva varit involverade i genom åren. Resultatet av vårt arbete, som vi närmast vill beskriva som en metarefleksion över vårt eget forsknings- och verksamhetsområde, redovisar vi i denna bok.

Det vi behandlar är formell utvärdering, vilken skiljer sig från den typ av bedömningar som vi gör av ting, händelser och handlingar i vårt dagliga liv och som kan beskrivas som informella utvärderingar. Utmärkande för en formell utvärdering är att den förväntas ske på ett systematiskt och genomtänkt sätt, vilket är en precisering som återkommer i de flesta definitioner som vi funnit i litteraturen. Utvärdering förstås därmed som en rationell process där man brukar vissa metoder och försöker att ha kontroll över de olika stegen i bedömningsarbetet. Krav på systematik finns även i den definition som ges i Nationalencyklopedin (NE) som beskriver utvärdering som en "sammanfattande term för metoder som syftar till en systematisk bedömning av resultaten och de mer långsiktiga effekterna av genomförda insatser."(www.ne.se).

Redan med denna korta redovisning kan vi konstatera att utvärdering är en mångfacetterad verksamhet, vilket innebär svårigheter att göra en kunskapsöversikt i traditionell mening där utgångspunkten är att det finns ett avgränsat kunskapsområde som kan och skall överblickas. Verkligheten är som bekant betydligt mer komplex än kartan, och när det gäller utvärderingar kan en kategorisering göras utifrån skilda grunder; i olika utvärderingsmodeller, utvärderingsansatser, utvärderingar inom skilda verksamhetsområden, utvärderingar från olika delar av världen, utvärderingar av skilda utbildningssystem, etc. (Franke-Wikberg & Lundgren 1980; Guba & Lincoln 1989; Furubo, Rist & Sandahl 2002; Karlsson, 1999; Kellaghan, Stufflebeam & Wingate 2003; Stern 2005). Detta gör att varje försök att avbilda utvärdering i form av en översikt eller en gemensam berättelse kan kritiseras på samma sätt som vetenskapsteoretiska översiktsverk brukar kritiseras. Likväl behöver vi en bild av utvärdering för kunna orientera oss i detta expanderande fält och inte minst för att kritiskt kunna granska de utvärderingar som redovisas. Nedan ska vi göra ett försök att ringa in ett antal aspekter som kan urskiljas inom utvärderingsområdet – aspekter som bildar utgångspunkter för vårt sätt att behandla utvärdering som företeelse och som styrinstrument i skola och undervisning.

1.2 Pedagogisk utvärdering

Vi lägger tonvikt på utvärdering inom utbildning och undervisning, vilket allmänt brukar kallas ”pedagogisk utvärdering”. Även om denna avgränsning innebär en viss precisering innebär den ingalunda att vårt undersökningsobjekt därmed har blivit enkelt överblickbart. På alla nivåer inom utbildningssystemet, från förskolan till den högre utbildningen vid universitet och högskolor, genomförs idag återkommande utvärderingar, och det är inom detta fortfarande relativt omfattande fält som vi söker generera kunskap. Det bör sägas att vi i den följande framställningen främst kommer att hålla oss inom grundskolans område.

Vid pedagogisk utvärdering står åtgärdsprogram, projekt eller produkter för uppfostran och undervisning i fokus. Det fastläggs i Franke-Wikbergs och Lundgrens allmänna definition av pedagogisk utvärdering som innebär: "Att värdera och uttala sig om resultatet av ett program för uppfostran och undervisning." (1980, s 10). Vad en sådan utvärdering faktiskt resulterar i sammanhänger med hur man som utvärderare besvarar grundläggande frågor om vad som skall utvärderas, i vilket syfte och för vem utvärderingen sker, samt hur och av vem den utförs. Författarna betonar att resultaten inte bara skall ställas mot uttalade mål. Utfallet ska också relateras till insatsens genomförande och förutsättningar. Därigenom ges en möjlighet att förklara de resultat som uppnåtts mot bakgrund av såväl förhållandena internt på skolan som i samhället. Författarna pläderar för en "teoriinriktad utvärdering" och menar att det gäller att utveckla en referensram för det objekt som är föremål för utvärdering. Utvärderaren skall söka förklaringar som kan ge mening åt resultatet och söka förstå vad som sker bakom det uppenbara och naturligt för handen varande. Forskning om utbildning och undervisning grundad på pedagogisk teoribildning blir här en viktig förutsättning för pedagogisk utvärdering (Kallós & Lundgren 1974). Ett ytterligare exempel på en pedagogisk forskare som har presenterat grundläggande arbeten i denna riktning är Urban Dahllöf (Dahllöf 1978) och hans utveckling av "ramfaktorteorin". (Se även Gustafsson & Selander 1994). Vi återkommer till dessa frågor längre fram i boken.

1.3 En historisk dimension

För att förstå utvärdering inom utbildning och undervisning i vår tid är det nödvändigt att ha kunskap om vår svenska utbildnings- och utvärderingshistoria. Frågor som rör den historiska framväxten av utvärdering tar vi upp i nästa kapitel.

Den folkundervisning som föregick den allmänna folkskola som växte fram från mitten av 1800-talet låg i händerna på kyrkan och prästerskapet. Ett centralt mål var att folket skulle tillägna sig den protestantiska läran. Hushållen skulle ansvara för själva

undervisningen, och kontrollen gjordes av prästerna vid återkommande husförhör, vilka vi har valt att betrakta som en tidig form av utvärdering. Med dagens språkbruk skulle prästerna kunna benämnas ”externa utvärderare” av hushållens ”undervisningsinsatser”. Med Upplysningen följde en annan syn på utbildningens roll i samhället. Kravet ökade på en samhällsordning baserad på vetenskapliga principer, individuella fri- och rättigheter och etablerade politiska system. Medborgarnas utbildning blev en uppgift för staten när riksdagen år 1842 fattade beslutet om en allmän folkundervisning organiserad som skola. För att kontrollera att de lokala aktörerna följde skolstadgans föreskrifter inrättades en skolinspektion, och så småningom ökade antalet statliga folkskoleinspektörer. Även dessa utförde sitt uppdrag som ”externa utvärderare”, nu med skolan som utvärderingsobjekt.

Utvärdering som begrepp infördes med reformeringen av det svenska skolsystemet som inleddes vid mitten av 1900-talet. I kapitel 3 beskriver vi hur utbildningsplanering var en central fråga för politikerna, och hur pedagogisk utvärdering under 1950- och 60-talet blev ett verktyg för en vetenskaplig organisation och ledning av utbildning och undervisning. Den pedagogiska utvärdering kom därvid att utvecklas till ett specialområde inom pedagogisk forskning och många pedagogiska forskare fick uppdrag som externa utvärderare av de reformprogram och utbildningsinsatser som riksdagen beslutat. Under 1970- och 80-talet vidgades perspektivet och utvärderingen av skolan sattes in i ett samhällspolitiskt och kulturellt sammanhang. I olika statliga och kommunala styrdokument framhålls kommunernas utvärdering av skolan som ett viktigt kunskapsområde. Idag har denna bild av pedagogisk utvärdering på olika nivåer i förvaltningsorganisationen komplicerats ytterligare genom att utvärdering av skolan har delats upp i en tydligare tillsyn och kontroll som en uppgift för Skolverket, medan ansvaret för utveckling har lagts på Myndigheten för skolutveckling. Under 1990-talet finns också en inriktning inom pedagogisk utvärdering som betonar utveckling, lärande och lokal förankring genom medverkan från olika berörda intressenter (Franke-Wikberg 1990; Karlsson 1995; Skolverket 1998).

1.4 Utvärdering som styrning och kunskapsområde

Utvärdering har också blivit en allt viktigare del i styrningen av skolan, vilket är en av de aspekter som vi kommer att uppmärksamma i bokens fjärde kapitel. Genom att se pedagogisk utvärdering i ljuset av förändringar av den politiska styrningen av skolan riktas uppmärksamheten på frågan om vem/vilka som har makt och inflytande vid utvärdering. Frågor av detta slag diskuteras mot bakgrund av den decentralisering som inleddes under 1980-talet och övergången från regelstyrning till mål och resultatstyrning. Vi granskar också det ökade inslaget av internationella jämförelser av skola och utbildning under 2000-talet.

Sedan utvärdering började ta form som praktik i USA under 1960-talet har den etablerats som profession. Det har skett genom utveckling av metoder och utvärderingsmodeller samt ett professionellt nätverk med egen organisation, tidskrift och etiska regelverk. Utvecklingen av en profession är dock en annan än den för en forsknings- och ämnesdisciplin. En disciplin stakar ut gränser för ett kunskapsområde i relation till andra forskningsområden och utvecklar ett koncept eller en självförståelse av sin verksamhet, samt en filosofi och logik för hur arbetet inom disciplinen skall ske. Även om utvärdering har praktiserats som profession under många år så är det först under det senaste decenniet som den börjat utvecklas som forskningsdisciplin, dvs. som ett eget kunskapsområde. Det sker genom att flera utvärderingsforskare utvecklar teorier, begrepp och metoder för att granska och förstå utvärdering som företeelse i samhället (se t.ex. Alkin 2004; House & Howe 1999; Scriven 1980; Shadish, Cook & Leviton 1991).

Statens krav på utvärdering av offentlig verksamhet har också inneburit att utvärdering som tillämpad forskning, vanligen kallad utvärderingsforskning – evaluation research (Scriven 1991) vuxit fram inom *skilda* vetenskapsområden och discipliner som pedagogik, statskunskap, ekonomi, socialt arbete, m.fl. Utvärderingsforskning uppvisar därför många olika sidor och inrymmer både likheter och olikheter. Hur utvecklingen av

utvärdering som forsknings- och kunskapsområde i Sverige ser ut kommer vi att se närmare på i kapitel 5 genom att analysera bidrag till utvärderingsforskning i form av svenska avhandlingar inom utvärderingsområdet. I det sista kapitlet gör vi en sammanfattande analys och diskussion med några tankar kring tendenser inom utvärdering i styrning av skola och offentlig verksamhet.

2 Tidiga former av utvärdering

Det är främst historien om utvärdering som instrument för politisk styrning och kontroll som vi skall berätta i detta kapitel. Syftet är att visa i vilket historiskt sammanhang en organiserad granskning för politisk styrning och kontroll uppkommer i Sverige. Vi vill också peka på några drivkrafter som vi menar har varit avgörande, och hur den expansion av utvärdering som vi ser idag kan förstås. Den historia vi berättar är tecknad i grova drag och utesluter därför en mängd detaljer som i sig kan vara av intresse, men inte bidrar till den förståelse vi här eftersträvar och som förutsätter en förenkling av i sig ytterst komplexa skeenden. De uppgifter som redovisas i den historiska tillbakablicken har huvudsakligen hämtats från sekundärkällor.

2.1 En vid tolkning av utvärderingsbegreppet

Den fråga som infinner sig är var en sådan historiebeteckning skall starta? Skall den exempelvis inkludera de tidiga försök till att följa upp och kontrollera folks läsförmåga och förståelse av de religiösa texterna som var en av kyrkans uppgifter redan under 1600-talet? Eller skall man ta avstamp vid mitten av 1900-talet då de moderna formerna av utvärdering uppstod? Svaret på frågan torde delvis ha att göra med hur man vill avgränsa begreppet utvärdering. Statsvetaren Evert Vedung (1998) inleder, i sin bok om utvärdering i politik och förvaltning, med att nämna historiska exempel på utvärdering redan från antiken när forskare och lärde togs till hjälp för att bistå hovet med råd och dåd i statsangelägenheter. Vedung säger vidare att "Utvärdering i politik och förvaltning handlar om att i efterhand upptäcka, kartlägga och bedöma offentliga åtgärders förvaltning och resultat i syfte att åstadkomma ökad självreflektion, fördjupad förståelse och bättre grundade beslut" (s. 19).

Om vi använder oss av en bred tolkning av utvärderingsbegreppet kan statens och kyrkans politik på 1600-talet i syfte att lära

folket att läsa ses som en form av offentliga åtgärders förvaltning och resultat som utvärderades genom den efterhandkontroll prästerna gjorde vid sina husförhör. Flyttar vi oss framåt i tiden finner vi exempel på hur statens tillsyn och kontroll av skola och undervisning tar fastare former genom inrättandet av skolinspektörer på 1860-talet. Även idag är tillsyn och kontroll en framträdande uppgift för Skolverket. Sedan några decennier har också utvärderingens roll för att främja utveckling och åstadkomma ökad självreflektion och fördjupad förståelse, som Vedung uttrycker det, blivit en viktig uppgift, och idag betonas gärna både kontroll och främjande som två sidor av utvärderingsmyntet.

Tillsyn och kontroll har således funnits som en funktion redan i tidiga former av utvärdering av skola och undervisning. Vill man göra en gradering kan man se dessa tidiga varianter som en enkel form av utvärdering till skillnad från de komplexa former som litteraturen numera ofta beskriver. Fördelen med att inkludera både de enkla och de komplexa formerna i definitionen är att det ger en möjlighet att tala om utvärdering med en lång historia och se på förändring över tid. Vi tror att det är viktigt för att förstå dagens utvärdering inom skola och undervisning.

2.2 Statens behov av styrning och kontroll

Behovet av instrument för politisk styrning och kontroll uppkom med framväxten av nationalstaten – en process som i Sverige inleddes med reformationen och övergången från katolicism till protestantism. Reformationen fick stöd av feodalherrar och enväldiga kungar som ville frigöra sig från påven och Vatikanstaten och bilda egna stadsstater eller nationalstater. I Sverige inleddes nationsbyggandet med Gustav Vasa som år 1523 valdes till svensk kung. Som kunglig överhöghet kan han sägas ha fått ”draghjälp” av den lutherska protestantismen som år 1527 genomdrevs vid riksdagen i Västerås.

I den katolska kyrkan var de heliga skrifterna skrivna på latin och därigenom endast tillgängliga för de skriftlärdade prästerna.

Luther och protestanterna ställde kravet att den enskilde själv skulle kunna ta del av vad Gud i sitt heliga Ord bjuder och befäller. I enlighet med principen – ”Skriften allena” – och betoningen av det allmänna prästadömet skulle individen själv kunna läsa och förstå förlåtelsens ord (Johansson 1981). De återkommande förhören av den enskildes läsning och förståelse av Ordet blev, liksom bikten i den medeltida katolska kyrkan, ett medel för ideologisk styrning och kontroll av folket.

2.2.1 Läskunnighet i kyrkans och statens intressen

I kyrkolagen 1689 föreskrevs att alla i befolkningen skulle läras att läsa i bok. Lagen var inte bara en påbjuden förordning utan tvingade även församlingsborna till efterlevnad då ett minimum av läsförmåga och katekeskunskaper sattes som kriterium för rätten att ingå äktenskap och fostra barn, att delta i församlingens nattvardsgemenskap och att uppträda som vittne vid tinget (Johansson 1981). Kyrkans strävan att ideologiskt styra och kontrollera församlingsundervisningen låg i linje med kungamaktens intressen att forma en nation.

Den period från omkring 1680 till 1730, när den protestantiska kyrkan initierade och organiserade en alfabetisering av folket, inföll i slutet av den s.k. Stormaktstiden (1611–1718). Sverige var denna tid ett utvecklade envælde med en klart utformad politik som syftade till att expandera territoriet och forma en nationell enhet. Av såväl praktiska som politiska skäl låg det i nationalstatens intresse att stödja utvecklingen mot en enhetlig religion och ett enhetligt språkbruk – det svenska modersmålet. Kyrkans läskampanj, som man kan se som ett tidigt exempel på ett centralt beslutat nationellt utbildningsprogram, kom på så sätt att bidra till en försvenskning av befolkningen i de områden som Sverige införlivat under sina många och långa erövringskrig.

Förutom den tidigare nämnda kyrkolagen 1689 tillkom andra förordningar som skulle gälla i hela kungariket Sverige. I en kunglig resolution år 1723 ålades således föräldrarna att: ”flitigt hålla sina barn till att lära läsa i bok och deras kristendomsstycken.” (Ibid, s. 202). I centrum för denna utbildningsreform stod hemmet. I enlighet med Hustavlans patriarkala världsbild

var hufadern överhuvud i hufhållsftåndet – en focial enhet fom vid denna tid även omfattade pigorna och drångarna på gården. Hufadern fculle anfvura för att alla fom ingick i hans hufhåll lärde sig läsa de heliga fkrifterna. Hemundervisningen fsköttes dock vanligen av någon av kvinnorna i hufhålllet. I de fall där föräldrarna själva inte kunde undervisa barnen fculle de sättas i lära hos fkolmästaren eller klockaren. Vid fekelstiftet 1700 var den kyrkliga läskampanjen i full gång, och ända till slutet av 1800-talet återfinns noteringar i förhörslängderna om resultatet av de insatser fom gjordes för att befästa läsförmågan och katekeskunskaperna hos församlingsborna.

Det var således en omfattande utbildningsreform för alfabetisering fom infördes i Sverige under 1700-talet. Insatserna utvärderades också kontinuerligt genom prästerna. Prövningen gjordes vid kyrkoförhör och vid hufförhör ute i kyrkorotarna, där man samlades hos någon av storbönderna. I kyrkböckerna förde prästen in uppgifter om: födsel och dop, lysning och vigsel, ut- och inflyttning, död och begravning. Och i förhörslängderna noterade han sin prövning och bedömning av den enskildes läsförmåga och katekeskunskaper.

2.2.2 *Vad bedömdes och hur?*

Det fom prästerna granskade och bedömde i sina ”utvärderingar” var i första hand att församlingsborna tillägnat sig den Lutherska läran. Läskampanjen var därför också helt och hållet knuten till de religiösa texterna, främst psalmboken och katekesen med Luthers förklaringar. Båda dessa böcker blev de mest spridda texterna under 1600-talet, och i båda fanns anvisningar till läsaren fom betonade: ”kraven på aktiv och engagerad läsning och tillämpning i livet.” (Johansson 1981, s. 202). Vid hufförhören, fom var en central del i den utvärderingsmetod fom man tillämpade, lades betoningen på kunskaperna i Ordet, vilket innebar att ”Den enskilde fculle kunna läsa och förstå förlåtelses ord.” (Johansson 1983, s. 102).

Dokumentationen fom skedde i form av prästernas anteckningar i förhörslängderna visar på stora variationer. Bedömningarna kunde också vara mer eller mindre formalise-

rade. Betygsgraderingar förkommer allmänt i förhörlängderna från mitten av 1700-talet, men betygssystem och symboler, som bestämdes av stiften, skiftar över landet. En förhørs- och communionbok från Tuna församling i Härnösands stift från år 1814 till år 1824 kan redovisas som exempel. (För en fördjupning se Johansson 1983, s. 101–105).

I kolumnen ”Läser och förstår” i förhörlängden anges här kriterierna för de tre centrala förmågor som utvärderades: Innanläsning och Begrepp, samt i en särskild kolumn Naturgåfvor. Efter en angiven skala skulle församlingsprästen vid varje förhörstillfälle göra en bedömning av den enskildes prestationer och med ett enkelt tecken sätta betyg på dennes kunskaper, färdigheter och förmågor. Det gjordes med hjälp av ett antal kriterier som angav en progression i 4 till 6 steg beroende på vad som bedömdes.

Tabell 1. Exempel på bedömningskriterier

<i>Vad bedöms: Mot vilka kriterier</i>	<i>Innanläsning</i>	<i>Begrepp</i>	<i>Naturgåfvor</i>
	Har begynt att läsa	Har vagt begrepp	Har trögt minne
	Läser svagt	Har någorlunda begrepp	Har svagt förstånd
	Läser någorlunda	Har försvarligt begrepp	Är lättjefull
	Läser försvarligt	Har godt begrepp	Har godt minne
	Läser med färdighet		Har godt förstånd
			Har flit

Den enskilde prästens uppfattning om vad som var svagt, någorlunda, försvarligt, godt eller trögt låg till grund för bedömningen och var helt avgörande för den enskilde individen. I dagens utvärderingsspråk kan man tala om en expertutvärdering där prästerskapet tolkade stiftets kriterier för vad som skulle bedömas och vad som var ett nödvändigt minimum för att bli bekräftad som fullvärdig församlingsmedlem. De individer som inte uppnått detta minimum utestängdes från möjligheterna att

ingå äktenskap och fostra barn, delta i församlingens nattvards-gemenskap och uppträda som vittne vid tinget. Prästerskapet hade rättigheterna att hålla husförhör, leda konfirmationsläsningen och med konfirmationen bekräfta de församlingsbor som tillägnat sig den minimikompetens som stiftet bestämt. Församlingsprästerna kom härigenom också att styra och kontrollera den sociala- och kulturella reproduktionen. Men nya ideologiska och politiska intressen började snart bryta in.

2.2.3 *Arbetslivet ställer nya kunskapskrav*

Under frihetstiden (1718–1772) fick den svenska naturvetenskapliga forskningen internationell betydelse och en starkare ställning i den akademiska utbildningen. Upplysningstidens idéer som ifrågasatte kyrkans överhet över den lärda skolan och den politiska medvetenhet som väcktes med den franska revolutionen kom till uttryck och fick sin spridning i den offentliga diskursen. Den agrara omvandlingen, den frambrutande merkantilismen och den begynnande industrialismen innebar en förändringsprocess som i grunden kom att omvandla det svenska samhället. Nya ideologiska och politiska drivkrafter bidrog till att förskjuta den protestantiska kyrkans hegemoniska makt, bryta det kungliga enväldet, upplösa ståndssamhället och bana väg för det moderna samhället och en borgerlig hegemoni. Den kyrkliga lästraditionen och prästerskapets inflytande skulle dock under lång tid framåt komma att behålla sitt grepp över uppfost-
ran, utbildning och undervisning.

Den agrara omvandlingen innebar en omstrukturering av böndernas ägor i syfte att rationalisera och effektivisera jordbruket. Strukturuomvandlingen medförde samtidigt att ett ökande antal egendomslösa lantarbetare blev tvungna att lämna landsbygden och söka sig till ett lönearbete i städerna och inom den framväxande industrin. Den enhet för barnens primära socialisation som hushållen utgjort kom härigenom att upplösas, eller i vart fall att försvagas som institution för undervisning. Föräldrarnas lagstadgade ansvar för den grundläggande undervisning, som skulle föregå konfirmationsläsningen för prästen, kunde inte fullgöras. Samhällsomvandlingen ställde också krav på nyttiga kunskaper i arbetslivet och för livet i det moderna industrisamhället, och

allt detta medförde ett växande behov av skolor och lärare för en undervisning som dels kunde ta hand om den grundläggande hemundervisningen i de fall föräldrarna inte förmådde, dels kunde bibringa den uppväxande generationen de kunskaper och färdigheter som svarade mot den nya tidens krav.

2.3 Folkskolan inrättas

Kampen för och emot en allmän folkundervisning organiserad som skola blev långvarig, och ett enhälligt beslut togs först vid riksdagen 1841. Ett år senare förelåg den första förordningen, 1842 års skolstadga undertecknad av kung Carl Johan Bernadotte. Stadgan föreskrev att det i varje socken och stadsförsamling skulle finnas en, helst fast folkskola, och i varje stiftsstad skulle det inrättas ett seminarium där de som ville ägna sig åt lärarkallet kunde undervisas och få övning i de läroämnen som hörde kallet till. Riksdagsbeslutet och skolstadgan utgjorde inte startskottet för tillkomsten av skolor i Sverige. Sådana hade sedan länge funnits i de sydligaste landskapen som tidigare tillhört Danmark. Beslutet innebar inte heller någon snabb utbyggnad av den lagstadgade folkundervisningen. I praktiken kunde det dröja betydligt längre innan en skolförlagd undervisning kom igång.

2.3.1 *Skolinspektörer som granskare av skolan*

Svårigheterna att få skolstadgan implementerad i praktiken ledde snart till krav både på en tillsyn av att församlingarna fullgjorde sin undervisningsplikt och en central styrning av den allmänna folkundervisningen. Under 1850-talet lades en rad riksdagsmotioner med förslag på en central organiserad inspektion av landets skolor och i juni 1861 förordnades tjugo inspektörer för åren 1861–1863. Endast en av dessa, C. J. Meijerberg som kallats till inspektör i Stockholms stad, hade inspektionsuppdraget som huvudsyssla. För övriga inspektörer var det en bisyssla vid sidan av annan verksamhet, de flesta som präster eller läroverkslärare (Svenska Folkskolans Historia III, s. 13–27).

Enligt instruktionen, som utfärdats av chefen för ecklesiastikdepartementet, skulle det:

äligga varje folkskoleinspektör att inom det honom anvisade området med noggrann uppmärksamhet följa folkskoleundervisningens gång, personligen besöka de inom området befintliga folkskolorna, inhämta kännedom om deras tillstånd och behov samt därom avge berättelse dels till domkapitlet, dels till ecklesiastikdepartementet. (SFS III, s. 14–15)

Inför den andra förordnandeperioden (1864–1866) uttalades i riksdagen önskemål om en årlig inspektion. Önskemålet bifölls och antalet inspektörer utökades nu till tjugoåtta. Endast en av dessa hade uppdraget som inspektör som huvudsyssla för de övriga var det en bisyssla. Fortfarande var prästerna flest i antal (14), därefter läroverkslärare (8) och seminarielärare (3), en konsistorienotarie samt en inspektör som var folkskollärare (SFH III, s. 17). Representationen av präster och läroverkslärare kom senare att ifrågasättas inte minst av folkskollärarna som under slutet av 1800-talet organiserade sig som en målmedveten yrkeskår och tilldelades en mycket aktad position i samhället (SFH III, s. 304–310).

Redan från början fanns en instruktion, utfärdad av chefen för ecklesiastikdepartementet, att följa för den som blev förordnad som inspektör. Under departementschefens ledning samlades inspektörerna också till återkommande möten där de delgav sina erfarenheter och tillsammans diskuterade angelägna frågor. Instruktionen gav dem en gemensam ”mall” att följa vid skolbesöken, och mötena syftade också till att de skulle nå en ”likhet både i åsikter och åtgärder”, (SFH III, s. 15) kanske något liknande det vi idag benämner ”gemensam referensram”.

Inspektörernas berättelser om folkskolorna i riket finns bevarade och är samlade i periodindelade volymer från respektive stift, alltifrån 1861 till och med 1904. Därefter följde årliga inspektörsberättelser från olika kontrakt och inspektörsområden fram till 1950. (Se förteckning i Andersson 1986a). Som exempel på inspektörernas åsikter om brister och vidtagna åtgärder kan vi läsa följande i inspektörsberättelsen för åren 1861–1863 från Härnösands stift:

Undervisningsmaterialen befanns vid en stor del skolor, åtminstone i vissa hänseenden, vara otillräcklig, äfven der den i andra var ganska riklig och väl bibehållen. Hvad som i allmänhet mest saknas, är undervisningsmaterial i naturkunnighet och en passande läsebok i tillräckligt antal exemplar, för ett ändamålsenligt bedrivande af öfningar i innanläsning. (Citat i Andersson 1986a, s. 56)

En beskrivning av hur läsundervisningen kunde bedrivas trots bristen på böcker ges i inspektörsberättelsen för åren 1864–1866:

Barnen ordnas i bänkar och förses med samma slags läseböcker. Ett mindre försigkommet barn sättes vid sidan af ett mer försigkommet, som hjälper det förra att följa med vid läsningen. Dessa två begagna samma bok. Härvid läser nu hvar barn för sig högt, under det de andra tysta följa med i sina böcker. Läsningen sker efter upprop af läraren, med täta ombyten för uppmärksamhetens vidmakthållande. (Citat i Andersson 1986a, s. 57)

I påföljande inspektörsrapport för perioden 1867–1868 nämns den åtgärd som inspektörerna initierat redan vid sitt första gemensamma möte 1862, där de hade framställt en begäran att staten skulle låta utarbета en lämplig läsebok för folkskolan som skulle tryckas i ett sådant antal och kunde inköpas till ett sådant pris att den blev tillgänglig i alla skolor. Några år senare påbörjades arbetet, och den första upplagan av ”Läsebok för folkskolan” utkom 1868. Läseboken kan beskrivas som ett universalläromedel då den inte bara innehöll texter för innanläsning utan också avhjälppte bristen på lämpliga läroböcker i historia, geografi och naturlära.

2.3.2 Inspektörsrapporterna standardiseras

Den av chefen för ecklesiastikdepartementet utfärdade instruktionen och de gemensamma mötena kan ha bidragit till att inspektionerna fått en likartad inriktning och gett inspektörsberättelserna en likartad utformning. Inspektörernas uttalade åsikter om och bedömning av verksamheten i de skolor de besökte liksom de åtgärder som de ansåg vara nödvändiga att vidta torde dock i hög grad vara grundade på deras egna personliga uppfattningar. Längre handlade det om själva skolhuset. I följande exempel från Härnösands stift för åren 1877–1881 redovisar inspektören en kvantitativ bedömning av skolornas standard och

indelar dem därvid i olika kvalitetskategorier: god, medelmåttig, icke tillfredsställande. Nedan ges exempel på hur han preciserade kriterierna:

Af de 119 skollokaler, som vid 1881 års slut voro färdiga, kunna 29 betecknas som goda, 78 medelmåttiga och 11 under medelmåttan. Under rubriken "god" har jag hänfört sådana, som hafva belysning från 2 sidor och ändamålsenliga pulpetbänkar, samt öfrigt äro inrättade i en ungefärlig öfverensstämmelse med normalritningarna. "Medelmåttiga" har jag kallat sådana, som äro byggda på det gamla sättet, så att salen har fönster på 3 väggar, men eljest äro snygga, rymliga och sunda. De flesta sådana hafva långbänkar. På några ställen finns dock pulpeter. De, som icke är tillfredsställande i fråga om snygghet, utrymme eller sundhet, äro hänförda till det tredje slaget. Ventilerna hafva temligen allmänt blifvit anbragta i lärosalarne. På de stationer, der egna lokaler icke finnas, utan skolan måste hållas i enskilda bostäder, är det naturligt, att stora svårigheter måste möta att erhålla lämpliga rum. Allmogem har visserligen i allmänhet ganska välbygdt, men som lärosalar äro rummen icke tjenliga. Ofta äro bänkarna af mycket dålig beskaffenhet. (Citat i Andersson 1986a, s. 80)

Att det fanns variation mellan skolorna är uppenbart, och en mer ingående analys av inspektörsberättelserna skulle sannolikt också visa på att det fanns stora variationer mellan olika stift och kontrakt. De källor som finns att tillgå tyder också på en stor variation både vad det gäller inspektörernas bedömningar och deras sätt att uppträda vid skolbesöken.

2.3.3 *Folkskollärarna kritiserar inspektionen*

De sekundärkällor som vi utnyttjat ger framför allt en bild av den framväxande folkskollärarkårens uppfattningar om inspektörerna. Som tidigare påpekats hade denna kår efter hand fått en allt starkare ställning. Runt om i landet hade lärarna bildat folkskolläraryöreningar, och 1880 slöt sig dessa samman i en riksomfattande organisation – Sveriges Allmänna Folkskolläraryörening (SAF). Den bild som folkskollärarna, i följande exempel, ger av de tidiga inspektörerna måste förstås mot denna bakgrund. Det som folkskollärarna särskilt ifrågasatte var inspektörernas bristande insikt om verksamheten i skolan. I ett uttalande av Fridtjuv Berg (1851–1916), som var folkskollärare, aktiv i SAF och liberal skolpolitiker, vid 1891 års riksdag var: "de flesta strängt taget blott dilettanter på folkskolans område." (SHF III, s. 27).

En av deltagarna vid folkskolläarmötet 1893 gav i sitt anförande följande bild av hur inspektionen i allmänhet gick till ute på landsbygden:

Sedan inspektören anlät till skolan, gör han sig vanligen underrättad om huru mycket som medhunnits under läsåret. Därefter tillsäger han läraren att examinera i något av de genomgångna ämnena. Det plötsliga uppträdandet av en främmande person i skolsalen, där barnen sällan se andra än läraren och varandra, hans myndiga och befallande ton m.m. framkalla hos barnen tank-spriddhet och förvirring. De avgivna svaren bliva därför virriga och osammanhängande, och för att än mer öka oredan inkastar inspektören en eller annan fråga. Efter att ha avhört en eller ett par lektioner begär inspektören uppgift å antal närvarande barn, antal gånger, som skolrådets ledamöter besökt skolan, etc., och så – reser han sin väg. Inom en och annan månad erhåller skolrådets ordförande en promemoria, vari inspektören meddelar, att han fann barnens kunskaper mindre tillfredsställande, att skolan borde förses med ändamålsenliga bänkar, att en rullgardin var söndrig och en annan snett upprullad o. s. v. Promemorian läses och lägges till handlingarna. (SHF III, s. 22–23)

Men det var inte all inspektion som gick till som i detta exempel. Det framgår av följande beskrivning.

Inspektören inträder i skolan vänlig och gemytlig. Han kommer sent på förmiddagen och vill endast höra, huru det är med färdigheten i innanläsning, finner den god, gör sig underrättad om huruvida läraren utfår sin lön ordentligt, huruvida han är belåten med bostaden o. s. v., gör en påhälsning i prästgården och finner skolväsendet i församlingen tillfredsställande. (SHF III, s. 22–23)

En beskrivning av folkskoleinspektören Meijerberg i Stockholms stad ger ytterligare en bild – bilden av en ytterst nitisk utövare av inspektörsämbetet. Meijerbergs anteckningarna i skolans diarium visade att han vid sina besök iakttog varje detalj: ”ett märke efter en snöboll på skolväggen, en smutsig fönsterruta, ej avdammade dörrposter, en oanständig inskrift på vedlåren” (SHF III, s. 57).

Motsättningarna mellan inspektör Meijerberg och lärarna blev också allt mer uttalade och fördes även offentligt i dagspressen. Lärarnas motstånd kan ses som en strid mot den många gånger auktoritära styrningen av undervisningen liksom kontrollen av skolan och den verksamhet som bedrevs. Deras motstånd var också riktad mot den formalism och extrema systematik som

föreskrevs för den grundläggande läsundervisningen och den mekanisering som kommit att karakterisera mycket av arbetet i Stockholms folkskolor under Meijerbergs ledarskap. Exemplet visar att den fråga som även nutidens utvärderare ställs inför – kritiken från dem som granskas och ett ifrågasättande av den externe utvärderarens kännedom om lokala förhållanden, ingalunda är nya frågor.

Vid tiden kring sekelskiftet 1900 började folkskollärarna också att inta rollen som det moderna samhällets kulturarbetare med ett uttalat intresse för vetenskap och lyhördhet för de borgerliga uppfostringsideal som formulerades av ledande kulturradikaler som Ellen Key, författare till essäsamlingen ”Barnets århundrade” som utkom år 1900. Den nya tidens ideal började så småningom sätta sina spår i den pedagogiska praktiken samtidigt som denna praktik blev allt mer reglerad och centralstyrd. Efter mångårig dragkamp mellan kyrkan och staten inrättades slutligen ett centralt ämbetsverk med benämningen – Kungl. Öfverstyrelsen för Rikets Allmänna läroverk år 1905. År 1914 fick även folkskoleväsendet en egen överstyrelse – Kungl. Folkskolöverstyrelsen.

Inspektörerna fortsatte att göra skolbesök och skriva berättelser om skolornas tillstånd i sina inspektörsområden. Som framgått grundade sig deras bedömningar av tillståndet i skolorna i hög grad på deras egna personliga uppfattningar. Med tiden torde de också ha tillägnat sig egen erfarenhetsbaserad norm för bedömning – en kvalitetsnorm byggd på jämförelser mellan olika skolor inom inspektionsområdet och en kategorisering efter kvalitetsmätt som bra, medelmåttiga eller icke tillfredsställande. Som metod för bedömning av skolhusets och undervisningsmaterielens standard kunde inspektionen kanske accepteras, men för en bedömning av genomförda undervisningsinsatser kunde den med rätta ifrågasättas, vilket också framgår av citaten från folkskolläarmötet 1893. Vetenskapliga metoder började nu också allt mer att utprövas för bedömning av enskilda individers mentala förmågor och utvärdering av genomförda undervisningsinsatser, och dessa metoder intresserade även lärarkåren.

3 Undervisning som vetenskap

Under slutet av 1800-talet utvecklades vetenskaper som psykologi och barnpsykologi som en gren av densamma, och under 1900-talets första årtionde inrättades pedagogik som akademisk disciplin. Liksom 1700-talets svenska naturvetenskapliga forskning fick internationell spridning och akademiskt erkännande, så fick den framväxande svenska forskningen i psykologi och pedagogik en internationell orientering och akademisk behörighet under början av 1900-talet. Naturvetenskapernas metodregler och kunskapskrav framstod vid denna tid som ideal, vilket förenklat kan beskrivas så att metoderna skulle ge tillförlitlig kunskap som gick att pröva empiriskt och bedömningar skulle ersättas med mätningar. Förklaringar skulle anges i termer av orsak – verkan, och kunskap uttryckas i lagbundenheter. Forskaren skulle vara objektiv och således inte låta sig styras av utomvetenskapliga värderingar (Wallén 1993).

3.1 Professorer i pedagogik inrättas

Tilltron till vetenskaplig rationalitet och uppfostran grundad på positivt given och empiriskt prövad kunskap, dominerade den inomvetenskapliga utvecklingen. Den tidigare nämnde skolpolitikern Fritjuv Berg lade vid riksdagen 1905, tillsammans med redaktören Emil Hammarlund, fram en motion om professorer i pedagogik. En första sådan inrättades i Uppsala, där Bertil Hammer (1877–1929) tillträdde som professor år 1910. Två år senare blev Axel Herrlin (1870–1937) professor i Lund, och år 1919 tillsattes Gustav Jaederholm (1882–1936) som professor i Göteborg. Bertil Hammer efterträddes år 1931 av Rudolf Anderberg (1892–1955), och sex år senare utsågs David Katz (1884–1953) till professor i Stockholm (Dahllöf 2002).

Den första generationen av professorer ansvarade för såväl pedagogik som psykologi då det sistnämnda ämnet saknades vid universiteten. Denna omständighet liksom professorernas

vetenskapliga skolning i differentiell psykologi och pedagogisk psykometri kom att bli avgörande för den svenska pedagogiska forskningens utveckling under lång tid framåt. Experimentell undersökningsdesign och statistiska analysmetoder blev länge de dominerade inslagen i de studerandes vetenskapliga skolning. När sedan pedagogisk utvärdering, under 1960-talet, etablerades som pedagogisk forskningsspecialitet var detta paradigm det förhärskande, och det blev inledningsvis förebildligt även för de forskare i pedagogik som engagerades för utvärdering av undervisningsprogram. Denna inriktning mot ett positivistiskt forskningsideal var inte unik för vare sig pedagogik eller pedagogisk utvärdering. Det var snarare den huvudfåra i vilken den samhälls- och beteendevetenskapliga forskningen löpte under denna tid (Lindberg & Lindberg, 1983).

3.1.1 Forskning och utvärdering får lärarnas stöd

Intresset för vetenskap kom tidigt till uttryck även inom lärarkåren, och vetenskapens nya landvinningar fick sin spridning vid lärarmöten och genom pedagogiska skrifter. Som exempel kan nämnas ett föredrag som hölls vid det Allmänna folkskollärarmötet 1898 av folkskolläraren och redaktören för tidskriften "Pedagogiska skrifter", J. Franzén, sedermera inspektör och ledamot i folkundervisningskommittén 1906. I sitt föredrag med titeln: "Om barnpsykologiska iakttagelser", konstaterar Franzén att:

Vi förstå nämligen nu tydligare än förr, att den kunskap, vi äga om den vuxna människans själslif, inte utan vidare låter tillämpa sig på barnet, utan vi måste göra just barnet till direkt föremål för våra studier under användning af den moderna vetenskapens metoder. Det är detta senare, som man under det senaste decenniet mångenstädes i landet påbörjat. Jag säger med afsikt påbörjat. Detta arbete, likaväl som hvarje annan vetenskaplig forskning, kräfer nämligen ett rundligt mått af arbetskraft, tid, tålmod och ihärdighet. Här som annorstädes måste den, som lägger hand till verket, vara förberedd på att åter och åter göra nya försök. Äfven oss åligger det att, så långt våra krafter räcka, taga del i detta forskningsarbete, så kärt det är oss att i grund lära känna just de barn, med hvilka var och en af oss har att skaffa. Många tecken tyda ock på att ett lifligt intresse för denna sak håller på att växa fram inom Sveriges folkskollärarkår, att denna kår icke vill hålla sig fjärran från detta för en rationell uppfostran så betydelsefullt arbete. (Citat i Andersson 1986a, s. 83)

I citatet ger Franzén uttryck för en uppfattning som vid slutet av 1800-talet blev allt mer uttalad i den offentliga debatten om utbildning och undervisning. En uppfattning som också vann gehör bland folkskollärarna. De tidigare nämnda Stockholmslärarna opponerade sig mot förhärskade pedagogiska ideal och ställde krav på ökad frihet att själva bestämma hur undervisningen skulle bedrivas. En rationell uppfostran och undervisning var den nya tidens pedagogiska ideal och en sådan uppfostran och undervisning förutsattes bygga på positiva kunskaper vunna med vetenskapliga metoder.

Enligt Franzén, som själv var insatt i den internationella skolforskningen, borde folkskollärarna lära sig de vetenskapliga metoderna så att de kunde undersöka de egna skolbarnen. Som internationellt kända forskare nämndes i detta sammanhang namn som Stanley Hall, Alfred Binet och Herman Ebbinghaus.² Sitt föredrag för de samlade folkskollärarna avslutade Franzén med följande klargörande:

Då jag talat om statistiska metoder, har jag icke därmed sagt, att barnets individualitet skall af läraren förbises. Man insamlar icke statistiskt material, för att man skall behandla alla barn efter en och samma schablon. Genom dessa metoder finner man visserligen likheten hos en del barn, men ock olikheten, ja individuella egendomligheter, som kräfva en speciell pedagogisk behandling, och som lämna oss en säker ledning för barnets behandling enligt dess individualitet. Ännu bättre ledning i detta ämne lämna individualitetsmetoderna. (Citat i Andersson 1986a, s. 84)

Som framgår av citatet skulle de statistiska metoderna dels peka ut det som generellt gällde för grupper av barn som gossar, flickor, åldersklasser, dels peka ut det individuellt avvikande. Individualitetsmetoderna var däremot särskilt användbara för lärarna, när det gällde att söka kunskap om den specifika individen och att anpassa undervisningen till dennes förutsättningar.

² Stanly Hall (1844–1924) amerikansk psykolog och pionjär inom barnpsykologin. Herman Ebbinghaus (1850–1909) tysk psykolog och grundare av minnespsykologin. Alfred Binet (1857–1911) fransk psykolog som utvecklade inteligenstest.

Citaten visar att man inom lärarkåren ville ta del av den framväxande barnpsykologiska forskningen. Med vetenskapliga metoder skulle lärarna själva också kunna söka kunskap om sina elever och anpassa undervisningsinsatserna efter deras individuella förutsättningar. I det sammanhanget kom differentialpsykologin och psykometrin att spela en allt viktigare roll med sina instrument för att mäta individens begåvningsnivå och personliga anlag.

3.1.2 *Intelligensmätning och test*

Den franske psykologen Alfred Binet (1857–1911) betraktas som pionjär inom experimental- och differentialpsykologin, och hans namn förknippas inte minst med experimentella undersökningar av barns intelligens. I samarbete med den franske läkaren Théodore Simon publicerade han en skala för intelligensmätning, i första hand för diagnostisering av begåvningshandikappade. Testmetoden fick snabbt internationell spridning, och i den tidigare nämnda tidskriftsserien ”Pedagogiska skrifter” 1915–1916 kunde läsarna ta del av Binets barnpsykologiska studier i svensk översättning. (NE, www.ne.se).

Med doktorsavhandlingen ”Undersökningar över intelligensmätningarnas teori och praxis”, som var en anpassning och utveckling av Binet och Simons intelligenstest, blev Gustav Axel Jaederholm (1882–1936) svensk pionjär inom testpsykologin. Han disputerade 1914 vid Lunds universitet och blev 1919 professor i filosofi och pedagogik vid Göteborgs högskola. Samma år som avhandlingen lades fram, publicerades också ett arbete om statistiska problem i samband med frågor om normalfördelning och ärftlighet som Jaederholm genomfört i samarbete med Karl Pearson (1857–1936), upphovsman till många statistiska standardmetoder och professor vid University College i London 1885–1933.

Under början av 1920-talet inrättades hjälpklasser i flertalet av landets städer, medan sådana endast undantagsvis anordnades i skoldistriktet i övrigt. De goda erfarenheterna av hjälpklasser, liksom möjligheterna att med intelligenstest göra säkrare

bedömningar av elevernas begåvningsnivå, ökade intresset för ett system med skilda begåvningslinjer eller klasser. Så inrättades exempelvis i Göteborg år 1927 en försöksverksamhet med en uppdelning av eleverna i hjälp-, normal- och elitklasser.

3.1.3 *Två olika referenspunkter*

Ur ett utvärderingshistoriskt perspektiv är det av intresse att se närmare på utvecklingen av test. Här kan två utvecklingslinjer urskiljas, den ena med rötter i differentialpsykologin och den andra med utgångspunkt i en bestämd kunskapsmassa och kunskapsstruktur. När den förstnämnda skolbildningen används för individutvärdering har det inneburit att test konstruerats för att mäta i vilken grad eleven uppnått de mål som operationaliserats. Den enskilde individens testresultat erhålls vanligen genom en jämförelse med andra individers resultat, vilket förutsätter att testet normerats med hjälp av data insamlade från en eller flera referensgrupper. Den enskilde individens testpoäng överförs till en skala i form av percentiler eller standardpoäng, vilka anger hur stor andel av referensgruppen som presterar bättre eller sämre. När det gäller utbildning har frågan om differentiering i skilda begåvningsgrupper och för urval till högre studier varit den centrala och testen därvid tänkta att tjäna som prognostiska instrument.

I den andra skolbildningen har syftet inte varit att särskilja eleverna vad det gäller deras allmänna förmågor utan att utvärdera hur de klarar vissa bestämda uppgifter och kunskapsområden. Kunskapsprov är bundna till vissa pågående eller avslutade undervisningsinsatser, exempelvis ett undervisningsprogram, ett utbildningsprojekt eller en studiekurs. Resultaten är tänkta att kunna ge både information om den faktiska prestationen i ett givet ämne och i viss mån även om individens framtida studieframgång. Då proven i hög grad begränsas till vissa bestämda kunskapselement, exempelvis basfärdigheter, hindras eleven från att visa prov på sin förmåga att nå kunskaper på högre nivåer i kunskapsstrukturen (jfr. Lundgren 1989, s. 81–104). För att gälla som pedagogisk utvärdering är det av betydelse att ett kunskapsprov speglar innehållet på samtliga målnivåer i ett program för uppfostran och undervisning.

Den amerikanske pedagogen och utbildningsforskaren Ralph W. Tyler (1902–1997), som från slutet av 1920-talet publicerat arbeten om utvärdering av studerandes kunskapsresultat och under 1960-talet ledde arbetet med att utveckla ett system för nationell utvärdering, hade bestämda synpunkter på vilken av de två huvudlinjerna inom test- och provutvecklingen som lämpade sig för utvärdering av pågående och slutförda utbildningsinsatser eller undervisningsprogram. I en artikel om programutvärdering uttrycker han sig på följande sätt:

The abstract numbers, percentiles on published tests, standard scores on college admission tests, were developed by psychometricians who sought to translate appraisals of human behaviour into terms used in measuring material objects and physical phenomena. They were not designed to help improve learning, and they have not facilitated the efforts of educators to improve education. The desire to compare student achievement led to the development of single abstract numbers as indices of learning in a field, for example, a single number to report achievement in English, or mathematics, or single for the grade point average. There are several possible learning objectives in every school subject. Rarely does a student attain these several objectives to an equal degree. It is more helpful to report achievements in each important objective rather than to aggregate the student's varied accomplishments into a single score. (Tyler 1991, s. 12)

Det är således olika kunskapsintressen som står i fokus vid dessa två huvudriktningar. Den ena vill veta vad eleven har presterat i förhållande till sin åldersgrupp. Det andra vill snarare veta vad eleven har lärt sig i förhållande till det innehåll som ingått i den aktuella kursen.³

3.1.4 Begreppet ”*evaluering*” etableras

Med en artikel av Tyler, som publicerades i tidskriften ”*Journal of Education*” 1942, fick begreppet ”*evaluation*” internationell

3 Skillnaden illustreras i det tidigare relativa och dagens målrelaterade betygssystemet i den svenska skolan. Det relativa betygssystemet byggde på den matematiska teorin om normalfördelning. Betyg sattes på en femgradig skala med en viss procentuell fördelning för hur stor andel av alla elever som skulle ha ett visst betyg. Systemet kritiserades för att inte säga något om elevernas egentliga kunskaper i ett ämne utan endast i förhållande till andra elevers. I det målrelaterade betygssystemet bedöms elevernas kunskaper i relation till specificerade kunskapskvaliteter i form av mål och kriterier som anges i kursplanerna för ett ämne.

spridning, blev i direktöversättning till svenska "evaluering" och försvenskades till "utvärdering". Uttrycket "utvärdering" infördes i den pedagogisk-psykologiska forskningen först med det reformarbete som inleddes under 1940-talet. Den pedagogisk-psykologiska forskning som därtills bedrivits var en forskning inriktad på att mäta och analysera enskilda individers begåvningsnivå, utveckla psykotekniska anlagsprövningar, utforma experimentella intelligensundersökningar och utpröva standardiserade kunskapsprov. Som exempel på de sistnämnda kan nämnas den testutveckling som inleddes av Carita Hassler-Göransson i början av 1920-talet, när hon konstruerade standardiserade test för mätning av rättstavningsförmågan hos elever i små- och folkskolan.

Förebilderna till dessa svenska mätinstrument hade hämtats från USA och utprövades för svenska förhållanden under professor Jaederholms ledning. Denne gjorde år 1921 en resa till USA för att studera dels amerikansk psykoteknik, dels den praktiska vägledningen av ungdomars yrkesval. Under 1920-talets första år insamlades vid psykologisk-pedagogiska institutionen i Göteborg ett utprövningsmaterial för kunskaps- och färdighetstest i räkning, rättskrivning och gehörsprov (Andersson 1986b). Om det första standardtestet skrev pedagogen Gottfrid Sjöholm år 1931 i tidskriften "Skola och samhälle":

Med rättstavningssskalan har läraren fått i sin hand ett medel att pröva sina skolbarns ståndpunkt i rättstavning, så att han kan se, om de kommit normalt långt, om de stå efter eller till äventyrs kommit före detta stadium, varpå barn i klassens ålder normalt böra befinna sig. Han kan se, om han bedrivit arbetet tillräckligt effektivt, om han måhända har släppt efter på kraven eller om han rent av ställt för stora krav. (Citat i Andersson 1986a, s. 151)

En viktig förutsättning för mätningen av elevernas stavningsförmåga var den fixerade normen. För varje ord som prövades fanns ett rätt svar, en norm som var fastlagd sedan stavningsreformen 1906. Utifrån återkommande standardprov 1923, 1925 och 1934 kunde en påtalig förbättring konstateras i elevernas rättstavningsförmåga. Enligt Hassler-Göransson förklarades de förbättrade resultaten främst av de mer effektiva undervisningsmetoder som numera praktiserades.

Bland dem som efterfrågade vetenskapliga undersökningar återfanns Sveriges Allmänna Folkskolläraryörening (SAF) som år 1941 tillsatte en särskild kommitté – ”Kommittén för SAF:s psykologisk-pedagogiska undersökningar”. Under Fritz Wigforss (1886–1953) vetenskapliga ledning påbörjades år 1942 arbetet med att utarbeta ett skolmognadsprov, det s.k. Kalmar-provet. Till utprövningen knöts en rad småskollärare som efter en introduktionskurs utprövade de aktuella proven i sin egen praktik. På grundval av resultaten utarbetades dels praktiska hjälpmedel med vars hjälp läraren kunde nå kännedom om den enskilde elevens förmåga, dels ett prov för en tidig differentiering. Wigforss underströk vikten av att inte ensidigt förlita sig på testresultaten. Alla prov borde således endast betraktas som hjälpmedel och som komplement till lärarens omdöme och andras bedömningar, t.ex. skolläkarens.

Om än som ett komplement kom skolmognadsproven att tjäna som instrument för urval av främst extremerna i en åldersgrupp, de som var mest respektive minst mogna att delta i nybörjarundervisningen. De förra kunde placeras i en högre klass, medan skolgången kunde uppskjutas eller anordnas på annat sätt för de senare. En sådan möjlighet till differentiering fanns inskriven i den gällande skolstadgan från 1921. En utvärdering av enskilda individers begåvning, mognad, anlag och intressen kunde nu läggas till grund för en rationell utbildningsplanering och organisering av undervisningen.

Det var kring begrepp som elevernas ”differentiering” och individens ”mognad” som tankarna kretsade. Med dagens utvärderingstermer ”input-process-output” kan det uttryckas så; att det gällde att få rätt input, dvs. att söka kunskap om den enskilde individens mognad för att kunna särskilja eleverna och placera ”rätt man på rätt plats”. Vad som var rätt grundades på kunskap om de förmågor, egenskaper eller anlag som individen redan bar med sig och som det gällde att upptäcka med rätt mätinstrument. Det var sedan skolans uppgift att förädla dessa anlag, snarare än att omdana det givna på något genomgripande sätt.

3.2 Rationella undervisningsmetoder

Den psykologisk-pedagogiska forskningen kring undervisningsmetoder skulle ge svar på hur undervisningen skulle bedrivas, och i enlighet med den förhärskande forskningstraditionen gällde även här laboratorieexperimentet som vetenskaplig metod. Som exempel kan nämnas Jon Naeslunds undersökning av läs- och inlärningsmetoder som genomfördes med experimentell design och en undersökningsgrupp som bestod av enäggstvillingar. Tvillingarna delades upp i två undervisningsgrupper som undervisades av en och samma lärare. Undervisningen bedrevs med två olika metoder, den analytiska läsemetoden i ena gruppen och den syntetiska läsemetoden i andra gruppen. Den förra metoden visade sig fungera bäst för de elever som redan hade en viss läsförmåga när de började skolan, medan den syntetiska metoden fungerade bäst för de elever som var nybörjare. Vinsten för den mer avancerade gruppen var dock så liten att Naeslund förordade den syntetiska metoden för samtliga elever. Denna metod kom också att vara den dominerande läsemetoden fram till början av 1970-talet (Naeslund, 1956).

Ser vi till den psykologisk-pedagogiska forskningens utvecklingen under 1900-talets första hälft kan det konstateras att "vetenskaplig ledning" etablerats som vägledande princip inte bara inom utbildningsområdet utan även inom andra samhällsområden. Med vetenskaplig ledning skulle det moderna samhället byggas – ett samhälle kännetecknat av rationalitet och effektivitet. Professor emeritus Torsten Husén beskriver denna tid på följande sätt:

Detta var en period då tron på rationella överväganden som en del av planering av reformer växte sig stark. Rationaliteten bestod inte bara av tron på vad man kunde göra med väl planerade samhälleliga interventioner utan också på värdet av beslutsunderlag levererat av forskningen. (Husén 1994, s. 17)

3.2.1 *Differentieringsfrågan*

Den kunskap och kompetens som byggts upp inom det psykologisk-pedagogiska forskningsområdet möttes vid början av 1940-talet av ökad efterfrågan från olika samhällsintressenter.

En sådan var Svenska arbetsgivarföreningen SAF. Bland efterfrågarna återfanns också 1940 års skolutredning som riktade sig direkt till landets fyra professorer i psykologi och pedagogik för att få vetenskaplig ledning i bl.a. differentieringsfrågan. En strid hade länge stått kring frågan om den optimala tidpunkten för en uppdelning av eleverna i skilda skolformer och på olika utbildningsvägar, en praktisk eller teoretisk. Förespråkarna för en gemensam bottenkola för alla menade att en differentiering skulle uppskjutas så länge som möjligt. Vissa steg mot en gemensam bottenkola togs under 1900-talets första hälft, och av 1946 års Skolkommision ställdes förslaget att de parallella skolformerna skulle ersättas med en för alla elever gemensam skolform upp till 16 år.

Förslaget antogs dock inte i sin helhet vid riksdagen 1950. Beslutet blev istället en kompromiss som innebar att en försöksverksamhet med enhetskola skulle inledas och att en jämförelse skulle göras mellan denna nya skolform och det gamla parallellskolsystemet. Först därefter skulle riksdagen ta ställning i frågan. Resultat från vetenskapliga undersökningar förutsattes här ge ett objektivet underlag för fortsatta beslut.

I Sverige inleddes från slutet av 1950-talet en intensiv satsning på forsknings- och utvecklingsarbete som kan avläsas i ett ökat antal forskningsprojekt och avhandlingar under 1960-talet och 1970-talets första hälft (Andersson, 1986b). Satsningarna på undervisningsprogram och självinstruerande undervisningsmaterial var särskilt stor under 1960-talet, som betraktas som något av en guldålder i svensk pedagogisk forskning (Wallin, 2005). Intresset för undervisningsprogram och självinstruerande undervisningsmaterial kan sättas i samband med genomförandet av den aktuella grundskolereformen som riksdagen fattade beslut om 1962. I reformarbetet inrymdes, som tidigare berörts, tanken om en för alla blivande medborgare gemensam grundskola med en undervisning i åldersmässigt sammanhållna klassgrupperingar så långt som möjligt.

3.2.2 *Individualisering inom klassens ram*

En långtgående individualisering inom klassens ram framhölls som medel och självinstruerande material och programmerad undervisning som praktiskt pedagogiska lösningar. Bakom de självinstruerande materialen och den programmerade undervisningen låg teorier hämtade från USA och undervisningsteknologin. Redan 1954 hade Burrhus. F. Skinner (1904–1990), i en amerikansk vetenskaplig tidskrift, presenterat en artikel med rubriken "Educational Technology". Några år senare påbörjades försök med att utforma undervisningsprogram med eller utan inlärningsmaskiner, vilket innebar: (1) att lärostoffet presenterades i små steg med stigande svårighetsgrad, (2) att studiegången var beroende av det svar som gavs och (3) att metoden såtillvida tillät en långtgående individualisering av inläringen. 1960 besökte Skinner Sverige och höll här seminarier vid några av de pedagogiska institutionerna (Wallin, 2005).

Med beteendeteknologi och rationellt uppbyggda undervisningsprogram skulle det bli möjligt att i allt fler ämnen anpassa undervisningen efter den enskilde elevens förutsättningar och behov. Och med dylika självinstruerande material och program skulle varje elev kunna gå fram i sin egen inläringstakt samtidigt som alla elever i klassen kunde hållas samman och bilda en enhet. Undervisningsteknologin tycktes såtillvida göra det möjligt att i Sverige förverkliga visionen om en demokratisk skola där alla skulle ha möjlighet att utveckla sina anlag så långt det var möjligt utan att skiljas ut ur klassgemenskapen. Denna tilltro till vetenskap som ledning för den pedagogiska praktiken skilde sig inte från den som fanns inom andra sektorer. Även inom arbetsliv, ekonomi och samhällsliv fanns en stark tilltro till vetenskaplig ledning och rationella lösningar för ökad produktivitet och effektivitet.

Med hjälp av forskare i pedagogik överfördes undervisningsteknologin till praktisk tillämpning i den svenska skolan. Det kanske mest uppmärksammade programmet blev här det s.k. IMU-systemet i matematik på högstadiet. Som exempel kan också nämnas mer avgränsade utvecklingsprojekt som "Ett försök med programmering av svensk språklära", "Effekter av självinstru-

erande och svårighetsgraderat läsmaterial” och ”Tyst läsning i hjälpklass – en undervisningsmetodisk studie med ett självinstruerande nivågrupperat läsmaterial.” (Andersson 1986b). Vid slutet av 1960-talet nådde den undervisningsteknologiska vägen också ut i lärarutbildning och lärarfortbildning. En ny läroplan förelåg både för grundskolan (Lgr69) och för gymnasieskolan (Lgy70), och vid introduktionen av dessa båda läroplaner kom både undervisningsplanering och skolans övergripande mål att stå i centrum. I både grundutbildning och fortbildning fick deltagarna utarbeta detaljerade planer med terminalmål och delmål uttryckta i beteendetermer.

3.2.3 *Utvärdering – ett verktyg för flera*

Utvärdering begränsades således inte enbart till vetenskapligt skolade forskare utan förutsattes i hög grad också bli ett verktyg för både byråkrater i skoladministrationen och för praktiker i skolan. Att med Benjamin Blooms taxonomi som förebild utarbeta hierarkiskt strukturerade studiegångar blev det centrala i planeringsarbetet.⁴ Inom Skolöverstyrelsen fanns sedan 1965 också en särskild arbetsenhet för provkonstruktion och utvärderingsfrågor (Skolöverstyrelsen, 1976). Centralt utarbetade måldokument kom senare att också bli en uppgift i det av Skolöverstyrelsen initierade MUT-projektet (Målbestämning och utvärdering) som pågick under åren 1972–1975. Som framgår av titeln skulle dessa dokument ligga till grund för utvärdering av skolans undervisning, men de skulle även ge underlag för målrelaterade prov och betygsättning (Skolöverstyrelsen, 1974). Avsikten var att beskriva alla skolans läroämnen på alla skolnivåer i termer av specificerade mål, men något egentligt resultat kom aldrig att redovisas. Professor emeritus Erik Wallin ger följande beskrivning av händelseförloppet:

In early memos and initial work, the project plans were founded in almost extremely Skinnerian conceptual ideas, demanding objectives stated in behavioural terms. In the course of the work there was a successive shift both in value and knowledge base. The conception of the function of objectives

4 Benjamin Bloom (1913–1999), amerikansk psykolog och pedagog knuten till Chicago University.

became different and accordingly the way of stating them too. The project, however, evoked strong feelings and the question of what it meant and would result in was brought up as an issue before the Swedish Parliament. In connection with this the project came to an end. (Wällin 2005, s. 451)

Att arbetet med att klassificera mål och precisera kriterier för olika betygsnivåer fick avbrytas kan ses som ett uttryck för den mer genomgripande inomvetenskapliga omprövning som vid början av 1970-talet fick genombrutt även i den svenska pedagogiska diskursen. Den grundmurade tilltron till tekniskt rationella lösningar och vetenskapliga metoder som svar på hur skolan skall styras och hur undervisning skall utformas kom nu att utmanas.

3.3 Nya perspektiv på utvärdering

Det man kan konstatera utifrån vår historiska tillbakablick är att utbildning är en av de samhällssektorer i Sverige där det tidigt fanns en stark tilltro till vetenskap och rationella lösningar på pedagogiska problem. Med efterkrigstidens ekonomiska tillväxt blev det möjligt att också genomföra utbildningsreformer som sedan länge varit politiska stridfrågor. Den sociala snedrekryteringen till högre utbildning lyftes fram som det centrala problemet, och en reformering av skolan blev ett politiskt uttalat krav som kunde drivas till beslut när socialdemokraterna, efter riksdagsvalet 1946, kom i regeringsställning. De svenska forskarna i pedagogik engagerades i kommittéer för att utreda frågor om skolsystemet, tidpunkten för en differentiering till olika utbildningsvägar och individuella förutsättningar för yrkesval och studieinriktning (Dahllöf 2002). Den pedagogiska vetenskapen under 1950–60-talet präglades av det som brukar kallas den sociala ”ingenjörskonsten” (Englund, 2004). Forskningen dominerades av statistiska och kvantitativa metoder och förväntades stå till statsmaktens tjänst i utvecklingen av skolan.

3.3.1 *Influenser från USA*

Även om den svenska psykologisk-pedagogiska forskningen redan från början haft en internationell orientering så var det från USA den nu främst kom att hämta nya impulser. "Evaluations for accountability" var i USA en allmän idé för utvärdering av sociala reformer (Dahllöf 1979). I USA började utvärdering som tillämpad vetenskap, eller "Evaluation Research", att expandera i slutet av 1950-talet när stora federala medel avsattes för sociala reformer. I följande citat från förordet till en antologi om utvärdering av utbildning beskrivs den amerikanska expansionsfasen på följande sätt:

However, 1965 was a watershed. President Johnson's "War on Poverty" focused attention on education, and massive funds were made available. And with the money came the demand for accountability, which slipped over into a demand for evaluation. Events moved quickly, and within a few years several academic groups had been established to study evaluation (or "evaluation research"), there were two learned societies (which later merged), and there were several journals. (McLaughlin & Phillips, 1991, s. x)

Den första artikeln i antologin är skriven av den tidigare nämnde Ralph W. Tyler och har titeln. "General Statement in Program Evaluation." Tyler inleder sin artikel med att konstatera att programutvärdering har använts för en rad olika syften men att begreppet "program" i många fall inte definieras. Program är rätt och slätt en etikettering, menar Tyler, och denna etikett betyder olika saker för olika personer som är involverade i programmet. Själv bygger han sina allmänna uttalanden om programutvärdering på erfarenheter från utvärderingar av undervisningsprogram, men dessa erfarenheter kan, menar Tyler, i väsentlig grad generaliseras till att gälla utvärdering även av andra sociala program som exempelvis 1960-talets "Poverty Program of the Great Society" (Ibid, s. 3-17).

Övriga artiklar i antologin är skrivna av amerikanska forskare som alla intagit en ledande ställning inom utvärderingsområdet under den senare delen av 1900-talet. Här återfinns således forskare som Michael Scriven, Robert E. Stake, Marvin C. Alkin, Thomas D. Cook, Robert F. Boruch, Elliott W. Eisner, Henry M. Levin, Carol Hirschon Weiss, Ernest R. House och Daniel

L. Stufflebeam (Ibid s.xiii–xiv.). I artiklarna ser dessa forskare tillbaka på utvecklingen inom utvärderingsområdet i USA under det senaste kvartssseklet. Nämnade forskare har i hög grad influerat svenska forskare som intresserat sig för pedagogisk utvärdering. Detta är i sig inte märkligt då ”social välfärd”, efter andra världskriget, sattes på den politiska agendan i både Sverige och USA.

3.3.2 *Att bedöma värde och nytta*

Utvärdering handlade om att med vetenskapliga metoder göra systematiska bedömningar av program för utbildning och undervisning (education) och uttala sig om värdet av eller nyttan med de undervisningsinsatser som genomförts. Ytterst handlade det om att ta fram ett vetenskapligt säkerställt underlag för politiska beslut och reformer. När forskarna engagerades att besvara dessa politiskt ställda frågor hade de en sedan tidigare grundlagd forskarkompetens som de kom att föra med sig in i det nya problemområdet, åtminstone inledningsvis. Förhållandena i Sverige och USA var här likartade, och i förordet till den tidigare nämnda antologin om utvärdering beskriver redaktörerna dem på följande sätt:

Whenever experts are set a challenging new task, it is natural for them to approach it initially in terms of techniques and concepts that worked well in the past. So it is no surprise that those who undertook evaluations in the period around 1965 first approached their work in terms of experience with social science research and more specifically with psychological, laboratory-style research. The true experiment was the favoured evaluation design: careful statement of objectives, followed by rigorous measurement and subsequent statistical analysis of the result, became the usual *modus operandi*. This general approach was given theoretical credibility by the early work of Ralph Tyler, who approached educational evaluation through a keen interest in the measurement of student gains in learning. (Mc Laughlin & Phillips, 1991, s. x)

Bland forskare och beställare av utvärdering hade denna tekniskt rationella syn på utvärdering som en fråga om experimentell design, mätning av tydliga mål och analyser av resultat med statistiska metoder således hög legitimitet. I slutet av 1960-talet började dock andra röster göra sig hörda, vilka lyfter fram alternativa perspektiv på hur utvärdering och pedagogisk forskning bör bedrivas.

Under 1970-talet och kommande decennier sker stora förändringar i synen på pedagogik som vetenskap. Uppmärksamheten riktas mer mot frågor om processer och hur interaktionen mellan eleven, läraren och samhället ser ut, vilket i sin tur ställer krav på mer kvalitativa metoder och mer kritiska perspektiv på skola och undervisning. Konsensusideologin utmanas av mer radikala teorier som uppmärksammar maktförhållanden, bristande jämlikhet och sociala orättvisor. Dessa förändringar bidrog till att utveckla pedagogiken som en samhällskritisk vetenskap (Englund 2004). Under 1980-talet får också pedagogiken en plats i utbildningspolitikens centrum genom det ökade intresset för didaktik och genom att läroplansfrågan formuleras tydligare som ett politiskt problem. Urvalet av kunskap och hur den görs begriplig för eleverna sätts in i ett socialt sammanhang och existerar inte som eviga sanningar att lära ut. Det som uppmärksammas är att mycket av den förändring som sker ligger i hur vi benämner vår omvärld, med vilka ord och begrepp. Ett tolkande perspektiv och talet om skolan blir minst lika viktigt som att hävda materiella förhållanden som bas för förståelsen av skola och undervisning. Värderingarnas roll blir allt mer central liksom frågan om vem som formulerar kriterier för vad som är god undervisning, varför och för vem, i det som Englund (2004, s. 42) träffande kallar "benämningens politik".

Den betydelse som utvärdering hade fått i samband med efterkrigstidens utbildningsreformer, de pedagogiska forskarnas engagemang i en rad olika utvärderingsuppdrag och den massiva kritik som riktats mot en ensidigt testinriktad programutvärdering, bidrog till att göra pedagogisk utvärdering till mer än en metod (Karlsson 1999). Med den förklaringsinriktade utvärderingsansats som de ledande pedagogiska forskarna förordade blev utvärdering också ett specialområde inom pedagogisk forskning (Kallós, 1981). Som sådant kom utvärdering att beröras av den vetenskapliga omprövning som ledde fram till en allt starkare kritik av den experimentella psykologin och beteendeteknologin som varit huvudinriktning i den psykologisk-pedagogiska forskningen. Vid utvärdering av program för utbildning och undervisning kom det dominerande produktionstänkandet att ifrågasättas och fokus att istället inriktas på processen.

3.3.3 *Koder, ramfaktorer och processer*

Redan vid slutet av 1950-talet hade den amerikanske språkvetaren Noam Chomsky (f.1928) riktat kritik mot Skinner och de behavioristiska förklaringarna av språkliga beteenden. Och denna kritik kom att spela en viktig roll för den teoretiska diskussionen inom psykologi och lingvistik under 1960-talet. Den inomvetenskapliga omprövningen innebar såtillvida ett brott med den dominerande behaviorismen, dvs. ett skifte av förebilder eller paradig. Intresset riktades nu mot Chomskys språk teorier och Jean Piagets (1896–1980) utvecklingsteorier. Vid denna tid uppmärksammades också den engelska språksociologen Basil Bernsteins⁵ teorier om skilda språkliga koder i olika sociala klasser. Dessa teoribildningar gav argument för ett synsätt som satte principerna om individualisering och självverksamhet i centrum på ett nytt sätt. Inläring uppfattades här som en aktiv process, där varje enskild måste få möjlighet att arbeta utifrån sina förutsättningar, på sina villkor och i sin utvecklings-takt.

En svensk studie som senare fick betydelse för den omprövning av programutvärdering som ovan berörts presenterades redan 1967 av professor Urban Dahllöf. Han granskade resultaten av den undersökning av försöksverksamheten med den nioåriga enhetsskola som pågick under åren 1950–1962 (Dahllöf 1967). Den aktuella undersökningen beskrivs av professor Husén på följande sätt:

Den undersökning som Nils-Eric Svensson svarade för var omfattande och ambitiöst upplagd. Den grep ett unikt tillfälle i flykten, nämligen den situation då Stockholms stad i Söderort fick försöksverksamhet med nioårig enhetsskola, medan det så kallade parallellskolsystemet med urval till läroverk och realskola i elvaårsåldern fortfarande bestod en tid i Norrort. Staden fick sålunda två olika skolstrukturer, en selektiv och en sammanhållen. En hel årskull elever följdes från årskurs fyra till årskurs nio. Syftet var att belysa vad som skedde med de begåvade respektive inte så särskilt begåvade eleverna i de bägge differentieringsmiljöerna, ty ”differentiering” var nyckelordet i den debatt om skolans struktur som då fördes. (Husén, 1994, s. 96)

5 Basil Bernstein (1945–2000) språksociolog verksam vid University of London. Se t ex. Barnstein 1975.

Enligt Husén klarade sig de väl begåvade eleverna ungefär lika bra i bägge miljöerna. Med hänvisning till Urban Dahllöfs senare omprövning av Svenssons undersökning påpekar Husén att dessa elever kanske t.o.m. klarar sig något bättre i selektiva miljöer. När det gäller elever från socialgrupp III, eller så kallade arbetarhem, tenderar dessa däremot att klara sig bättre i den odifferentierade miljön än i den differentierade. Urban Dahllöfs omprövning av resultaten i Svenssons undersökning av kunskapsresultaten i olika elevgrupperingar publicerades i slutet av 1967. Dahllöfs reanalys, som hade sin empiriska bas i de kursplaneundersökningar som han tidigare genomfört i samverkan med 1957 års skolberedning, beskriver han på följande sätt:

Här visade en förnyad analys, att de kunskapsprov som använts i den officiella utvärderingen (Svensson 1962) hade låg innehållsvaliditet, genom att de underskattade det innehåll som hade behandlats i de mer segregerade skolorna. En helt annan resultatbild framträdde, när man tog hänsyn till de mellanliggande undervisnings- och inlärningsprocesserna både i form av metodik och tidsåtgång. Avgörande begränsningar utgjordes inte bara av elevgrupperingen utan också av tiden till förfogande enligt timplanen samt frånvaron av en mer långtgående individualisering och läromedel för att underlätta den. (Dahllöf 2002, s. 96–97)

Dahllöfs kritik riktades mot sättet att utvärdera på systemnivå. Han underströk därvid behovet av att även söka information om själva undervisningsprocessen och påvisade vilken betydelse de så kallade ramfaktorerna har för vad som över huvudtaget var möjligt att uppnå med genomförda undervisningsinsatser. För att bli rättvisande som effektivitetsmått måste också kunskapsproven återspegla vad undervisningen faktiskt innehållit. De utvärderingar som genomförts var, enligt Dahllöf, rent deskriptiva och saknade uttalade förklaringar till uppnådda resultat. Att Svensson inte funnit några resultatskillnader i sin tidigare utvärdering men att tydliga skillnader framträdde i den senare omvärdering hade, enligt Dahllöf, sin förklaring:

Eftersom de använda testen endast täckte de mest elementära basfärdigheterna, som dessutom hade lärts ut i lägre årskurser, kom skillnaderna i stället att synas ifråga om tid som behövde ägnas åt olika moment vid tämligen konventionell klassundervisning, med olika grupperingar av eleverna. (Dahllöf 1989, s 19)

Det principiellt viktiga vid en utvärdering på systemnivå är, menar Dahllöf, att så långt som möjligt söka kartlägga och förstå samspelet mellan lärarna och den aktuella elevgruppen i själva undervisningsprocessen och därvid beakta de ramar som begränsar vad som är möjligt att göra och de villkor som gäller för genomförda undervisningsinsatser. När man känner till elevgruppens sammansättning, de ramar som begränsat förutsättningarna och därtill vet var tyngdpunkten legat i den genomförda undervisningen, är det möjligt att förutsäga vilka kunskaper, färdigheter och attityder som är möjliga att uppnå. Det är också lättare att kontrollera i de prov som lärarna själva konstruerar. Förlitar man sig däremot bara på generellt konstruerade test, vars innehållsvaliditet inte är känd, uppkommer istället en känsla av främlingskap. Möjligheterna att förklara utfallet och upptäcka de svaga punkterna i själva undervisningsprocessen går dessutom förlorade, menar Dahllöf (1989).

Det s.k. Cambridge-manifestet från december 1972 är också ett exempel på den kritik som riktades mot utvärdering av traditionella undervisningsprogram och stora sociala reformprogram. Manifestet utformades av ett antal utvärderingsforskare vid ett symposium i Cambridge, däribland britter som Stephen Kemmis, Helen Simons, David Hamilton and Malcolm Parlett och amerikaner som Robert Stake, Louis Smith, Elliot Eisner, Egon Guba, Yvonna Lincoln och Ernest House (Greene 2006). Enligt de forskare som undertecknat manifestet hade programutvärdering på det hela taget inte levt upp till det som förespeglats. Detta dels för att betydelsen av undervisningsprocesser underskattats, dels på grund av att betydelsen av förändringar i de studerandes beteenden hade överskattats. I manifestet kritiserades det då gängse sättet att se utvärdering som begränsad till en jämförelse mellan mål och resultat. Istället ville forskarna lyfta fram det som sker i ”den svarta boxen”, dvs. i klassrummet met och den praktiska verksamheten, för att förstå och förklara vad som leder till det uppnådda resultatet. (För resonemang om praktikbaserad utvärdering, se Karlsson 2003a).

3.3.4 *Teoriinriktad utvärdering*

Hur utvärdering bör bedrivas blev den stora frågan under 1970-talet. Inledningsvis fokuserades valet av metod, och den akademiska striden tycktes i första hand stå mellan försvararna av kvantitativa metoder och förespråkarna för kvalitativa metoder. Men även behovet av en teoretisk tolkningsram efterlystes. I enlighet med vad som hade uttryckts i det ovan nämnda Cambridge-manifestet var de ledande forskarna eniga om att de dominerande modellerna för programutvärdering som hade använts ensidigt fokuserat på förändringar i de studerandes beteenden. Som Urban Dahllöf redan tidigare hade påpekat måste utvärderaren också samla in data som möjliggjorde en analys av "hela sambandet mellan omgivande ramfaktorer, processer och resultat" (Dahllöf 2002, s. 97. Se även Dahllöf 1989). Men för att göra detta förutsattes samtidigt både andra metoder för datainsamling och pedagogisk teori för en vetenskaplig analys.

Vid övergången till 1980-talet introducerade de båda pedagogikprofessorerna Sigbrit Franke och Ulf P. Lundgren en alternativ ansats inom pedagogisk utvärdering som de benämnde "teoriinriktad utvärdering." Utgångspunkten för denna ansats är att varje pedagogisk utvärdering så långt det är möjligt skall utgå från en teori om det objekt eller de fenomen som skall studeras. Författarna menar att utvärderingsarbetet skall bedrivas inom en uttalad teoretisk referensram för utvärderingens objekt, exempelvis ett utbildningsprogram. Utvärderaren, i det här fallet kallad forskaren, skall utforma en teoretisk ram för programmet ifråga som omfattar såväl de kunskaper som finns att tillgå som de uppfattningar och antaganden som programmet är uttryck för. Teorin skall ta sin utgångspunkt i antaganden om utbildningens samhälleliga förankring, och strävan bör vara att "gå bakom det synliga" för att kritiskt granska vad som påverkar olika skeenden och våra uppfattningar om det som sker. Argumentet är att man sällan till fullo kan förstå och förklara sociala och pedagogiska problem på samma nivå som problemen uppfattas spontant.

Det är först med ett makro- eller samhällsteoretiskt perspektiv som det kan bli möjligt att förstå och förklara vad som sker i utbildningen, säger författarna. De strukturella villkoren i samhället antas utgöra begränsningar för utbildningens utformning och innehåll och påverkar skeendet ända ned till klassrumsnivå. Om man skall kunna klargöra hur undervisningsprocessen styrs och utformas är det dessutom nödvändigt att också utreda och beakta de interna villkor som gäller. Författarnas slutsats är därför att den pedagogiska processen måste studeras och analyseras i relation till såväl dess yttre villkor som dess inre förutsättningar (Franke-Wikberg & Lundgren 1980).

Denna starka betoning av att gå utanför den lokala kontexten och att sätta in det som utvärderas, dvs. utbildning och pedagogiska verksamheter, i ett samhälleligt sammanhang betonades allt mer i texter om utvärdering från 1980-talet och framåt. Inte minst kom utvärdering nu att kopplas till och analyseras som en del av den politiska styrningen av skola och undervisning.

4 Utvärdering och skolans styrning

Det går att urskilja olika styrmedel inom offentlig verksamhet. Å ena sidan kan man tala om "direkt styrning" i form av order, lagar, förordningar, riktlinjer, mål- och resultatstyrning, etc. Å andra sidan kan man tala om "indirekt styrning" för stöd till lokal utveckling i form av kombination kunskapsspridning, exempelvis som ett resultat av utvärdering och forskning. Enligt Pettersson och Wallin (1995) har statens styrning av skolan under de senaste decennierna förskjutits från direkta styrmedel, till en ökad användning av indirekta och informativa styrmedel. Utvärdering spelar i det sammanhanget en viktig roll som information på alla nivåer i systemet (staten, kommunerna, skolledningen och lärarna) och kan påverka såväl verksamhetens innehåll som metoder. Utvärdering stannar därmed inte vid att enbart vara en systematisk bedömning och uttala värdeomdömen, utan tjänar också som instrument för att leda skolan. Som vi inledningsvis klargjort betraktar vi utvärdering som instrument för politisk styrning och kontroll och tar i den följande framställning vår utgångspunkt i den förskjutning mot informativa och indirekta styrmedel som Pettersson och Wallin pekar på.

4.1 Minskad central styrning

Vid slutet av 1960-talet framstod det allt tydligare att den nio-åriga grundskolan inte hade uppnått de mål som den förknippats med. Den programmerade undervisningen hade inte hållit vad den lovat, och läget i grundskolan framställdes närmast som kaotiskt i de skolmotioner som förelåg vid vårriksdagen 1970. Enligt statsutskottet borde problemen utredas, och riksdagen fattade i enlighet med utskottets förslag ett enhälligt beslut om att tillsätta en statlig utredning, vilken de sakkunniga gav benämningen: "SIA – Utredningen om Skolans inre arbete" (SOU 1974:53).

I direktiven till SIA-utredningen framhåller föredragande statsrådet nödvändigheten av att rikta resurser mot specifika behov

och inte att som hittills söka generella lösningar på praktisk pedagogiska problem. Generella reformer kunde nämligen missgynna de elever som inte förmådde utnyttja de förbättrade möjligheterna. I det slutbetänkande som utredningen redovisade 1974 bekräftades också att generella undervisningsmetoder, färdigställda undervisningsprogram och programmerade läromedel inte hade förmått möta åtskilliga elevers behov av en varierad undervisningsmetodik. Och de pedagogiska lösningar som prövats hade inte heller kunnat tillgodose dessa elevers behov av en mer personlig kontakt med lärarna.

I stället för nationellt påbjudna skolreformer, centralt utformade lösningar för den pedagogiska praktiken och en centralt reglerad resursfördelning, skulle skolan nu anpassas till de lokala förhållandena. Den lokala skolans ansvarsområde skulle vidgas genom att integreras med den kommunala fritidsverksamheten. Därtill skulle en friare resursanvändning gälla för hela det lokala skolväsendet. SIA-utredningens och den efterföljande propositionens förslag kom dock att till stor del "läggas på is." Vid valet hösten 1976 förlorade socialdemokraterna sin mångåriga regeringsställning, och i deras ställe inträdde en borgerlig trepartiregering med moderaterna, folkpartiet och centern. Under den borgerliga regeringstiden, som sträckte sig över två valperioder, ändrades sammansättningen vid tre tillfällen. Efter valet hösten 1982 intog socialdemokraterna åter regeringsställning.

Förändringen av den nationella läroplanen för grundskolan, som gjordes efter regeringsskiftet 1976, innebar inget egentligt brott med SIA-utredningens förslag. Den proposition, som den borgerliga regeringen presenterade 1978/79, anknyter snarare till den tidigare utredningen – en anknytning som skolminister Birgit Rodhe klargör i sitt tal inför riksdagen i början av juni 1979. Med hänvisning till att varje förändring av grundskolan berör över en miljon barn och ungdomar konstaterar hon att:

Vägledande måste därför vara vad som är bra och angeläget för barnen och de unga, vilka förändringar som är väsentliga för deras skull. SIA-reformen hade den utgångspunkten, och läroplansförslaget är ett led i en vidareutveckling av denna. (...) I den mån problemen kan lösas inom skolan måste det ske genom träget, konstruktivt vardagsarbete inom varje skola i vårt land. Denna läroplan med dess

krav på grundläggande kunskaper och färdigheter, dess möjligheter till friare studieval, är, det är min övertygelse, en god grund för det konstruktiva vardagsarbetet i skolan. (Rodhe 1997, s. 99)

Den aktuella läroplanen – Lgr80 – består av två delar. I den första delen, som inrymmer förordningstexten, återfinns mål och riktlinjer samt timplaner och kursplaner. Den andra delen utgörs av kommentarmaterial – ett material som skulle tjäna som en del i den löpande förnyelse av skolan som staten avsett. Det var också tänkt att inspirera personalen vid utformningen av den lokala arbetsplan som varje rektorsområde, enligt förordningen, skulle utarbeta. Ett av de kommentarmaterial som Skolöverstyrelsen publicerade har titeln ”Utvärdering i klass, arbetsenhet, rektorsområde och kommun” (Skolöverstyrelsen, 1983).

Den nationella läroplanen lade fast mål och riktlinjer för verksamheten i skolan samtidigt som den gav ett friutrymme för en lokal anpassning. SIA-utredningen och Lgr80 innebar såtillvida en förskjutning vad det gäller styrningen av skolan, dvs. en viss decentralisering av befogenheter och ansvar. I detta första steg överfördes ansvaret för det lokala utvecklingsarbetet och lärarnas fortbildning från centrala politiska organ till lokala. Under rubriken ”Lokalt ansvar och utvärdering” skriver man i kommentarmaterialet:

Kraven på utvärdering – vilka som skall göra den och vad den skall riktas in på – förändras när den centrala styrningen minskar och den lokala friheten att forma skolverksamheten ökar. Den enskilda skolan skall inte längre bara så nära som möjligt följa regler som finns. Det är i stället angeläget att använda de mer allmänt hållna reglerna på det sätt som man själv tycker är bäst – med hänsyn till skolans mål och de förutsättningar och behov som den egna skolan har. (...) Det är med andra ord nödvändigt att man i varje skola fortlöpande arbetar med utvärdering och att man gör utvärdering själv. För den enskilda skolan finns därigenom utrymme att successivt pröva sig fram till vilka lösningar som fungerar bäst. (Skolöverstyrelsen, 1983, s. 15–16)

Att utvärdera själv i klass, arbetsenhet, rektorsområde och kommun var av annan karaktär än den utvärdering av reformer av skolan i stort som Skolöverstyrelsen och länskolnämnder skulle ansvara för. De stora organisatoriska reformerna var slutförda

efter införandet av den allmänna nioåriga grundskolan och den organisatoriskt sammanhållna gymnasieskolan. Ett första steg mot en reformering av den statliga skoladministrationen hade tagits i samband med läroplansreformen, Lgr80. Ansvaret för skolutveckling och lärarfortbildning hade därmed överförts till kommuner och skolor lokalt.

4.1.1 Lokalt utvecklingsarbete – motorn i reformarbetet

En reformering av det inre arbetet i enlighet med SIA-utredningens förslag återstod dock att genomföra och dessa reformsträvanden aktualiserades när socialdemokraterna åter fick bilda regering efter valet hösten 1982. Med denna socialdemokratiska regering tog man på nytt upp reformeringen av skolans inre arbete, men denna gång utifrån en helt ny strategi. Den tidigare strategin kan beskrivas som ett flerårigt utredningsarbete, omfattande försöksverksamhet, implementering via lärarutbildning och lärarfortbildning och utvärdering i efterhand. (Det kanske vore mer korrekt att tala om uppföljning av reformernas implementering än om utvärdering eftersom det knappast handlade om att ompröva några beslut, snarare om att kontrollera att de genomfördes). Istället avsattes nu särskilda medel för att stödja och stimulera det lokala utvecklingsarbetet. Den första särskilda satsningen tillkom våren 1983 när riksdagen beslutade om medel, motsvarande 30 miljoner kronor för lokal skolutveckling på grundskolans lågstadium under en fyraårsperiod. Under det sista året (1986/87) kunde de särskilda medlen användas även för åtgärder på mellanstadiet, vilket därefter tilldelades särskilda medel för lokal skolutveckling under ytterligare två år. Under senare delen av 1980-talet blev det högstadiets tur att bli föremål för en särskild statlig satsning. Samma strategi brukades också för en reformering av gymnasieskolan i samband med den försöksverksamhet som genomfördes i denna skolform under åren 1985–1989.

Nämnda, politiskt beslutade, satsningar utgör alla en ekonomisk resurs som för en begränsad tidsperiod "öronmärks" för en bestämd nivå eller stadium i skolsystemet och för en viss form av verksamhet – lokal skolutveckling. Detta projektorganiserade utvecklingsarbete skiljer sig såtillvida från det lokala utvecklings-

arbete som förespråkas i SÖ:s kommentarmaterial. Det utvecklingsarbete som initieras och genomförs lokalt syftar inte till att ge generella lösningar som kan överföras till och tillämpas vid andra skolor i riket. Som framgår av citatet ovan från kommentarmaterialet är tanken snarare att ett lokalt utvecklingsarbete skall utgå från de behov och problem som finns i den lokala skolans verksamhet. Det skall inriktas på att lösa dessa och lokalt finna vägar att nå de nationellt formulerade mål som anges i de aktuella läroplanerna och kursplanerna.

De inriktningar som föreslås eller prioriteras i samband med de särskilda satsningarna kan däremot ses som centralt initierade förslag till lösningar på problem som uppmärksammat på nationell nivå eller angetts i tidigare reformförslag, i detta fall SIA-utredningen och den efterföljande regeringspropositionen 1976. I direktiven för den s.k. lågstadiesatsningen anges följande förslag på inriktningar: (1) att förbättra samverkan mellan förskola och skola genom bl.a. fortbildningsinsatser, (2) att öka kontakterna med föräldrarna för att förbättra samverkan mellan hem och skola och (3) att utveckla arbetssättet på lågstadiet genom t.ex. försök med årskurslös undervisning (Andersson, 1996).

I direktiven för mellanstadiesatsningen anges att projekt med inriktning på undervisningen i matematik och svenska bör prioriteras, och för högstadiesatsningen föreslås att utvecklingsprojekten skall inriktas på elevaktiva arbetsformer och arbetssätt samt på samverkan mellan olika ämnen, i första hand inom de samhällsorienterande- och naturorienterande ämnesblocken. Det program som Skolöverstyrelsen utarbetade för försöksverksamheten på gymnasieskolan avgränsade följande huvudområden: studieorganisationen, den inre arbetsmiljön, praktik och samarbete mellan gymnasieskola och vuxenutbildning (Andersson & Åsemar 1990).

4.1.2 Från reform till programutvärdering

Det bör understrykas att den strategi för skolans styrning, som tas i bruk under 1980-talet, innebär en viktig förändring i utvärderingens roll. Från att tidigare främst ha varit ett instrument för utvärdering av nationella skolreformer, skall utvärdering

nu brukas för bedömning av lokal skolutveckling – något som närmast kan beskrivas som programutvärdering i mindre skala. Utvärdering förutsätts med andra ord ge besked om insatsen fungerar som den var tänkt, om den skall fortsätta, förändras eller läggas ned. Uppgiften blir på så sätt mer lik den form av programutvärdering som beskrivs i den amerikanska litteraturen.

Vid samtliga nämnda satsningar genomfördes också externa utvärderingar, vars resultat redovisats i ett stort antal rapporter. I detta sammanhang skall vi inte närmare redogöra för dessa utvärderingar utan följa de två huvudlinjer för utvärdering av lokal skolutveckling som kan urskiljas i det som ovan redovisats. Den ena gäller den interna utvärdering som man på den lokala skolan själv förutsätts göra av sin verksamhet. Den andra innefattar externa utvärderingar som uppdragits till forskare vid pedagogiska institutioner att genomföra. I det följande skall vi koncentrera oss på de senare och återkommer till den lokala självvärderingen längre fram i detta kapitel.

4.1.3 Försök med teoriinriktad utvärdering

I den externa utvärdering av Lågstadiesatsningen gjordes försök att tillämpa en utvärderingsansats där processen följdes i två steg under de fyra år som satsningen pågick. I det första steget gjordes varje år en kartläggning av innehållet i de projekt som tilldelats medel och av de principer som länskolnämnderna tillämpat vid fördelningen av de särskilda medel som riksdagen beslutat. I ett andra steg gjordes ett urval av projekt för direkta fältstudier där data om processen dokumenterades med olika tekniker som observationer, strukturerade intervjuer, informella samtal och skriftlig information via dagböcker som de projektansvariga förde under projektarbetets gång (Andersson 1992; Rönnerman 1993). I samband med försöksverksamheten i gymnasieskolan genomfördes på liknande sätt fallstudier vid ett urval av skolor i landet (Andersson & Åsemar, 1990).

I enlighet med den teoriinriktade utvärderingsansats som tillämpades i dessa studier var syftet med den externa utvärderingen att söka förstå och förklara varför verksamheten i ett lokalt projekt utvecklades som den gjorde och varför resultaten blev som

de blev. Med denna ansats är det, som tidigare framhållits, nödvändigt att inte bara söka information om genomförandeprocesserna utan även om förutsättningarna, både interna och externa, vid de skolor som valts ut för fältstudier. Projektverksamheten lokalt måste i sin tur förstås i ett vidare socialt sammanhang och effekterna av den särskilda satsningen som helhet analyseras och tolkas utifrån den redovisade teoretiska referensramen.

Att följa processer är tidskrävande och direktstudierna i den pedagogiska praktiken begränsades därför till ett antal skolprojekt och försöksskolor där projektverksamheten följdes, med hjälp av ovan nämnda metoder, under den period som den särskilda satsningen pågick. På så sätt skulle det vara möjligt att som forskare generera kunskap om det objekt som var föremål för utvärdering. Men alla utvärderingsuppdrag är tidsbegränsade och som utvärderare förväntas man redovisa sina resultat inom den avsatta tidsramen. Det kan göra att tiden för ett systematiskt kunskapsbyggande inom valda problemområden kan vara begränsad och en fortsatt kunskapsutveckling dessutom svår att driva vidare i nya utvärderingsuppdrag med andra utvärderingsobjekt. Särskilda satsningar är däremot en strategi som regering och riksdag fortsatt att bruka för att stödja och stimulera lokal skolutveckling, och denna indirekta strategi kan ses som en ny form av centralstyrning – en indirekt reformstrategi som med ett internationellt uttryck har fått benämningen ”back-door centralism.”

Med grundskolereformen 1962 och gymnasieskolereformen 1968 hade de stora organisatoriska förändringarna av skolsystemet avslutats och intresset riktades mot en reformering av skolans inre arbete i såväl grund- som gymnasieskolan. Som framgått var problematiken nu en annan. Medan reformeringen av det svenska skolsystemet hade grundats på en stark tilltro till den vetenskapliga expertisen och generella pedagogiska lösningar vändes tilltron nu till lokala lösningar och lokalt inflytande. Den starkt centraliserade skolan skulle söka sina former och metoder i den lokala pedagogiska praktiken, och decentraliseringsfrågan aktualiserades därmed utifrån ett lokalt perspektiv, samtidigt som den parades med en fortsatt politisk ambition att central-

styra den svenska skolan. Det första steget i en reformering av den statliga skolförvaltningen förordades, som framgick, av SIA-utredningen och genomfördes med den efterföljande läroplansreformen (Lgr80). I detta steg i förvaltningsreformen överfördes befogenheter och ansvar för det lokala utvecklingsarbetet och lärarnas fortbildning från central nivå till lokal. För gymnasieskolans del handlade det också om att anpassa utbildningarnas inriktning och innehåll till lokala icke offentliga aktörers krav och intressen. (Jfr Lundgren 1989).

Det avslutande steget i förvaltningsreformen togs vid övergången till 1990-talet och detta reformsteg präglas liksom det som togs vid övergången till 1980-talet av den förändrade reformstrategi som tidigare berörts – en strategi som förenklat kan beskrivas som en förändring inifrån och nedifrån. I budgetpropositionen 1988/89 motiverar utbildningsminister Bengt Göransson den förändrade reformstrategin på följande sätt:

Den traditionella metoden att förändra skolan, med stora reformer som beslutas centralt och som genomförs överallt samtidigt, fungerar enligt min mening mindre väl när det gäller att åstadkomma förändringar av arbetssätt och innehåll i skolan. Sådana förändringar är svåra att åstadkomma genom centrala beslut. Skall vi kunna få till stånd en utveckling i detta avseende måste tyngdpunkten för förändringsarbetet ligga lokalt. (Citat i Andersson 1996, s. 32)

I samma budgetproposition framhålls att en minskad detaljstyrning och en ökad frihet för dem som arbetar i skolan kan vara en viktig åtgärd för att åstadkomma ett mer positivt utvecklingsklimat i skolan och en större förändringsberedskap. Den lokala utveckling som förväntades komma till stånd förutsattes samtidigt ligga i linje med de nationellt formulerade målen.

4.1.4 Decentralisering, målstyrning och utvärdering

Den bild vi hittills tecknat visar att skolan som institution för social och kulturell reproduktion alltid har varit politiskt styrd och kontrollerad. I likhet med Lindensjö och Lundgren (2000) har vi framhållit vikten av ett historiskt perspektiv, och vi anknyter här till deras resonemang om olika former av styrning. Med utgångspunkt i Dahllöfs ramfaktormodell, som vi tidigare berört, skriver dessa författare:

Grovt kan vi ange tre slag av eller ”verktyg” för styrning. Ett är juridisk reglering. Ett annat verktyg är att styra med resurser, dvs. en ekonomisk styrning, och en tredje är att styra ideologiskt, dvs. att styra mål, innehåll och resultat. Dessa tre former för styrning är naturligtvis relaterade till varandra. Ekonomisk liksom ideologisk styrning kräver en juridisk reglering. (Lindensjö & Lundgren 2000, s. 25–26)

I detta sammanhang används begreppet målstyrning – ett begrepp som hör samman med den aktuella förvaltningsreformen. Förskjutningen av befogenheter och ansvar från centrala politiska organ till lokala brukar allmänt beskrivas som en övergång från regelstyrning till målstyrning. I detta sammanhang bör det dock påpekas att de olika slagen av statlig styrning – ekonomisk, juridisk, ideologisk – dvs. pengar, lagar och mål, inte är varandra helt uteslutande. Samtliga tre styrinstrument förekommer parallellt över tid men får olika tyngdpunkt i skilda tider. (Jfr. Marklund 1984. Se även Du Rietz, m.fl. 1987).

Under 1990-talet kom uttrycket mål- och resultatstyrning mer allmänt att brukas för den ansvarsfördelning mellan stat och kommun som riksdagen beslutade 1991. Beslutet innebar att kommunerna som arbetsgivare nu tilldelades det fulla huvudmannskapet för skolan och dess personal. Reformeringen av den statliga skoladministrationen innebar samtidigt att Skolöverstyrelsen, Statens Institut för Läromedelsgranskning (SIL), Länskolnämnderna och Fortbildningsnämnderna avvecklades, och ett nytt statligt ämbetsverk för skolväsendet – Skolverket – inrättades. Verkets huvudsakliga uppgift skulle vara att utveckla skolan och att svara för uppföljning och utvärdering av verksamheten i den lokala skolverksamheten. Medan reformeringen av den statliga skoladministrationen å ena sidan innebar att kommunerna och de lokala skolorna fick ett större friutrymme att själva utforma den pedagogiska praktiken inom givna ramar, så innebar förvaltningsreformen å andra sidan en ökad statlig kontroll genom utvärdering. Professor Ulf P. Lundgren, som deltagit i den s.k. styrberedningen, beskriver läget vid övergången till det förändrade styrsystemet på följande sätt:

Utvärdering torde vara ett av de starkaste instrumenten för styrning av utbildning. Om en central utvärdering genomförs i ett system som förändras mot ökad decentralisering kan detta innebära att en ny form av central styrning

kommer att införas. Om ansvarsfördelningen mellan politiker och professionella är diffus – vilken den nu är – finns risken att decentraliseringen kommer att möjliggöra en marginell lokal anpassning av resurser till en utbildning som innehållsmässigt har en stark central styrning. (Lundgren 1989, s. 90)

Citatet kan ses som en hypotes inför övergången till 1990-talet. Vad utvärdering faktiskt skulle komma att bli efter förvaltningsreformen var vid denna tid en öppen fråga och kan även nu, cirka femton år efter denna reform, vara svår att klart urskilja. Utvärdering som instrument för politisk styrning kommer givetvis alltid att vara föränderlig och ta gestalt på olika sätt i olika tider. Olika former av utvärdering kommer också att finnas samtidigt, vilket gör att utvärdering som företeelse inte entydigt kan beskrivas eller definieras. I det följande skall vi försöka att teckna konturerna av det förlopp som kan urskiljas efter den s.k. kommunaliseringen 1991 och väljer då att i vår framställning särskilja utvärdering på tre nivåer: den nationella, den lokala och den internationella nivån. Framställningen kommer att till vissa delar att knyta an till tidigare avsnitt.

4.2 Utvärdering på nationell nivå

Som tidigare framhållits fick Skolöverstyrelsen redan den 1 juli 1965 en särskild arbetsenhet för provkonstruktion och utvärderingsfrågor. Inom projektet ”Mål och utvärdering” (MUT) gjordes under början av 1970-talet också ett försök att utveckla ett mer generellt system för mål- och utvärdering i skolans alla ämnen, men detta försök fick som framgång, läggas ned. I en skrift om Skolöverstyrelsens pedagogiska forsknings- och utvecklingsarbete från 1976 återfinns ett avsnitt med rubriken ”Utvärdering” där man skriver om utvärderingens uppgifter att:

Under senare år har allt oftare uttalats önskemål att skolan och skolreformerna skall utvärderas från mer allmänna och utbildningspolitiska utgångspunkter. En sådan utvärdering kan inte begränsas till elevresultat i enskilda ämnen utan måste avse mer övergripande utbildningsresultat. Har den nya skolan givit de unga en god förberedelse för rollen som samhällsmedborgare och yrkesutövare? Har utbildningen främjat jämlikhet och samverkan? etc. (Skolöverstyrelsen 1976, s. 41-44)

4.2.1 *Fem professorer kritiserar utvärderingen*

Trots dessa påpekande framgår det av den aktuella texten att konstruktion av prov av olika slag och för olika ämnesområden vid mitten av 1970-talet alltjämt dominerade Skolöverstyrelsens utvärderingsinsatser. Drygt tio år senare och inför det andra steget i förvaltningsreformen, som förväntades innebära att utvärdering fick en ökad betydelse, kvarstod denna betoning av kunskapsprov som metod av särskild vikt för bedömningen av den lokala skolans insatser. Fem professorer i pedagogik samlade sig nu till en gemensam kritik av det program för nationell utvärdering av skola och vuxenutbildning som Skolöverstyrelsen presenterat hösten 1987 (Skolöverstyrelsen 1987-09-11). I förordet till den skrift som professorerna publicerade på eget initiativ fastslås att ett program för nationell utvärdering förvisso är nödvändigt, att kunskaps- och färdighetsprov förvisso utgör en självklar del i en nationell utvärdering och att en utveckling av instrument för prövning av kunskaper och färdigheter därför är högst lovvärd, men den utformning som Skolöverstyrelsen presenterat har sina brister, framhåller professorerna:

Vi uppfattar dock att resultatmätningar på elevnivå tilldelas alltför stor tyngd i förhållande till kunskaper om undervisningens ramar och process och att det i Skolöverstyrelsens program knappast klargörs hur data skall analyseras och tolkas och hur och till vad de skall användas. Allvarlig kritik mot hur Skolöverstyrelsen har utformat programmet har också framförts av forskare som är direkt involverade i den nationella utvärderingen, t ex för att angivna syften är otydliga och för att intresset för hur resultaten skall återkopplas till lokal nivå är ytterst svagt. Sammantaget tvivlar vi som medverkar i denna skrift på att utvärderingen såsom den skisserats ger underlag för att värdera skolan och för att fatta väl underbyggda beslut om hur den skall förändras. (Franke-Wikberg 1989, s. 8)

Den insikt om utvärdering som forskare i pedagogik utvecklat och framfört under 1960- och 1970-talen kommer till uttryck i professorerna skrift: Och när Skolverket startade sitt utvärderingsarbete hösten 1991 kom mycket av denna insikt att också att få genomslag i de utvärderingar som Skolverkets tjänstemän genomförde under 1990-talets första hälft. Utvärdering för kunskapsuppyggnad stod uppenbart i förgrunden framför utvärde-

ring av måluppfyllelse och kontroll. I ett internt arbetsmaterial, där en översikt av Skolverkets utvärderingsprogram presenterats, formuleras det på följande sätt:

Utvärdering inom ramarna för Skolverkets arbete innebär fördjupad analys inom valda problemområden. Den sker inte främst som en fortlöpande verksamhet, utan områden för utvärdering väljs därför att analys visar att fördjupad kunskap är önskvärd, utifrån resultat av uppföljning eller utifrån andra signaler om att problem föreligger. Utvärdering, syftar till kunskap som innebär förståelse av, ökad insikt i och förklaringar av olika förhållanden i skolverksamheten. Den är ytterst relaterad till de mål och grundläggande principer som gäller för verksamheten och har det dubbla motivet att utgöra kontroll och ge underlag för utveckling. (Skolverket 1993-09-19, s. 1)

Parallellt med Skolverkets utvärderingsarbete, som avsåg att dels ge information riktad till huvudmän, dels ge underlag för resultatredovisning riktad till riksdag och regering, fortsatte de särskilda satsningarna på tidsbegränsade utvecklingsprojekt. Under den borgerliga regeringsperioden 1992–1995 genomfördes en satsning på åtgärder mot främlingsfientlighet och rasism under ledning av fyra statliga myndigheter. Myndigheternas uppdrag, att fördela pengar till utvecklingsprojekt och därvid söka samordna sina insatser, övergick efter ett par år till en nationell kommitté. Regeringsuppdraget var nu att leda arbetet med den ungdomskampanj som pågick under åren 1994–1996 och som avsåg att bidra till ungdomars förståelse för ett mångkulturellt Europa (Andersson, 1995). Under den borgerliga regeringsperioden tillkom Stiftelsen för Kunskap- och Kompetensutveckling (KK-stiftelsen) som initierade en omfattande skolsatsning i syfte att stödja och stimulera utvecklingen av informationsteknologi i grund- och gymnasieskolans undervisning (Andersson & Grensjö, 2002).

KK-stiftelsens skolsatsning, som avslutades vid slutet av 1990-talet, kompletterades med en omfattande nationell satsning på lärarnas kompetensutveckling. Det så kallade ITiS-program-

met,⁶ som drevs av en delegation inom utbildningsdepartementet, pågick under åren 1999–2002. Redan hösten 1997 riktade sig departementet och dåvarande statsrådet Ylva Johansson också direkt till landets skolor med en inbjudan att söka medel för utvecklingsprojekt inriktade på att utveckla lärarnas yrkesroll och finna nya former för samverkan med universitet/högskola och inom ramen för regionala utvecklingscentra. Denna satsning har fått sin förlängning i ett nationellt projekt med benämningen ”Attraktiv skola” som inleddes 2001 och avslutades våren 2006. Om detta projekt skriver den socialdemokratiska regeringen i sin utvecklingsplan 2001/02:

Syftet med skolutvecklingsprojektet Attraktiv Skola, som sedan 1998 drivs av Lärarförbundet, Lärarnas Riksförbund, Sveriges skollärförbund, Svenska Kommunförbundet, Utbildningsdepartementet och Skolverket i samverkan, är att utveckla skolan mot högre kvalitet och göra skolan till en attraktiv arbetsplats. Ett viktigt mål för projektet Attraktiv Skola är att utveckla karriärtjänster. Projektet Attraktiv Skola ger lärare särskilda möjligheter att påverka skolutvecklingen just i sin kommun. (Regeringens skrivelse 2001/02:188, s. 61)

4.2.2 *Det goda exemplets strategi*

Liksom de särskilda satsningar som genomfördes under 1980-talet har de ovan nämnda satsningarna varit föremål för externa utvärderingar. Under 1990-talet blev skriftlig dokumentation och informationsspridning ett mer uttalat krav som ställdes på de aktörer som tilldelats medel för lokala utvecklingsprojekt (Andersson, 1995). Vad som utvecklats i skolprojekt vid ett antal lokala skolor och kommuner skulle ges spridning och komma alla andra kommuner och skolor i landet till godo. ”Goda exempel” blev snart ett nyckelbegrepp som användes i allt fler sammanhang. Exempelen var tänkta att bilda modeller att efterlikna eller att jämföra sig mot och tjänade såtillvida som ”måttstock”, dvs. ett standardmått för utvärdering eller bench-marking som är ett vanligt uttryck idag.

6 ITiS Delegation for ICT in Schools. Stockholm: Ministry of Education and Science.

Utvärderingarna av de särskilda satsningarna och de många skolprojekt som genomförts i anslutning till dessa har pågått vid sidan av den uppföljning, tillsyn och utvärderingar som Skolverket samtidigt har haft uppdraget att genomföra. Hur denna mångfald av informationskanaler har påverkat den fördjupade analys och strävan till kunskapsuppbyggnad, som kommer till uttryck i citatet från Skolverkets utvärderingsprogram, saknar vi uppgifter om. Det vi kan konstatera är dock att Skolverket sedan tillkomsten 1991 har omorganiserats i flera omgångar och att utvärderingsprogrammets kunskapsbyggande funktion har tonats ned medan verkets tillsyns- och kvalitetsgranskande regeringsuppdrag har markerats allt starkare.

I det nya Skolverkets huvuduppgift ingår, som framgått, såväl nationell som internationell utvärdering. Båda innebär fördjupade studier och omfattande analyser som syftar till att beskriva, förklara och värdera olika företeelser inom barnomsorg och skola. Ett kunskapsintresse kommer således till uttryck, men detta intresse riktas mot att lösa problem, tillgodose behov och ge underlag till förbättringar. För en mer kritiskt granskande utvärderingsansats i syfte att generera kunskap om villkoren för förändring, som är en förutsättning för att en skolutveckling ska ske, finns däremot inte något uttalat intresse. Utvärdering tycks dessutom fortsättningsvis bli en mer marginell uppgift för det nya Skolverket. I praktiken finns inte heller, efter den senaste omorganisationen, någon särskild utvärderingsenhet inom myndigheten.

4.2.3 Central tillsyn och lokal skolutveckling separeras

I regeringens utvecklingsplan 2001/02 sägs att de nationella tematiska kvalitetsgranskningar som genomförts av statliga skolinspektörer tjänat som ett värdefullt komplement till Skolverkets reguljära tillsyn, uppföljning och utvärdering (Utbildningsdepartementet 2002). I sin utvecklingsplan föreslår regeringen att verkets uppgifter fortsättningsvis skall renodlas och särskiljas i en myndighet för skolutveckling och en myndighet för kvalitetsgranskning och tillsyn. Detta förslag antogs av riksdagen och verkställdes i mars 2003, när Myndigheten för skolutveckling

inrättades och tilldelades uppgiften att initiera lokal skolutveckling och stödja kompetensutveckling. Båda dessa uppgifter har sedan det första steget i förvaltningsreformen 1980 legat på kommunerna som har huvudansvaret för den lokala skolans kvalitetsutveckling. Redan hösten 2000 fick Skolverket regeringens uppdrag att utveckla formerna för ett mer aktivt utvecklingsstöd. Den ändrade inriktningen innebar att ”medel för lokal skolutveckling tillförs till följd av ’uppsökande verksamhet’ och i mindre utsträckning efter ansökan.” (Regeringens skrivelse 2001/02:188, s. 26). Att inrätta en ny skolutvecklingsmyndighet motiverades av att staten mer aktivt skulle kunna verka för en likvärdig skola i hela landet för alla barn och studerande. En mer aktiv myndighetsroll formuleras också i regeringens utvecklingsplan:

Det statliga utvecklingsstödet kan bättre stimulera och stödja det lokala utvecklingsarbetet genom en myndighet som har som sin huvuduppgift att främja utveckling. Genom att inte behöva svara för mer traditionella förvaltande uppgifter får en sådan myndighet stor frihet att mer förutsättningslöst verka inspirerande och pådrivande. (Regeringens skrivelse 2001/02:188, s. 30)

Det regeringsuppdrag som tilldelades det nya Skolverket som inrättades efter omorganisationen 2003 skulle innefatta uppföljning, utvärdering, granskning och tillsyn. Verkets huvuduppgift borde, enligt regeringen, vara att granska utbildningens kvalitet i skolor och kommuner och att utöva tillsyn över den verksamhet som bedrivs lokalt. Nationell likvärdighet anförs även här som grundmotiv, och en renodling av myndighetens uppdrag föranleddes av att: ”ett decentraliserat system förutsätter inte bara att staten sätter nationella mål utan också att staten systematiskt följer upp och granskar verksamheten med utgångspunkt i skollag och fastställda mål.” (Ibid, s. 27).

Den renodling av uppgifter som följt med den senaste omorganisationen har uppenbart inneburit en förskjutning tillbaka mot en starkare central styrning. Den lokala frihet inom givna ramar som framhållits som vägledande princip för skolans styrning har efterhand urholkats genom att styrningen utvidgats till att gälla mål- och resultatstyrning. Mål att uppnå liksom betygskriterier har preciserats i kursplanerna, och prov – såväl nationella som

internationella – har fått en ökad betydelse som kvalitetsmått och jämförelsegrund. Utvärderingens funktion tycks också ha fått en annan tyngdpunkt. Den fördjupade analys och strävan att söka förstå och förklara resultaten och de långsiktiga effekterna av genomförda insatser har tonats ned, medan en granskning och bedömning mot gällande lagar och nationellt formulerade mål har fått en starkare betoning.

På Skolverkets hemsida på webben anges den roll som verket idag har att fylla i det svenska utbildningssystemet: ”att ange mål för att styra, att informera för att påverka och att granska för att förbättra.” (www.skolverket.se 2006). De tre att-satserna preciseras under respektive rubrik. I texten under den tredje rubriken – att granska för att förbättra – nämns dels nationell utvärdering, dels internationell utvärdering. Nationell utvärdering, som innebär att beskriva, förklara och värdera olika företeelser inom barnomsorg och skola, förutsätts både ge fördjupade kunskaper och underlag till medvetna förbättringar på alla nivåer i skolsystemet. De nationella utvärderingarna förutsätts vara till gagn för alla aktörer i deras strävan att utveckla och förnya verksamheten och även tjäna som faktaunderlag för dem som vill delta i den skolpolitiska debatten. (Jfr. Karlsson Vestman, m.fl 2007).

I Skolverkets myndighetsuppdrag ingår att ansvara för att Sverige deltar i internationella jämförande kunskapsbedömningar. Dessa internationella utvärderingar syftar dels till att undersöka hur väl svenska elever klarar sig jämfört med elever i andra länder, dels till att undersöka på vilken nivå den svenska skolans nationella mål ligger i förhållande till skolor internationellt. Vilka områden som skall utvärderas läggs inte fast i något formulerat utvärderingsprogram utan bestäms utifrån de problem som uppkommer och de behov som uppmärksammas i den officiella statistiken över utbildningssektorn eller av utbildningsinspektionen.

Kvalitetsgranskning och tillsyn är som framgått idag Skolverkets huvuduppgift. Mot bakgrund av den historia vi redovisat kan det hävdas att utvärdering som instrument för politisk styrning och kontroll tonats ned vid övergången till 2000-talet, och nygamla

instrument som utbildningsinspektion och kunskapsprov kommer åter i bruk. Medan pedagogisk expertis under 1900-talet har haft stort inflytande över utformningen av den pedagogiska utvärderingen, har tjänstemän inom den statliga förvaltningen nu kommit att dominera. Vad detta skifte kan komma att betyda återstår att följa genom empiriska studier. Förhåller det sig så att de akademiska forskarna får träda tillbaka och ge plats för den vanliga människan som utvärderare i rollen som brukare och medborgare. Eller är det kanske snarare så att tjänstemän inom stat och kommun, underställda regeringen, kommer att bli dominerade?

4.3 Utvärdering på lokal nivå

Självvärdering som verktyg för lokal skolutveckling tillkom när det första steget mot en ökad decentralisering togs och Lgr80 infördes som nationellt styrdokument. Med denna förordning följde en förskjutning från regelstyrning till målstyrning. Det innebar att befogenheter och ansvar för den lokala skolans utvecklingsarbete och lärarnas fortbildning överfördes från central till lokal nivå. Förutsättningarna för en övergång till målstyrning var kända och fanns uttalade i de politiska dokument som förgick reformering av den statliga skoladministrationen och implementeringen av den nya läroplanen för grundskolan.

I den socialdemokratiska regeringspropositionen om "Skolans utveckling och styrning" 1988/89 konstateras, att de nationellt formulerade målen aldrig kan uttryckas på ett enkelt och helt entydigt sätt då de inrymmer en kompromiss mellan olika politiska uppfattningar. De kan inte heller preciseras så långt att de direkt anger vilka kunskaper och färdigheter som eleverna skall tillägna sig. En sådan precisering måste göras på den lokala skolnivån, där det konkreta arbetet skall planeras och utföras. För detta arbete har skolans personal det professionella ansvaret och rektor tilldelas ett särskilt ansvar för det pedagogiska ledarskapet. Som viktiga förutsättningar för en ökad målstyrning framhåller regeringen vikten av information, lärarutbildning och fortbildning. Lärarna och skolans personal nämns som den

självlart viktigaste gruppen och skolledarna utpekades som en nyckelgrupp (Andersson 1996). Med övergången mot ökad målstyrning följde också kravet på kommunala skolplaner och lokala arbetsplaner.

Vi har tidigare berört det kommentarmaterial som Skolöverstyrelsen tog fram i anslutning till Lgr80. Kommentarmaterialet var inte bindande och syftet med kommentarmaterialet var inte att ge en direkt handledning utan tänkt att tjäna som underlag för diskussion i klass, arbetsenhet, rektorsområde och kommun. De frågor som fokuseras är: varför ett ökat arbete med utvärdering är naturligt och nödvändigt, varifrån modeller kan hämtas för arbetet med utvärdering i grundskolan och vad detta arbete kan användas till, dvs. vilken nytta man på lokal nivå kan ha av utvärdering i skolan (Skolöverstyrelsen, 1983).

4.3.1 Krav på utvärderarkompetens

De standardiserade test och utvärderingsmodeller som utvecklades under 1900-talet hade tagits fram av pedagogisk expertis, och denna expertis engagerades också för utvärderingsuppdrag. De pedagogiska forskarna var vetenskapligt skolade och som externa utvärderare framstod de som garanter för en systematisk och objektiv bedömning. Att utvärdera själv i klassen, arbetsenheten, rektorsområdet och kommunen hade inte samma legitimitet, och varken skolans personal eller kommunens tjänstemän hade en skolning som kunde jämföras med forskarnas. Förväntningarna på skolorna var däremot högt ställda, vilket också avspeglas i det kommentarmaterial som Skolöverstyrelsen tagit fram. I det inledande kapitlet med rubriken "Utvärdering är något mer än att ge prov och sätta betyg" beskrivs utvärdering exempelvis med nyckeltermerna som: granska, undersöka, utforska, stanna upp, "ställa sig utanför", betrakta ur ett annat perspektiv, reflektera, sätta in verksamheten i ett vidare sammanhang, ifrågasätta, vara kritisk, var självkritisk och bygga upp kunskap. Uppräkningen av allt vad det kan innebära att utvärdera själv i den lokala pedagogiska praktiken ger bilden av en ytterst omfattande kollektiv utvärderingsverksamhet, och den kontinuerligt pågående verksamheten: "har intimt samband

med arbetet med lokala arbetsplaner och arbetet med att på eget initiativ och ansvar utveckla verksamheten i den riktning som läroplanen anger, dvs. genomföra lokalt utvecklingsarbete.” (Skolöverstyrelsen, 1983, s. 7).

Men uppräknigen av allt som den lokala utvärderingen kan vara ger också uttryck för mer än en hög målsättning. Den uttrycker en vision om att den decentraliserade skolan skall ge utrymme för lärarnas och skolledningens egna professionaliseringssträvanden. Att förverkliga visionen om en framgångsrik professionalisering som skulle kunna ge de praktiskt verksamma pedagogerna ökad status och legitimitet innebär om möjligt en än mer genomgripande ”omskolning” än den som krävs för att genomföra den aktuella reformen. En ”omskolning” för professionell kompetens förutsätter nämligen en mer långsiktig process som inrymmer kunskaper både från akademisk forskning och erfarenheter från den pedagogiska praktiken (Andersson, 1996). Den kompletteringsfortbildning som Skolöverstyrelsen planerade vid slutet av 1980-talet kom dock inte att fullföljas då andra steget i förvaltningsreformen genomfördes, Skolöverstyrelsen lades ned och kommunerna fick det fulla ansvaret som huvudman för skolan och dess personal.

När det andra steget i förvaltningsreformen skulle verkställas hade den svenska ekonomin, som redan under 1970-talet började stagnera, väsentligen försämrats, och kommunerna ställdes inför stora besparingskrav. Att hålla budgeten blev nu viktigare än att leda och bedriva pedagogiskt utvecklingsarbete. När lågstadiesatsningen beslutades 1982 förkom tidningsrubriken: ”Det våras för lågstadiet.” Denna rubrik föranleddes av att pengar, som använts som förstärkningsresurs på högstadiet, nu återfördes till lågstadiet. Att söka projektmedel från de särskilda satsningarna har fortsatt att vara en möjlighet för kommunen och skolan lokalt att i någon mån vidga sina ekonomiska ramar. Att starta skolprojekt som tillför medel kan såtillvida också bli mer attraktivt än att engagera sig i det reguljära utvecklingsarbetet på skolan. Som tidigare framhållits har de tidsbegränsade utvecklingsprojekten bedrivits som en särskild insats. Projektpengarna har satt en viss ”guldkant” på vardagsarbetet men i många fall

också lett till en gruppering av personalen. Att starta skolprojekt och ansöka om medel från en särskild satsning är, som framgått, ett frivilligt åtagande och långt ifrån alla på en skola vill delta i detta arbete. Många visar inte heller något intresse för att ta del av resultaten (Andersson, 2002).

De externa utvärderingar som gjorts i samband med de särskilda satsningarna har inneburit att forskare, som genomfört dessa utvärderingar, har fått möjligheter att följa och ta del av de pedagogiska processerna och tillsammans med dem som deltar i projektarbetet lära av praktiken. De kvalitativa metoder som fått acceptans efter 1970-talets vetenskapliga paradigmskifte har understött en sådan kunskapsprocess, och skiftet i synsätt avspeglas också i det strategiprogram för skolforskning som Skolverket presenterade för perioden 1993/94–1995/96. Följande beskrivning är hämtad från detta program:

Av de studier av svensk skolforskning som Skolverket låtit göra framgår att den genom åren dominerats av den s.k. FoU-modellen. Även om denna skall vila på grundläggande forskning har den i sina resultat och sin förlängning en ”teknologisk” innebörd: den leder till ”produkter” i form av nya metoder, tekniker, arbetsätt och rekommendationer som inom skolområdet förmedlats genom t.ex. fortbildning av lärare. (...) De nämnda studierna visar också samtidigt att denna ursprungligen naturvetenskapligt/tekniskt grundade modell för forskning och utveckling under 1980-talet alltmer ifrågasatts. Den har ersatts av föreställningar och modeller som betonar forskningens uppgift att förstå och vinna insikt i t.ex. undervisningsprocesser i skolan och i det sammanhang i vilka de utformas. (Citat i Andersson 1996, s. 38)

4.3.2 Utvärdering för reflektion och lärande

Det kunskapsintresse som vi funnit i Skolverkets utvärderingsprogram, finner vi således också i den strategi som formuleras för verkets forskningsprogram. Uppfattningen om teknik som medel för att nå de nationella målen ersätts här av uppfattningen om kunskap som tillåter reflektion och egna ställningstagande. Den lokala utvärderingen utgör i detta sammanhang ett viktigt instrument i det som Petterson och Wallin (1995) gett benämningen ”informativ styrning”. Som framgått hade pedagogiska forskare sedan länge hävdad nödvändigheten av en processinriktad och kunskapsgenererande utvärderingsansats som förutsätter

dels information om de pedagogiska insatsernas förutsättningar och genomförande, dels fördjupade analyser.

I de externa utvärderingar, som under 1980-talet genomfördes i anslutning till de särskilda satsningarna, brukades en sådan ansats. De begränsningar som var nödvändiga för att kunna genomföra fältstudier i samband med lågstadiesatsningen och fallstudier i samband med försöksverksamheten på gymnasieskolan har tidigare berörts. Även om studierna genomfördes i strategiskt utvalda skolor begränsades bilden till de aktuella fallstudieskolorna. För en bedömning av skolsystemet i sin helhet krävs att uppgifter som samlas in om förutsättningarna för och undervisningsprocesser i den pedagogiska praktiken har en större bredd och är fullt jämförbara. En samordning mellan olika nivåer i systemet är därför nödvändig och lokal självvärdering framstod här som ett nödvändigt komplement till Skolverkets och Länsskolnämndernas utvärderingar. I de fem pedagogikprofessorernas egen skrift om skolan och utvärderingen uttrycker Urban Dahllöf det på följande sätt:

Under alla förhållanden står det klart, att en lokalt bedriven, öppen och självkritisk utvärdering, inriktad på att ge informationsunderlag till förbättringar, är en helt nödvändig förutsättning för en levande och utvecklingsbar skola. Utan en sådan dynamisk bas framstår det som ganska ointressant att bedriva utvärdering på det nationella planet framför allt om man hoppas på att den vägen på sikt kunna avläsa några väsentliga förbättringar av undervisningens situation och resultat. (Dahllöf, 1989, s. 26)

Under 1980-talet kom många pedagogiska forskare att intressera sig för didaktik och praktisknära skolforskning. (Se t. ex. Marton 1986; Balnkertz 1987). Dessa områden betonades inte minst i den nya lärarutbildning för grundskolan som inrättades 1988/89. Pedagogisk utvärdering fick i detta sammanhang också en mer praktisk pedagogisk inriktning. Utvärdering som verktyg för lokal utveckling av utbildning/undervisning och för professionell utveckling betonades framför utvärdering som instrument för kontroll (Franke-Wikberg 1990; Andersson 1991; Pettersson & Wallin 1995). Under 1990-talets första hälft växte intresset för praktisknära skolforskning i samarbete mellan forskare och praktiker. Förutsättningarna för både den kunskapsutveckling

som fanns uttalad i Skolverkets utvärderingsprogram och intresset för en ökad professionalisering av skolans personal tycktes kunna realiseras, även om något slutgiltigt idealtillstånd inte förespeglades (Andersson, 1999; Andrae Thelin, 1999). Liksom de konstruktivistiska teorier om lärande som aktualiserades under denna tid tänktes utveckling av både utbildning och profession som en process och utbildningssystemet som en lärande organisation med en öppen kommunikation mellan aktörer på olika nivåer i systemet (Andersson, 1996; Pettersson & Wallin 1995; Kemmis, 1998).

4.3.3 *Kvalitet och marknadstänkande*

De strävanden till samverkan mellan pedagogisk forskning och pedagogisk praktik, som tog fart under 1990-talet, fick viss support genom den särskilda satsning som utbildningsdepartementet inbjöd Sveriges skolor till hösten 1997 och som syftade till att utveckla lärarrollen i samarbete med universitet och högskolor. Som tidigare påpekats präglades denna tid samtidigt av hårda ekonomiska besparingar. Ett marknadstänkande med nyckelbegreppet "kvalitet" fick sin spridning och kom att brukas i allt fler sammansättningar som: kvalitetskontroll, kvalitetssäkring, kvalitetsutveckling och kvalitetsbedömning. Kvalitetsbegreppet har sedan länge förknippats med varor och tjänster och kunders subjektiva bedömningar. Under 1990-talet etablerades det i den offentliga diskursen om den svenska förvaltningsorganisationen och satte avtryck i såväl regeringens "Utvecklingsplan för förskola, skola och vuxenutbildning – kvalitet och likvärdighet" (1996/97: 112) som i "Förordningen om kvalitetsredovisning inom skolväsendet" (1997:702). Förenklat kan händelseförloppet beskrivas på detta vis.

Vid början av 1990-talet gör staten en drastisk ekonomisk nedskärning och kommunerna drabbas av sparbetning. Ett marknadstänkande breder ut sig samtidigt med en nyliberal ideologi och NPM (New Public Management) som modell för styrning får gehör.⁷ Medan lokal utvärdering hittills inneburit en utvärdering mot de nationella målen följer nu en tid av utvärdering mot brukarna. Enkätundersökningar blir vanligt förekommande på alla nivåer, och brukarnas uppfattningar redovisas i absoluta procentsatser oavsett hur stor eller liten undersökningsgruppen varit. Opinionssiffror redovisas och sprids genom media som vore de exakta bilder av verkligheten. Med mediabevakning och informations spridning via nätet växer utvärdering som företeelse till en bölja som sköljer över hela världen. Och i allt detta informationsbrus gäller det att synas – annars finns man inte!

Sättet att formulera sig i aktuella planer ändras från konkreta beskrivningar av innehållet/kursen (i den mån nu sådan förekom) till allt mer vittsyftande mål som beskriver vad skolan långsiktigt strävar mot – mål som utlovas vara uppfyllda på kort sikt, respektive längre sikt. Beskrivningarna riktar sig framför allt mot brukarna, och skolans ledning inriktar sig på s.k. SWOT-analys av verksamhetens styrkor, svagheter, möjligheter och hot (Åkerblom, 1994). Kvalitetsutmärkelser från organisationer som exempelvis SIQ – Svenska institutet för kvalitetsutveckling får i detta sammanhang sin betydelse som trofé likt filmens Guldbagge, musikens Guldplatta och sportens Bragdmedalj. Att vinna en utmärkelse är viktigt, men att nomineras har också sin betydelse då det gäller att bli uppmärksam, inte bara nationellt utan helst också internationellt.

7 NPM var ursprungligen inte utformat för att implementeras. Uttrycket myntades av Christofer Hood 1990 (Hood, 1990) som funnit att en rad likartade reformer har genomförts i olika OECD-länder och satte etiketten NPM på denna trendförskjutning. Begreppet har sedan brukats och fått spridning till organisationer och aktörer som argumenterat för en reformering inom offentlig sektor och NPM och blivit en modell för att jämföra och bedöma olika länders förvaltnings som brukas för bl.a. utvärdering. Se vidare Sahlin-Andersson 2000.

En institutionalisering av kvalitetstänkandet kom igång med stöd av SIQ och genom Kommunförbundet som kvalitetsorganisation. Utmärkelsen svensk skola delades ut första gången 1995, och därefter har SIQ fortsatt sina aktiviteter inriktade mot skolan i samråd med representanter för de olika parterna. SIQ erbjuder två verktyg för kvalitetsutveckling: Utmärkelsen Svensk Kvalitet och Förbättringsinstrumentet svensk skola. Med hjälp av dessa verktyg tar man på den enskilda skolan fram en självvärderingsrapport som både kan användas för det lokala förbättringsarbetet och för att identifiera och lyfta fram goda exempel. Att söka kvalitetsutmärkelsen, som delas ut varje år, kan också göra skolan uppmärksammas dels genom den utomstående examinator som granskar självvärderingsrapporten och redovisar en återföringsrapport, dels vid den offentliga utdelningen av utmärkelsen som sker i samband med en större festlighet. Att en skola uppmärksammas och dess verksamhet synliggörs offentligt kan likställas med marknadsföring, och på marknaden gäller principerna om konkurrens som drivkraft för förbättringar och valfrihet för kunderna (Karlsson 1997).

Det krav på kvalitetsredovisningar som infördes i förordningen 1997 avsåg att fylla liknade syften. De skulle tjäna som förbättringsunderlag, möjliggöra insyn och ökat inflytande, ge information till föräldrar och elever vid valet av skola och ligga till grund för politiska beslut på kommunal nivå. Som vägledande principer gäller här likvärdighet och ansvar. Motivet till den aktuella förordningen anges i de Allmänna råd om kvalitetsredovisning inom skolväsendet som Skolverket publicerade 1999. I denna text återfinns följande formulering:

Regering och riksdag har bedömt att värdering och utveckling av kvaliteten i skolsystemet behöver förstärkas så att de blir självklara inslag. Staten har därför nyligen skärpt dessa krav på alla ansvarsnivåer. Nationellt betonas Skolverkets kvalitetsgranskande roll och för skolor och skolhuvudmän har kravet på årliga kvalitetsredovisningar införts. Motivet är att ännu tydligare än hittills markera det gemensamma ansvaret för kvalitet och likvärdighet. (Skolverket, 1999)

I Utvecklingsplanen 2001/02 konstateras att skolornas kvalitetsarbete och kvalitetsredovisningarna blivit bättre. År 2000 fanns det dock fortfarande många skolor som inte uppfyllde kravet på

att göra kvalitetsredovisningar, och i många av de kvalitetsredovisningar som inlämnats fanns påtagliga brister. För att underlätta kvalitetsarbetet föreslår regeringen att kravet på arbetsplan och skolplan tas bort och att kvalitetsarbetet koncentreras i ett enda dokument – en föreskriven kvalitetsredovisning. Vad som föreskrivs preciseras i det pressmeddelande som utbildningsdepartementet skickade ut vid slutet av maj 2002. Här formuleras i korthet vad ett förstärkt kvalitetsarbete innebär för landets skolor och Skolverkets uppdrag i anslutning till de föreskrivna kvalitetsredovisningarna:

Varje skola ska minst en gång per år utarbeta en kvalitetsredovisning där man utvärderar och bedömer den egna verksamheten i förhållande till de nationella målen samt redovisa vilka åtgärder man vidtar om målen inte nås. Statens krav på skriftlig dokumentation renodlas till kvalitetsredovisningen genom att övriga krav på skolplan och arbetsplan tas bort. Kvalitetsredovisningar införs också i förskolan och skolbarnomsorgen. Skolverket får i uppdrag att utarbeta ett antal gemensamma resultatmått som skall ingå i alla skolors kvalitetsredovisningar. De ska spegla hela skolans uppdrag och omfatta målen såväl i läroplanen som i kursplanerna. (Utbildningsdepartementets pressmeddelande 2002-05-27, s. 2–3)

I ett senare pressmeddelande återkommer utbildningsdepartementet till Skolverkets uppdrag. Av detta meddelande framgår det att Skolverket skall komplettera regelverket och dels slå fast ett övergripande syfte med kvalitetsredovisningen, dels förtydliga bestämmelserna så att det framgår: ”vad redovisningen närmare skall innehålla beträffande obligatoriska kvalitetsområden och kvalitetsindikatorer.” (Utbildningsdepartementet 2004-07-09). En precisering avser att skapa kvalitetsmått som kan stödja det lokala utvecklingsarbetet och leda till att kvalitetsredovisningarna blir mer enhetliga och jämförbara. Kravet på ökad jämförbarhet och enhetlighet förstärks också av en ökad användning av övergripande internationella och nationella utvärderingar av skolan.

4.4 Utvärdering på internationell nivå

Internationella jämförelser av olika länders skolsystem tillkom när utvärdering fick sin spridning under 1960-talet och forskare i pedagogisk psykologi började utveckla instrument för de s.k. IEA-undersökningarna.⁸ Professor emeritus Torsten Husén, som under sexton år var ordförande för IEA beskriver bakgrunden till de internationella undersökningarna på följande sätt:

Detta var vid den tid man i Sovjetunionen sköt upp Sputnik, vilket innebar en chock för amerikanerna. De såg i denna prestation ett uttryck för att den teknologi som åstadkommit detta underverk måste grunda sig på en överlägsen naturvetenskaplig undervisning i skolor och högskolor. (...) Det kan diskuteras i vad mån man på detta sätt kan mäta nationers prestationer inom teknologi och naturvetenskap. Hur som helst: Sputnik blev en symbol och utgjorde en faktor i ett konglomerat av omständigheter kring 1960, som ledde till ett genombrott för den komparativa pedagogiken, "comparative education." (Husén 1994, s. 102)

Liksom i USA har kunskapsmätningar i matematik och naturvetenskap fått stor betydelse för värderingen av skolans undervisning. Om IEA-undersökningarna har visat att svenska elevers provresultat är svaga jämfört med elever i vissa andra länder eller att de svenska provresultaten försämrats sedan den föregående mätningen, så har talet om "kunskapsbrister" och "kris i skolan" uppkommit. Detta har särskilt gällt kunskapsresultat som rör matematik och svenska, och påtalade brister har, som framgått, lett till både statliga utredningar och särskild satsningar för att komma till rätta med problemen. Idag går de internationella utvärderingarna under benämningar som TIMSS och PISA.

På Skolverkets hemsida beskrivs att TIMSS – Trends in International Mathematics and Science Study, som är en utvärdering av elevers kunskaper i matematik och naturkunskap i årskurs 4 och 8. Utvärderingen görs i förhållande till ländernas

8 IEA – The International Association for the Evaluation of Educational Achievement – är en internationell organisation som genomför studier för att jämföra länders skolsystem, där bland andra Sverige ingår.

nationella läroplaner ur vilka man sammanfattar kriterierna för utvärderingen. Totalt deltar cirka 60 länder i TIMSS under 2007. Undersökningen bedrivs av IEA. En annan stor internationell utvärdering av skolan är PISA – Programme for International Student Assessment. Utvärderingen omfattar ett 40-tal länder och genomförs av inom OECD (Organisation for Economic Co-operation and Development). Det som studeras är i vilken grad respektive lands utbildningssystem rustar femtonåringar att möta framtiden. Genom olika prov undersöks elevernas förmågor inom fyra kunskapsområden: matematik, naturvetenskap, läsförståelse och problemlösning.

Det finns en viktig skillnad mellan TIMSS och PISA när det gäller vad man bedömer och kriterierna för detta. I stället för att utvärdera elevernas kunskaper mot *läroplanens mål*, vilket TIMSS gör, avser PISA att mäta kunskaper och färdigheter relaterade till *vardagslivet* som bedöms vara av betydelse i det vuxna livet. Det *livslånga lärandet*, att eleverna fortsätter att lära sig under hela livet, betonas. Stor vikt läggs i PISA-undersökningen vid elevernas förmågor att sätta kunskaper i ett sammanhang.

Med hänvisning till den PISA-undersökning som redovisades år 2000 skriver regeringen i sin Utvecklingsplan 2001/02:

Sverige har nått relativt goda resultat i matematik och naturvetenskapliga ämnen i internationella undersökningar de senaste åren. Bilden är emellertid inte entydig, bl.a. visar den senaste PISA-undersökningen att intresset för matematikämnet i Sverige ligger något under genomsnittet för OECD-länderna medan resultaten ligger något över. Regeringen har högre ambitioner än så. (Regeringens skrivelse 2001/02:188, s. 37–38)

Som åtgärd skriver regeringen att den avser att tillsätta en delegation med uppgift att utarbeta en handlingsplan för matematik från förskolan till högskolan. Resultaten från PISA-undersökningen 2004 visade att de svenska elevernas resultat försämrats, medan Finland intog en ledande position. De svenska resultaten framhölls som alarmerande och länge fördes i massmedia en intensiv diskussion om orsakerna till de försämrade resultaten. Brister i lärarutbildningen påtalades och även dessa blev alarmerande när Högskoleverket våren 2005 presenterade sin utvärdering av lärarutbildningen (Högskoleverket, 2005).

4.4.1 Kritisk granskning av TIMSS och PISA

Att politiker och myndigheter behöver information som underlag för att kunna fatta beslut om resurser, stifta lagar, utfärda förordningar och skolreformer är självklart, och utvärdering spelar i dessa sammanhang en viktig roll. Med internationella utvärderingar som TIMSS och PISA vill politiker och myndigheter också skaffa sig information om hur det egna landets skolsystem ligger till vad gäller utbyggnad, kostnader, elevernas kunskapsresultat etc. Frågan är om den kunskap som produceras genom dylika utvärderingar är tillförlitlig, väl underbyggd och verkligen ger ett underlag som håller för internationella jämförelser.

Av tidigare avsnitt har det framgått, att pedagogiska forskare sedan länge kritiserat det ensidiga bruket av kunskapsmätningar vid utvärderingar på systemnivå. I det redovisade citatet skriver professor Torsten Husén också, att det kan diskuteras i vad mån prestationer på nationell nivå kan mätas med de prov som brukas i de internationella utvärderingarna. Ändock har resultat från kunskapsmätningar och jämförelser mellan skolor och skolsystem stort nyhetsintresse, och resultaten efterfrågas inte minst av politikerna. Men även om svagheter med kunskapsmätningar som verktyg för jämförelser påpekas av forskare och diskuteras i de utvärderingsrapporter som redovisas har statiska uppgifter en vidunderlig förmåga att övertyga och framstå som objektivt gripbara sanningar. Och de brukas inte sällan som ”hårda fakta” i den politiska debatten.

Den norske pedagogikprofessorn, Svein Sjøberg (2005), har granskat några av de aktuella utvärderingarna TIMSS och PISA. De båda utvärderingarna har vissa likheter men kännetecknas också av viktiga skillnader. Som nämndes ovan mäter TIMSS hur eleverna presterar i förhållande till en konstruerad läroplan, medan PISA mäter mot kriterier som konstruerats av medverkande forskare. Ett problem med TIMSS läroplansperspektiv är att man konstruerat en generell läroplan baserad på alla länder som deltar i undersökningen. Sjøberg ifrågasätter om det finns någon gemensam ”skolkunskap” som är identisk i t.ex. Egypten, Ghana, USA, Iran, Indonesien, Sverige och Norge. Det är knap-

past heller önskvärt att samma ämneskriterier gäller för alla deltagande länder och att resultaten skall kunna mätas med samma prov. För att kunna jämföra mellan länderna blir konsekvensen att man rensar ut alla olikheter och väljer sådana uppgifter som kan finnas i alla läroplaner, påpekar Sjöberg. De uppgifter som ingår i provet utgör en minsta gemensam nämnare som representerar en konservativ typ av läroplan. Det som väljs ut blir med andra ord det som man i alla berörda länder kan känna igen och som okritiskt tas som givna sanningar. Provet får också en konserverande effekt eftersom det i många länder styr vad man betraktar som viktigt att lära ut (man lär för testet). Kort sagt baserar sig en internationell läroplan av detta slag på kompromisser och politiskt innehåll, konstaterar Sjöberg.

I PISA-undersökningen, som mäter mot konstruerade kriterier, skall uppgifterna vara knutna till konkreta och aktuella situationer i dagens samhälle. Förutom problemet med att avgöra vad som är representativt för "dagens samhälle" i de olika deltagande länderna, är det också ett problem att uppgifterna är hemliga. Forskarna vill inte att provuppgifterna skall bli kända i förväg. Man vill förhindra en undervisning inriktad på att lära för testet. Samtidigt gör bristen på öppenhet att det inte går att föra en kritisk debatt kring de antaganden som uppgifterna baseras på. Det finns endast ett "rätt" svar, vilket Sjöberg ifrågasätter efter att han granskat de prov som har genomförts och uppgifterna därigenom blivit kända. Sammanfattningsvis kan det hävdas att internationella utvärderingar ger upphov till problem i form av alltför långtgående generaliseringar och därtill kännetecknas av bristande förankring i de kontextuella förhållandena i varje land. Som vi tidigare påpekat kan dessa förhållanden förklara skillnader i de deltagande ländernas provresultat.

4.4.2 En nyliberal marknadsfilosofi

I en intervju i Pedagogiska magasinet 2005 kommenterar Michael Uljens, professor vid Åbo Akademi, internationella jämförande studier som PISA och TIMSS (Mathiasson, 2005). Enligt Uljens ingår PISA i en strategi som syftar till att harmonisera de olika nationernas utbildningssystem med varandra.

Harmonisering och standardisering anses nödvändigt för att underlätta det ekonomiska samarbetet. Denna bakomliggande utbildningspolitiska filosofi i PISA har nästan inte alls lyfts fram, konstaterar han. I stället tenderar diskussionerna att landa i en oro över landets egna dåliga resultat eller grannländers goda resultat. Uljens menar att PISA och TIMSS är exempel på en marknadsorienterad och nyliberal utbildningsfilosofi. Genom att ange medeltal som möjliggör rangordning provocerar man fram en konkurrens men också en politisk dialog mellan de deltagande länderna. Undersökningen säger ingenting om hur bra ungdomar uppnår det som respektive lands skolsystem strävar att utveckla och preciserar inte heller vad exempelvis medborgerlig kompetens är. Skulle man lägga fast en sådan norm och låta den utgöra måttstocken, skulle kanske flertalet länder klara normen och på sikt kanske till och med alla deltagande länder, anser Uljens. Därför är man inom OECD väldigt noga med att inte ange en sådan norm utan man söker i stället att trimma in nationerna i en konstant hetsjakt. Alla länder måste hela tiden bli bättre men oavsett hur bra de blir så får de aldrig bli färdiga.

4.4.3 Kombination av extern och intern utvärdering

Det är inte ovanligt att centrala makthavare å ena sidan framhåller att internationella jämförelser är viktiga och att man behöver sådana för att avgöra hur det egna landets skola och utbildningspolitik ligger till. Å andra sidan uppmanar många till försiktighet vid sådana jämförelser och framhåller vikten av att beakta de unika omständigheterna och de lokala skillnaderna vid tolkningen av resultaten. Man kan se det som ett försök att balansera mellan centrala och lokala intressen vid utformningen av utvärdering av skola och undervisning. Ett exempel som kan tolkas som ett sådant balansförsök mellan centralistiska och decentralistiska strävanden ges inom ramen för Europasamarbetet och utarbetandet av politiska åtgärder inom den allmänna utbildningen. I ett sådant sammanarbete redovisade en kommitté redan våren 2000 en rapport med titeln: "Utbildningens kvalitet i skolorna – Sexton kvalitetsindikatorer." (Europeiska Kommissionen, 2000). Två av dessa indikatorer rör området "Tillsyn över utbildningen" och en av dem gäller: "Utvärdering

och styrning av utbildningen.” Vad utvärdering inom EU kan vara och vilka former av utvärdering det kan handla om beskriver kommittén på följande sätt:

Utvärdering kan antingen vara intern (självvärdering) eller extern, eller en kombination av båda. Båge utvärderingsformerna inverkar på resurserna och utbildningen. Därför är de flesta länder i Europa på jakt efter den bästa och mest produktiva kombinationen av de två utvärderingsformerna. Eftersom både den externa utvärderingen och självvärderingen är viktiga informationskällor bör man låta dem komplettera varandra. (Ibid, s. 41)

I rapporten antyder kommittén att det fortsatt viktigaste arbetet är att samla in uppgifter från medlemsländerna som åskådliggör fördelningen mellan intern och extern utvärdering och vilken roll var och en av dessa spelar för styrning och utvärdering av skolsystemet.

År 2001, ett år efter det att kommittérapporten hade presenterats och när Sverige var EU:s ordförandeland, hölls en konferens i Karlstad, där frågan om fördelningen eller det ömsesidiga växelspelet (reciprocal) mellan extern och intern utvärdering fokuserades.⁹ Från denna konferens finns ett stort material redovisat dels i form av en konferensrapport, dels som separata papper som behandlar olika teman. I sitt inledningsanförande anknöt David Coyne, director för utbildningsområdet inom EU-kommissionen, till de sexton indikatorerna och pekade på vilken betydelse de kunde ha som riktninggivare för de fortsatta diskussionerna och vad de kan resultera i. Han uttrycker det på följande sätt:

I think that this is the type of debate that we should have, in our discussions on quality. Indicators can be very useful as tool to bring our policies and practice closer together, not just our rates and ratios. It is, in my view, useless to show that one country “scores” better than others in some fields of quality, without analysing the reasons and offering ways to learn from that country.

Indicators and benchmarks are not only there to close performance gaps

9 Swedish Ministry of Education and Science: Quality in Education: a European perspective. Conference “The meaning of Quality in Education – the reciprocal relations of power-sharing, learning and efficiency” Karlstad (S), 2–4 April, 2001.

between countries, but also to close communication gaps. In other words, I believe that when we start debating and comparing quality issues in education, the process counts perhaps more than the results. It is about agreeing on terminology, on concepts, finding common ground, speaking common language. The aim is to provide policy makers with reference points and to identify issues which need to be investigated further, and to suggest alternative routes to policy goals. (Ibid, s. 6–7)

Citatet ger uttryck för att internationella utvärderingar med gemensamma indikatorer kan vara användbara som medel att harmonisera de deltagande ländernas utbildningspolitik. Samtidigt betonas att tolkningen av resultaten kräver en djupare förståelse av de lokala förhållandena i respektive land. En slutsats blir därför att utvärdering behövs på olika nivåer i det utbildningssystem som nu samordnas och formas i samarbetet mellan EU:s medlemsländer.

Ser man på hur de två svenska statliga verken inom utbildningssektorn, Skolverket och Högskoleverket, tacklar dessa frågor så kan man konstatera att för båda gäller en modell som kombinerar externa utvärderingar med interna. Som framgått av tidigare avsnitt har Skolverket omorganiserats, dess uppdrag renodlats och under den senaste tioårsperioden inriktats framför allt på kvalitetsgranskning och tillsyn. Detta medan utvärdering numera handlar om fördjupade studier och kunskapsbedömningar, vilka båda avser att generera kunskap. I Högskoleverkets utvärderingsmodell används begreppen "självvärdering" och "expertbedömning" (Karlsson & Kuivanen 2000; Karlsson, Andersson & Lundin 2002), och Skolverket kombinerar "kvalitetsgranskning" med "kvalitetsredovisning". I båda fallen gäller, att det har skett en tyngdpunktsförskjutning från en inriktning på kvalitetsarbete mot en betoning av kvalitetsbedömning, vilket kan tolkas som ett uttryck för att myndigheternas värderande roll har markerats.

I Skolverkets uppdrag ingår även fortsättningsvis att ta in uppgifter som beskriver hur skolverksamheten är organiserad, hur mycket den kostar och vilka resultat som uppnås. Uppgifter av detta slag, som utgör den officiella statistiken över utbildningssektorn, låter sig lätt ordnas och ger inte sällan tidningar under-

lag för publicering av rankinglistor. Idag tycks dock publicerade rankinglistor mer fungera som ett snabbt signalsystem för eleverna vid deras val av skola. Men valfriheten har ett också pris för skolor och kommuner då ”skolpengen” följer med eleven. Så i spåren av detta marknadsspel följer både nedläggning av skolor och segregation av eleverna inte bara på kommunal- och nationell nivå utan också internationellt. I detta sammanhang spelar internationella utvärderingar liksom de graderade betyg från universitet och högskolor, som förhandlas på Europanivå inom ramen för den s.k. Bologna-processen, en viktig roll.¹⁰

10 Tillsammans med 45 andra europeiska länder deltar Sverige i ett mellanstatligt samarbete som fått benämningen Bologna-processen. Syftet är att främja rörlighet, möjligheterna till anställning och Europas konkurrenskraft som utbildningskontinent. Målsättningen är att skapa ett gemensamt europeiskt område för högre utbildning Våren 2006 skall den svenska regeringen fatta beslut i frågor som rör Bolognaprocessen och 2010 skall ”the European Higher Education Area” vara en realitet. (Se vidare: Regeringskansliets faktablad: Sammanfattning av regeringens proposition 2004/05: 162. Ny värld – ny högskola. Stockholm: Utbildnings- och kulturdepartementet UO5.127* juni 2005.)

5 Forskning om utvärdering

I detta kapitel fokuserar vi på utvärdering som kunskapsområde och vi vidgar här vårt undersökningsfält till fler tillämpningsområden än pedagogisk utvärdering. Den fråga vi ställer är hur *forskning* förhåller sig till utvärdering. För att få ett underlag som gör det möjligt att besvara den frågan har vi valt att granska de *svenska avhandlingar* som anger utvärdering, utvärdera eller evaluation som sökord eller som har orden som en del av avhandlingens titel.

5.1 Metod och material

En informationssökning i olika databaser som bokförlag, Googles, etc., med de sökord vi har nämnt ovan ger många träffar, och de publikationer som listas omfattar ett brett och skiftande material som speglar utvärderingsfältet i stort. Förutom vetenskapliga publikationer återfinns även ett stort antal utvärderingar som genomförts inom olika verksamhetsområden. Några exempel är Högskoleverkets utvärderingar av nationella ämnen och program och Skolverkets utvärderingar av skolor samt utvärderingar inom kommuner och landsting. Genom att välja ut svenska avhandlingar har vi begränsat oss till en liten del av alla de publikationer som man får fram med våra sökord. Vårt urval är såtillvida begränsat vilket givetvis kan kritiseras. Vi menar dock att vår avgränsning är en gångbar metod för att få en representativ bild av utvärderingsforskning. Syftet med vårt urval är att ge underlag för att teckna denna bild, inte att presentera en fullständig kartläggning.

Vi har använt oss av Libris biblioteksdatabas med sökord som vi anger ovan och med en avgränsning till publikationstypen avhandlingar. Vår informationssökning den 2005-10-05 gav 160 träffar. Sedan vi exkluderat de avhandlingar som producerats utanför Sverige och det fåtal arbeten som är licentiatuppsatser, samt senare har kompletterat med de avhandlingar som publi-

cerats under 2006 återstod 108 publikationer i det urval som vi kommer att behandla i detta kapitel. Dessa avhandlingar representerar lärosäten i hela landet, och gäller olika discipliner och tillämpningsområden; medicin, ekonomi, förvaltning, pedagogik, m.fl.

Det bör påpekas att vårt undersökningsmaterial innehåller såväl forskning *om* utvärdering som redovisningar *av* olika utvärderingsstudier där det också kan föras en diskussion kring utvärdering som företeelse. Detta betyder, att även om vi försöker koncentrera oss på avhandlingar med utvärdering som uttalat kunskapsobjekt, så går det inte alltid att dra någon knivskarp gräns mot studier som i huvudsak tillämpar utvärdering som metod inom ett verksamhetsområde, t.ex. sjukvård, barnomsorg, utbildning etc.

Vi har valt att behandla vårt undersökningsmaterial i kronologisk ordning efter publiceringsår och strukturerar den följande framställningen utifrån en indelning i decennier, från 1970-talet och framåt. Bland de publikationer som vi identifierat inom respektive decennium har vi sedan gjort ett urval utifrån de principer som vi ovan beskrivit, dvs. valt sådana verk som rubricerats som avhandlingar. I detta sammanhang bör det noteras att ett avhandlingsarbete i allmänhet är ett projekt på 4–5 år eller längre. Det gör att användningen av avhandlingar, som uttryck för en aktuell vetenskaplig debatt om utvärdering, kan vara begränsad. Frågor som tas upp i de olika avhandlingarna avspeglar inte sällan en forskningsinriktning som var giltig flera år tidigare. Analysen av avhandlingsarbetenas innehåll som uttryck för utvärderingsdebatten under olika decennier, måste ses mot den bakgrunden.

De avhandlingarna som valts ut och som det refereras till i den följande texten finns redovisade i en separat referenslista sist i boken. Siffran inom klammer i texten anger avhandlingens decennium och placering i denna referenslista.

5.2 1970-tal – *Forskning om utvärdering formas*

På 1970-talet presenteras de första avhandlingarna som rör utvärdering. Vi har funnit elva avhandlingar från olika discipliner från denna tid. I vår uppräkningslista nedan anges vilka utvärderingsobjekt som behandlas i dessa arbeten.

I statsvetenskap utvärderas 1971-års länsförvaltningsreform [70:7] och i astronomi utvärderas objektivitetsprismaspektra av tidiga stjärnor med hjälp av dator [70:1]. Inom ekonomi finner vi en utvärdering av företagskonkurser som analyserar hela policykedjan: problem, analys, utvärdering och åtgärder [70:4]. I psykologi kommer en avhandling som utvärderar arbetsanpassning [70:6] och en annan som granskar riktlinjer för analys, datainsamling och utvärdering i beteendeterapi [70:8]. I sociologi utvärderas alternativa produktionssystem vid Saab-Scania i Trollhättan [70:9], och i pedagogik finner vi en avhandling med utvärderingstema som handlar om "Test-wiseness" och provkonstruktion [70:10] och en som utvärderar ett kollektiv utifrån ett organisationsperspektiv [70:11].

Uppräkningen visar att utvärdering redan vid denna tidpunkt förekommer inom olika discipliner som representerar skilda fakultetsområden; humaniora, naturvetenskap, samhällsvetenskap och beteendevetenskap. Mångfalden gör att utvärdering snarast kan beskrivas som mångdisciplinärt forskningsfält. En närmare granskning visar att de avhandlingar som läggs fram vid olika ämnesinstitutioner grundas på genomförda utvärderingar eller empiriska studier, och utvärdering kan därför också definieras som ett tillämpningsområde. Det primära, när det gäller utvärdering, är med andra ord att tillämpa vetenskapliga metoder. Utvärderingsobjekten, dvs. det som utvärderas, skiftar däremot, och mångfald kan därför sägas vara kännetecknande för utvärderingsfältet.

Som framgår av vår uppräkningslista handlar det under 1970-talet i hög grad om att utpröva metoder för utvärdering. Men bland de arbeten som listas återfinns också några som mer explicit tar upp mer övergripande utvärderingsprinciper och frågor av mer

allmän karaktär. I vårt undersökningsmaterial från denna tid kan vi urskilja följande problemområden:

- Utvärdering som systematisk verksamhet
- Metoder och socialpolitisk påverkan från USA
- Målens och referensramens betydelse för utvärderingens inriktning

5.2.1 *Utvärdering som systematisk verksamhet*

Med den mångfald som kännetecknar utvärderingsfältet måste ”utvärdering” med nödvändighet ges en allmän definition. Vad utvärdering faktiskt innebär beror, som vi tidigare framhållit, på vad som utvärderas och hur man som utvärderare ser på sitt utvärderingsobjekt. I en avhandling om utvärdering i pedagogik [70:3] av Franke-Wikberg och Johansson (1975) brukas ”pedagogisk utvärdering” som precisering. Författarna understryker att deras avhandling gäller pedagogisk utvärdering av undervisning och inte utvärdering enligt organisationsteoretiska eller ekonomiska modeller. Det kan också uttryckas så att de avgränsar sitt utvärderingsobjekt eller tillämpningsområde till att gälla undervisning. Att fokusera på undervisningen som sådan eller på själva undervisningsprocessen var, som framgått, inte så vanligt i tidigare utvärdering av undervisningsinsatser. Snarare fokuserade man på studenternas prestationer och relaterade till mål uttryckta i beteendetermer.¹¹ I sin avhandling preciserar Franke-Wikberg och Johansson innebörden i utvärderingsbegreppet på följande sätt:

Utvärdering av undervisning är en jämförelse mellan observationer av och intentioner med undervisningen varvid informationsinhämtning samt beskrivning, analys och tolkning erfordras i syfte att möjliggöra en välinformerad bedömning av hur undervisningen fungerar. (Franke-Wikberg & Johansson, 1975, s 16)

¹¹ Jämför till exempel med 50-talsperspektivet genom Ralph Tyler: ”Evaluation is the process for determining the degree to which changes in behavior are actually taking place.”, ett perspektiv som fortfarande förekommer inom delar av den internationella utvärderingsdebatten.

I definitionen identifieras fyra aktiviteter som är utmärkande för och ingår i en utvärdering. Här nämns datainsamling, beskrivning, analys och tolkning i syfte att möjliggöra en välinformerad bedömning.

Vårt intryck är att det fortfarande är bilden av den neutralt forskaren som presenterar objektiva fakta som sätter sin prägel på vad pedagogisk utvärdering syftar till. När författarna klargör sin syn på utvärderarens uppgift ifrågasätter de samtidigt detta förhärskande synsätt. De formulerar sig på följande sätt:

Enligt vår mening är den väsentliga uppgiften för utvärderaren att ge en så god bild som möjligt av undervisningen ifråga, vilket inbegriper att relevant information insamlas och att undervisningen belyses ur olika utgångspunkter och aspekter i avsikt att försöka tolka och förstå den bild som framkommit. Oavsett vem som utför bedömningen torde underlaget härför visa sig vara tillräcklig i vissa avseenden och otillräcklig i andra. Det är således ibland nej ej alltid motiverat att avge värdeomdömen. I de fall det görs av utvärderaren bör utgångspunkter för ställningstaganden klargöras så att såväl premisser som bedömningar blir möjliga att ifrågasätta. (Franke-Wikberg & Johansson, 1975, s 15)

Som tidigare berörts fanns bland pedagogiska forskare vid början av 1970-talet en uttalad uppfattning om att pedagogisk utvärdering förutsätter pedagogisk teori och en uttalad teoretisk referens- eller tolkningsram.

Franke-Wikberg och Johanssons beskriver sina undersökningar som i huvudsak "explorativa och om metodprovande" (s 7) och rör undervisning inom tre universitetsämnen. Författarna ifrågasätter utvärderingens kontrollfunktion och framhåller istället vikten av att de som berörs av utvärderingen medvetandegörs och ges utrymme för reflektion. Synen på utvärderingens medvetandegörande funktion får konsekvenser för vem som förväntas värdera hur undervisningen fungerar. Författarna menar att bedömningen skall göras av studierektor i samråd med lärare, studerande och studievägledare, och de förespråkar såtillvida ett inifrånperspektiv vid utvärdering av undervisning. Ett sådant perspektiv gäller också den form av utvärdering som senare får benämningen "självvärdering". (Jfr. Holmer 2003).

5.2.2 *Politik och påverkan från USA*

Även Bell (1975) som har skrivit en avhandling om utvärdering av sociala program [70:2], för ett resonemang kring vad som kännetecknar utvärdering. Han framhåller att utvärdering ingår i en större beslutsgång och är en metod för att avgöra om ett planlagt program eller en insats uppnår de mål som anges. Bell ger en allmän beskrivning av vad utvärdering är och konstaterar att utvärdering innebär att medvetet ta ställning till frågan om ett program fungerar i enlighet med i förväg uttalade avsikter eller intentioner. En liknande allmän definition återfinns i citatet ovan från Franke-Wikbergs och Johanssons avhandling. Som inledningsvis påpekades återkommer en sådan allmän formulering också i de flesta definitioner som vi funnit i litteraturen utöver avhandlingar på området.

Utvärdering i Sverige avspeglar influenser från USA konstaterar Bell och pekar på kopplingen till såväl forskningstraditionen som den politiska utvecklingen i USA under 1960-talet. Intresset för utvärdering uppstod i samband med den förnyelse av socialpolitiken som initierades under Kennedys och framförallt Johnsons presidentskap. Bell pekar på kopplingen till den "New Deal-anda" som präglade förnyelsen och skriver: "Här finner vi den reformistiska tanken i kromosomformat" (s 31). Den nya given i den amerikanska politiken grundas på principen om social rättvisa och övertygelsen att politik kan fungera som ett medel för att uppnå centralt uppsatta mål liksom att kunna handla utifrån ett program eller en policy.

Att analysera grundprinciperna för ett företags eller en organisations handlande är viktigt, och policyanalys är kanske mest känd i säkerhetspolitiska sammanhang. Robert McNamara, som var försvarsminister under både Kennedys och Johnsons presidentskap, medverkade till att införa policyanalys som instrument i den federala förvaltningen. Den ideologi som policyanalys är ett uttryck för inrymmer tron på möjligheten att rationellt kunna granska, planera och förvalta offentliga verksamheter och institutioner. Grundtanken är att offentlig verksamhet istället för att "växa fram" kan styras och utvecklas mot tydliga mål och med generella medel som kan beslutas aktivt. Utvärdering

tillkommer i detta sammanhang och brukas för att granska de olika komponenterna i policyprocessen, dvs. hela kedjan av: mål, insats, genomförande, resultat och effekter.

Dokumenterade erfarenheter av utvärdering från efterkrigstidens reformperiod uttrycker, enligt Bell, besvikelser av olika slag. En besvikelse var att utvärdering inte uppfyllde de metodologiska krav som kunde ställas på samhällsforskning. Kritik riktades mot att utvärdering inte gjordes enligt en experimentell design, och en annan kritik gick ut på att utvärderingar ofta fick alltför kortsiktiga tidsperspektiv. Studierna var inte heller kontinuerliga så att komplexiteten i sociala program och reformer kunde kopplas till en mer teoretisk förklaring och förståelse av social förändring.

En annan besvikelse kom från samhällsforskare som, i en tid av reformistisk "New Deal-anda" såg sin roll som "aktivister", och som kritiserade utvärdering för att vara ett konservativt verktyg. Nixonregeringen på 1970-talet innebar ett slut på de stora sociala reformer, och i detta sammanhang framstod utvärdering som något positivt. Bell anger två skäl för detta: Utvärdering kunde å ena sidan visa att programmet inte haft någon effekt, vilket passade Nixonregeringens reformpolitik. Å andra sidan kunde en utvärdering också på olika grunder kritiseras och avfärdas som beslutsunderlag om den var för kritisk. Utvärdering kunde således tjäna skilda politiska intressen och på så sätt utgöra ett populärt verktyg i händerna på makthavarna.

De amerikanska erfarenheterna bör, enligt Bell, göra oss medvetna om att det i denna beslutsmodell inte finns någon självskreven intressegemenskap mellan de som styr och de som påverkas. Bell framhåller att utvärderingsforskaren måste vara mycket noga med sin egen ställning. Det är därför viktigt att den som arbetar med utvärderingsforskning tar frågan om: "utvärdering i vems intresse" som sin utgångspunkt.

Vi har i tidigare kapitel berört influenserna från USA på svensk pedagogisk forskning och då särskilt på pedagogisk utvärdering som specialområde inom denna ämnesdisciplin.

Undervisningsprogram och programutvärdering har sedan 1960-talet fått en allt vidare spridning och brukas idag allmänt i den pedagogiska diskursen. Som framgår av de tidigare resonemangen talar man om program av olika slag, exempelvis undervisningsprogram, programmerad undervisning och programutvärdering. Vi har också nämnt Ralph Tyler som myntade begreppet "Evaluation" och som under 1960-talet var ledande inom området utvärderingsforskning. Enligt Tyler är "program" närmast en etikett som kan klistras på skilda typer av insatser eller reformer, således även sådana som "the Poverty Program of the Great Society in the 1960's." I USA handlar programmen om att bekämpa fattigdomen i samhället. I Sverige handlar det om ett politiskt beslut om en nioårig grundskola för alla och undervisningsprogram som medel för en individualiserad undervisning, samt om att tillskapa en demokratisk skola. När Bell hävdar att utvärderingsforskaren måste vara mycket noga med sin egen ställning och ta frågan: "utvärdering i vems intresse" som sin utgångspunkt, blir medvetenheten om de politiska sammanhangen av särskild betydelse. (Jfr. Karlsson Vestman & Conner 2006, Karlsson Vestman 2006).

5.2.3 *Målen och referensramens betydelse*

I en avhandling av de båda författarna Dahlgren och Lindkvist (1976), som lagts fram inom ekonomisk historia [70:3], inleds traditionsenligt med en definition av begreppet "utvärdering". Författarna gör en åtskillnad mellan utvärdering och utvärderingsforskning och det är det senare som de främst intresserar sig för. Med termen "evaluation research" ansluter de sig till en gren av utvärderingsteori som betonar vikten av teoretiska perspektiv och medvetenhet om ramar vid tillämpningen av utvärdering, eller m.a.o. utvärderingspraxis. Enligt författarna har "evaluation research" inte utvärderingen som forskningsobjekt. I stället menar de att utvärderingspraxis måste göras vetenskaplig genom att vedertagna forskningsmetoder används för att garantera en objektiv och korrekt bedömning. Utvärderingsforskning tolkas därmed som en fråga om att tillämpa vetenskapliga metoder och inte som forskning om utvärdering.

Dahlgren och Lindkvist diskuterar några problem inom utvärdering och tar upp frågan om målens dubbla funktion att både ange i vilka dimensioner information skall insamlas och att vara den måttstock mot vilken den insamlade informationen skall värderas. Att mål ges denna centrala funktion kritiserar av en del som hävdar att officiella mål ofta är för allmänt hållna för att kunna utgöra en måttstock i mätteknisk bemärkelse. Författarna varnar för att en ensidig strävan efter teknisk mätbarhet kan reducera målens dubbla funktioner. Det som författarna kallar organisatoriska mål är uttryck för grundläggande synsätt och principer som förmedlar en idé om organisationens funktion och identitet. Dessa mål kan ses som "cultural entities" eller symboler som legitimerar organisationens handlande. Sådana mål är sällan tekniska. Snarare är de organisatoriska målen ofta grundade på en dominerande referensram som får karaktären av ett objektiva faktum, vilket dock sällan artikuleras och som aktörerna därför till stor del är omedvetna om. Detta medför att legitimiteten i den situationstolkning, utifrån vilken beslutsfattare agerar, måste problematiseras. Här kan, enligt Dahlgren och Lindkvist, utvärderarens uppgift vara att kritiskt granska den situationstolkning som i ett givet program framstår som naturlig eller meningsfull handling.

Författarnas resonemang om målens dubbla funktion eller karaktär kan illustreras med den aktuella läroplanen (Lp094) som anger kunskapsmål för den grundläggande utbildningen i sin helhet och som särskiljer mellan två typer av mål: "mål som skolan och undervisningen i ämnet skall sträva mot" och "mål som skall uppnås av alla elever." Dessa två typer av mål grundas på skilda synsätt och principer eller cultural entities med Dahlgrens och Lindkvists benämning. Vilken typ av mål som betonas beror på skillnader i grundläggande uppfattningar om hur elevernas lärande skall organiseras och undervisningen skall utformas (Andersson 1994). Dahlgrens och Lindkvists slutsats att utvärderarens uppgift kan vara att kritiskt granska den verklighetstolkning som i ett givet program framstår som naturlig eller meningsfull handling, är således alltså relevant.

5.3 1980-tal – Rationalitet och nytta ifrågasatt

Som framgått är inte utvärderingsforskning ett entydigt begrepp. Vi har pedagogisk utvärdering som huvudfokus. Sådan utvärdering rör uppfostran, utbildning och undervisning och förutsätter, som vi tidigare understrukit, pedagogisk teori. Vi tror samtidigt att det är viktigt att redovisa avhandlingsarbeten även från andra ämnesområden för att se hela spektrumet av de frågor som utvärderas. Uppräkningen av avhandlingstitlar från 1980-talet faller inom olika forskningsområden: medicin och omvårdnad, samhällsplanering och ekonomi, utbildning och undervisning. Även under denna tidsperiod fokuseras forskningen på utvärdering som tillämpningsområde och av de sammanlagt fjorton avhandlingar som publicerats från denna period med utvärdering som nyckelord, har vi funnit följande avhandlingar och teman inom olika ämnesdiscipliner.

I medicin och omvårdnadsforskning finner vi fyra avhandlingar som rör olika problem inom vårdområdet. En avhandling redovisar en utvärdering av testmetod för handens greppfunktion [80:1], en annan rör verksamheten vid Rygginstitutet i Sundsvall [80:2]. En gäller hörapparatutprovning [80:6] och en annan jämför friska personers utvärdering av patientinformation med patienternas utvärdering [80:14]. I ekonomi och samhällsplanering kommer ett par avhandlingar. Den ena handlar om utvärdering av nytta och kostnad i offentligt beslutsfattande inom trafikområdet [80:5] och den andra om utvärdering av företagsdemokratien och MBL [80:8]. I pedagogik finner vi sex avhandlingar med olika utvärderingsobjekt. En handlar om skolan som organisation [80:3], en annan om utvärdering av skolläroverutbildning [80:4] och en tredje handlar om granskning av inläring vid auditiv och visuell presentation [80:7]. Vidare återfinns en utvärdering av elevsituation och elevroll vid förberedelse och införande av samlad skoldag i en kommun [80:22] och en utvärdering av de studerandes ideologiska formning i fyra högskoleutbildningar [80:12], samt en utvärdering av relationen skola och arbete utifrån två empiriska studier av försök med att förstärka denna koppling [80:13].

Dessa arbeten är i huvudsak empiriska utvärderingar, men det finns också avhandlingar som har utvärdering som kunskapsobjekt. Vi ska ta upp ett par exempel nedan kring teman som:

- Utvärderingens värdemässiga aspekter
- Makt, legitimitet och rättvisa, kunskap och användning

5.3.1 *Utvärderingens värdemässiga aspekter*

Segraeus (1987) ser på utvärdering som ett värdebaserat politiskt instrument [80:10]. Hon menar att värdemässiga och politiska aspekter återfinns i utvärderingen på alla nivåer i organisationen då syftet med den institutionella utvärderingen är att ge underlag för politiska beslut. Uppdragsgivaren eller beställaren av utvärderingen är ofta politiskt ansvarig, och rapporten riktas också till politiskt ansvariga som mottagare. Det får konsekvenser för den som åtar sig utvärderingsuppdraget, enligt Segraeus. Hon hävdar att accepterandet av vissa målsättningar som kriterier för utvärdering också innefattar ett accepterande av underliggande premisser och att det är ett politiskt ställningstagande att ställa upp på ett programs målsättningar och sätt att arbeta.

Segraeus avhandling rör utvärdering av behandlingsmetoder inom institutionell narkomanvård och hör således till området: medicin och omvårdnadsforskning. Behandlingsforskningen innehåller ojämlika relationer mellan forskare, behandlare och patient, framhåller hon. Även då utvärderingsidealet innebär en situation av öppna relationer, dvs. att forskaren inte långvarigt allierar sig med en part, finns det i behandlingssituationer ett inbyggt maktförhållande mellan vårdare och vårdad (behandlare och klient). På samma sätt har utvärderaren genom sitt uppdrag möjlighet att påverka hela verksamhetens framtid och utvärderarens allianser blir därför av avgörande betydelse. Segraeus förespråkar ett "underifrånperspektiv" och att "forskningen skall skapa förståelse för de maktsvagas perspektiv." (s 12).

När det gäller utvärderingens karaktär urskiljer hon två tydliga inriktningar: utvärdering som forskning respektive utvärdering som utredning. I den förstnämnda betonas utvärderingsforskningens mer vetenskapliga syfte att bidra med en helhetssyn,

generera teori och bygga upp kunskap baserad på empiriska data, medan den senare är inriktad på "åtgärdsrelevans" som betonar manipulerbara faktorer. Skillnaden mellan de båda inriktningarna är inte endast en fråga om vetenskaplig metod, utan handlar även om kunskapsintresse; att utvärderingsforskningen utvecklas som en del av kunskapsfältet inom den disciplin där den är placerad, säger Segraeus, och hävdar att om utvärdering ska kallas forskning och inte utredningsverksamhet, så räcker det inte med att förfinna den metodiska arsenalen. Hon tycker inte heller att det duger med att vara betjänter till beslutsfattare och planerare. Snarare måste forskningen på ett mer fristående sätt värdera olika samhällsinsatser och dess effekter på kort och lång sikt för olika grupper. På så sätt är det inte mycket som skiljer utvärderingsforskning från annan forskning, konstaterar hon.

En jämförelse mellan exempelvis Dahlgrens och Lindqvists (1976) [70:5] och Segraeus ståndpunkt visar således på olika sätt att betrakta utvärderingsforskning: som en (vetenskaplig) metod inom utvärderingsfältet, respektive som en metautvärdering eller forskning *om* utvärdering.

5.3.2 *Rättvisa, kunskap och användning*

Lander (1987) [80:9], som betraktar utvärdering som ett självständigt projekt, gör i sitt avhandlingsarbete en genomgång av den utvärderingsdebatt som förts under 1970- och 1980-talet kring så kallade programutvärdering. Han strukturerar utvärderingsdebatten enligt två huvudlinjer: Den ena rör makt, legitimitet och rättvisa, och inom den andra, som rör kunskapsbildning, återfinns diskussioner om utvärderingens användare, användning och användbarhet. Lander beskriver utvärdering som en verksamhet som är inriktad på att studera målinriktad verksamhet, vilket innebär att studera verksamhetens värdebas och rationalitet samt villkor och konsekvenser av dessa. Han utvecklar denna ståndpunkt genom att hävda att värdena identifierar vad som anses eftersträvansvärt, medan rationaliteten med hjälp av värden och andra föreställningar skapar mening kring det som eftersträvas och ger en handlingsinriktning åt strävan.

För Lander är värdebasen något som inryms i utvärderingens undersökningsobjekt och tycks inte på förhand given. När det gäller utvärdering är rationalitet ett av de nyckelord som han diskuterar ingående. Rationaliteten tar sig enligt Lander olika former: professionell, ekonomisk och administrativ rationalitet, men också i form av en rationalitet inom områden som politik (politisk/ideologisk rationalitet), moral (moralisk rationalitet) och vetenskap (vetenskaplig rationalitet). För varje sådant område utpekar Lander också olika områden för analys av principerna för handling och meningsskapande: För politik handlar det exempelvis om världsbild, intresse och strategi och för vetenskap om sanning, validitet, metod och stringens.

Exemplen visar att det under 1980-talet sker en problematisering av utvärderingens värdegrund. Inom pedagogisk utvärdering med sin ofta starka koppling till skolans offentliga mål, blir det viktigt att hantera problem som handlar om utifrån vilka principer eller rationalitet som politiker, administratörer, lärare, föräldrar och elever formulerar sina krav och förväntningar på skolan.

5.4 1990-tal – Intressenter, styrning och makt

Under 1990-talet ökar antalet avhandlingar som rör utvärdering och omfattar närmare ett fyrtiotal arbeten/publikationer. Fortfarande är det utvärdering som tillämpning eller vetenskaplig metod som dominerar. Endast en handfull av studierna har utvärdering som kunskapsobjekt. Även här inleder vi med att redovisa avhandlingarna genom att ange studiernas inriktning i tabell 2 nedan. Det kan vara på sin plats att påminna om att utvärderingstemat inte alltid är särskilt uttalat eller genomarbetat i många av dessa arbeten. Ofta definieras en avhandling som ”utvärdering” när man använt någon teknik som förknippas med utvärdering som vetenskaplig metod. Inom det ämnesdidaktiska fältet återfinns exempelvis flera avhandlingar som till sitt innehåll mycket litet berör eller anknyter till utvärdering eller utvärderingsforskning, trots att det anger utvärdering som ett nyckelord. I vårt urval hittar vi från detta decennium avhandlingar från följande forskningsområden och teman:

Tabell 2: Teman inom utvärderingsforskning under 1990-talet

Medicin, hälsovård, kriminalvård, socialt arbete – 9 teman

-
- | | |
|--|--|
| - Medicinsk rådgivning på vårdcentraler [90:2] | - Chefsutbildning för läkare [90:12] |
| - Modeller för äldreomsorg [90:3] | - Socialt förändringsarbete [90:37] |
| - Personalutbildning i äldreomsorg [90:5] | - Utvärdering i socialt arbete [90:1] |
| - Feministiskt perspektiv - teknisk mätning [90:7] | - Utvärdering av familjeterapi [90:38] |
| - Jämförande av metoder i smärtlinning [90:11] | |

Teknik, byggforskning, samhällsplanering – 6 teman

-
- | | |
|---|---|
| - Överglasning av stora byggnadsvolymer [90:14] | - Utvärdering av bostadsområden [90:26] |
| - Bedömning av byggforskning [90:24] | - Bedömning av kvalitet i boendet [90:27] |
| - Metoder vid exploateringsplanering [90:29] | - Utvärdering av färdtjänst [90:32] |

Ekonomi – 6 teman

-
- | | |
|--|---|
| - Modeller för mätning av skogsbestånd [90:33] | - Servicemöten vid charterresor [90:17] |
| - Belöningar i offentlig verksamhet [90:15] | - Användning av biobränsle [90:21] |
| - Organisation och personalenkäter [90:19] | - Kommunala verksamhetsmått [90:28] |

Datavetenskap, arbetsvetenskap, psykologi, sociologi, statsvetenskap – 6 teman

-
- | | |
|---|--|
| - Kvalitet i forskning [90:6] | - Metoder vid presidentval [90:30] |
| - Interaktion mellan människa och dator [90:10] | - Utvärderingens teori och praktik [90:31] |
| - Implementering och utvärdering [90:20] | - Metoder för modellutformning [90:34] |

Utbildning och undervisning – 11 teman

-
- | | |
|---|---|
| - Relativa betyg – teori och historia [90:4] | - Ungdomar utan gymnasieskola [90:18] |
| - Matematik ur elev- och lärarperspektiv [90:8] | - Kriterieproblemet vid utvärdering [90:23] |
| - Gymnasiekulturer [90:9] | - Taldatorbaserade läsprogram [90:25] |
| - Utbildning i Kina och studentprestationer [90:13] | - Utvärdering av förskolan [90:35] |
| - Förändrad styrning av skolan [90:16] | - Läsdebatten som kontrovers [90:36] |
| - Utvärdering av sjuksköterskeutbildning [90:22] | |

Nedan kommenteras ett antal avhandlingar som publicerats under 1990-talet och som mer explicit diskuterar utvärdering som teori och praktik. Teman och frågor som behandlas rör:

- Rationalitet och institutionalisering
- Externt och internt perspektiv
- Styrningsfrågan
- Kriterier och intressenter
- Maktrelationer ur ett feministiskt perspektiv

5.4.1 *Rationalitet och institutionalisering*

Diskussionen om utvärdering som rationalitet fortsätter under 1990-talet. Åberg (1997) [90:31] analyserar utvärderingens legitimitet. Han delar i grunden Landers (1987) [80:9] syn på utvärdering som rationell verksamhet. Enligt Åberg är utvärdering något mer än att beskriva ett kunskapsinnehåll. Han ser utvärdering som en särskilt organiserad verksamhet för produktion och bedömningar. Vidare är utvärdering något som görs i efterhand och gäller konkreta utvärderingsobjekt. Syftet är att göra en bedömning av verksamhetens eller insatsens värde. Bedömningen grundas på ett antal uttalade eller outtalade värdepremisser som är avgörande för vad man som bedömare värderar vara gott eller inte gott. Det kan röra sig om konkreta mål eller om högre ideal som rättvisa, jämlikhet eller delaktighet.

Utvärderingen får sin legitimitet genom rationaliteten den är uttryck för säger Åberg och han kopplar rationaliteten till olika handlingar som att producera kunskap och att systematiskt göra en bedömning. En skillnad kan göras mellan utvärdering i en mer allmän betydelse, något som människan alltid ägnat sig åt, och utvärdering i en specifik situation, där värderingen blir en verksamhet ”i sig”. Detta uppstår när värderandet institutionaliseras och blir en verksamhet ”för andra”, något som utförs med specifika metoder och av specialister. Åberg menar att utvärderande ”för andra” uppstår genom att man börjar se rationalistiskt på sociala problem och inrättar serier av rationella planerade och specialiserade åtgärder – ”program” – för att lösa problemen.

5.4.2 *Externt och internt perspektiv*

Att skilja mellan utvärdering som sker med ett utifrånperspektiv och utvärdering med inifrånperspektiv är en viktig poäng i Eriksson och Karlssons (1990) avhandling [90:1]. Utvärdering *av*, i detta fall socialt arbete, betecknar enligt författarna ett utifrånperspektiv som karakteriseras av att utvärderaren har distans till verksamheten. Utvärderingen sker i efterhand och är tillfällig. De som beställer sitter i beslutsposition och utvärderingens syfte är att ge beslutsunderlag eller generell kunskap. Som

kontrast till detta ställer Eriksson och Karlsson utvärdering i socialt arbete som sker med ett inifrånperspektiv på verksamheten (där utvärderaren ofta är del i verksamhet) och som görs för att direkt påverka arbetet. Författarna menar att metoderna för denna slags utvärdering är outvecklade.

Eriksson och Karlsson diskuterar också förhållandet mellan utvärdering och utvärderingsforskning och ser den senare som forskning *om* utvärdering i syfte att ”producera för praktiken relevant kunskap och bidra i metodutvecklingen” (s 17). Utvärdering som tillämpning eller metod kan, menar författarna, placeras på en skala med två poler: ”vetenskapliga metoder” respektive ”praktikanpassade metoder”. Forskning *om* utvärdering är en metaposition som ”innebär utveckling, granskning och förändring av utvärderingens praktik” (s 17).

5.4.3 *Styrning och decentralisering*

Som vi i ett tidigare kapitel har understrukit så spelar styrningsfrågan en central roll för utvärderingens funktion. Styrningsfrågan tas också upp av Johansson och Johansson (1994) [90:16]. I sin avhandling går författarna igenom olika perspektiv på och definitioner av utvärdering. Fokus ligger på styrningen av skolan genom officiella dokument på olika nivåer – nationella måldokument, kommunala skolplaner och lokala arbetsplaner och hur de mål som formuleras kommer att konkretiseras på lägre nivåer. Enligt författarna är utvärdering ett centralt inslag i en ny styrstrategi för skolan under 1980-talet och början av 1990-talet. Utvärdering är med andra ord en del av ökad målstyrning. Enligt författarna förutsätter målstyrning att det sker utvärdering, medan däremot det motsatta – utvärdering förutsätter målstyrning – inte gäller.

Johansson och Johansson framhåller att utvärdering under 1980-talet kommit att knytas till styrning av skola och utbildning med effektivitet och produktivitet som vägledande principer för handling. Därvid har utvärdering ofta kommit att förknippas med negativa värden som kontroll och förtryck. De analyserar officiella dokument och finner två huvudmotiv för utvärdering:

kontroll och förändring. I det betänkande och den proposition som föregick beslutet om en ny nationell läroplan (Lp094) menar de att kontrollmotivet har förstärkts. Författarna ger belägg för sitt påstående genom uttryck som: ”mätning av resultat, resultatbedömning, grad av måluppfyllelse och utvärderingsbara mål” (s 113).

I den fortsatta diskussionen fokuserar de på utvärderingens styrfunktion och menar att utvärdering kan ha en stark styrförmåga i sig själv på alla nivåer. Vid målstyrning av skolan kan man fråga om det är målen eller utvärderingen (såväl den genomförda som den föreskrivna, t ex proven) som är mest styrande. Det finns mycket som talar för det senare, anser författarna. Skolans mål är många gånger beskrivna på ett övergripande och allmänt sätt och de tolkningar och preciseringar som kan göras i en utvärdering kan bli riktninggivande i verksamheten för att uppnå högre måluppfyllelse. Författarna frågar huruvida denna förskjutning mot mätbarhet genom utvärdering också innebär en förändring i tolkningen av målen från det socialt värdefulla till det mätbara.

Som vi framhållit har utvärdering, med den ökade decentraliseringen, blivit ett allt starkare instrument för politisk styrning och kontroll. Samtidigt har andra former av centralstyrning tagits i bruk, t.ex. de särskilda statliga satsningarna, som med den förändrade styrstrategin blivit ett medel för indirekt styrning – en centralstyrning nerifrån och inifrån. Under den senaste tioårsperioden har ytterligare en nivå – EU – tillkommit, och dessutom börjar ännu en nivå utkristalliseras i form av Europas framväxande regioner. Det visar att den politiska styrningen ställts inför allt mer komplexa förhållanden där nya styrformer och förskjutningar mellan olika nivåer i styrsystemet kommer att utformas och brukas, något som också gäller nya former av utvärdering. Ett exempel som vi har pekat på är de internationella proven som styr inte bara vad som händer på nationell nivå utan också kan påverka vad som faktiskt kommer att undervisas i den lokala skolan.

5.4.4 *Kriterier och intressenter*

Diskussionen om styrning är kopplad till frågor om vilka aspekter eller insatser som skall utvärderas samt vilka kriterier för bedömning som skall gälla. Ett exempel på ett mer detaljerat instrument med standardiserade skalor har gjorts av Andersson (1999) som i sin avhandling [90:35] har utvärderat en skala för bedömning av kvalitet i förskolan, kallad ECERS-skalan, vilken innehåller sju huvudkriterier som är målrelaterade. Andersson diskuterar olika problem som finns med denna typ av mätinstrument, inte minst då valet av kriterier på den bedömnings-skala som används. Andersson frågar vilka värderingar som styr definitionerna av de olika skalstegen och om det finns någon prioritering eller rangordning av de olika kriteriernas betydelse inom de olika bedömningsområden som skalan fokuserar på. Är exempelvis omdömet ”utmärkt” av samma dignitet inom olika bedömningsområden som ”Bygglek”, ”Föräldrainflytande”, ”Sand och vatten”, ”Stämning/atmosfär” eller ”Förståelse för andra kulturer och könsroller”? Olika föräldrar, barn och lärare kanske har olika syn på vilka aspekter som väger tyngst. Vad gör man med resultaten om man i denna fråga är oenig? Kanske fungerar skalan bäst som en konsumentupplysning till vilket var och en själv tar ställning, resonerar Andersson.

Kriteriefrågan och utvärderingens relation till intressenterna är frågor som också behandlas av Karlsson (1995). I avhandlingen [90:23] undersöker han hur man kan väga de olika berörda gruppernas önskemål och kriterier mot varandra i en utvärdering. Vem tar man hänsyn till när konflikter uppstår mellan olika prioriteringar av kriterier. Med utgångspunkt från ett konkret arbete med intressentutvärdering visar Karlsson på det demokratiska perspektivet vid utvärdering. Det är en aspekt av utvärdering som sällan brukar vara explicit uttryckt i de traditionella definitionerna av utvärdering. Karlsson visar också att intressentperspektivet har en politisk dimension. En intressentutvärdering kan leda till en upplevd delegitimering av det politiska mandatet genom att:

Utvärderingen uppfattas bli en tummelplats för stridiga viljor, konflikter och intressemotsättningar som hindrar politikerna från att gå vidare och själva ta beslut utifrån sina politiska mandat. (Karlsson 1995, s 167)

Intressentmodellen kan även fungera legitimerande genom att politiska makthavare får möjlighet att peka på att man valt en utvärderingsmodell som representerar användarna, dvs. väljarna och att man är beredd att ta hänsyn till deras åsikter. Frågan är hur olika intressenter i praktiken kan föra fram sina åsikter och göra sitt perspektiv gällande. Råder lika villkor för olika intressenter? För att utvärderingen inte helt ska kunna åsidosätta samhällsfrågor och maktperspektiv bör den, enligt Karlsson, använda kriterier utöver de traditionella och rationellt baserade – mål, behov, nytta, effektivitet – och även inkludera värden som rättvisa och jämlikhet.

Ett pluralistiskt perspektiv med flera tänkbara aktörer som intressenter vid utvärdering finns även i en avhandling [90:4] av Andersson (1991) där en studie om relativa betyg redovisas utifrån ett historiskt perspektiv. Han placerar betygssättningen inom en vidare ram än den mättekniska och betonar i synnerhet lärarens maktposition i relation till eleven. Andersson noterar att eleverna är minst representerade i betygsdebatten och konstaterar också att lärare, föräldrar och elever i betygsfrågan har svårare att få igenom sina åsikter än företrädare för tunga politiska organisationer och företrädare för arbetslivet.

Andersson visar hur synpunkter från arbetsgivarorganisationen (SAF) fått ökad genomslagskraft. I diskussionen om betyg fanns, i början av 1990-talet, förslag om att införa betygskriterier på personlighetsaspekter som initiativförmåga, kreativitet och samarbetsförmåga i ett målrelaterat betygssystem. Andersson framhåller att detta skulle leda till ytterligare problem i den maktkontext som betygen omges av. En viktig fråga i detta sammanhang torde komma att gälla skolans eventuella rätt att hävda vissa personlighetsegenskaper hos eleven, trots att föräldrar och elev kan ha avvikande uppfattningar, förutspår Andersson. I sin avhandling pekar han också på den ökande betydelse som frågor om styrning och kontroll kan komma att få, medan andra betygsfunktioner, som rör information, motivation och urval, gradvis kan få minskad betydelse.

5.4.5 *Feministiskt perspektiv och maktfrågan*

I flera av de avhandlingar som vi har kommenterat i detta kapitel har frågan om ojämlika maktrelationer lyfts fram. Det görs också i en avhandling [90:7] av Morgall (1991) där författaren, utifrån ett genusperspektiv, diskuterar hur ojämlika maktrelationer kommer att påverka utvärderingsresultatet. I sin avhandling visar hon hur kvinnors röster exkluderats inom det utvärderingsområde som benämns "Technology Assessment" (TA) och som definieras som "försöket att analysera och förstå konsekvenserna av en viss teknik eller teknologisk trend" (s 23) och där målet är att man bättre ska kunna förutse behov av reglering.

Morgall kritiserar denna typ av teknisk utvärdering för att utgesig för att vara en politiskt neutral verksamhet. Hon pekar på konsekvenserna för genus genom att visa att valet att fokusera enbart på fysiska variabler är exkluderande och utgör ett legitimt sätt att ignorera genus. Hon ser ett behov av ett särskilt feministisk förhållningssätt till TA, (och utvärdering av teknisk natur) i synnerhet när det handlar om teknologi för födande och reproduktion. När det gäller just detta område pekar hon på att patriarkala relationer avspeglas genom att tillgången på preventivmedel kontrolleras och därmed behålls kontrollen över den kvinnliga reproduktionen.

5.5 2000-tal – Forskning med pluralism

Antalet avhandlingar med utvärdering som nyckelord och centralt begrepp fortsätter att öka under 2000-talet. Redan efter ett halvt decennium har ett drygt 40-tal studier publicerats inom skilda forskningsområden. Jämfört med 1970-talet, när utvärderingsforskning etablerades i Sverige, innebär det fyra gånger så många avhandlingsarbeten. I vårt urval av finner vi under detta decennium avhandlingar inom följande forskningsområden och teman:

Tabell 3: Teman inom utvärderingsforskning under 2000-talet

Medicin, hälsovård, socialt arbete – 12 teman

- | | |
|--|---|
| - Behandling barns synfel [00:3] | - Självuppskattning vid friskvård [00:24] |
| - Utvärdering av mammografi [00:5] | - Anpassning till boende och ADL [00:26] |
| - Metoder för höftledsoperationer [00:6] | - Utvärdering av index i socialtjänst [00:31] |
| - Klinisk bedömning vid handkirurgi [00:7] | - Social och ekonomisk utvärdering [00:32] |
| - Arbetsterapi och smärtlindring [00:9] | - Bedömning av psykoterapi [00:33] |
| - Ängslan vid psykologiska test [00:16] | - Dokumentation i socialtjänsten [00:36] |

Teknik, byggforskning, samhällsplanering – 6 teman

- | | |
|---|--|
| - Upplevelsen av ljudsystem [00:11] | - Cykelfält och säkerhet [00:21] |
| - Trafiksäkerhetsåtgärders effekter [00:18] | - Energi och inomhusklimat [00:41] |
| - Strategier för miljöarbete [00:22] | - Utvärdering av industriventilation [00:42] |

Ekonomi – 7 teman

- | | |
|---|--|
| - Prestationer på internationella mässor [00:4] | - Ranking och ledarskapsutbildning [00:29] |
| - Sociala program - arbetsmarknad [00:14] | - Valfärd och jämlikhet [00:38] |
| - Vuxenutbildning - arbetsmarknad [00:23] | - Institutionella förändringar - lönepolitik [00:45] |
| - Företags konkurrenskraft [00:35] | |

Datavetenskap, arbetsvetenskap, informatik, statsvetenskap – 5 teman

- | | |
|---|---|
| - Skrivsårigheter vid datoranvändning [00:15] | - Lokal utveckling av folkbibliotek [00:30] |
| - Utvärdering av info-system [00:17] | - Utvärdering i staten och högre utbildning [00:27] |
| - IT-användning och nytta [00:19] | |

Utbildning och undervisning – 15 teman

- | | |
|---|---|
| - Bedömningar i gymnasiet [00:1] | - Validering av bedömningar [00:28] |
| - Bedömning av kognitiva förmågor [00:2] | - ICT utvärdering och lärande [00:34] |
| - Utbildningsplanering [00:10] | - Bedömning av test i matematik [00:37] |
| - Språkundervisning och läromedel [00:8] | - Projektutvärdering i Tanzania. [00:39] |
| - Läraryrket [00:12] | - Gymnasiet, betyg och prov [00:40] |
| - Frihet och kontroll i 6-årsverksamhet [00:13] | - Bedömning i historiskt perspektiv [00:43] |
| - Redskap för analys av skolan [00:20] | - Kvalitet och styrning av skolan [00:44] |
| - Examensarbeten i läraryrket [00:25] | |

De avhandlingstitlar som listats antyder att utvärdering i huvudsak alltjämt är ett vetenskapligt tillämpningsområde. Avhandlingarna avspeglar också både den inomvetenskapliga förändring som fått genomslag i det senmoderna samhället och de

samhällsproblem som står i förgrunden i början av sekelskiftet 2000. Följande övergripande problemområden, som diskuteras i den fortsatta framställningen, ger en bild av de tendenser som vi urskiljer i de avhandlingsarbeten som lagts fram under de första sex åren av det nya seklet:

- Mätbarhet av miljö och energi
- Konkurrens och ekonomisk jämförbarhet
- Nya vetenskapliga inriktningar
- Intressentperspektiv och politisk styrning
- Granskning av kunskap och kvalitet

5.5.1 *Mätbarhet av miljö och energi*

I den mer tekniskt orienterade utvärderingslitteraturen finns ett jämförande bredare perspektiv på utvärdering som verkar växa fram under 2000-talet. Ett exempel på hur tekniskt orienterade utvärderingar visar på dilemman som finns genom att peka på flera, ibland motstridiga variabler som bör vägas in kan exemplifieras av Nilsson (2003) [00:22]. I en studie av cyklisten och biltrafik inför hon flera parametrar och utvärderar såväl cyklisters säkerhet som deras förmåga att tävla med biltrafik. Att ta hänsyn till flera parametrar kan tyckas självklart. Nilssons val illustrerar dock ett exempel på hur utvärderaren problematiserar sitt uppdrag och vidgar dess perspektiv. Nilsson argumenterar för sitt val genom att hänvisa till en studie av cykelns konkurrenskraft gentemot bilen som talar till cykelns fördel, medan däremot resultaten, när det gäller cyklisters säkerhet, är betydligt dystrare. Här finns således en risk för att cyklister invaggas i falsk trygghet, menar Nilsson. Samtidigt framhåller hon att denna dubbla bild visar på styrkan med att ta in flera variabler i bedömningen. En utvärdering som endast hade gjorts med hänsyn till cykelns konkurrenskraft gentemot bilen hade visat att cykelfälten var mycket lyckade, medan utvärdering endast med hänsyn till cyklisters säkerhet hade lett till slutsatsen att cykelfälten inte hade den önskvärda effekten.

Två avhandlingar i vårt urval rör komplexa problem i dagens högteknologiska samhälle där fokus har flyttat från massproduktion till hållbar utveckling. Energiförsörjning och miljö, i såväl

bostäder som industrianläggningar är här aktuella problemområden och utvärderingsforskningen handlar fram för allt om att utpröva ansatser och tekniker som är tillförlitliga och ger ett säkert underlag för bestämda förbättringsinsatser, vilka på sikt skall leda till optimala lösningar. Den ena avhandlingen inom teknikområdet är skriven av Karlsson (2006) [00:41] och handlar om en flerdimensionell ansats för att utvärdera inomhusklimat i lågenergibyggnader. Den andra avhandlingen är skriven av Lindhe (2006) [00:42] som studerar problem med dåliga arbetsmiljöer inom industrin. Han beskriver att det redan vid slutet av 1800-talet noterades att typografer inom grafisk industri drabbades av sjukdom och tidig död orsakad av höga halter av blydamm i luften. Olika åtgärder för att komma till rätta med problemen provades, och liknande problem har sedan dess varit aktuella och ett angeläget område inom arbetsmiljöforskning. Lindhe utvärderar hur olika system för ventilation fungerar och med vilken effektivitet.

5.5.2 *Konkurrens och ekonomisk jämförbarhet*

För och nackdelar med ranking är den fråga som Wedlin (2004) [00:29] diskuterar i sin avhandling. Han granskar utvecklingen av internationella rankinglistor bland handelshögskolor mot bakgrund av teorier om formationer av fält – en teori som utvecklats av Pierre Bourdieu. Med denna teoretiska referensram vill Wedlin visa hur rankinglistor kommer att fungera som arenor för att kartlägga ett landskap av handelshögskolor och positionera de medlemmar som finns i detta fält i relation till varandra. På detta sätt skapas en gemensam identitet bland medlemmar som ingår i fältet, men också en legitimering av vad som utgör en god utbildning. Wedlins arbete illustrerar hur klassifikation i sig har en funktion att utforma organisationsfält och skapa identitet. Utvärderingen blir i detta sammanhang en del i den pågående debatten för att bestämma värde, genom att föra upp debatten på en metanivå. En tanke som prövas är att värde inte går att bestämma utifrån en på förhand definierad kriterielista, utan är just det som utformas genom utvärderingsprocessen som leder fram till en rankinglista, vilken i sin tur ger hållpunkter för vad som kan anses utgöra god standard.

Även Jess (2005) [00:32] diskuterar en relativisering av utvärderingen i sitt försök att anpassa socioekonomiska modeller för utvärdering av socialt arbete. Genom att jämföra olika metoder och deras utfall på några olika program kan Jess studera om en flerdimensionell utvärdering, som använder sig av olika metoder, kan ge bättre förståelse än en enskild utvärderingsansats. Jess menar att förståelse för konsekvenser av olika skillnader i mätteknik och utformning av olika metoder är nödvändig för att det skall vara möjligt att nå en korrekt värdering/bedömning. Hon grundar, liksom Wedlin, sin avhandling och därmed utvärderingen i en teoretisk bas. I Jess fall handlar det om olika syn på utilitarism med referenser till Mills, Bentham, Rawls m.fl.

Ekonomisk tillväxt och rättvis fördelning av resurser är motto för det jämlika och demokratiska samhället. Detta motto brukar också förknippas med den svenska välfärdsmodellen. Som vi berört i tidigare avsnitt möjliggjorde industriutvecklingen efter det andra världskriget en ekonomisk tillväxt som kunde brukas för sociala reformer inte bara i Sverige utan också som bistånd till utvecklingsländer. Rättvis fördelning är ett moraliskt dilemma i vårt individualistiska samhälle och diskuteras ofta i massmedia. Principen ”åt var och en efter behov av var och en efter förmåga” är svår att tillämpa i praktiken, och en utvärdering av en specifik fördelningsmodell kan därför vara en lösning. En avhandling som rör denna problematik är Bångman (2006) [00:38]. Hon diskuterar olika kriterier för att avgöra hur en rättvis fördelning ska se ut. Exempelvis en inkomstfördelning som uppfyller krav på rättvisa. Hon granskar en modifierad (viktad) Cost-benefit modell som ett alternativ vid utvärdering av välfärd. Avhandlingen berör ett centralt problem i utvärderingsforskning om vilka kriterier som kan väljas för bedömning.

I det högteknologiska samhället gäller också en global marknadskonkurrens. Liksom på börsen, som handlar om marknadsprognoser, hittar vi en utvärderingsforskning inom företagsekonomi som är inriktad på att göra prognoser om regionala företags internationella konkurrenskraft. Ahlström Söderling (2006) [00:35] har utvärderat en undersökningsmetod för att bedöma regionala företags förutsättningar för internationell konkurrens-

kraft. Konkurrens förutsätts vara en drivkraft för förbättring inte bara inom företag på den globala marknaden utan också inom avgränsade verksamhetsområden inom såväl privat- som offentlig sektor. En förändrad policy inrymmer tänkta effekter och utvärdering har sedan länge varit självklar i sådana sammanhang. Ahlström Söderling konstaterar att den studerade modellen (Best Business Practice) kan vara användbar om man håller i minnet att den ger en mycket förenklad bild av verksamheten. Ett konstaterade som kan sägas om de flesta modeller som genom att mäta processer på ett fåtal variabler (i detta fall sex processvariabler) försöker visa något om komplicerade verksamheters kvalitet.

5.5.3 *Nya vetenskapliga inriktningar*

Skolan, liksom sjukvården och socialtjänsten är komplexa verksamheter med insatser av många olika slag som förutsätter professionella kompetenser på vetenskaplig grund. En professionell kompetens inom dessa verksamhetsområden är i hög grad också erfarenhetsbaserad och grundas på ett lärande av den egna praktiken – lärdomar som under de senaste decennierna beskrivits som ”tyst kunskap”. För att synliggöra professionell kompetens har praktikerna stimulerats att dokumentera och ge spridning åt sina erfarenheter. Som tidigare framhållits ställdes exempelvis, i samband med den särskilda satsning på skolutveckling med IT som genomfördes med medel från KK-stiftelsen, krav på de lärare som beviljades projektmedel att de skulle ge spridning åt sina erfarenheter. Detta förutsatte både självvärdering och en systematisk dokumentation men förverkligades långt ifrån i alla projektgrupper.

Alexandersson (2006) [00:36] tar upp frågan om implementering av systematisk dokumentation för verksamhetsutveckling inom socialtjänstens verksamhetsområde. Hennes utgångspunkter är att granska den metod för dokumentation som prövats i socialtjänsten för att systematiskt kunna bedöma resultatet av olika insatser. Hon konstaterar att den studerade metoden är optimal om den kan genomföras fullt ut, men ställer den kritiska frågan om detta är en realistisk ambition. Baksidan kan bli en standar-

disering av det sociala arbetet och metoden kan riskera att göra socialarbetaren mindre lyhörd för de behov som ligger utanför dokumentationsmetodens frågor och kriterier.

Att vara lyhörd för det lokala och specifika i situationen är en fråga som även Zetterlund (2004) [00:30] diskuterar. Hennes avhandling handlar om hur lokal utvärderingspraxis, kopplad till programutvärderingar för förändring, kan relateras till institutionella förhållanden. Zetterlund tar utgångspunkt i hermeneutiska och fenomenologiska forskningsprinciper och drar slutsatsen att utvärdering är ett komplext fenomen, där tre huvudsakliga mönster eller strategier kan identifieras: (1) participatory evaluation, (2) rational-managerial evaluation, och (3) pragmatic-political evaluation. Zetterlund menar att utvärderingspraxis skiljer sig avsevärt från det vetenskapliga expertideal som den ofta baseras på. Hon menar vidare att utvärdering inte kan ses som ett neutralt och tekniskt redskap. Snarare producerar och reproducerar den institutionella identitetens legitimitet.

Ett annat exempel på ett skifte av perspektiv från det objektiva till ett mer subjektivt ideal för forskarens och utvärderarens roll representeras av Germeten (2002) [00:13]. Hon undersöker hur undervisningsbegreppet konstrueras hos lärare för 6-åringar. Underlaget för studien är i form av fältobservationer, *case studies* och intervjuer, vilket är ett vanligt metodval i många av de undersökta avhandlingarna där utvärderingar ofta utgör en del av arbetet. Germeten gör en poäng av skillnaden mellan olika forskningstraditioner i form av den ”traditionella” som forskar på objekt, och de hermeneutiska, fenomenologiska, kritiska och poststrukturella ansatserna med fokus på subjektet. Hon ifrågasätter flera av de vanliga ståndpunkterna vid traditionell utvärdering i positivistisk tradition om att vara objektiv, rationell etc. Mot objektet ställer hon subjektet, det objektiva språket ersätts av ett subjektivt. Mot rationaliteten ställer hon intuitionen. Hennes teoretiska utgångspunkt är Michael Foucault och hans uttolkares och efterföljares teorier kring makt och maktrelationer. Detta kommer också att påverka Germetens syn på forskarens roll och framställningssätt genom medvetenheten om att också detta kan ingå i en maktpositionering.

Germeten placerar sig i en postmodern och dekonstruktionistisk tradition, i bl.a. Paul Ricoeurs efterföljd, och pläderar för en medveten perspektivförskjutning där utvärderaren/forskaren har möjlighet att välja en utgångspunkt för att peka på en viss aspekt. Sanningsanspråk ses som subjektiva och situationsanpassade. Vilken forskning och metod som ger resultat, beror på det avsedda syftet. I slutänden är val av forskningsinriktning för Germeten ett riskfyllt etiskt val. Utvärderaren gör sig därmed till del av undersökningen, genom en medveten positionering av såväl undersökningsobjekten som sin egen person som ett görande och handlande subjekt. Detta påverkar också framställningssättet menar hon i följande citat:

Det är inte ovanligt att läsa forskning som framställer det akademiska språket som neutralt, där 'man' eller 'en' presenterar forskning om 'någon annan'. I den här avhandlingen har jag positionerat mig själv i förhållande till de subjekt och teman jag har skrivit om genom att konsekvent använda jag-form när jag (be)skriver eller dryftar. Detta för att markera mig själv tydligare som ett handlande och *görande* subjekt. För det andra handlar det om att följa sin intuition. Eftersom rationalitet har betraktats som den viktigaste forskaregenskapen, har intuition [som sans] blivit översatt i akademisk tradition. [vår övers]. (Germeten, s 182)

Germetens resonemang leder över till Tornbergs (2000) [00:8] avhandling som kan sägas utgöra del av en kritik mot tidigare rådande kunskaps- och vetenskapssyner. Tornberg betonar också en stark teoretisk bas med utgångspunkt hos bl.a. Michael Foucault och dennes diskussioner om diskursens förhållande till makt. Liksom Germeten anknuter hon till filosofen Cherryholmes (1988) syn på vetenskap med ett pragmatiskt synsätt som innebär att man kan överge sökandet efter den ultimata sanningen, utan att detta måste leda till en total relativisering. Värden är inte på förhand givna utan delar i samtalet.

Uppmaningen är att hålla diskussionen levande om vad som kan ha värde, vilket förutsätter ett ständigt pågående samtal från många olika utgångspunkter, ständiga omprövningar och ett ständigt nytänkande när det gäller konsekvenserna av våra handlingar, säger Tornberg. Hon menar att ytterst innebär det pragmatismiska förhållningssättet en perspektivförskjutning dels från epistemologi med sina frågor om vad som är kunskap och

hur kunskap uppstår, till frågor om moraliska relationer mellan människor. Förhållningssättet innebär också en förskjutning bort från definitionsfrågor om vad som menas med förnuft, sanning och framsteg i absoluta termer, mot något man skulle kunna kalla den ständigt pågående diskussionen om "det goda livet" för människorna i en demokratisk gemenskap. Denna perspektivförskjutning får också konsekvenser för synen på undervisning, från metodutveckling till problematisering av vissa diskursers roll i klassrummet som möjlig demokratisk mötesplats.

5.5.4 *Intressentperspektiv och politisk styrning*

Vi har redan tidigare pekat på hur intressenternas inträde på utvärderingens arena gjort politik och maktfrågor synliga. Henningson-Yousif (2003) [00:20] tar upp intressenternas perspektiv i sin studie av skolan och söker finna redskap eller metoder för att kunna organisera och analysera vad olika aktörer riktar intresset mot när det gäller skolan. Hon vill skapa analysredskap för att undersöka och diskutera hur olika aktörer, i de pedagogiska processer som rör skolan, resonerar om dessa processers sammanhang och innebörd. Tidigare försök till kategoriseringar genom ideologi, livsåskådning, klass m.m. räcker inte för att ge en bild av hur man fokuserar, menar författaren. Det går inte att utifrån dessa kategoriseringar förutsäga hur en viss aktör i skolan kommer att ta ställning i specifika frågor. Henningson-Yousif har i stället funnit i sina intervjuer med olika aktörer i utbildningssystemet, då framför allt i skolan, att möjligheterna till kombinationer av synpunkter och analyser tycks vara oändliga.

Med utgångspunkt från en fenomenologisk "livsvärldsteori" urskiljer hon tre olika processer: personliga, pedagogiska och samhälleliga processer. Var och en av dessa kan i sin tur analyseras utifrån tre olika aspekter. Den första är en *metaaspekt* som uppmärksammar en aktuell process – t.ex. pedagogiska insatser – kan påverka förhållanden bortom den egna processen. Den andra är en *kärnaspekt* som handlar om de konkreta inslagen i den process som analyseras – t.ex. pedagogikens organisation och innehåll. Den tredje är en *förhållningsaspekt* som handlar

om hur man själv agerar och blir behandlad i sin roll som lärare eller studerande i relation till andra aktörer. Med dessa försök att analysera aktörernas livsvärld utifrån personliga, pedagogiska och samhällsliga processer rör sig Henningsson-Yousif på en annan nivå än den som gäller vid en intressentutvärdering där syftet är att besvara frågan vad intressenterna vill. Hennes fråga kan snarast omformuleras till att försöka förstå *varför* man har en viss vilja, dvs. anta ett metaperspektiv på aktörsperspektivet.

Diskussionen om vem som ska ha makten över utvärdering är en aspekt av det politiska. En annan är utvärdering som instrument för politisk styrning. Att utvärdering har en politisk koppling diskuterades redan på 1980-talet men uppmärksammas under 2000-talet i diskussionen om utvärdering som styrinstrument och ifrågasättandet av om utvärdering kan vara värderingsfri. Det blir allt vanligare att se utvärdering, inte som en opartisk verksamhet, utan som kopplad till statlig (dvs. ytterst politisk) verksamhetsutövning. Staten beställer och finansierar utvärderingar, och den bestämmer vad som skall utvärderas och hur detta sker. I sin avhandling ger Gröjer (2004) [00:27] exempel på detta perspektiv och betraktar utvärdering som ett instrument för att noggrant bedöma genomförandet och resultatet av offentlig politik. Utvärderingens uppgift är att ge underlag för nya rationella beslut, men tillägger hon; ”denna både idealiserade och teoretiska utgångspunkt kan för vissa te sig som verklighetsfrämmande men den passar väl in i den rationalistiska styrningsprocess som länge varit en del av förvaltningspolitiken” (s 9).

5.5.5 *Granskning av kunskap och kvalitet*

Av de elva avhandlingar med utvärderingsperspektiv som har publicerats under 2006 kan vi konstatera att fem av dessa avhandlingar faller inom området ”Utbildning och undervisning” varav flera handlar om granskning av kunskap, kvalitets-säkring, prov och betyg. Även ett politiskt styrningsperspektiv finns med som en utgångspunkt.

Nytell (2006) [00:44] studerar den senmoderna idén om ”kvalitet” som invaderat utbildningsområdet under 1990-talet. Han utgår från ett politiskt styrningsperspektiv och visar hur staten från senare delen av 1990-talet allt fler återgått till en ökad centralstyrning genom att i förordningen föreskriva årliga kvalitetsredovisningar från de lokala skolorna och återkommande kvalitetsgranskningar av Skolverkets inspektörer. Denna förskjutning från en betoning på decentralisering tillbaka mot en ökad centralisering har vi berört i ett tidigare avsnitt, och den bild vi tecknar i vår framställning får belägg av det avhandlingsarbete som Nytell presenterar.

Hur kunskapsbedömningar i skolan har förändrats över historisk tid är temat för Lundahl (2006) [00:43]. Han diskuterar hur kunskapsbedömning sker på flera nivåer och ”re/producerar ett vetande som formar en skolverklighet” (s 409). Han beskriver hur den historiska utvecklingen påverkar uppfattningar om läroplaners syfte, innehåll och funktion. Vidare hur den konkreta styrningen och kontrollen av utbildning, särskilt då kunskapsbedömning, har betydelse för den pedagogiska praktikens innehåll. Han tar också upp vad som händer i den konkreta undervisningsprocessen när kunskapsbedömning blir en del i skolans vardag. Ett exempel är att betygssättning inte endast blir ett mått på elevernas prestationer utan även på lärarens insatser. Något som visar den komplexa situation som utvärdering och betygssättning ställs inför. De förändringar som Lundahl lyfter fram ger också stöd åt den bild av utvärdering som vi tecknat i tidigare avsnitt. Det lokala friutrymme för lärare och skolledare som förväntades bli realitet med den förvaltningsreform som inleddes 1980 har gradvis krympt. Preciserade betygskriterier och nationella prov innebär en starkare centralstyrning av vad som faktiskt undervisas och vilka kunskaper som provas och betygssätts.

Som framgått av tidigare avsnitt har prov och betyg sedan länge brukats för kunskapsmätningar och prestationsbedömningar i skolan och det har återkommande förts en kritisk debatt kring hur betyg och prov ska utformas, samt hur ofta de ska användas. Som instrument för systemutvärdering har de kritiserats då de

inte belyser vare sig förutsättningarna för eller själva undervisningsprocessen. Frågan om prov och betyg är dock ständigt aktuell inom diskussionen om bedömning. Under 2006 kom också ett par avhandlingar som fokuserar på prov och betyg och som därför skulle kunna definieras som forskning om utvärdering.

Prov och betyg på gymnasiet är objekt för kritisk analys i en avhandling av Korp (2006) [00:40] som gör en studie om betyg, nationella prov och social reproduktion. Hon tar sin utgångspunkt i Bourdieus habitusbegrepp och visar att betyg på de ämnen som ger allmän behörighet sätts utifrån skilda dispositioner (olika klass- och genusmönster) och att betyg och nationella prov på så sätt differentierar eleverna i skilda positioner på de teoretiska respektive de praktiska programmen. Studien, som baseras på intervjuer med lärare och elever på fyra skolor och fem program och med rektorer från de olika skolorna, är ett av många forskningsarbeten som belyser hur skolan bidrar till den sociala reproduktionen.

Även Boesen (2006) [00:37] har frågan om kunskapsmätning i fokus. Närmare bestämt utgör undersökningsobjektet kunskapsmätningar, eller test som är den benämning som Boesen använder. Avhandlingen är framlagd inom ämnet matematik, och kan såtillvida definieras som en ämnesdidaktisk avhandling. Titeln antyder att matematisk kreativitet är en kompetens av värde att uppnå, och författaren inriktar sin studie på att jämföra nationella test med lärarkonstruerade test, förklara skillnader mellan de båda testen och undersöka effekterna.

5.6 Trender inom utvärderingsforskning

Inledningsvis konstaterade vi att utvärdering som företeelse idag har fått en allt vidare utbredning i en rad olika sammanhang. Som framgår av vår redovisning har vetenskapliga arbeten i form av avhandlingar inte bara vuxit i antal utan även utvecklats till specialområden inom flera forskningsdiscipliner. Avhandlingsmaterialet ger också belägg för påståendet att utvär-

dering, sedan begreppet infördes som akademiskt specialområde, i huvudsak har varit och alltjämt är ett vetenskapligt tillämpningsområde. Tillämpningen sker inom vissa verksamhetsområden eller praktiker och rör framför allt insatser inom teknik, sjukvård, socialtjänsten och utbildning/undervisning.

Som framgått av tidigare avsnitt har rationalitet, effektivitet och nytta varit centrala begrepp alltsedan 1960-talet och dessa principer har motiverat den expansion av utvärdering som vi kan avläsa i det växande antalet avhandlingar. Det har med andra ord legat i statens intresse att få veta om kostnaderna för politiskt beslutade och i praktiken genomförda insatser verkligen varit värda pengarna ("value for money"). Merparten av redovisade avhandlingsarbeten är, som vi visat, en tillämpning av vetenskapliga metoder för en systematisk bedömning av särskilda insatser eller program. Som externa utvärderare har forskare med kompetens inom sin ämnesdisciplin eller sitt forskningsområde fått uppdraget att genomföra en aktuell utvärdering. Då kravet på objektivitet och neutralitet varit vägledande principer har det varit självklart att utvärderingsuppdrag i huvudsak givits till externa utvärderare. För forskare och deras institutioner har utvärderingsuppdrag på så sätt tillfört både pengar och gett underlag för avhandlingsarbeten.

Vår analys av avhandlingsarbeten visar att utvärderingsforskning följer och anpassas till de s.k. paradigmskiften som den vetenskapliga forskningen mer generellt genomgår. De följer också de forskningsinriktningar som har legitimitet inom respektive ämnesdisciplin, och olika tillämpningsområden kan därför komma att utveckla skilda tillämpningsprofiler. Efterhand som utvärderingsforskning blivit etablerad har den på så sätt blivit allt mer mångfacetterad, vilket också avspeglas i det material vi analyserat. Mångfald kännetecknar idag inte bara de vetenskapliga metoderna utan även de teoretiska utgångspunkterna, valen av perspektiv och principerna för värdering. Med all denna postmoderna mångfald blir det underlag för beslut, som en utvärdering förutsätts ge, ytterst mångtydigt och svårtolkat. Mot den bakgrunden blir krav på handfasta mätinstrument, tillförlitliga (valida) utvärderingsmodeller och generaliserbarhet åter aktu-

ella, något som vi också kan urskilja bland de avhandlingar som lagts fram under år 2006. Flera av dessa titlar pekar på att utvärdering idag rör aktuella problem i ett globalt samhälle – problem som är mycket komplexa. Angivna titlar tyder också på att utvärderingsforskningen inriktas mer på att utpröva rationella modeller och tekniker för en systematisk bedömning av genomförda insatser inom specifika verksamhetsområden.

En tendens inom utvärderingsforskningen under 2000-talets första hälft är en starkare betoning av teoretisk förankring parad med lokal utvärderingspraxis. Tendensen är inte ny men har accentuerats. Utvärderingen utgår från en uttalad teoretisk referensram baserad på teori inom den disciplin där avhandlingen lagts fram, t.ex. pedagogisk teori, statvetenskaplig teori etc. Ytterligare en tendens är att utvärdering är en mångdisciplinär verksamhet där olika perspektiv får möjlighet att komplettera varandra och där olika ansatser och tekniker anses nödvändiga för att få bra resultat.

Utvärdering som politiskt redskap lyfts också fram som problemområde med diskussionen om utvärdering som en ickeneutral verksamhet. Även frågan om påverkan från sammanhanget liksom frågan om utvärderingens relativitet står allt mer i fokus. Det är vanligt att man tar avstånd från den traditionella synen på utvärdering som objektiv och rationell och att mer betona ett hermeneutisk, tolkande ansats vid utvärdering. På 70-talet började forskarna inom utvärderingsområdet att intressera sig inte enbart för kunskapsmätningar utan också för den pedagogiska processen. Man nöjde sig inte med att redovisa de uppnådda resultaten utan ville också förstå och förklara utfallet av genomförda eller pågående insatser. Man brukade gärna kvalitativa metoder och hämtade data från avgränsade fall, vilka dock inte är generaliserbara på samma sätt som kvantitativa surveyundersökningar. Idag verkar tendensen vara att en del forskare fortsätter på denna väg och drar ut konsekvenserna ytterligare i riktning med postmodernt tänkande och idéer/teorier om hur kunskap konstrueras.

Samtidigt finns en trend där det mätbara återigen prioriteras och politikerna frågar efter ökad jämförbarhet för att stimulera till konkurrens med nationell och internationell ranking. Utvärderingsforskningen tycks bli allt mer inriktad på att utpröva och tillämpa effektiva modeller för mätning av komplexa fenomen. Prov och betyg blir åter den optimala metoden men riktas nu mot elever istället för system. Och vad eleven lär och presterar blir ytterst lärarnas ansvar. Den som har frihet och möjlighet att välja, väljer den bästa läraren, förskolan, skolan, gymnasiet, lärosätet för sina barn och ungdomar. En uppgift för forskning om utvärdering blir i detta sammanhang att undersöka och kritiskt granska inte bara utvärderingspolitiken utan också det som faktiskt sker när utvärdering får genomslag i praktiken.

6 Sammanfattning och kommentarer

I detta kapitel är syftet att ge en sammanfattning med kommentarer av den beskrivning av vårt undersökningsobjekt – utvärdering – som vi redovisat i bokens fem kapitel. Framställningen, som med nödvändighet kommer att innehålla en viss upprensning från tidigare kapitel, avser att ge en komprimerad bild av hur utvärdering som företeelse har vuxit fram och förändrats fram till idag. I denna grovt tecknade bild fokuseras problem som vi funnit vara framträdande i olika skeden i historien. Vi ser utvärdering som instrument för politisk styrning och utifrån detta perspektiv sammanfattar vi vår redovisning av utvärderingens historia, i första hand speglar den det utbildningspolitiska fältet.

6.1 Från präster till skolinspektörer

Som framgått av den historiska tillbakablicken låg kontrollen av den tidiga folkundervisningen i händerna på den protestantiska kyrkan och dess prästerskap. Målet var att folket skulle tillägna sig den Lutherska läran som med reformationen blev den rätta läran. Ansvar för undervisningen låg på hushållen och att dessa fullgjorde sin uppgift kontrollerades av prästerna vid återkommande husförhör. Kontrollen gällde resultaten, dvs. att husfolket tillägnat sig de föreskrivna texterna och kristendomens centrala budskap som detta uttolkades i den lutherska katekesen. Med moderniteten, som växte fram ur Upplysningen, följde en växande kritik av kyrkan och prästerskapets ledande ställning. Allt starkare krav ställdes på en samhällsordning baserad på vetenskapliga principer, individuella fri- och rättigheter och etablerade politiska system. Riksdagsbeslutet om en allmän folkskola 1842 innebar dock inget omedelbart brott med den kyrkliga hegemonin. Skolstadgan föreskrev allmän undervisningsplikt, vilket innebar att varje stadsförsamling och landsbygdssocken skulle inrätta och ansvara för en allmän folkundervisning organiserad som skola. Den grundläggande undervisningen, som skulle

föregå konfirmationsläsningen för prästen, kunde alltjämt skötas av hemmen men skulle då kontrolleras.

Undervisningens innehåll och genomförande blev högst olika beroende på de lokala förutsättningarna i olika delar av landet. För att kontrollera att de lokala aktörerna följde skolstadgans föreskrifter, och för att granska situationen i den organiserade skolundervisningen, inrättades en skolinspektion under ecklesiastikministerns ledning. Inledningsvis var inspektörsuppdraget i huvudsak en bisyssla för präster men blev efterhand en huvudsyssla och en profession för ett växande antal statliga folkskoleinspektörer.

6.1.1 En skola på vetenskaplig och rationell grund

Med utbyggnaden av folkskolan etablerades också en lärarprofession som så småningom organiserade sig fackligt. Många inom lärarkåren fick ett betydande inflytande som politiker och ledamöter i olika undervisningskommittéer. Många deltog också aktivt med föredrag vid lärarmöten och artiklar i pedagogiska tidskrifter. På så sätt gavs spridning åt nya pedagogiska idéer som fokuserade den individuella utvecklingen och förordade en undervisning på vetenskaplig grund. Inrättandet av professurer i pedagogik i början av 1900-talet kom att bestämma inriktningen på den pedagogiska forskningen. Differentialpsykologi och pedagogiska mätningar kom länge att vara dominerande kunskapsområden.

Förenklat kan det hävdas att utvärderingens kunskapsintresse under denna tid främst var inriktat på att mäta individuella prestationer på standardiserade test. Bedömningen av resultat gjordes mot fastlagda skalmått och fördelningen redovisades på olika begåvningsnivåer, anlagsinriktningar och standards. Den individuella utvecklingen sågs som inifrånstyrd och det gällde för läraren att lära känna, odla och förädla det bästa hos varje enskild elev. Detta organiska synsätt stämde väl överens med industrisamhällets rationella arbetsorganisation där principen om "rätt man på rätt plats" skulle vara vägledande. Tro och subjektiva bedömningar måste förkastas då objektivitet och veten-

skaplig ledning skulle prägla det moderna samhället. I detta samhälle och med dessa vägledande principer blev samhälls- och beteendevetenskaperna något av en ingenjörskonst med vars hjälp framtiden skulle kunna förutsägas och prognoser utfärdas. Utvecklingen, såväl individens som samhällets, skulle planeras rationellt och styras mot allt högre nivåer, större enhetlighet och ökad harmoni. Utvärdering blev en del i denna planering och styrning.

6.1.2 Centralstyrning och utvärdering stärks

Vid mitten av 1900-talet förlängdes den allmänna folkundervisningen till ett nioårigt obligatorium. Utbildningssystemet fick en enhetlig struktur över hela landet och den direkta politiska styrningen av skolan stärktes. Innehållet i undervisningen reglerades i den nationella läroplanen. Antalet undervisningstimmar i respektive skolämne angavs i timplanen och innehållet i undervisningen föreskrevs i kursplanerna. Det stoff som skulle undervisas preciserades ytterligare i de statligt godkända läromedlen och de vetenskapligt utprovade undervisningsprogrammen.

För att förstå styrningen av utbildning kan situationen analyseras med utgångspunkt från den engelske sociologen Basil Bernsteins resonemang om relationen mellan tre ”informations-system”. Han menar att formell utbildning och kunskap kan ses som realiserade dels via läroplaner, dels via pedagogiska insatser eller processer, samt genom utvärdering. Fritt översatt säger Bernstein (1971 s. 47) att läroplanen definierar vad som räknas som giltig kunskap, pedagogiken definierar vad som räknas som tillförlitliga metoder att överföra denna kunskap, och utvärdering slutligen definierar vad som räknas som en giltig realisering av denna kunskap i den aktuella lärandesituationen. Utvärdering får på så sätt en legitimerande funktion

När dessa tre informationssystem är starkt styrande kännetecknas det av att: Läroplanen anger detaljerade mål och kunskaper att uppnå, metoderna för det pedagogiska arbetet är preciserade, och att utvärderingen är utförlig och mätande. Med den utgångspunkten kan det hävdas att den svenska läroplan (cur-

riculum), som infördes med grundskolan 1962, utmärktes av stark klassifikation och inramning. Varje skolämne hade sitt föreskrivna innehåll och sin reglerade undervisningstid. För undervisningens genomförande (pedagogy) förordades vissa vägledande principer och generellt giltiga undervisningsmetoder som förmedlades i lärarutbildningen och genom lärarfortbildning. För att få information om utfallet av genomförda reformer och undervisningsinsatser föreskrevs en statligt organiserad utvärdering (evaluation).

Begreppet utvärdering infördes i skolan under 1950-talet och blev en uppgift för Skolöverstyrelsen. Inledningsvis användes traditionella kunskapsprov för utvärdering på systemnivå, men denna ansats kom snart att bli ifrågasatt. Som specialområde inom pedagogik kom utvärdering också att beröras av den vetenskapliga omprövning som följde under 1960- och 1970-talet. Den experimentella psykologin och beteendeteknologin, som varit huvudinriktning inom den psykologisk-pedagogiska forskningen, blev allt mer kritiserat vid denna tid. Vid utvärdering av program för utbildning och undervisning kom också det dominerande produktionstänkandet att mäta elevernas prestationer att omprövas, och intresset kom att även inriktas på själva genomförandeprocessen.

Under 1970-talet riktades allt starkare kritik mot den svenska grundskolan. Reformeringen av skolsystemet hade inte lett till de mer långsiktiga mål som den förknippats med. Studier visade att arbetarklassens barn alltjämt var underrepresenterade i den högre utbildningen. I en rad motioner vid vårriksdagen 1970 beskrevs läget i grundskolan som krisartat, och en stor andel av eleverna bedömdes ha bristande basfärdigheter när de lämnade grundskolan. De påtalade problemen ledde till ett ökat intresse för en samhällskritisk forskning och teoretiker som den franske sociologen Pierre Bourdieu och den tidigare nämnde Basil Bernstein. Med utgångspunkt från den schweiziske psykologen och pedagogen Jean Piagets teorier om inlärning och utvecklingsfaser förordades en pedagogik anpassad till de enskilda elevernas behov och intressen. De individuella behoven kunde inte avgöras på central nivå, och frågan om hur undervisningen

skulle organiseras och genomföras måste därför beslutas av den lokala skolan och kommunen. Därmed inleddes en ny strategi för skolans politiska styrning.

6.2 Från regelstyrning till målstyrning

Med SIA-utredningen om skolans inre arbete följde en omprövning av den starka centrala styrningen av skolan och undervisningen och en reformering av den statliga förvaltningen inleddes. De första stegen i förvaltningsreformen togs vid övergången till 1980-talet. Då skedde implementering av en ny läroplan (Lgr80), och ansvaret för den lokala skolans utveckling och lärarnas kompetensutveckling överfördes till kommunerna. Förvaltningsreformen brukar förenklat beskrivas som en övergång från regelstyrning till målstyrning, vilket i praktiken innebär att de nationellt formulerade målen skulle konkretiseras lokalt för att bli utvärderingsbara. Att utvärdera mot mål innebär, som Ralph Tyler påpekade redan vid mitten av 1900-talet, en operationalisering i form av strukturerade kursplaner eller studiegångar. Den uppgiften hade legat på experter knutna till Skolöverstyrelsen och på läromedelsförfattare. Nu ställdes de lokala aktörerna inför uppgiften att själva tolka de nationella styrdokumentet, precisera målen och utarbeta konkreta skolplaner och lokala arbetsplaner. En sådan förskjutning i ansvarsfördelningen mellan stat och kommun ökade kraven på skolans personal och på en medveten pedagogisk ledning, samt på en lokal utvärderingsverksamhet.

6.2.1 Utvärdering för utveckling

Vid det andra steget i förvaltningsreformen, som inleddes 1991, hade utvärdering för skolutveckling fått visst gehör bland pedagogiska forskare. Ett intresse för kunskapsbyggande kom också till uttryck i den nyinrättade skolmyndighetens (Skolverket) utvärderingsprogram. Istället för att "lära organisationen" skulle ledningen nu "lära av organisationen", vilket i sin tur förutsatte information, grundad på självvärdering från olika nivåer i styrsystemet. Den hierarkiska toppstyrningen skulle ersättas med

en organisation i form av nätverk och skolverksamheten skulle utformas lokalt i dialog mellan olika intressenter. Brukarnas inflytande och kontinuerliga utvärdering var principer som skulle gälla i hela utbildningssystemet.

I praktiken innebar förvaltningsreformen en radikal förändring av lärarnas och skolledningens arbetsvillkor. Reformen ställde dem både under kommunalt huvudmannaskap och inför ett nationellt uppdrag som krävde betydligt mer av egna initiativ och ansvarstaganden. Uppdraget begränsades inte längre till att som enskild lärare realisera politiskt fastlagda mål med hjälp av centralt utprovade medel. Nu skulle man kollektivt, både i skolan och kommunen, söka sig fram till den organisation och de arbetsätt som fungerade i praktiken. De kraftiga ekonomiska besparingarna under början av 1990-talet resulterade i ändrade förutsättningar för såväl utvecklingen av den pedagogiska praktiken, som för lärarnas egen kompetensutveckling. När åter en ny läroplan (Lp094) och (Lpf94) antagits av riksdagen och politikererna även fattat beslutet att införa ett nytt kursrelaterat betygssystem, fanns det uppenbara brister i den kunskapsbas som förutsattes för en professionell implementering av en ny läroplan och ett nytt betygssystem.

6.2.2 Indirekt centralstyrning med fokus på resultat

När förväntade effekter av politiska reformer uteblir eller låter vänta på sig fattas inte sällan beslut om nya insatser, vilka i sin tur kan visa sig bli kontraproduktiva. Ett exempel är de satsningarna i form av särskilda medel för lokal skolutveckling som vi beskrev i kapitel fyra. Kommuner och skolor skulle här själva avgöra om de ville ansöka om projektmedel och även bestämma vilka projekt de ville genomföra. I regeringens direktiv gavs förslag på inriktningar eller vissa prioriterade områden. Skrivningen skulle inte uppfattas som en förordning, men kom ändå indirekt att styra både vilka projekt som initierades och vilka som tilldelades medel. De särskilda satsningarna kan såtillvida betraktas som en ny form av indirekt centralstyrning. I praktiken kom de att bidra till en likriktning av det lokala utvecklingsarbetet. Det framgår av de externa utvärderingar vilket bl.a. påtalats i en rapport från Riksrevisionsverket (1998).

Några tänkbara förklaringar till att det blev så kan nämnas. En är att de särskilda satsningarna kom att erbjuda ett tillskott i den lokala budgeten i en tid med bristande ekonomiska resurser. Intresset för att verkligen implementera en förändring i den lokala skolan stod inte alltid i proportion till de statliga medlen som tillfördes. Ett annat problem var att projekten var tidsbegränsade och ytterst sällan inordnade i en långsiktig strategi för lokal skolutveckling. Detta gav upphov till kortvariga insatser och resulterade i ett bristfälligt underlag för kunskapsbyggande. Dessutom var de system för styrning och utvärdering som kommunerna förutsattes utveckla, för att själva synliggöra den lokala skolverksamheten, i många fall bristfälliga. En återgång till en starkare central styrning blev tydlig dels genom att målstyrning utvidgades till mål- och resultatstyrning, dels genom att mål som alla elever skall ha uppnått när de lämnar grundskolan preciserades i kursplanerna.

Återgången till en starkare politisk centralstyrning kan tolkas som en bristande tilltro till lärarnas och skolledarnas professionalism. Återgången kan också tolkas så att politisk vilja prioriteras före professionell pedagogisk ledning. Lindensjö och Lundgren (2000) uttrycker detta drastiskt genom att påpeka att ”professionsstyrning i strikt mening ter sig direkt stötande för den politiska ledningen” (s.166). Utvärdering blir i detta sammanhang mindre inriktad på att förstå och förklara utfallet av genomförda insatser och mer på att bedöma om skolan når uppställda mål och förväntade resultat.

6.2.3 *Kvalitetsredovisning ersätter utvärdering*

Vid slutet av 1990-talet krävde regeringen en skärpning av skolans uppdrag och införde en förordning om en årlig kvalitetsredovisning från kommunerna. Efter att skolpolitiken under flera decennier handlat om att bygga ut en ”skola för alla”, blir frågan om kvalitet i utbildningen allt viktigare i början av 2000-talet. De stora nedskärningarna under 1990-talet för att ”sanera rikets finanser” ökade missnöjet bland medborgarna med de offentliga tjänsternas kvalitet, vilket även gällde skolan. Som lösning lanserades en marknadsbaserad styrningsmodell med förebild från det privata näringslivet. Offentlig verksamhet blev en tjäns-

temarknad där ”produkterna” skulle kvalitetsdeklarerars så att brukaren – kunden – skulle kunna välja skola, sjukvård, äldreomsorg etc. Utvärdering blir i detta sammanhang en viktig del i legitimeringen av olika valalternativ vilket i sin tur leder till en ökad användning av utvärdering som styrinstrument såväl inom utbildningsområdet som inom andra politikområden.

I slutet av 1990-talet introducerades nya strategier för skolans styrning och 1997 kommer en förordning som ålägger kommunerna och skolan att upprätta kvalitetsredovisningar som led i en kontinuerlig uppföljning och utvärdering. (SFS 1997:702). Kvalitet går från att vara ett vardagsbegrepp till att få en innebörd av centralt styrningsbegrepp (Nytell, 2006). Utvärdering som begrepp och uppdrag kom i mångt och mycket att ersättas av kvalitetsgranskning som beskrivning av en huvuduppgift för myndigheter underställda regeringen (Karlsson, 1997). Samtidigt avvisades tanken om ett fristående utvärderingsinstitut. I utvecklingsplanen 2001/02, klargör den sittande socialdemokratiska regeringen sin ståndpunkt på följande sätt:

Regeringen har den bestämda uppfattningen att uppgifter som syftar till att uppnå nationell likvärdighet på utbildningsområdet skall skötas av myndigheter underställda regeringen. Det innebär att generella förvaltningsregler och principer tillämpas, vilket bidrar till rättssäkerhet och att kraven på offentlighet och insyn tillgodoses. Regeringen avvisar därför bestämt de motioner i riksdagen som föreslagit fristående kvalitetssäkringsinstitut för skolan, som skall stå fritt i förhållande till regering och riksdag. Förslaget har också avvisats av en bred majoritet i riksdagen. Tillsyn och granskning av en stor mål- och resultatstyrd samhällssektor som förskola, skola och vuxenutbildning har inslag av myndighetsutövning som inte kan överlåtas till privaträttsliga organ eller andra fristående institut. (Regeringens skrivelse 2001/02:188, s. 27)

Det nya Skolverkets huvuduppgift omfattar idag såväl nationell som internationell utvärdering. Båda innebär fördjupade studier och omfattande analyser som syftar till att beskriva, förklara och värdera olika företeelser inom förskola och skola. Ett kunskapsintresse kommer således till uttryck som riktas mot att lösa problem, tillgodose behov och ge underlag till förbättringar. Däremot ges inte utrymme för en mer kritiskt granskande utvärderingsansats i syfte att generera kunskap om vad som faktiskt pågår i den pedagogiska praktiken och vad som är möjligt att

genomföra under rådande betingelser. Utvärdering för kunskapsbyggande tycks snarare bli en mer marginell uppgift för det nya Skolverket som efter den senaste omorganisationen inte längre har någon särskild utvärderingsenhet inom myndigheten.

Övergången från en utvärdering inriktad på kunskapsbyggande för skolutveckling till kvalitetsgranskning och kvalitetsredovisning kan också förstås utifrån ett ideologiskt perspektiv. Under 1990-talet får ett marknadstänkande sin spridning tillsammans med en nyliberal ideologi och styrningsstrategi (New public management, se 4.3.3 ovan). Lokal utvärdering involverar nu brukarna som i återkommande enkätundersökningar får ge sin bedömning av kvaliteten på de tjänster som erbjuds. Särskilt viktigt blir detta i ett skolsystem där brukarna själva kan välja skola och allt fler alternativ finns etablerade.

Sättet att formulera sig i lokala planer tenderar därför också att ändras: Från beskrivningar och bedömningar av en lokal skolverksamhet, till formuleringar av visioner om vad den specifika skolan långsiktigt strävar mot. Visioner som utlovas vara uppfyllda på kort sikt och sådana som kommer att uppnås på längre sikt. Den visionära bilden av skolan tjänar som marknadsföring, och kvalitetsutmärkelser av olika slag kan därvid tjäna som förstärkning då det gäller att bli uppmärksam, inte bara nationellt utan helst också internationellt. Listor över enskilda skolors, länders och lärosätens rangordning får stor uppmärksamhet i massmedia och leder till långtgående spekulationer om framtiden. Det skrivs och talas om den svenska skolans bristande kvalitet och farhågor väcks återigen om "a nation at risk". Förskjutningen tillbaka mot nationella och internationella utvärderingar i form av prov och betyg kan ses som en anpassning till EU men kan också förstås som ett ökat behov av eller intresse för internationella jämförelser.

6.3 Forskning och debatt

I den tidiga utvärderingsforskningen under 1960-70-talet handlar frågorna om utvärdering som systematisk verksamhet och dess koppling till forskarsamhället, om influenser från USA och om målens och referensramens betydelse för utvärderingens inriktning, samt om metodologiska problem. Målrelaterad utvärdering spelade en central roll i den tidiga utvärderingsforskningen, och forskare som intresserade sig för pedagogisk utvärdering hämtade i huvudsak sin empiri inom ramen för statliga utredningar och utvärderingsuppdrag. Detta kan ses som uttryck för att utvärdering och utvärderingsforskning (liksom pedagogisk forskning kan tilläggas) traditionellt har varit nära kopplad till staten och det politiskt administrativa systemets behov av utvärdering som underlag för beslut.

Den metodologiska stridsfrågan i den tidiga utvärderingsforskningen, särskilt på internationell nivå där USA var ledande, ställde kvantitativa metoder mot kvalitativa (Cook & Reichardt, 1985). I grunden handlade motsättningarna om skilda kunskapsintressen och olika vetenskapsteoretiska forskningstraditioner. Begrepp som positivism, hermeneutik, fenomenologi och kritisk teori fungerade som vattendelare i debatten. I linje med en positivistisk rationellt inriktad vetenskapssyn låg betoningen på test och kunskapsprov även om operationaliseringen av mål uttryckta i beteendetermer problematiserades. Trots den kritik som forskare riktade mot det ensidiga bruket av kunskapsmätningar var det under 1970-talet alltså vanligt att analysera resultaten av genomförda insatser utifrån olika variabler och med statistiska metoder beräkna olika faktorerens förklaringsvärde. Kunskap om statistiska beräkningar hade också sedan länge ingått i grund- och forskarutbildning i pedagogik.

6.3.1 *Ramar och processer i fokus*

Snart vidgades perspektivet och allt fler pedagogiska forskare som intresserade sig för utvärderingsforskning argumenterade för en bredare samhällskritisk ansats i studier och utvärdering av utbildning. Pedagogikprofessor Urban Dahllöf framhöll redan i

slutet av 1960-talet nödvändigheten att vid utvärdering på systemnivå inte bara mäta kunskapsresultat utan även samla information om själva genomförandeprocessen. Han visade hur faktorer som tid, pengar och undervisningsgruppens sammansättning sätter ramar och begränsar vad som är möjligt att uppnå. Utfallet eller de uppnådda provresultaten måste därför relateras till såväl förutsättningarna som själva genomförandeprocessen. En sådan analysmodell gör det möjligt att förklara samspelet mellan olika ramfaktorer och förstå varför en del skolor lyckas uppnå målen medan andra misslyckas. Modellen förutsätter samtidigt tid för och kompetens i, att bruka kvalitativa metoder för insamling och analys av data.

Frågan om grundskolans möjligheter att nå sina mål analyserades också och pedagogiska forskare använde Basil Bernsteins modell för läroplansanalys där frågor som stod i fokus var förskjutningen mellan stark klassifikation och inramning respektive svag klassifikation och inramning av undervisningens innehåll. Relationen mellan skola och samhälle studerades och för analys av social- och kulturell reproduktion återopades Pierre Bourdieu och hans ”kapitalbegrepp” och teorier om sociala fält, positioner och dispositioner. Även osynlig pedagogik och den dolda läroplanen var uttryck som infördes i pedagogisk forskning. En annan teoretiker var Michael Foucault, filosof och idéhistoriker, som väckte intresse bland pedagogiska forskare för sina historiska studier om disciplinering och maktutövning.

Intresset för kritisk teori tenderade samtidigt att förskjuta den pedagogiska diskursen mot ett makroperspektiv, vilket kritiserades av skolans praktiker som ställde sig frågande till teoriernas betydelse för de konkreta problem som de mötte i klassrummet. Perspektiv blev ett begrepp för sociologiskt orienterade utbildningsforskare, och idén om aspektskeende bidrog till att utmana den vedertagna sanningen om att vetenskaplig forskning grundas på en neutral position vid val av problem och vid analyser och bedömningar.

6.3.2 *Politik, värderingar och legitimering*

Under 1980-talet ökade intresset för utvärderingens värdemässiga och politiska aspekter. Även frågor som rör makt, legitimitet och rättvisa aktualiserades. Kunskap och kunskapsanvändning, dvs. vad man lär sig av utvärdering och hur utvärdering kommer till nytta, blev nu också frågor på agendan. Värderingsfrågan förs vidare i diskussionerna om forskarnas etik och behovet av etiska regler. Utvärderaren skulle nu betrakta sitt objekt utifrån olika perspektiv och analysera inbyggda maktförhållanden i undervisnings- och vårdsituationer. Frågan om utvärderaren som advokat eller ombud för de maktsvaga har sedan dess funnits med i den utvärderingsteoretiska diskussionen, och argument för att utvärderaren aktivt skall ta värdemässig ställning har uttalats allt tydligare (Karlsson 2002, 2003b). Samtidigt har ledstjärnan för många utvärderingsforskare fortsatt att vara principerna om objektivitet, neutralitet och rationalitet. För dessa utvärderare är målsättningen att ge mening åt resultaten men förhålla sig neutrala i sina bedömningar. Uppgiften är att leverera ett säkerställt underlag för beslut utan att befatta sig med besluten eller påverka politiken.

Frågan om utvärderingens legitimering börjar problematiseras under 1990-talet. Tidigare skedde legitimering genom hänvisning till utvärderarens forskarkompetens och forskarsamhällets rationalitet. Denna form av legitimering står fortfarande högt i kurs för myndigheternas syn på formerna för utvärdering. Samtidigt ställer ökad brukar- och kundbaserad utvärdering nya krav på legitimering. Här blir det viktigt att visa att utvärderingen tar hänsyn till olika behov, värderingar, viljor och handlingsmönster som kommer till uttryck på lokal nivå. Legitimering måste ske i närkontakt med brukare och utvärderaren måste göra beskrivningarna av det som utvärderas så att brukarna kan känna igen sig i den bild som ges. Frågor om val av standarder, kriterier och vilka intressenter som skall medverka i utvärderingen blir också viktiga, liksom hur man kan väga samman olika intressenters resultat och vilka som ges röst i en sådan process.

Samtidigt med denna förändring förväntas utvärdering också ge underlag för beslut, visa om handlingar/insatser/program är nyttiga och effektiva och inte minst för att visa om insatserna ger värde för pengarna. Det gör att mätbarhet i kvantifierbara termer aktualiseras som ett starkt krav hos beställare av utvärdering. Politikerna efterlyser statistik som visar i siffror om olika åtgärder och insatser får effekter, vilket i sin tur kan leda till att efterfrågan på databaser ökar hos statliga myndigheter.

6.3.3 Standardisering och ranking i marknadens tjänst

Under det nuvarande decenniets första del har det dykt upp en del nya frågor inom utvärderingsfältet. Internationell ranking mellan olika länders skolsystem är en sådan fråga. Information om utbildningens resultat blir allt viktigare att redovisa och en stark trend är utvecklingen av utvärdering som del i en övernationell styrning, såväl inom som utanför EU. En målsättning är att åstadkomma en ökad standardisering. Inom exempelvis högre utbildning pågår bolognaprocessen som avser att göra alla kurser jämförbara varhelst den enskilde studerande läser dem i Europa och (väst)världen. Sett i detta perspektiv finns det en risk att internationell utvärdering reduceras till ett administrativt standardiseringsinstrument.

Enligt den allmänna skolplikten ska alla barn gå i grundskolan och anvisas en plats av kommunen. I linje med ett ökat marknadstänkande har möjligheter att välja skola införts. Föräldrar och barn ställs inför en valsituation och inför beslut om skola vill man veta att verksamheten uppfyller kvalitetskrav och även staten måste kontrollera skolornas kvalitet. För att ge besked om kvaliteten blir utvärderingens uppgift allt mer inriktad på att övertyga om det rätta valet av skola och mindre på att göra en kritisk granskning. I varje fall inte mer kritisk än det som producenten kan vilja visa för att få legitimitet i kundens ögon, vilket inte sällan innebär att klädsamt meddela en viss kritik tillsammans med besked om att felen nu har rättats till.

6.4 Riktningar för framtida utvärdering

En avgörande fråga när man ska försöka svara på hur framtidens utvärdering kommer att gestalta sig är hur makten över utvärdering kommer att se ut. Vem ska bestämma vad som ska utvärderas och av vem? Som vi har visat kan utvärdering av skolan i hög grad påverka hur resurser fördelas och olika berörda aktörer har ett intresse av att få sina behov tillgodosedda. Det betyder att utvärdering, även med en till synes neutral modell, måste förhålla sig till en politisk agenda, där inte bara kvantifierbara aspekter som ekonomi och antal studieplatser är centrala. Även värdemässiga, etiska och moraliska aspekter som social rättvisa och jämlikhet är exempel på viktiga värden som måste beaktas, vilket också framhålls i den nationella läroplanstexten om skolans värdegrund (Karlsson & Andersson 2002).

Mot den bakgrunden blir det intressant att närmare reflektera kring vilka former av utvärdering som kan tänkas bli dominerande under de närmaste åren. Vi ska avslutningsvis försöka peka på några tendenser som vi tycker oss se både i den utvärderingsforskning som vi har studerat och i den mer allmänna debatten om skola och utbildning.

6.4.1 *Kvalitetsgranskning som informellt styrmedel*

Under det senaste årtiondet har utvärdering i form av kvalitetsgranskning och kvalitetsbedömning fått en ytterst stark ställning som informativt styrmedel. Samtidigt har de tidigare modellerna inte försvunnit ur bilden. Nygamla utvärderingsmodeller som inspektion, har väckts till liv och såväl nationella som internationella prov har åter tagits i bruk. Vad som granskas av inspektörerna, provens innehåll och vilka bedömningskriterier som preciseras, ger en bild av vilka kunskaper, normer och värderingar som anses vara viktiga och riktiga. Preciseringar av detta slag ger måhända en tydligare bild av vad som prioriteras än den bild som framgår av den nationella läroplanens mer allmänt hållna formuleringar. Och sådana prioriteringar kan, liksom diskursiva formuleringar, ses som uttryck för inflytande och makt.

Som vi har visat har pedagogisk expertis under 1900-talet utövat stort inflytande såväl på utformningen av utbildning och undervisning som på genomförandet av utvärderingar av utbildningsreformer och undervisningsinsatser. Under senare decennier har brukarnas inflytande inom utbildning, undervisning och utvärdering kommit i förgrunden. De senaste tre åren har tjänstemän inom den statliga förvaltningen fått som sin huvuduppgift att granska och bedöma kvaliteten i verksamhet som bedrivs lokalt. Vad detta skifte faktiskt kan komma att betyda återstår att se. En fråga av intresse att försöka besvara är: Förhåller det sig så att de akademiska forskarna får träda tillbaka och ge plats för den vanliga människan som utvärderare. Eller är det kanske snarare så att tjänstemän inom stat och kommun, underställda regeringen, är de som kommer att bli dominerade som utvärderande aktörer?

6.4.2 Internationalisering och standardisering

Ett problem som aktualiserats på senare år är hur medborgarna kan få insyn i vad som sker i byråkratin på nationell nivå och EU-nivå. Hur kan de se att skattemedlen används på rätt sätt? Frågor av detta slag ökar efterfrågan på utvärdering. Med EU:s olika program och bidrag följer också ökade krav på en typ av utvärdering som har stora likheter med revision. Syftet blir att granska och kontrollera användningen av offentliga medel på olika nivåer i systemet. Detta sker vanligtvis enligt standardiserade mallar, beslutade och administrerade av nationella och internationella organ och utförs av konsulter eller fristående utvärderingsinstitut. Även forskare finns med som uppdragstagare i större utvärderingsprojekt. Här blir det viktigt att också ge utrymme för en kritiskt granskande utvärdering som även ger medborgarna möjlighet till insyn och kunskap om vart skattemedel tar vägen och att de används till avsedda syften.

En riktning som blivit allt tydligare under senare år är att utvärdering har blivit en del i en övernationell styrning, där inte minst utbildning är ett centralt problemområde. I dag erbjuds utbildning (f. n. oftast högre utbildning) på en internationell marknad och många studenter genomför hela sin utbildning,

eller delar av den, utanför hemlandet. När politikerna vill öka medborgarnas möjligheter till rörlighet mellan länder blir frågan om vad olika examina är värda ett påtagligt problem att finna lösningar på. Här är nyckelordet validering, vilket kan ses som en utvärdering av vad olika utbildningar och examina är värda. För att genomföra validering bör kriterierna helst vara standardiserade om modellen ska vara administrativt optimal för utbildningsbyråkratin. Det gör att frågan om hur utbildning inom EU kan samordnas och hur standardisering och kontroll kan bedrivas effektivt hamnar på en övergripande politisk nivå. En uppgift för utvärdering skulle här kunna vara att kritiskt granska vilka effekter denna harmonisering och standardisering har för olika grupper i samhället.

6.4.3 Behov av en kunskapsgenererande utvärdering

En aktuell riktning inom utvärdering, särskilt inom medicin och socialt arbete, är att åtgärder ska vara evidensbaserade eller bevisbara. Oscarsson (2006) säger att debatten i Sverige främst rör evidensbaserad medicin med resultat från randomiserade studier som ideal kunskapskälla. Men han menar att metoden även bör kunna tjäna som förebild inom det sociala arbetets praktik. Detta är ett sätt att betona vikten av kunskapsbaserad utvärdering. Andra vill, som vi har sett i vår genomgång av olika traditioner inom utvärdering, se en bredare och mer samhällskritisk utgångspunkt för utvärderingens kunskapsgenerering. Man vill se alternativ till den utvärdering som sker inom en tydlig uppdragsram där utvärderarens utrymme för egna frågor är begränsat och där uppdragsgivaren har problemformuleringsprivilegiet. Istället betonar man vikten en mer fristående, kritiskt granskande utvärdering där syftet mer handlar om kunskapsgenerering, än om kontroll. Det kan exempelvis gälla att kritiskt granska modellerna för styrning där mål förväntas bli omsatta i resultat som kan utvärderas. Vi känner igen den tankegången från traditionell beslutsteori som utgår från att prat i "pyramidens" topp leder till beslut och handling på den operativa nivån i pyramidens botten. En alternativ tanke enligt Brunsson (2002) är att prat, beslut och handling inte sällan är skilda från varandra. Lindensjö och Lundgren (2000) har pekat på att medan

man på formuleringsarenan pratar om sina mål, så för aktörerna på realiseringsarenan sin måldiskussion utifrån sina problem och utformar handlingar som inte alltid går i samma riktning som det centrala pratet på formuleringsarena ger uttryck för, dvs. – den tänkta inriktningen.

Med det senare synsättet räcker det inte för en utvärdering att mäta om centrala mål har realiserats genom att utvärdera målluppfyllelse. För att förstå och bedöma vad som sker behövs snarare en mer kunskapsinriktad utvärdering som försöker förstå hur aktörer på olika nivåer i en organisation resonerar och utformar sina strategier för att hantera de problem man möter på sin nivå, exempelvis hur rektorer och lärare löser pedagogiska problem i sin praktik i skolan och i klassrummet. Detta ställer i sin tur krav på en utvärdering som, oavsett om den görs centralt eller lokalt, förmår vara innovativ och kritiskt granskande och som kan vaska fram guldgruvorna som är värda att tillvarata. Det handlar också om att uppmärksamma de värderingsmässiga implikationer som gäller för varje utvärdering. Ett sådant synsätt öppnar vägar för en mer nyanserad diskussion om att utforma en framtida utvärdering av skolan med legitimitet inom *såväl* praktik och vetenskap *som* politik.

Referenslitteratur

Alkin, M. C. (2004). *Evaluation Roots*. London: Sage.

Andersson, I. (1986a). *Läsning och skrivning. En analys av texter för den allmänna läs- och skrivundervisningen 1842–1982*. Akad. avh. Pedagogiska institutionen, Umeå universitet.

Andersson, I. (1986b). *Läsning och skrivning som pedagogiska problemområden – en översikt av pedagogiska arbeten och psykologisk- pedagogisk forskning i Sverige 1905–1980*. Stencil. Pedagogiska institutionen, Umeå universitet.

Andersson, I. (1991). Utvärdering – ett instrument för pedagogisk reflektion och professionell utveckling. I *Didaktisk Tidskrift*, 1–2, 1991, s. 39–45.

Andersson, I. M. (1992). *Lågstadiesatsningen – en riktad resurs för förnyelse och förändring*. (Stencil). Pedagogiska institutionen, Umeå universitet.

Andersson, I. M. (1994). Hur kan utökade respektive förändrade kunskapsmål tänkas påverka utbildningens kvalitet? I SOU 1994:45. *Grunden för livslångt lärande. En barnmogen skola*. Betänkande av Utredningen om förlängd skolgång. Stockholm: Utbildningsdepartementet, s. 105–121.

Andersson, I. M. (1995). *Insatser mot främlingsfientlighet och rasism i Sverige 1993–1994. En extern utvärdering*. Slutrapport. (Stencil). Pedagogiska institutionen, Umeå universitet.

Andersson, I. (1996). *Lärarfortbildning och kompetensutveckling i ett historiskt perspektiv – en policyanalys*. Pedagogiska rapporter nr 49/1996. Pedagogiska institutionen, Umeå universitet.

Andersson, I. M. (1999). Skola och profession i utveckling – lokala skolprojekt i samarbete mellan lärare och forskare. I *Lärarutbildning och forskning i Umeå*, 2/1999, s. 7–19.

Andersson, I. M. (2002). *Till frågan om spridning. Reflektioner kring KK-stiftelsens skolsatsningar*. Pedagogiska rapporter, 68/2002. Pedagogiska institutionen, Umeå universitet.

Andersson, I. & Åsemar, C. (1990). På försök... Gymnasieskola i utveckling. I Skolöverstyrelsen, *Från fall till fall. 12 forskare om gymnasieskolans utveckling*. Skolöverstyrelsens rapporter – planering, uppföljning och utvärdering, R 90:2. Stockholm: HLS Förlag, s. 79–95.

Andersson, I. M. & Grensjö, B. (2002). *KK-stiftelsens satsning på de mindre skolutvecklingsprojekten*. Stockholm: Stiftelsen för Kunskaps- och Kompetensutveckling.

Andrae Thelin, A. (red) (1999). *Forskningen i skolan, Skolan i forskningen. Ett möte på lika villkor*. Dokumentation från forskningsymposium i Sundsvall 15–16 april 1999. Stockholm: Skolverket och Mitthögskolan, Centrum för Lärarutbildning och Skolutveckling.

- Asplund, J. (1970). *Om undran inför sambället*. Kalmar: Argos Förlag AB.
- Balnckertz, H. (1987). *Didaktikens teorier och modeller*. Översättning och introduktion av Gerd Arfwedsson. Stockholm: HLS Förlag.
- Bernstein, B. (1971). On the Classification and Framing of Educational Knowledge. In Young, M.F. D. (Ed.): *Knowledge and Control. New Directions for the Sociology of Education*. London: Collier-Macmillan.
- Bernstein, B. (1975). *Class, Codes and Control Volume 3*. London: Routledge & Kegan Paul.
- Björndal, C. R. (2005). *Det värderande ögat: observation, utvärdering och utveckling i undervisning och handledning*. Stockholm: Liber.
- Brunsson, N. (2002). *The Organization of Hypocrisy – Talk, Decisions and Action in Organizations*. Malmö: Liber.
- Cherryholmes, C.H. (1988). *Power and Criticism: Poststructural Investigations in Education*. London: Teacher College Press.
- Cook, T. D. & Reichardt, C. S. (Ed) (1985). *Qualitative and Quantitative Methods in Evaluation Research*. London: Sage.
- Dahllöf, U. (1967). *Skoldifferentiering och undervisningsförlopp*. Stockholm: Almqvist & Wiksell.
- Dahllöf, U. (1978). *Curriculum Evaluation, Frame Factors, and Teaching for Mastery; Reprints of three articles*. Uppsala reports in education, 2. Pedagogiska institutet, Uppsala universitet.
- Dahllöf, U. (1979). The Classroom Complexities Behind the Test Scores: some Swedish experiences. In L. Stenhouse. (Ed). *Accountability, Educational analysis*, Vol 1(1), s. 59–62.
- Dahllöf, U. (1989). Tredje gången gillt: En ny start för svensk utvärdering! I Franke-Wikberg (red.). *Skolan och utvärderingen. Fem professorer tar ordet*, s. 13–33. Stockholm: HLS Förlag.
- Dahllöf, U. (1999). Det tidiga ramfaktorteoretiska tänkandet. En tillbakablick. I *Pedagogisk Forskning i Sverige: På återbesök i ramfaktorteorin*, Årg 4 (1), s. 5–29.
- Dahllöf, U. (2002). Uppbyggnadstid, omgivningsfaktorer och pedagogikämnets forskningsutveckling. Några notiser från nära grannar i väst. I Uljens, M. (red). *Pedagogiken i Finland 150 år*. Pedagogiska fakulteten, Helsingfors universitet, s. 83–105.
- Du Rietz, L. M. fl. (1987). *Ansvarsfördelning och styrning på skolområdet. Ett beredningsunderlag utarbetat inom utbildningsdepartementet*. DsU 1987: 1. Stockholm: Utbildningsförlaget.
- Eliasson, R., Levin, C., Meeuwisse, A., & Sunesson, S. (red.). (1990). *Den värderande blicken. Om utvärdering i socialt arbete*. Lund: Studentlitteratur.
- Englund, T. (2004). Nya tendenser inom pedagogdisciplinen under de tre senaste decennierna. *Pedagogisk forskning i Sverige*. Årg 9 (1), s. 37–49.

- Europeiska Kommissionen. (2000). *Europeisk rapport om utbildningens kvalitet i skolorna: Sexton Kvalitetsindikatorer. Rapport grundad på arbetet inom kommittén för kvalitetsindikatorer*. Generaldirektoratet för utbildning och kultur, maj 2000.
- Franke-Wikberg, S. & Lundgren, U. P. (1980). *Att värdera utbildning Del I. En introduktion till pedagogisk utvärdering*. Stockholm: Wahlström & Widstrand.
- Franke-Wikberg, S. & Lundgren, U. P. (red). (1981). *Att värdera utbildning Del 2. En antologi om pedagogisk utvärdering*. Stockholm: Wahlström & Widstrand.
- Franke-Wikberg (red). (1989). *Skolan och utvärderingen. Fem professorer tar ordet*. Stockholm: HLS Förlag.
- Franke-Wikberg, S. (1990). *En strategi för utvärdering och lokal utveckling av utbildningskvalitet*. Arbetsrapport 81/1990. Pedagogiska institutionen, Umeå universitet.
- Furubo, J-E., Rist, R. C., & Sandahl, R. (2002). *International Atlas of Evaluation. Part I-II*. London: Transaction Publishers.
- Greene, J. C. (2006). "Evaluation, Democracy, and Social change", In I. Shaw, J. C. Greene, and M. M. Mark (eds.), *Handbook of Evaluation: Policies, Programs, and Practices*, (pp 118-140). London: Sage.
- Guba, E. G. and Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. London: Sage Publications.
- Gustafsson, C. & Selander, S. (1994). *Ramfaktorteoretiskt tänkande. Pedagogiska perspektiv. En vänbok till Urban Dabblöf*. Uppsala: Pedagogiska institutionen, Uppsala universitet.
- Holmer, J. (2003). *Självvärdering som stöd för personal- och verksamhetsutveckling*. Göteborg: Göteborgs universitet, Institutionen för arbetsvetenskap.
- Hood, C. (1990). A Public Management For All Seasons. *Public Administration*, Vol 69, s 3-19.
- House, E. R., & Howe, K. R. (1999). *Values in Evaluation and Social Research*. London: Sage.
- Husén T. (1994). *Skola och universitet inför 2000- talet. En Utbildningsforskares perspektiv*. Stockholm: Bokförlaget Atlantis AB.
- Högskoleverket (2005). *Utvärdering av den nya lärarutbildningen vid svenska universitet och högskolor*. Stockholm: Högskoleverkets rapportserie 2005:17 R.
- Johansson, E. (1981). Den kyrkliga lästraditionen i Sverige. En konturteckning. Ur nordisk kulturhistoria: *Läskunnighet och folkbildning före folkskoleväsendet. Mötesrapport III vid Nordiska historikermötet 1981*. Jyväskylä 1981, s.193-224.
- Johansson, E. (1983). *Kyrkböckerna berättar*. Stockholm: Liber Förlag
- Kallós, D. (1981). Anteckningar om utvärdering och pedagogisk forskning. I Franke-Wikberg, S. och Lundgren, U. P. (red). *Att värdera utbildning Del 2. En antologi om pedagogisk utvärdering*, s. 289-309. Stockholm: Wahlström & Widstrand.

- Kallós, D. och Lundgren, U. P. (1974). Evaluering av universitetsundervisning som pedagogiskt vetenskapligt problem. *INU* 1974:20, s. 47–55.
- Karlsson, O. (1995). *Att utvärdera – mot vad? Om kriterieproblemet vid intressentutvärdering*. Stockholm: HLS Förlag.
- Karlsson, O. (1996). *Petrus. Slutrapport i projekt för utveckling av program för utvärdering och uppföljning av barnomsorg i Västmanlands län*. Eskilstuna: Mälardalens högskola.
- Karlsson, O. (1997). Från utvärdering till kvalitetssäkring. En kritisk granskning av utvecklingen inom utvärderingsfältet. I *Röster om den svenska barnomsorgen*, s. 113–128. Stockholm: Socialstyrelsen, SoS-rapport 1997:23.
- Karlsson, O. (1998). ”Två tu sen ett” – Utvärdering av sammanslagningen mellan Mälarsjukhuset och Kullbergsga sjukhuset. Eskilstuna: Mälardalens högskola, CVF, skriftserie A: nr 2.
- Karlsson, O. (1999). *Utvärdering – mer än metod! Tankar och synsätt i utvärderingsforskning*. Stockholm: Svenska kommunförbundets skriftserie Åjour nr 3.
- Karlsson, O. (2002). Program Evaluation in Europe: Between Democratic and New Public Management Evaluation. In T. Kellaghan, D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*, p. 422–432. Dordrecht: Kluwer Academic Publishers.
- Karlsson, O. (2003a). Behovet av praktikbaserad utvärdering. I Christina Tors Hugosson (red.), *Värdera och utvärdera*, s. 48–59. Stockholm: Lärarförbundets förlag.
- Karlsson, O. (2003b). Democratic Evaluation: An Arena for Dialogue and Learning? In T. Schwandt, & P. Haug (Eds.), *Evaluating Educational Reforms: Scandinavian Perspectives. Evaluation and Society, Vol 2*. p. 133–147.
- Karlsson, O. (2004). *Evaluation as learning. A study of social worker education in Leningrad County*. Eskilstuna: Centre for Welfare research & Malardalen Evaluation Academy, Report 8A.
- Karlsson Vestman, O. (2006). Relationen utvärdering och politik. *Sociologen* 4/2006, Tema: Utvärdering, s. 24–27.
- Karlsson, O. & Kullberg, C. (1998). *Slutrapport och granskning av ”Triaden-projektet”. En samverkan kring utvärdering i Eskilstuna kommun, Västerås stad och Örebro kommun*. Eskilstuna: Mälardalens högskola och Örebro universitet.
- Karlsson, O. & Juhlin-Svensson, A-C. (1999): *Utvärdering av IT-projekt i Kristianstad län*. Eskilstuna: Mälardalens högskola, Centrum för Välfärdsforskning.
- Karlsson, O. & Kuivainen, H. (2000). *Att utvärdera kvalitet i högre utbildning. Internationella jämförelser av kvalitetsgranskning*. Stockholm: Högskoleverket, Rapport 2000:15AR.
- Karlsson, O., Andersson, I. M. & Lundin, A. (2002). *Metautvärdering av Högskoleverkets modell för kvalitetsbedömning av högre utbildning*. Stockholm: Högskoleverket. Rapport 2002:20 R.

- Karlsson, O. & Andersson, M. I. (2002). Utvärdering av utbildning – både politik och vetenskap. *Framtider nr 3, 2002*, s. 25–29.
- Karlsson Vestman, O. & Conner, R. F. (2006). The Relationship Between Evaluation and Politics. In I. F. Shaw, J. C. Greene and M.M. Mark (Eds.), *Handbook of Evaluation. Policies, Programs and Practices*, p. 225–242. London: Sage.
- Karlsson Vestman, O., Sedigh, M., Månsson, N. & Jäder, J. (2007). *Ett välstämt piano – till rätt melodi? En metautvärdering av Skolverkets nationella utvärdering 2003*. Eskilstuna: Mälardalens Högskola, Studies of Social Sciences, 1/2007.
- Kellaghan, T., Stufflebeam, D. L., & Wingate, L. A. (2003). *International Handbook of Educational Evaluation*. London: Kluwer Academic Publishers.
- Kemmis, S. (1998). Interrupt and say: Is It Worth Doing? Communicative action, developments inspired by Jürgen Habermas. Interview in Lifelong Learning in Europe (*LLinE*) 3/1998.
- Lindberg, G. & Lindberg, L. (1983). *Pedagogisk forskning i Sverige 1948 – 1971. En explorativ studie av inom- och utomvetenskapliga faktorer*. Umeå: Umeå universitet, Pedagogiska institutionen.
- Lindensjö, B. och Lundgren, U. P. (2000). *Utbildningsreformer och politisk styrning*. Stockholm: HLS Förlag.
- Lundgren, U. P. (1989). Att värdera utbildning och att styra den. I Franke-Wikberg, S. (red.) *Skolan och utvärderingen. Fem professorer tar ordet*. Stockholm: HLS Förlag, s. 86–90.
- Marklund, S. (1984). *Skolan förr och nu. 50 år av utveckling*. Stockholm: Utbildningsförlaget.
- Marklund, S. (1998). *Skolväsendets centrala ledning. Från Gustav Vasa till Skolverket*. Årsbok i svensk undervisningshistoria. Årgång LXXVIII 1998 Vol. 187.
- Marton, F. (1986). *Fackdidaktik. Volym I. Principiella övervägande. Yrkesförberedande ämnen*. Lund: Studentlitteratur.
- Mathiasson, L. (2005). "PISA – en maktfaktor i skoldebatten." *Pedagogiska magasinet*, Nr 02/ 2005.
- MacLaughlin, M. W., & Phillips, D. C. (Ed.) (1991). *Evaluation and Education: At Quarter Century*. Chicago: The University of Chicago Press.
- Nationalencyklopedin (NE) (2005). www.ne.se/ alt Höganäs: Bra Böcker Bokförlag.
- Naeslund, J. (1956). *Metodiken vid den första läsundervisningen*. Akad. avh. Uppsala universitet.
- Nytell, H. (2006). *Från kvalitetsidé till kvalitetsregim. Om statlig styrning av skolan*. Akad. avh. Uppsala Studies in Education No 114. Uppsala: Uppsala universitet, Pedagogiska institutionen.
- Oscarsson, L. (2006). Evidenskravet och socialt arbete. *Socionomen 4/2006, Tema: Utvärdering*, s. 31–34.

- Pettersson, S. & Wallin, E. (1995). Utvärdering och utveckling. I *Analysen och utvärderingar 1994-1995 - resultat - kommentarer perspektiv*. Stockholm: Skolverket, s. 81-87.
- Regeringens skrivelse (2001/02:188). *Utbildning för kunskap och jämlikhet - regeringens utvecklingsplan för kvalitetsarbetet i förskola, skola och vuxenutbildning*.
- Regeringskansliets faktablad: Sammanfattning av regeringens proposition 2004/05: 162. *Ny värld - ny högskola*. Stockholm: Utbildnings- och kulturdepartementet.
- Riksrevisionsverket (1998). *Lokal skolutveckling - statens roll och ansvar*. RRV 1998:26. Stockholm: Riksrevisionsverket.
- Rodhe, B. (1997). 360 dagar som skolminister. I Richardson, G. (red). *Spjutspets mot framtiden? Skolministrar, riksdagsmän och SÖ-chefer om skola och skolpolitik*. Årsböcker för svensk undervisningshistoria. Årgång LXXVII 1997. Vol.185.
- Rönnerman, K. (1993). *Lärarinnor utvecklar sin praktik. En studie av åtta utvecklingsprojekt på lågstadiet*. Akad. avh. Umeå universitet, Pedagogiska institutionen.
- Sahlin-Andersson, K. (2000). *Transnationell reglering och statens omvandling. Granskningsambällets framväxt*. Stockholm: SCORE Centrum för forskning om offentlig sektor.
- Scriven, M. (1980). *The Logic of Evaluation*. Inverness, CA: Edgepress.
- Scriven, M. (1991). *Evaluation Thesaurus*. Forth Edition. London: Sage Publications.
- SFS 1997:702. *Förordning om kvalitetsredovisning inom skolväsendet*. Stockholm: Utbildningsdepartementet.
- Shadish, W. R., Cook, T. D. & Leviton, L. C. (1991). *Foundation of Program Evaluation*. London: Sage.
- Sjöberg, S. (2005). "TIMSS och PISA - brickor i det politiska spelet." *Pedagogiska magasinet*, Nr 2/2005.
- Skolverket (1998). *Vägar till lokal arbetsplan*. Stockholm: Skolverket. Liber Distribution. Best. nr. 98:413
- Skolverket (1999). *Skolverkets allmänna råd om kvalitetsredovisning inom skolväsendet*. Allmänna råd (1999:1). Stockholm: Skolverket. Liber Distribution, Publikationstjänst.
- Skolverket 1993-09-19. *Utvärderingsprogrammet - en översikt 1*. Internt arbetsmaterial, s. 1.
- Skolöverstyrelsen (1974). *Mål för undervisningen i grundskolan och gymnasieskolan. Riktlinjer för målarbetet inom projektet Målbestämning och utvärdering, MUT*. Aktuellt från Skolöverstyrelsen, 1974.
- Skolöverstyrelsen (1976). *Skolöverstyrelsens pedagogiska forsknings- och utvecklingsarbete*. Stockholm: Liber läromedel. Utbildningsförlaget.

Skolöverstyrelsen (1983). *Läroplan för grundskolan, Lgr 80, Kommentarmaterial: Utvärdering i klass, arbetsenhet, rektorsområde och kommun*. Stockholm: SÖ:s publikation Läroplaner 1983:2.

Skolöverstyrelsen (1987–09–11). Innehåll och former för en nationell utvärdering av skola och vuxenutbildning. I S. Franke-Wikberg. (red.) (1989). *Skolan och utvärderingen. Fem professorer tar ordet*. Stockholm: HLS Förlag. Bilaga, s.1–26.

SOU 1974:53. Utredningen om skolans inre arbete SIA.

SOU 2005:104. Expertrapporter från 2005 års katastrofkommission. *Sverige och tsunamin – granskning och förslag*.

Stern, E. (2005). *Evaluation Research Methods. Part I–IV*. London: Sage.

Svenska Folkskolans Historia III: *Det svenska folkundervisningsväsendet 1860–1900. (SFS III)*. Stockholm: Albert Bonniers Förlag.

Svensson, N-E. (1962). *Ability grouping and scholastic achievement*. Stockholm Studies in Educational Psychology 5. Stockholm: Almqvist & Wiksell.

Tyler, W. R. (1991). General Statement on Program Evaluation. In Mc Laughlin, M. and Phillips, D. C. (red). *Evaluation and Education: At Quarter Century. Ninetieth Yearbook of the National Society for the Study of Education*. Chicago, Illinois: The University of Chicago Press.

Utbildningsdepartementet 2004–07–09. Uppdrag till Statens skolverk avseende kvalitetsredovisningen inom skolväsendet.

Utbildningsdepartementet 2002–05–27. Pressmeddelande. Utbildning för kunskap och jämlikhet – Regeringens utvecklingsplan för kvalitetsarbete i förskola, skola och vuxenutbildning.

Wallén, Göran. (1993). *Vetenskapsteori och forskningsmetodik*. Lund: Studentlitteratur.

Wallin, E. (2005). The Rise and Fall of Swedish Educational Technology 1960–1980. In *Scandinavian Journal of Educational Research*,. 49(5), s. 437–460.

Vedung, E. (1998). *Utvärdering i politik och förvaltning*. Lund: Studentlitteratur

Åkerblom, H. (1994). *Vilka ska vi finnas till för? Ledarskapsprocessen LOTS. Ett sätt att tänka och leva utifrån och in och inifrån och ut*. Uppsala: LOTS AB.

Avhandlingar i utvärdering 1970–2006

1970-tal

1. Kullberg, Lars (1973). *Utvärdering av objektivitetsprismespektra av tidiga stjärnor med hjälp av dator*. Stockholms universitet, Astronomi.
2. Bell, Robert (1975). *Att utvärdera sociala program*. Stockholms universitet, Pedagogiska institutionen.
3. Franke-Wikberg, Sigbrit Johansson, Martin (1975). *Utvärdering av undervisning: en problemanalys och några empiriska studier på universitetsnivå*. Umeå Universitet, Pedagogiska institutionen.
4. Kedner, Gösta (1975). *Företagskonkurser: problem – analys – utvärdering – åtgärder*. Lunds universitet, Företagsekonomiska institutionen.
5. Dahlgren, Jörgen & Lindkvist, Lars (1976). *Programorienterad utvärdering*. Lunds universitet, Företagsekonomiska institutionen.
6. Forslin, Jan (1978). *Arbetsanpassning: utvärdering av en psykologisk modell mot bakgrund av två empiriska studier*. Stockholms universitet, Psykologiska institutionen.
7. Hagström, Bo (1978). *1971 års länsförvaltningsreform: en utvärdering*. Lunds universitet, Statsvetenskapliga institutionen.
8. Rönnerberg, Sten (1978). *Beteendeanalys: riktlinjer för analys, datainsamling och utvärdering i beteendeterapi*. Stockholms universitet, Pedagogiska institutionen?
9. Karlsson, Ulf (1979). *Alternativa produktionssystem till lineproduktion: en utvärdering av produktionssystem i karossvärkstad vid Saab-Scania i Trollhättan*. Göteborgs universitet, Sociologiska institutionen.
10. Nilsson, Ingvar (1979). *Test-wisens och provkonstruktion: några studier med tonvikt på effekter av instruktionen och uppgifternas utformning på svarsbeteendet*. Umeå universitet, Pedagogiska institutionen.
11. Stensmo, Christer (1979). *Kollektivet Trollängen: en utvärdering på organisationsnivå*. Uppsala universitet, Pedagogiska institutionen.

1980-tal

1. Sollerman, Christer (1980). *Handens greppfunktion: analys och utvärdering samt en ny testmetod*. Göteborgs universitet, Institutionen för Ortopedisk kirurgi.
2. Åberg, Jonas (1980). *Hur framgångsrik är ryggrehabilitering?: utvärdering av verksamheten vid Rygginstitutet i Sundsvall*. Stockholm: Karolinska institutet.

3. Berg, Gunnar (1981). *Skolan som organisation: en analys av skolans organisatoriska struktur i ett förändringsperspektiv: slutrapport från SÖ-projektet "Modeller för lokal utvecklingsplanering och utvärdering"*. Uppsala universitet, Pedagogiska institutionen.
4. Hultman, Glenn (1981). *Organisationsutveckling genom ledarutbildning: en utvärdering av skolledarutbildningens första utbildningsomgångar*. Linköpings universitet, Pedagogiska Institutionen.
5. Ahlstrand, Ingemar (1983). *Nytta och kostnad i offentligt beslutsfattande – med exempel från trafikområdet*. Stockholm: KTH, Institutionen för trafikplanering.
6. Ringdahl, Anders (1984). *Hörapparaturprovning: en utvärdering baserad på mätning av insättningsförstärkning*. Göteborgs universitet, Medicinska fakulteten, Institutionen för oto-rhino-laringologi.
7. Welwert, Claes (1984). *Läsa eller lysna?: redovisning av jämförande undersökningar gjorda åren 1890–1980 rörande inläring vid auditiv och visuell presentation samt ett försök till utvärdering av resultaten*. Lunds universitet, Department of educational research.
8. Aspling, Anders (1986). *Företagsdemokratin och MBL: en empirisk och organisationsteoretisk utvärdering av en arbetsrättsreform*. Stockholm: Handelshögskolan, Ekonomiska forskningsinstitutet.
9. Lander, Rolf (1987). *Utvärderingsforskning – till vilken nytta?: pedagogiska röster och två exempel från forskningsfältet skolutveckling*. Göteborgs universitet, Institutionen för pedagogik.
10. Segraeus, Vera (1987). *Institutionell narkomanvård, organisation och innehåll: metodproblem vid utvärdering av social behandling för missbrukare*. Uppsala universitet, Sociologiska institutionen.
11. Svensson, Bengt, E. (1988). *Mot samlad skoldag: studie kring elevsituation och elevroll vid förberedelse och införande av samlad skoldag i en kommun*. Lunds universitet, Department of educational research.
12. Zetterström, Bo-Olof (1988). *Sambället som föreställning: om studerandes ideologiska formning i fyra högskoleutbildningar*. Umeå universitet, Pedagogiska institutionen.
13. Jonsson, Christer (1989). *Om skola och arbete: två empiriska studier av försök med en förstärkt skola-arbetslivskoppling*. Umeå universitet, Pedagogiska institutionen.
14. Maerker, Margot (1989). *Healthy persons' evaluation of patient information: a comparison with evaluation by patients*. Uppsala University, Faculty of Social Science.

1990-tal

1. Eriksson, Bengt, Karlsson, Per-Åke (1990). *Utvärderingens roll i socialt arbete*. Göteborgs universitet, Institutionen för socialt arbete.
2. Marklund, Bertil (1990). *Medicinsk rådgivning på vårdcentraler och apotek: omfattning, innehåll och utvärdering med särskild tonvikt på vårdcentralernas telefonrådgivning*. Göteborgs universitet, Allmänmedicin.
3. Sundman, Lars (1990). *Utvärdering av Hudiksvallsmodellen för äldreomsorg*. Uppsala universitet, Institutionen för socialmedicin.
4. Andersson, Håkan (1991). *Relativa betyg: några empiriska studier och en teoretisk genomgång i ett historiskt perspektiv*. Umeå universitet, Pedagogiska institutionen.
5. Bauer Alfredson, Betty (1991). *Förändringsarbete i äldre vården: utvärdering av personalutbildning inom institutionsbunden och alternativ äldre vård*. Lunds universitet, Pedagogiska institutionen.
6. Hemlin, Sven (1991). *Quality in science: researchers' conceptions and judgments*. Göteborgs universitet, Psykologiska institutionen.
7. Morgall, Janine Marie (1991). *Developing technology assessment: a critical feminist approach*. Lunds universitet, Sociologiska institutionen.
8. Löthman, Anna (1992). *Om matematikundervisning – innehåll, innebörd och tillämpning: en explorativ studie av matematikundervisning inom kommunal vuxenutbildning och på grundskolans högstadium belyst ur elev- och lärarperspektiv*. Uppsala universitet, Pedagogiska institutionen.
9. Malmgren, Gun (1992). *Gymnasiekulturer: lärare och elever om svenska och kultur*. Lunds universitet, Department of comparative literature.
10. Wang, Eric Min-yang (1992). *Usability evaluation for human-computer interaction (HCI)*. Luleå Tekniska högskolan, Institutionen för Arbetsvetenskap.
11. Werner, Mads Utke (1992). *Assessment of neuromuscular blockade by accelerometry: a method comparison study during anesthesia*. Lunds universitet, Department of Anesthesiology and Intensive Care.
12. Wåglund, Monica (1993). *Orientering mot ledarskap: analys av en chefsutbildning för läkare*. Uppsala universitet, Department of Education.
13. Zhao, Shangwu (1993). *Chinese science education: a comparative study of achievement in secondary schools related to student, home and school factors*. Stockholms universitet, Institutet för internationell pedagogik.
14. Öman, Robert (1993). *Överglasning av stora byggnadsvolymer: en tvärvetenskaplig utvärdering av ett köpcentrum*. Stockholm: KTH, Institutionen för byggnadsteknik.
15. Bergqvist, Erik (1994). *Belöningar och prestationer i offentlig verksamhet: en utvärdering av fyra fall inom Stockholms läns landsting*. Stockholms universitet, Institutet för kommunal ekonomi, Företagsekonomiska institutionen.

16. Johansson, Bertil, Johansson, Britt (1994). *Att styra eller inte styra: en kritisk granskning av den svenska grundskolans förändrade styrning med inriktning mot planeringsdokument*. Lunds universitet, Pedagogiska institutionen.
17. Larsson Mossberg, Lena (1994). *Servicemöten och deras betydelse vid charterresa*. Göteborgs universitet, Företagsekonomiska institutionen.
18. Murray, Åsa (1994). *Ungdomar utan gymnasieskola: en uppföljningsstudie från 13 till 24 års ålder*. Stockholms universitet, Institutionen för internationell pedagogic.
19. Paul, Ann-Sofi (1994). *Organisationsutveckling genom personalenkäter: en personalekonomisk utvärdering: personalenkäter som mätinstrument och utvecklingsunderlag: kostnader, effekter, relevans och utvärderingsmetoder granskas i tre empiriska studier*. Stocksolms universitet, Företagsekonomiska institutionen, Personalekonomiska institutet.
20. Björkemarken, Mariann (1995). *Implementeringsanalys som komplement vid utvärdering: en fråga om perspektiv och förklaring*. Göteborgs universitet, Sociologiska institutionen.
21. Hillring, Bengt (1995). *Forest fuel systems utilising tree sections: system evaluation and development of methodology*. Garpenberg: Sveriges lantbruksuniversitet.
22. Kapborg, Inez (1995). *Evaluation of Swedish nursing education and professional practice*. Stockholms universitet, Lärarhögskolan.
23. Karlsson, Ove (1995). *Att utvärdera – mot vad?: om kriterieproblemet vid intressentutvärdering*. Stockholms universitet, Lärarhögskolan.
24. Eriksson, Jan (1996). *Forskning, utveckling, nytta: om nytto- och relevansbedömningar av BFR-stödd forsknings- och utvecklingsverksamhet*. Stockholm: KTH.
25. Dahl, Iréne (1997). *Orator verbis electricis: taldatorn en pedagogisk länk till läs- och skrivefärdighet: utprövning och utvärdering av taldatorbaserade träningsprogram för elever med läs- och skrivsvårigheter*. Umeå universitet, Pedagogiska institutionen.
26. de Laval, Suzanne (1997). *Planerare och boende i dialog: metoder för utvärdering*. Stockholm: KTH, Institutionen för arkitektur och stadsbyggnad.
27. Johansson, Rolf (1997). *Utvärdering av bostadsbus i bruk: om att bedöma kvalitet i boendet*. Stockholm: KTH, Institutionen för arkitektur och stadsbyggnad.
28. Ramberg, Ulf (1997). *Utformning och användning av kommunala verksamhetsmätt*. Lunds universitet, Företagsekonomiska institutionen.
29. Smideman, Bengt (1997). *Planera med samverkan: utveckling och utvärdering av Planeringsboksmetoden vid detaljplanering med exploateringssamverkan*. Stockholm: KTH, Institutionen för arkitektur och stadsbyggnad.
30. Stolpe, Claus (1997). *Metoder vid presidentval: en utvärdering*. Vasa: Åbo Akademi, Samhälls- och vårdvetenskapliga fakulteten.
31. Åberg, Jan-Olof (1997). *Det rationella och det legitima: en studie av utvärderings teori och praktik*. Göteborgs universitet, Sociologiska institutionen.

32. Knutsson, Stig (1998). *Funktionshindrades värdering av färdtjänst*. Stockholm: KTH, Institutionen för infrastruktur och samhällsplanering.
33. Gustavsen, Hans Gustav (1998). *Volymtillväxten och övre höjdens utveckling i talldominerade bestånd i Finland: en utvärdering av några modellens validitet i nuvarande skogar*. Joensuu: Metsäntutkimuslaitos, Joensuun tutkimusasema.
34. Lif, Magnus (1998). *Adding usability: methods for modelling, user interface design and evaluation*. Uppsala universitet, Institutionen för materialvetenskap.
35. Andersson, Mona (1999). *The early childhood environment rating scale (ECERS): as a tool in evaluating and improving quality in preschools*. Stockholms universitet, Lärarhögskolan.
36. Hjärme, Anita (1999). *Kan man bli klok på läsdebatten?: Analys av en pedagogisk kontrovers*. Uppsala universitet, Pedagogiska Institutionen.
37. Nyström, Siv (1999). *Socialt förändringsarbete: en fråga om att omförhandla mening*. Stockholms universitet, Lärarhögskolan.
38. Sundelin, Johan (1999). *Intensive family therapy: a context for hopes put into practice*. Lunds universitet, Department of Child and Youth Psychiatry.

2000-tal

1. Andersson, Per (2000). *Att studera och bli bedömd: empiriska och teoretiska perspektiv på gymnasie- och vuxenstuderandes sätt att erfarastudier och bedömningar*. Linköpings universitet, Institutionen för pedagogik och psykologi.
- 2., Carlstedt, Berit (2000). *Cognitive abilities - aspects of structure, process and measurement*. Göteborgs universitet, Institutionen för pedagogik och metodik.
3. Ek, Ulla (2000). *Children with visual disorders: cognitive development, developmental disorders and consequences for treatment and counselling*. Stockholms universitet, Psykologiska institutionen.
4. Hansen, Kåre (2000). *Prestasjoner på messer: en studie av utstillere på internasjonale matvaremesser*. Göteborgs universitet, Företagsekonomiska institutionen.
5. Nyström, Lennarth (2000). *Assessment of population screening: the case of mammography*. Umeå universitet, Medicinska fakulteten.
6. Olsson, Ola (2000). *Alternative techniques in trochanteric hip fracture surgery: clinical and biomechanical studies on the Medoff sliding plate and the twin book*. Lunds universitet, Institutionen för rörelseorganens sjukdomar.
7. Rosén, Birgitta (2000). *The sensational hand: clinical assessment after nerve repair*. Lunds universitet, Institutionen för medicin, kirurgi och ortopedi.
8. Tornberg, Ulrika (2000). *Om språkundervisning i mellanrummet – och talet om ”kommunikation” och ”kultur” i kursplaner och läromedel från 1962 till 2000*. Uppsala universitet, Pedagogiska institutionen.

9. Müllersdorf, Maria (2001). *Needs assessment in occupational therapy: studies of persons with long-term/recurrent pain*. Uppsala universitet, Institutionen för folkhälso- och vårdvetenskap.
10. Westlund, Elvy (2001). *Undervisnings- och utbildningsplanering – för vem?: en studie av ingenjörutbildning vid en svensk högskola*. Uppsala universitet, Pedagogiska institutionen.
11. Berg, Jan (2002). *Systematic evaluation of perceived spatial quality in surround sound systems*. Luleå tekniska universitet, Musikhögskolan.
12. Eklund, Stig (2002). *Lärarytelse – mer än ämnen*. Umeå universitet, Pedagogiska institutionen.
13. Germeten, Sidsel (2002). *Grenser for undervisning?: frihet og kontroll i 6-åringerens klasserom*. Stockholms universitet, Lärarhögskolan.
14. Larsson, Laura (2002). *Evaluating social programs: active labor market policies and social insurance*. Uppsala universitet, Nationalekonomiska institutionen.
15. Magnuson, Tina (2002). *Assessment of writing difficulties and evaluation of computerized writing support*. Stockholm: KTH, Institutionen för tal, musik och hörsel.
16. Persson, Roger (2002). *Trait anxiety as a determinant of psychological test results*. Lunds universitet, Institutionen för psykologi.
17. Säfstén, Kristina (2002). *Evaluation of assembly systems: an exploratory study of evaluation situations*. Linköpings universitet, Institutionen för konstruktions- och produktionsteknik.
18. Wikström, Per-Erik (2002). *Utvärderingsinstrument för uppföljning av system- och säkerhetseffekter av trafiksäkerhetsåtgärder i tätort: tillämpade i Storuman*. Luleå tekniska universitet, Institutionen för samhällsbyggnadsteknik.
19. Fagerström, Birgitta (2003). *IT-användning: en analysmodell för IT-nytta*. Luleå tekniska universitet, Institutionen för Industriell ekonomi och samhällsvetenskap.
20. Henningsson-Yousif, Anna (2003). *Skolperspektiv: utveckling av redskap för analys av politikers, lärares och elevers resonemang om skolan*. Uppsala universitet, Pedagogiska institutionen.
21. Nilsson, Annika (2003). *Utvärdering av cykelfälts effekter på cyklisters säkerhet och cykelns konkurrenskraft mot bil*. Lunds Tekniska Högskola, Institutionen för teknik och samhälle.
22. Nilsson, Susanna (2003). *Miljöstrategier för det lilla grannskapet: utveckling och utvärdering av strategier för miljöarbete och robusthet genom två fallstudier i bostadsområdet från 1950-60-talen*. Stockholm: KTH, Institutionen för infrastruktur.
23. Stenberg, Anders (2003). *An evaluation of the adult education initiative relative labour market training*. Umeå universitet, Institutionen för nationalekonomi.

24. Emmelin, Maria (2004). *Self-rated health in public health evaluation*. Umeå universitet, Institutionen för Folkhälsa och klinisk medicin.
25. Emsheimer, Peter (2004). *Hur utveckla vetenskaplighet?: Utvärdering av examensarbetena i den särskilda lärarutbildningen (SÅL). Våren 2004*. Stockholms universitet, Lärarhögskolan.
26. Fänge, Agneta (2004). *Strategies for evaluation of housing adaptations: accessibility, usability and ADL dependence*. Lunds universitet, Institutionen för hälsa, vård och samhälle.
27. Gröjer, Anette (2004). *Den utvärdera(n)de staten: utvärderingens institutionalisering på den högre utbildningens område*. Stockholms universitet, Statsvetenskapliga institutionen.
28. Nyström, Peter (2004). *Rätt mått på prov: om validering av bedömningar i skolan*. Umeå universitet, Pedagogiska institutionen.
29. Wedlin, Linda (2004). *Playing the ranking game: field formation and boundary-work in European management education*. Uppsala universitet, Företagsekonomiska institutionen.
30. Zetterlund, Angela (2004). *Att utvärdera i praktiken – en retrospektiv fallstudie av tre program för lokal folkbiblioteksutveckling*. Göteborgs universitet, Enheten för biblioteks- och informationsvetenskap, Borås.
31. Engström, Christer (2005). *Implementering och utvärdering av Addiction Severity Index (ASI) i socialtjänsten*. Umeå universitet, Institutionen för psykologi.
32. Jess, Kari (2005). *Att räkna med nytta: samhällsekonomisk utvärdering av socialt arbete*. Stockholms universitet, Socialhögskolan, Institutionen för socialt arbete.
33. Nygren, Marianne (2005). *Rorschach comprehensive system variables and assessment for psychodynamic psychotherapy*. Stockholms universitet, Psykologiska institutionen.
34. Roos, Bertil (2005). *ICT and formative assessment in the learning society*. Umeå universitet, Pedagogiska institutionen.

Avhandlingar under 2006

35. Ahlström Söderling, Ragnar (2006). *Regionala företags förutsättningar för internationell konkurrenskraft: utvärdering av en undersökningsmodell*. Stockholms universitet, Företagsekonomiska institutionen.
36. Alexandersson, Karin (2006). *Vilja, kunna, förstå: om implementering av systematisk dokumentation för verksamhetsutveckling i socialtjänsten*. Örebro universitet, Institutionen för beteende- social- och rättsvetenskap.
37. 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
38. Bångman, Gunnel (2006). *Equity in welfare evaluations: the rationale for and effects of distributional weighting*. Örebro universitet, Institutionen för ekonomi, statistik och informatik.
39. Cars, Mikiko (2006). *Project evaluation in development cooperation: a meta-evaluative case study in Tanzania*. Stockholm University, Department of Education.
40. Korp, Helena (2006). *Lika chanser i gymnasiet?: en studie om betyg, nationella prov och social reproduktion*. Malmö högskola, Lärarutbildningen.
41. Karlsson, Fredrik, (2006). *Multi-dimensional approach used for energy and indoor climate evaluation applied to a low-energy building*. Linköping University, Division of Energy Systems, Department of Mechanical Engineering.
42. Lindhe, Jonas. (2006). *Utvärdering och optimering av industriventilationsanläggningar*. Lunds universitet, Avdelningen för installationsteknik, Institutionen för bygg- och miljöteknologi.
43. Lundahl, Christian (2006). *Viljan att veta vad andra vet: kunskapsbedömning i tidigmodern, modern och senmodern skola*. Uppsala universitet, Pedagogiska institutionen.
44. Nyttell, Hans (2006). *Från kvalitetsidé till kvalitetsregim. Om statlig styrning av skolan*. Uppsala universitet, Pedagogiska institutionen.
45. Söderström, Martin (2006). *Evaluating institutional changes in education and wage policy*. Uppsala universitet, Nationalekonomiska institutionen.

SERIERN MONOGRAFIER OCH
FORSKNING I FOKUS

MONOGRAFISERIEN

Glenn Hultman och Cristina Hörberg

Kunskapsutnyttjande

Ett informellt perspektiv på hur kunskap och forskning används i skolan

Ingegerd Muncio

Genomförande

Vem tolkar beslut och vem ser till att reformer blir mer än ord?

Britt Hallerdt

Studieresultat och social bakgrund

– en översikt över fem års forskning

Kjell Granström och Charlotta Einarsson

Forskning om liv och arbete i svenska klassrum

– en översikt

Ingrid Pramling Samuelsson och Ulla Mauritzson

Att lära som sexåring

En kunskapsöversikt

Birgitta Sahlin

Matematiksvårigheter och svårigheter när det gäller koncentration i grundskolan

En översikt av svensk forskning 1990-1995

Erik Wallin

Gymnasieskola i stöpsleven – då, nu, alltid

Perspektiv på en skolform

Mats Börjesson

Om skolbarns olikheter

Diskurser kring ”särskilda behov” i skolan – med historiska jämförelsepunkter

Hans Ingvar Roth

Den mångkulturella parken

– om värdegemenskap i skola och samhälle

Ulla Forsberg

Jämställdhetspedagogik

– en sammanställning av aktionsforskningsprojekt

Jens Pedersen

Informationstekniken i skolan

En forskningsöversikt

Peder Haug

Pedagogiskt dilemma

Specialundervisning

Moira von Wright

Genus och text

När kan man tala om jämställdhet i fysik läromedel?

Tullie Torstenson-Ed och Inge Johansson

Fritidshemmet i forskning och förändring

En kunskapsöversikt

Thomas Tydén och Annika Andræ Thelin (Red.)

Tankar om lärande och IT

En forskningsöversikt

Ulla Riis (Red.)

IT i skolan mellan vision och praktik

En forskningsöversikt

Pia Williams, Sonja Scheridan och Ingrid Pramling Samuelsson

Barns samlärande

En forskningsöversikt

Monica Söderberg Forslund

Kvinnor och skolledarskap

En kunskapsöversikt

Mats Ekholm, Ulf Blossing, Gösta Kåräng, Kerstin Lindvall
och Hans-Åke Scherp

Forskning om rektor

En forskningsöversikt

Ingemar Emanuelsson, Bengt Persson och Jerry Rosenqvist

Forskning inom det specialpedagogiska området

En kunskapsöversikt

Ulla Lind

Positioner i svensk barnpedagogisk forskning

En kunskapsöversikt

Karin Rönnerman

Vi behöver varandra

En utvärdering

Irene Rönnerberg och Lennart Rönnerberg

Minoritets elever och matematikutbildning

En litteraturoversikt

Maj Asplund Carlsson, Ingrid Pramling Samuelsson och Gunni Kärrby

Strukturella faktorer och pedagogisk kvalitet i barnomsorg och skola

Karin Wallby, Synnöve Carlsson och Peter Nyström

Elevgrupperingar

En kunskapsöversikt med fokus på matematikundervisning

Gunilla Härnsten

Kunskapsmöten i skolvärlden

Exempel från tre forskningscirkel

Elisabet Öhrn

Könsmönster i förändring

En kunskapsöversikt om unga i skolan

Rolf Helldin

Specialpedagogisk forskning

En kritisk granskning i ett omvärldsperspektiv

Anna Klerfelt

Var ligger forskningsfronten?

67 avhandlingar i barnpedagogik under två decennier, 1980-1999

Louise Limberg, Frances Hultgren och Bo Jarneving

Informationssökning och lärande

En forskningsöversikt

Jan-Eric Gustafsson och Eva Myrberg

Ekonomiska resursers betydelse för pedagogiska resultat

En kunskapsöversikt

Björn Eriksson, Odd Lindberg, Erik Flygare, Kristian Daneback

Skolan – en arena för mobbning

FORSKNING I FOKUS

- 1) Eva R Fåhræus och Lars-Erik Jonsson
Distansundervisning
Mode eller möjlighet
- 2) Ingrid Nilsson
Fristående skolor
– internationell forskning 1985-2000
- 3) Ingegerd Tällberg Broman, Lena Rubinstein Reich och Jeanette Hägerström
Likvärdighet i en skola för alla
Historisk bakgrund och kritisk granskning
- 4) Birgit Lendahls Rosendahl och Karin Rönnerman
Handledning av pedagogiskt yrkesverksamma
– en utmaning för skolan och högskolan
- 5) Bernt Gustavsson
Vad är kunskap?
En diskussion om praktisk och teoretisk kunskap
- 6) Eva Johansson
Möten för lärande
Pedagogisk verksamhet för de yngsta barnen i förskolan
- 7) Gunnel Colnerud och Robert Thornberg
Värdepedagogik i internationell belysning
- 8) Peder Haug
Om kvalitet i förskolan
Forskning om och utvärdering av förskolan 1998-2001
- 9) Per Andersson, Nils-Åke Sjösten och Song-ee Ahn
Att värdera kunskap, erfarenhet och kompetens
Perspektiv på validering
- 10) Lars Holmstrand och Gunilla Härnsten
Förutsättningar för forskningscirklar i skolan
En kritisk granskning
- 11) Joakim Ekman och Sladjana Todosijević
Unga demokrater
En översikt av den aktuella forskningen om ungdomar, politik och skolans demokrativärden

- 12) Staffan Selander (Red.)
Kobran, nallen och majjen
Tradition och förnyelse i svensk skola och skolforskning
- 13) Helena Korp
Kunskapsbedömning
– hur, vad och varför
- 14) Tullie Torstenson-Ed
Ungas livstolkning och skolans värdegrund
- 15) Gunnar Berg och Hans-Åke Scherp (Red.)
Skolutvecklingens många ansikten
- 16) Anders Marner och Hans Örtegren
En kulturskola för alla
– estetiska ämnen och läroprocesser i ett mediespecifikt och medieneutralt perspektiv
- 17) Therese Welén
Kunskap kräver lek
- 18) Mikael Alexandersson och Louise Limberg
Textflytt och sökslump
Informationssökning via skolbibliotek
- 19) Håkan Jenner
Motivation och motivationsarbete
– i skola och behandling
- 20) Malene Karlsson
Kunskap om familjedaghem
- 21) Monica Rosén, Eva Myberg och Jan-Eric Gustafsson
Läskompetens i Skolår 3 och 4 – en jämförelse mellan 35 länder
Progress in International Reading Literacy Study (PIRLS)
- 22) Jan-Eric Gustafsson och Monica Rosén
Förändringar i läskompetens 1991-2001
En jämförelse över tid och länder
- 23) Sangeeta Bagga-Gupta
Literacies and Deaf Education
A theoretical analysis of the international and Swedish literature

- 24) Helena Ahl
Motivation och vuxnas lärande
En kunskapsöversikt och problematisering
- 25) Pia Björklid
Lärande och fysisk miljö
En kunskapsöversikt om samspelet mellan lärande och fysisk miljö i förskola och skola
- 26) Gunilla Höjlund, Inger Göhl och Elisabeth Hultqvist
Utbildningsmodeller och läroprocesser i brytningen mellan skola och arbetsliv
En kunskapsöversikt
- 27) Annika Andrae Thelin och Karl Jan Solstad
Utbildning i glesbygd – samspel eller konflikt?
En kunskapsöversikt
- 28) Claes Nilholm
Inkludering av elever ”i behov av särskilt stöd” – vad betyder det och vad vet vi?
En kunskapsöversikt
- 29) Ann-Kristin Boström och Birgitta Lidholt (red.)
Lärares arbete. Pedagogikforskare reflekterar utifrån olika perspektiv
En antologi från en konferens anordnad av Myndigheten för skolutveckling
- 30) Inger Hensvold
Elevaktiva arbetsmodeller och lärande i grundskolan
En kunskapsöversikt
- 31) Monika Vinterek
Individualisering i ett skolsammanhang
- 32) Camilla Hällgren, Lena Granstedt och Gaby Weiner
Överallt och ingenstans. Mångkulturella och antirasistiska frågor i svensk skola
En kunskapsöversikt
- 33) Kjell Granström (red)
Forskning om lärares arbete i klassrummet
- 34) Eva Johansson
Förskolebarns moral
– i forskning och pedagogisk praktik

I DENNA BOK PRESENTERAS hur utvärdering inom skolan, har vuxit fram och förändrats fram till idag. I det agrara samhället är församlingspräster viktiga aktörer i kontrollen av den lagstadgade läs- och kristendomsundervisningen. Med folkskolestadgan 1842 får folkskoleinspektörerna uppgiften att kontrollera att kommunerna fullgör sina lagstadgade skyldigheter att bygga skolhus och tillsätta godkända lärare.

I början av 1900-talet differentieras blivande medborgarna till olika uppgifter i det framväxande industrisamhället genom intelligenstest och pedagogiska mätningar. Syftet är att på vetenskaplig grund göra urval. I mitten av 1900-talet bedöms statliga reformer och program inom utbildningsområdet genom systematiska utvärderingar. I slutet av århundradet ser vi resultaten av reformeringen av det svenska förvaltnings-systemet. Utvärdering brukas inledningsvis för insyn i och information om verksamheten på olika nivåer i systemet men blir efterhand ett allt starkare verktyg för en centralstyrning av verksamheten i skolan.

I slutet av boken kan vi konstatera, att marknadstänkandet har fått allt större genomslag, och att kvalitet blivit en tvingande regim även inom utbildningssektorn. Åter blir kvalitetsgranskning en huvuduppgift och kunskapsmätningar av olika slag får allt större betydelse. Något som också håller på att växa sig allt starkare är utvärdering som en del i en övernationell styrning, där inte minst utbildning är ett centralt område. Pedagogisk utvärdering handlar allt mer om att mäta och värdera för att jämföra såväl elever som skolor och utbildningssystem.

OVE KARLSSON VESTMAN är professor i pedagogik vid Mälardalens högskola, Institutionen för samhälls- och beteendevetenskap. Han har e-postadress ove.karlsson@mdh.se

INGER M. ANDERSSON är docent i pedagogik vid Uppsala universitet, Institutionen för lärarutbildning. Hon har e-postadress inger.m.andersson@bredband.net



MYNDIGHETEN FÖR
SKOLUTVECKLING

www.skolutveckling.se

