

Utvärdera för utveckling

– om utvärdering av skolpolitiska reformer

*Slutbetänkande av Utredningen om förbättrade
resultat i grundskolan*

Stockholm 2014



STATENS OFFENTLIGA
UTREDNINGAR

SOU 2014:12

SOU och Ds kan köpas från Fritzes kundtjänst. För remissutsändningar av SOU och Ds svarar Fritzes Offentliga Publikationer på uppdrag av Regeringskansliets förvaltningsavdelning.

Beställningsadress:
Fritzes kundtjänst
106 47 Stockholm
Orderfax: 08-598 191 91
Ordertel: 08-598 191 90
E-post: order.fritzes@nj.se
Internet: www.fritzes.se

Svara på remiss – hur och varför. Statsrådsberedningen (SB PM 2003:2, reviderad 2009-05-02)
– En liten broschyr som underlättar arbetet för den som ska svara på remiss.
Broschyren är gratis och kan laddas ner eller beställas på
<http://www.regeringen.se/remiss>

Textbearbetning och layout har utförts av Regeringskansliet, FA/kommittéservice.

Omslag: Elanders Sverige AB.

Tryckt av Elanders Sverige AB.
Stockholm 2014

ISBN 978-91-38-24080-9
ISSN 0375-250X

Till statsrådet och chefen för Utbildningsdepartementet

Regeringen beslutade den 24 maj 2012 att ge en särskild utredare i uppdrag att dels bedöma när de åtgärder som har vidtagits från 2007 och framåt för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan kan förväntas få avsedd effekt, dels utreda vilka relevanta data som behövs för att i ökad utsträckning kunna utvärdera de reformer som genomförs på skolområdet, och vid behov lämna författnings- eller andra förslag, samt belysa hur reformer på skolområdet bör genomföras för att underlätta utvärdering av dem.

Per Thullberg anställdes fr.o.m. den 1 juni 2012 som särskild utredare.

Ämnesrådet Maria Caryll och departementsrådet Kjell Nyman, Utbildningsdepartementet, förordnades att fr.o.m. den 25 september 2012 ingå i utredningen som experter. Maria Caryll entledigades fr.o.m. den 18 mars 2013. Samma dag förordnades ämnesrådet Kerstin Hultgren, Utbildningsdepartementet, som expert i utredningen.

Som sekreterare i utredningen anställdes fr.o.m. den 1 juni 2012 Eva Wallberg och fr.o.m. den 11 juni 2012 Mats Miljand.

Utredningen överlämnade i maj 2013 delbetänkandet *Det tar tid – om effekter av skolpolitiska reformer* (SOU 2013:30).

Utredningen överlämnar härmed slutbetänkandet *Utvärdera för utveckling – om utvärdering av skolpolitiska reformer* (SOU 2014:12).

Utredningens uppdrag är härmed slutfört.

Stockholm i mars 2014

Per Thullberg

/ Eva Wallberg
Mats Miljand

Innehåll

Sammanfattning	11
1 Inledning.....	23
1.1 Bakgrund	23
1.2 Uppdraget mer i detalj.....	24
1.3 Inriktning och avgränsning	25
1.4 Utredningens arbete	27
1.5 Betänkandet.....	27
2 Om evidensinformerad politik och praktik	29
2.1 Evidens – ett begrepp med flera betydelser.....	29
2.2 Evidens i politik och praktik	30
2.2.1 Evidens i politiskt beslutsfattande.....	31
2.2.2 Evidensinformerad snarare än evidensbaserad.....	33
2.2.3 Styrningsperspektiv och utvecklingsperspektiv.....	36
2.3 Forskning och politik – en problematisk relation	39
3 Om utvärdering	41
3.1 Vad menar vi med utvärdering?	41
3.1.1 Skilda tolkningar av begreppets innebörd.....	42
3.1.2 Några avgränsningar	45
3.2 Olika metodologiska ansatser	49
3.2.1 Effektutvärdering.....	50
3.2.2 Teoribaserad utvärdering	52

3.2.3	Utvärdering av komplexa system.....	54
3.2.4	Systematiska översikter	56
3.3	Sammanfattande bedömning	58
4	Dagens utvärderingssystem.....	61
4.1	Det nationella uppföljnings- och utvärderingssystemet	61
4.1.1	Uppföljning och utvärdering.....	62
4.1.2	Inspektion och revision	73
4.1.3	Forskning	76
4.1.4	Hur kommer reformutvärderingen in?.....	78
4.2	Sammanfattande bedömning	79
5	Reformutvärdering – historik och nuläge.....	81
5.1	En historisk tillbakablick	81
5.1.1	1900-talets mitt – en period av stora utredningar	81
5.1.2	En bild från 1970-talet	82
5.1.3	Den fortsatta debatten om den nationella utvärderingen.....	83
5.1.4	Övergången till mål- och resultatstyrning i början av 1990-talet	85
5.1.5	Sammantaget – ingen reformutvärdering att tala om	88
5.2	Var står vi i dag?	89
5.2.1	Reformutvärdering fortfarande ovanlig.....	89
5.2.2	Skolverkets utvärderingsuppdrag.....	90
5.3	En internationell utblick	95
5.4	Sammanfattande bedömning	97
6	Reformutvärdering – strategi för framtiden.....	101
6.1	Reformutvärdering och implementering	101
6.2	Reformstrategier och utvärdering	106

6.3	Regeringens och regeringskansliets arbete	110
6.3.1	OECD om det centrala regeringsarbetet i Sverige ...	110
6.3.2	Kopplingen mellan reformstrategier och reformutvärdering i regeringskansliets arbete	111
6.4	Reformutvärderingens roll bör stärkas	115
7	Ramverk för utvärdering av skolan – för bättre överblick och styrning	119
7.1	OECD om behovet av ramverk för utvärdering av utbildningen	119
7.2	Förslag om ett nationellt ramverk för utvärdering av skolan	123
8	Försöksverksamhet – för ökad evidens vid reformer	127
8.1	Användningen av försöksverksamhet	127
8.2	Försöksverksamhet som strategi i andra länder	131
8.3	Försöksverksamhet stärker förutsättningarna för reformutvärdering	132
8.4	Etiska överväganden i samband med försöksverksamhet	133
8.5	Förutsättningarna för försöksverksamhet i skollagen	135
8.6	Möjligheterna att bedriva försöksverksamhet bör tillvaratas	140
9	Systemet för att följa kunskapsresultat – utvecklingsbehov	143
9.1	Dagens nationella system för kunskapsbedömning	143
9.1.1	Nationella prov och betyg – lämpar sig inte som grund för trendmätning	144
9.1.2	Internationella studier – de enda instrument som i dag kan mäta kunskapsresultat över tid	149

9.2	Användningen av nationella prov i andra länder	151
9.3	Tidigare förslag om rullande stickprovsbaserad utvärdering av elevernas kunskapsresultat.....	152
9.4	Förslag till utveckling av systemet för att mäta trender	153
10	Data för utvärdering – utvecklingsbehov	159
10.1	Den nationella statistiken om skolan – en infrastruktur för systemförståelse	159
10.1.1	Kort om skolstatistikens framväxt.....	159
10.1.2	Några grundläggande begrepp.....	161
10.1.3	Skolverkets statistikansvar	162
10.1.4	Dagens statistiksystem	164
10.1.5	Uppgiftslämnarbördan	168
10.2	Regelverket för hantering av data.....	170
10.2.1	En omfattande lagstiftning.....	170
10.2.2	Pågående utredningar.....	174
10.3	Förslag och bedömningar om behovet av data för utvärdering och utvärdering av reformer	177
10.3.1	Några utgångspunkter	178
10.3.2	Individdata.....	179
10.3.3	Uppgifter om undervisningens förutsättningar och genomförande	181
10.3.4	Uppgifter om undervisningens resultat.....	191
10.3.5	Kostnader och elevresultat	195
10.3.6	Utveckling, modernisering och resurser	197
10.3.7	Inga förslag om författningsändringar.....	198
10.3.8	Avslutande diskussion	199
11	Organisering för stärkt utvärdering	203
11.1	Dagens aktörer i Sverige	203
11.2	Internationell utblick	205
11.3	Andra röster om organiseringen av utvärdering.....	208
11.3.1	Att nå ut och nå ända fram (SOU 2009:94)	208

11.3.2	Institutionaliseringen av utvärdering – ett fenomen i tiden	209
11.3.3	OECD om organisering och styrning av policyrelevant utvärdering.....	211
11.4	Förslag om ett råd med dubbla funktioner	212
12	Konsekvenser.....	221
	Referenser	225
Bilagor		
Bilaga 1	Kommittédirektiv 2012:53.....	243
Bilaga 2	Tilläggsdirektiv 2013:73.....	249
Bilaga 3	Kunskap och skolreform – En studie om underlag för och utvärdering av reformer på skolområdet av Mats Björnsson.....	251
Bilaga 4	Exempel på uppföljningar och utvärderingar genomförda av Skolverket på regeringens uppdrag åren 2002–2013	283
Bilaga 5	Internationell utblick – med fokus på organiseringen av nationell utvärdering	293

Sammanfattning

I maj 2012 beslutade regeringen att tillkalla en särskild utredare med uppdrag¹ att dels bedöma när de åtgärder som har vidtagits från 2007 och framåt för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan kan förväntas få avsedd effekt, dels analysera förutsättningarna för utvärdering av reformer. Utredningen redovisade i maj 2013 delbetänkandet *Det tar tid – om effekter av skolpolitiska reformer* (SOU 2013:30). Den andra delen av uppdraget redovisas i detta slutbetänkande från utredningen.

Arbetet med detta betänkande och utredningens delbetänkande har stärkt oss i uppfattningen att det är långt ifrån enkelt att förändra skolan genom strukturella reformer. Det tar tid innan effekter uppstår så att de kan bli mätbara på nationell nivå. Dessutom är de effekter man kan förvänta sig ofta inte så stora som man kanske önskat. Detta innebär inte att man kan avstå från att försöka att förbättra förutsättningarna för skolan. Det understryker dock behovet av att beslut om reformer vilar på en ordentlig kunskapsgrund. För varje åtgärd som vidtas bör det tydligt framgå på vilket sätt åtgärden är tänkt att påverka skolan, och vad som styrker denna slutsats. Forskning och utvärdering har här en central roll både inför att reformer ska utformas och vid utvärderingen av dem. Utvärdering av reformer kan i sin tur resultera i kunskap som kan ligga till grund för evidensinformerade beslut inför eventuella framtida reformer.

I det följande redovisas hur vi kommit fram till och underbyggt denna slutsats kapitel för kapitel samt redogörs för våra mer konkreta bedömningar och förslag.

¹ Dir. 2012:53, dir. 2013:73.

Bakgrund

Inledning (kapitel 1)

I Sverige har frågan om hur utvärderingskulturen inom skolområdet kan förstärkas – på både lokal och nationell nivå – diskuterats länge. Det betyder inte att frågan alltid befunnit sig i centrum för den skolpolitiska debatten eller varit föremål för politiska initiativ. Vi uppfattar direktiven till utredningen som att Sverige nu behöver lägga en ökad vikt vid frågor som har med utvärdering av skolan att göra. Denna ökade betoning på utvärdering delar Sverige med andra länder.

Utredningens uppdrag är vittomfattande och berör, som vi ser det, hela uppföljnings- och utvärderingssystemet för skolan liksom organiseringen av detsamma. Även om utredningen i första hand har ett övergripande nationellt perspektiv på utvärdering av reformer går det inte heller att bortse från den nationella uppföljningens och utvärderingens samspel med den lokala nivån och verksamhetsnivån. Vidare har reformutvärdering en nära relation till implementering av reformer, dvs. det som belystes i utredningens delbetänkande.

Om evidensinformerad politik och praktik (kapitel 2)

Evidens är ett begrepp som fått en allt större spridning. Begreppet kan syfta på tre olika företeelser:

- *Evidensbaserad politik* innebär att beslutsfattare, på olika nivåer och inom olika samhällsområden, så långt möjligt bör fatta beslut som baseras på kunskap.
- *Evidensbaserad praktik* syftar till att den dagliga verksamheten inom ett område ska vila på systematiskt utvecklad kunskap.
- Evidens kan också användas i en mer metodologisk betydelse där evidensbaserad politik och praktik avser en särskild *utvärderingsmetodik* som bygger på s.k. meta-analyser och systematiska kunskapsöversikter.

Utredningen föredrar att använda begreppet *evidensinformerad* i stället för evidensbaserad. En grundläggande skillnad mellan exempelvis evidensbaserad och evidensinformerad politik är *vad som kan betraktas som belägg eller bevis* (evidens). Vid utformningen av evi-

densinformerad politik kan utöver forskningsresultat även användas andra systematiskt insamlade och kvalificerade underlag.

Frågan om evidens i politiken kan dessutom ses ur två perspektiv – ett styrningsperspektiv och ett utvecklingsperspektiv. Vilket perspektiv som dominerar får avgörande betydelse för vilket slags utvärderingssystem som kommer att byggas upp.

Om utvärdering (kapitel 3)

Det kan vara svårt att orientera sig rätt bland alla de olika teorier och utgångpunkter som finns om utvärdering och det är en grannliga uppgift att bestämma hur en utvärdering ska genomföras. Det finns samtidigt stora möjligheter att genomföra utvärdering, även i de fall förutsättningarna i förstone inte verkar vara optimala. Det finns dock ett behov av ökad metodisk medvetenhet och förmåga hos både beställare och utförare av utvärderingar.

Gemensamt för god utvärdering är att det finns

- en systematisk datainsamling som ligger till grund för analys, tolkningar och bedömningar,
- en metodik vid behandlingen av data som uppfyller rimliga krav på validitet och reliabilitet, samt
- en värdering (av utvärderingsobjektets nytta, effekter, påverkan eller motsvarande) i förhållande till något slags måttstock (riktvärden, kriterier eller referenspunkter).

För att möta olika krav och förutsättningar kan utvärderingar vara av en mängd olika slag: effektutvärdering, programteoretisk utvärdering, processutvärdering, etc. Valet av metod bör göras med utgångspunkt i själva objektet för utvärderingen och i de frågeställningar som ska utvärderas. Interventioner i skolan har i de flesta fall en komplex karaktär, något både reformstrategi och val av metoder för utvärdering bör förhålla sig till. Kombinationer av olika metoder kan vara en god strategi för att fånga olika aspekter av en mer eller mindre sammansatt verklighet. Kvantitativt och kvalitativt inriktade metoder bör inte ställas mot varandra.

Utvärdering kan syfta till att svara på om en viss åtgärd fått några effekter och vilka dessa i så fall har varit – positiva, negativa, avsedda eller oavsedda. Den kan även syfta till att upptäcka samband mellan olika företeelser och om möjligt bidra till att finna

orsaker (förklaringar) till dessa. En utvärdering kan vidare syfta till fördjupad systemkunskap, till exempel genom att öka förståelsen för varför en viss reform eller åtgärd får olika genomslag i olika verksamheter. En formativ utvärdering kan bidra till att utveckla en reform genom att ge underlag för justeringar av reformen. Slutligen kan nationell utvärdering även utformas för att fungera utvecklingsstödjande på lokal nivå och i verksamheten.

Dagens utvärderingssystem (kapitel 4)

Hur det befintliga utvärderingssystemet är utformat – med uppföljning och utvärdering, men även viss tillsyn och granskning samt viss forskning – har betydelse inte bara för systemkunskapen om skolan, utan även som underlag för utvärdering av reformer.

Med en bred tolkning av utvärderingsbegreppet finns det många aktörer inom utvärderingssystemet i Sverige. Skolverket har en central roll genom att svara för den officiella statistiken om skolväsendet, det nationella provsystemet och de internationella studierna. Utvärderingar av skolan kan på regeringens uppdrag genomföras av Skolverket, Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU), Statskontoret, det offentliga utredningsväsendet samt av forskare och andra utvärderingsorgan. Skolinspektionens kvalitetsgranskningar och Riksrevisionens granskningar har utvärderande karaktär, men genomförs med andra syften. Riksrevisionen har dessutom en helt självständig roll i förhållande till regeringen.

Systemet består i dag av ett stort antal olika aktiviteter som i vissa fall kompletterar varandra. I andra fall är gränsdragningen mellan olika aktiviteter inte helt tydlig, som t.ex. mellan Skolverkets utvärdering och Skolinspektionens kvalitetsgranskning. Det finns också exempel där de syften som tillskrivs ett specifikt instrument knappast är möjliga att infria med ett och samma instrument. Det gäller för de nationella proven. Slutligen har systemet luckor, dvs. det finns frågor som det inte kan besvaras på ett tillfredsställande sätt. Det gäller t.ex. hur skolans resultat och måluppfyllelse utvecklas över tid (se vidare nedan).

Övergripande om reformutvärdering

Reformutvärdering – historik och nuläge (kapitel 5)

Reformer som genomförts under de senaste decennierna har föregåtts av utredningar med ett ofta begränsat utrymme för att på ett mer djuplodande sätt kunna utreda den aktuella frågeställningen, t.ex. genom att initiera utvärdering och forskning eller genom att i förväg pröva tänkta lösningar genom försök eller stegvis införande av reformer. Samtidigt finns en ökad ambition om att de politiska besluten ska vara evidensinformerade.

Utredningen menar att skolpolitiska reformer skulle vinna på att beslutas efter en mer noggrann prövning. För att en sådan ambition ska kunna infrias är det önskvärt att beslutsprocesserna ges mer tid än i dag, så att beslut i högre grad kommer att kunna baseras på kunskap. I ett allt mer komplicerat och komplext samhälle kommer politiker att behöva utgå mer från evidens vid formandet av sin politik, för att öka chanserna att reformerna ska leda till avsedda resultat.

Sverige saknar vidare en tradition att utvärdera reformer på skolans område, på det sätt vi sett exempel på i vissa andra länder. Så långt bakåt i tiden vi har kunnat överblicka saknas i Sverige exempel på ordentliga reformutvärderingar på skolans område som genomförts på initiativ av regering eller riksdag. De exempel som finns gäller främst utvärderingar som genomförts på initiativ av oberoende forskare. Inte heller för de viktiga reformer som genomförts under de senaste åren, dvs. sedan 2007, finns något utvärderingsprogram. De få regeringsuppdrag som nu är aktuella avser reformer från början 1990-talet, dvs.

- uppdraget till en särskild utredare att utvärdera eventuella kausala effekter av kommunaliseringen på elevernas studieresultat, läraryrkets status, lärarnas och rektorernas arbetsuppgifter, samt likvärdigheten i skolväsendet som nyligen redovisats i betänkandet *Staten får inte abdikera – om kommunaliseringen av den svenska skolan* (SOU 2014:5), och
- uppdraget till IFAU att utvärdera 1990-talets stora reformer på skolområdet, mer specifikt för grundskolan och gymnasieskolan.

Reformutvärdering – strategi för framtiden (kapitel 6)

En planerad och systematisk utvärdering av reformer är enligt utredningens uppfattning nödvändig för att säkerställa en både bred och djup systemkunskap om skolan. Inte minst är det viktigt för det politiska ansvarsutkrävandet att det går att få veta om den förda politiken har avsedda effekter och om de allmänna medlen allokeras på ett effektivt sätt. Utvärdering av reformer ger också kunskap som kan bidra till att öka genomslaget för reformen respektive signalera om något håller på att gå fel. Slutligen kan reformutvärdering bidra med kunskap inför kommande reformer.

Utformningen av reformer och utvärderingen av dessa står i ett ömsesidigt förhållande till varandra. Beslut om reformstrategi påverkar förutsättningarna för utvärdering. Samtidigt kan ett beslut om en utvärderingsstrategi, om det fattas tidigt inför beslutet om en reform, påverka såväl utformningen som implementeringen av denna. Frågan om reformutvärdering bör därför prioriteras redan i samband med planeringen av reformer och andra mer genomgripande åtgärder. I bästa fall kan utvärderingen bidra till genomförandet av reformstrategin genom att fungera formativt i förhållande till den nationella beslutsnivån och utvecklingsstödande på verksamhetsnivån.

Det finns behov av att stärka inslaget av evidens i regeringens och Utbildningsdepartementets arbete med att förbereda och genomföra beslut om reformer på skolans område. Utvärderingen av reformer har då en central roll. Frågor som är viktiga att ta ställning till är t.ex. om reformen lämpar sig för försök eller stegvis införande (vilket ökar möjligheterna för ett evidensinformerat beslutsfattande), vilken typ av metodik som är lämplig för utvärdering av reformen, vilken roll utvärderingen är tänkt att ha, etc. För denna typ av överväganden behövs kvalificerad utvärderingskompetens.

Mer konkreta bedömningar och förslag

Ramverk för utvärdering av skolan – för bättre överblick och styrning (kapitel 7)

Utredningen föreslår att ett nationellt ramverk för utvärdering av skolan ska fastställas av regeringen, förslagsvis i form av en skrivelse till riksdagen. Bakgrunden är bl.a. det stora antalet aktörer

som finns på den nationella arenan för utvärdering av skolan i dag. OECD har föreslagit ett liknande ramverk. Långsiktighet när det gäller utvärderingen av skolan är centralt såväl med hänsyn till användbarhet och kvalitet som ur ett demokratiperspektiv. Elever och föräldrar, lärare, rektorer, politiker, media, allmänhet – dvs. alla – ska kunna förvänta sig en kontinuitet när det gäller informationen om den svenska skolan.

Ett nationellt ramverk för utvärdering av den svenska skolan och förskolan bör tydliggöra det övergripande syftet med uppföljningen och utvärderingen samt

- vilka delar systemet består av för att följa upp, utvärdera, granska och beforska skolan,
- vilka syften de olika delarna av systemet har och vad de förväntas ge svar på,
- hur dessa olika delar förhåller sig till varandra, samt
- vilka aktörer som har ansvar för de olika delarna och hur dessa förväntas samverka.

Av ramverket bör också framgå vilken roll utvärderingen har i samband med större reformer. Ett sådant nationellt ramverk bör utgöra en över tid stabil grund utifrån vilken den svenska skolan kan utvärderas. Vi föreslår också att det nationella ramverket ska kompletteras med regelbundna planer för utvärdering av skolan. Av dessa planer bör framgå regeringens övergripande prioritering för utvärderingen för en överblickbar period, t.ex. tre år.

Försöksverksamhet – för ökad evidens vid reformer (kapitel 8)

Försöksverksamhet respektive stegvis införande av reformer är sätt att underlätta utvärdering av reformer som i ökad utsträckning bör tillvaratas. Kunskaper och erfarenheter som vinnas genom denna typ av försök kan bidra till att lägga en grund för långsiktigt bärkraftiga beslut.

Genom att pröva i liten skala kan man undvika misstag i full skala. Försök kan därför lämpa sig i de fall utfallet av en åtgärd inte i förväg kan förutses genom resultat från tidigare forskning och utvärdering. I ett försök är det också möjligt att vid behov justera och samtidigt lära av de erfarenheter försöket ger innan det eventuellt permanentas. Genom att försöksverksamhet ger förutsätt-

ningar för att skapa en kontrollgrupp möjliggörs dessutom användningen av utvärderingsdesigner som annars kan vara svåra att få till stånd. Det finns dock även andra metoder att skapa kontrollgrupper än genom försök.

Vid försök måste hänsyn tas till effekterna för berörda elever. Enskilda elevers rättssäkerhet får inte påverkas negativt. Ett försök får inte heller allvarligt påverka likvärdigheten. Ett försök i liten skala är dock enklare att justera än en intervention i full skala. De etiska övervägandena är således lika viktiga när en reform ska genomföras i full skala utan föregående försök, som i samband med en försöksverksamhet.

Nuvarande bestämmelse om försöksverksamhet i skollagen (2010:800) behöver ses över. Enligt denna bemyndigas regeringen eller den myndighet som regeringen bestämmer att få meddela föreskrifter om försöksverksamhet inom skolväsendet. I sådana föreskrifter får undantag göras från *organisatoriska bestämmelser* i skollagen. Vad som avses med organisatoriska bestämmelser är oklart. Det är därför angeläget att det utreds hur försöksverksamhet ska regleras i skollagen i framtiden.

Systemet för att följa kunskapsresultat – utvecklingsbehov (kapitel 9)

Trendmått av god kvalitet utgör ett viktigt underlag, bland andra, för att det ska vara möjligt att följa resultatet av skolreformer. I dag är det endast de internationella studierna som har en sådan utformning och kvalitet att de kan utgöra underlag för att följa kunskapsutvecklingen över tid i den svenska skolan, dvs. för att följa trender. Varken betygen eller de nationella proven möjliggör denna typ av trendmätning.

Betygssystemet är inte tillräckligt stabilt över tid för att fungera väl för trendmätning. Systemet har varit föremål för återkommande förändringar under de senaste decennierna. Det målrelaterade betygssystemet har dessutom, i varierande utsträckning över både tid och för olika ämnen, varit föremål för betygsinflation. En invändning mot de nationella proven är att proven inte är stabila över tid. För att vara det behöver varje nytt prov innehålla ett antal uppgifter som ingått i tidigare prov, s.k. ankaruppgifter. Framför allt svarar inte de nationella proven mot kravet på att varje år innehålla ett tillräckligt stort antal uppgifter för att täcka det aktuella ämnets

innehåll. Om man i stället genomför stickprovsundersökningar kan en provkonstruktion användas som fungerar bättre för trendmätning. Genom att låta ett begränsat antal elever (ett urval) svara på ett begränsat antal, men delvis olika, frågor som inte görs offentliga, kan det aktuella ämnets innehåll täckas in.

Det är enligt utredningens uppfattning ett problem att vi i dag är hänvisade till de internationella undersökningarna för trendmätning. De mäter endast utvecklingen i vissa ämnen och täcker endast vissa av de mål som återfinns i de svenska styrdokumenterna. Vi föreslår därför att ett system med stickprovsbaserade undersökningar för trendmätning ska komplettera de internationella studierna. Systemet bör ha en bredd med avseende på både de ämnen och de kunskapskvaliteter som mäts.

Det är emellertid inte lämpligt att nu endast lägga till ytterligare ett instrument i form av stickprovsundersökningar till ett redan hårt belastat provsystem utan att samtidigt utreda konsekvenserna av detta för övriga delar av systemet. Vi föreslår därför att en utredning tillsätts med uppdrag att i ett sammanhang utforma systemet för trendmätning genom stickprovsundersökningar och utreda hur det framtida systemet för utvärdering av kunskapsresultat ska vara utformat i sin helhet. Det gäller framför allt de nationella proven som i dag har många olika syften. Arbetet bör inriktas på att klargöra vilka syften som kan och ska knytas till systemet, vilka instrument som bör knytas till respektive syfte och vilka aktörers informationsbehov de olika instrumenten ska svara mot.

Data för utvärdering – utvecklingsbehov (kapitel 10)

Utökade och förbättrade registerdata kan underlätta möjligheterna att utvärdera reformer. För utvärdering av reformer behövs dock ofta även riktade uppgiftsinsamlingar, med frågeställningar relevanta för den aktuella reformen.

Det nationella uppföljningssystemet har en potential att utvecklas vad gäller ny uppgiftsinsamling, utökad analys av befintliga data från uppföljningssystemet och de internationella studierna samt det sätt på vilket informationen presenteras på. Vid utvecklingen bör ställning tas till vilka uppgifter som lämpar sig för att samla in i form av registerdata, vilka uppgifter som bör utgöras av direktinsamlade data och ingå i det nationella uppföljningssystemet samt vilken information som inte lämpar sig för att ingå i detta. Detta är

bedömningar som löpande bör göras av Skolverket som statistikansvarig myndighet.

Det har ingått i utredningens uppdrag att bedöma behovet av vilka relevanta data som behövs för att utvärdera reformer. Våra bedömningar avser informationsbehov avseende exempelvis undervisningens förutsättningar och undervisningens genomförande. Vidare avser bedömningarna informationsbehov om undervisningens resultat samt om sambandet mellan kostnader och resultat. I kapitel 10 redovisas dessa bedömningar område för område.

Vi föreslår vidare att Skolverket ges möjlighet att i enlighet med verkets hemställan kunna samla in individdata för skolan i sin helhet, dvs. även för modersmålsundervisning och skolgång i specialskola, särskola och sameskola. Detta bland annat mot bakgrund av att målen för specialskolan och sameskolan ska motsvara dem som gäller för grundskolan

För att det ska vara möjligt att utveckla och modernisera det nationella uppföljningssystemet behöver denna infrastruktur förstärkas finansiellt.

Med anledning av att det för närvarande pågår två utredningar som arbetar med frågor kopplade till hur förutsättningarna för registerbaserad forskning kan förbättras, lämnas inte här några författningsförslag. Det är emellertid viktigt att regeringen i detta sammanhang bevakar de långsiktiga förutsättningarna för den s.k. UGU-databasen.

Organisering för stärkt utvärdering (kapitel 11)

Ansvar för att fatta strategiska beslut om utvärderingen av skolan måste ligga på central politisk nivå. Det är t.ex. regeringen som måste ta ansvar för att reformer utvärderas och att detta sker med den bredd och med de olika perspektiv som skissats ovan. Mängden aktörer som kan ha en roll vid utvärderingen av skolan och av reformer på skolområdet pekar på ett behov av en stärkt nationell styrning.

Mot denna bakgrund föreslår utredningen att ett *Råd för utvärdering av skolan* inrättas. Rådet föreslås få följande två huvuduppgifter:

1. Löpande bistå regeringen i utvärderingsfrågor avseende skolväsendet. Det innebär t.ex. att

- utarbeta förslag till ett nationellt ramverk för utvärdering,
 - lämna förslag till nationell plan för utvärdering av skolväsendet, samt
 - inom ramen för en sådan plan vid behov kunna initiera policyrelevant forskningsbaserad utvärdering.
2. Vid större och genomgripande reformer inom skolväsendet bistå regeringen med
- förslag till program för reformutvärdering,
 - att vid behov initiera forskningsbaserad reformutvärdering,
 - att samordna utvärderingsinsatserna i det fall flera aktörer är involverade i reformutvärderingen (forskare, utvärderingsinstitut, myndigheter etc.),
 - att sammanfatta resultat från olika aktörers utvärderingar, samt
 - att sammanfatta identifierade behov av ytterligare stöd till, eller justeringar av, den aktuella reformen.

Förslaget innebär oförändrade uppgifter för myndigheterna på skolområdet. De uppgifter som rådet är tänkt att svara för, och de utvärderingar som det kan komma att föreslå eller initiera, genomförs i allt väsentligt inte i dag.

Rådet föreslås bli organiserat som en kommitté inom Regeringskansliet, med ett begränsat kansli. Rådets ledamöter, dvs. de som ska ha beslutsbefogenhet i rådet, föreslås vara forskare som representerar olika forskningsdiscipliner. I övrigt föreslås rådet bestå av sakkunniga från olika delar av skolsamhället, t.ex. Utbildningsdepartementet, berörda myndigheter och skolhuvudmän, samt företrädare för lärare, rektorer och elever. På detta sätt skapas ett forum för de strategiska frågorna om utvärderingen av den svenska skolan där de som har de största kunskaperna om behovet av, och förutsättningarna för, utvärdering samlas.

Konsekvenser (kapitel 12)

I direktiven till utredning har regeringen inte angett att det ska göras några specifika konsekvensbeskrivningar. Utredningen innehåller dock ett begränsat antal förslag som har föranlett oss att i korthet diskutera vissa ekonomiska konsekvenser av förslagen.

Kostnaden för en grundbemanning för det föreslagna *rådet för utvärdering av skolan* bör röra sig om cirka 2–2,5 miljon kronor per år, inkl. kostnader för sekretariat, ledamöter och sakkunniga. Kostnaderna för rådet blir naturligtvis avsevärt högre i de fall rådet ges särskilda uppdrag i anslutning till en specifik reform. Detta måste dock beräknas särskilt vid varje enskilt reformtillfälle. Det bör vara självklart att varje reform får sin finansiering, inte endast avseende kostnaderna i verksamheterna, utan också för implementering och utvärdering.

Kostnaderna för att införa ett *stickprovsbaserat system för att följa kunskapsutvecklingen i skolan* bör beräknas av den utredning som föreslås för att utreda frågan om stickprovsbaserade undersökningar i ett sammanhang tillsammans med övriga delar av systemet för utvärdering av kunskapsresultaten, dvs. även det nationella provsystemet.

Den finansiella förstärkning som utredningen bedömer behövs för att *förstärka och utveckla Skolverkets statistik- och uppföljningssystem* har både permanenta och tillfälliga inslag. För att systemet ska kunna omfatta fler insamlingar än i dag, vilket krävs till följd av de senare årens reformer, och samtidigt bibehålla den nuvarande kvalitetsnivån, behövs enligt Skolverket en permanent förstärkning med 10 miljoner kronor per år. För att utveckla och modernisera systemet behövs dessutom, enligt verket, tillfälliga förstärkningar med 10 miljoner kronor för 2014, 7 miljoner kronor för 2015 och 7 miljoner kronor för 2016. Utredningen bedömer att detta är lågt räknat. En särskild anslagspost för statistik- och uppföljningssystemet skulle tydligare än i dag kunna uttrycka statens ambitioner på området.

Slutligen bedömer utredningen att det är önskvärt att det s.k. UGU-projektet vid Göterborgs universitet ges stabila ekonomiska förutsättningar.

1 Inledning

I maj 2012 beslutade regeringen att tillkalla en särskild utredare med uppdrag¹ att i ett första steg bedöma när de åtgärder som har vidtagits från 2007 och framåt, för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan, kan förväntas få avsedd effekt. Utredningen redovisade i maj 2013 denna första del i delbetänkandet *Det tar tid – om effekter av skolpolitiska reformer* (SOU 2013:30). I ett andra steg ska utredaren analysera förutsättningarna för utvärdering av reformer².

1.1 Bakgrund

Utvärdering av skolpolitiska reformer kan syfta till att ge besked om huruvida en viss åtgärd fått några effekter och vilka dessa i så fall har varit – positiva, negativa, avsedda eller oavsedda. Utvärdering kan även syfta till att upptäcka samband mellan olika företeelser och om möjligt bidra till att finna orsaker (förklaringar) till dessa. En utvärdering kan vidare syfta till fördjupad systemkunskap, t.ex. genom att öka förståelsen för varför en viss reform eller åtgärd får olika genomslag i olika verksamheter. Utvärdering av reformer kräver med andra ord god tillgång till data av både kvalitativt och kvantitativt slag.

I Sverige har frågan om hur utvärderingskulturen inom skolområdet kan förstärkas – både på lokal och nationell nivå – varit aktuell inte minst sedan decentraliseringen av skolan och det mål- och resultatstyrda styrsystemet infördes vid 1990-talets början. Det betyder inte att frågan alltid befunnit sig i centrum för den skolpolitiska debatten eller varit föremål för politiska initiativ. Av bilaga 3 framgår att så inte har varit fallet. Utredningen uppfattar

¹ Dir. 2012:53 (bilaga 1).

² Dir. 2013:73 (bilaga 2).

dock direktiven som att Sverige nu behöver lägga en ökad vikt vid frågor som har med utvärdering av skolan att göra. Denna ökade betoning på utvärdering delar Sverige med andra länder, bl.a. med våra nordiska grannländer och övriga OECD-länder.

Det finns flera skäl till att många länder fäst ökad vikt vid utvärdering av skolans verksamhet och prestationer i allmänhet och elevernas kunskapsresultat i synnerhet. Utbildningens betydelse för sysselsättning och tillväxt betonas.³ Den internationella konkurrensen om högutbildad arbetskraft har ökat liksom kraven på att avsätta medel för finansiering av offentlig verksamhet ska användas effektivt.⁴ Slutligen har föräldrars krav på information om skolors och de egna barnens resultat ökat, i Sverige bl.a. mot bakgrund av det fria skolvalet.

Bakom det ökade intresset för uppföljning och utvärdering av kunskapsresultat har funnits, och finns, uppfattningen att uppföljning och utvärdering kan utgöra en drivkraft för utveckling och förbättring av skolans prestationer och resultat.

1.2 Uppdraget mer i detalj

Enligt direktiven ska utredaren i detta slutbetänkande

- utreda vilka relevanta data som behövs för att i ökad utsträckning kunna utvärdera reformer inom de obligatoriska skolformerna, dvs. grundskolan, grundsärskolan, specialskolan och sameskolan,
- om det behövs, lämna författnings- eller andra förslag för att öka tillgången till sådana data, och
- belysa hur reformer inom de obligatoriska skolformerna bör genomföras för att underlätta utvärdering av reformerna.

Enligt direktiven har utvärdering av reformer och politiska beslut under lång tid spelat en ganska undanskymd roll inom utbildningsområdet. Enligt regeringen innebär dock det omfattande reformarbetet inom sektorn att behovet av utvärderingar har förstärkts.

³ Se t.ex. Hanushek, Eric och Ludger Woessmann (2010) *The cost of Low Educational Achievement in the European Union*. European Expert Network on Economics of Education. EENEE Analytical Report No. 7.

⁴ Se t.ex. OECD (2013) *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*.

Områdets politiska och ekonomiska vikt medför dessutom, enligt regeringen, att behovet av beslutsunderlag av hög kvalitet ökar. Direktiven pekar också på att det även internationellt inom EU, OECD och i enskilda länder pågår en diskussion om hur användningen av utvärderingar kan öka och hur utformningen av offentlig verksamhet i högre grad ska kunna baseras på kunskap.

1.3 Inriktning och avgränsning

En utgångspunkt för arbetet med uppdraget är, som utredningen ser det, regeringens intentioner att utformningen av offentlig verksamhet i högre grad än tidigare ska kunna basera sig på tillförlitlig kunskap (evidens). En annan är att utbildningen enligt skollagen (2010:800) ska vila på vetenskaplig grund och beprövad erfarenhet.

Utredningen definierar reform på skolans område som en förändring av bärande delar av skolsystemet som är reglerade i lag eller förordning. Det som sägs i detta betänkande om reformutvärdering avser utvärdering av sådana reformer, men även utvärdering av olika politiskt initierade åtgärder av betydelse, om än av annan mer tillfällig karaktär, t.ex. satsningar på läs-skriva-räkna eller på matematik och naturvetenskap.

I flera avseenden har vi kunnat konstatera att vårt uppdrag, trots det primära fokuset på utvärdering av reformer, är vittomfattande och berör hela uppföljnings- och utvärderingssystemet för skolan.

Även om vi har uppfattat att uppdraget i första hand har ett nationellt perspektiv på utvärdering av reformer går det exempelvis inte att bortse från den lokala nivån och verksamhetsnivån. En nationell utvärdering har, och bör ha, betydelse även för huvudmannens och verksamhetens utvärdering och utveckling, och behöver utformas med hänsyn taget till detta. Man kan utgå från att reformer som utvärderas ytterst syftar till att utveckla verksamheten lokalt. Det är därför viktigt att belysa sambanden mellan systemets olika nivåer. Den lokala nivån och verksamhetsnivån bidrar dessutom med underlag för den nationella nivåns utvärdering.

Fokuset på reformer och utvärdering skulle kunna leda till slutsatsen att det räcker att enbart belysa förutsättningarna för just särskilda utvärderingar av reformer. Vi menar dock att reformutvärdering i vissa fall kan genomföras integrerat med det befintliga uppföljnings- och utvärderingssystemet och är under alla omstän-

digheter, i de allra flesta fall, beroende av information från detta. Uppföljnings- och utvärderingssystemet som helhet – med statistik, trendstudier, internationella studier, utvärderingar och till viss del även granskning och forskning – har därför getts en central roll i vårt arbete. Det är också i bl.a. detta system som de relevanta data som behövs för utvärdering av reformer kan återfinnas.

Reformutvärdering har en nära relation till implementering av reformer, dvs. till de slags frågor som utredningen belyste i sitt delbetänkande. Vi har därför knutit an till den modell för implementering som presenterades i delbetänkandet och diskuterar utvärderingens roll inför en reform, under pågående implementering och efter att en reform institutionaliserats. Det innebär att reformutvärdering, enligt vår uppfattning, kan ha olika karaktär. Den kan t.ex. vara formativ i ett skede och summativ i ett annat. Metodmässigt innebär detta att det finns många olika sätt att se på utvärdering, vilket redogörs för i kapitel 3.

Ytterligare en fråga som har att göra med hur reformer genomförs gäller hur de organiseras med avseende på utvärdering. Det handlar bland annat om hur reformutvärderingens oberoende kan säkerställas och om hur forskning och forskarkompetens kan knytas till reformutvärderingen.

Sammantaget har det visat sig att direktiven väcker en rad viktiga frågor om uppföljnings- och utvärderingssystemet och dess funktion – generellt och i samband med reformutvärdering – samt om organiseringen av både utvärdering generellt och av reformutvärdering specifikt. Utredningen konstaterar att det både inom akademien och bland olika utvärderingsinstitut och nätverk av utvärderare länge pågått en diskussion om värdet av olika utvärderingsansatser. Diskussionen handlar bland annat om vad som kan och inte kan betecknas som utvärdering, samt om vilka slags metoder som kan ligga till grund för tillförlitliga utvärderingar. Utredningen konstaterar att såväl OECD⁵ och Europeiska kommissionen⁶ har en bred syn på utvärdering som ger utrymme för en variation av design och metoder. Utredningen ansluter till detta breda synsätt i kapitel 3.

De frågor som detta betänkande behandlar är slutligen av en så generell karaktär att det inte har varit meningsfullt att begränsa framställningen till endast aspekter av utvärdering av reformer som

⁵ Se t.ex. OECD (2007) *Evidence in Education. Linking research and policy* och OECD (2013) *Synergies for Better learning. An international perspective on evaluation and assessment*.

⁶ Se t.ex. www.ec.europa.eu/europeaid/evaluation/methodology/methods/mth_att_en.htm

avser grundskolan och motsvarande skolformer, även om vi med hänsyn till direktiven har haft ett huvudfokus på dessa skolformer. Våra bedömningar och förslag avser dock skolväsendet i sin helhet.

1.4 Utredningens arbete

Liksom inför delbetänkandet har utredningen haft ambitionen att samla kunskap från många olika aktörer, genom omfattande litteraturstudier och genom möten med personer i Sverige och andra länder.

Centralt för utredningen har varit den grupp av medarbetare från berörda myndigheter, dvs. Statens skolverk, Statens skolinspektion och Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU), som utredningen träffat vid flera tillfällen. Utredningen har dessutom fått hjälp med underlag från myndigheterna och slutligen samrått med dem. Inför slutfasen av arbetet har utredningen även haft kontakt med Vetenskapsrådet.

Utredningen har dessutom sammanträffat med flera nationella forskare och andra experter inom området.

Inför slutbetänkandet har utredningen även gjort en internationell utblick. Vi har fördjupat oss särskilt i litteraturstudier av utvärdering och utvärdering av reformer på skolans område i Danmark, Canada, England, Finland, Nederländerna, Norge, Nya Zeeland, Skottland och USA. Vi har dessutom besökt Danmark, Finland, Norge och Skottland och på plats kunnat diskutera frågeställningarna med företrädare för ministerier och/eller berörda myndigheter.

Utredningen har sammanträffat med Sveriges Kommuner och Landsting (SKL) för att diskutera uppdraget.

Vi har slutligen träffat en grupp med företrädare för Lärarförbundet, Lärarnas Riksförbund och Skolledarförbundet för att diskutera hur de verksamma ser på utvärdering och reformer.

1.5 Betänkandet

Efter denna inledning är betänkandet disponerat i tre huvuddelar:

I *kapitel 2–4* ges en bakgrund avseende

- evidensinformerad politik och praktik,

- utvärdering, samt
- dagens utvärderingsystem.

I *kapitel 5–6* avhandlas frågan om reformutvärdering övergripande, i *kapitel 5* med avseende på historik och nuläge och i *kapitel 6* med avseende på strategi för framtiden. *Kapitel 6* är således ett centralt kapitel i utredningen.

Kapitel 7–11 behandlar sedan, med *kapitel 6* som utgångspunkt, ett antal mer konkreta frågor. Det gäller

- behovet av ett ramverk för utvärdering,
- användningen av försöksverksamhet,
- behovet att utveckla systemet för att följa kunskapsutvecklingen,
- behovet av data för utvärdering, samt
- behovet att stärka organiseringen för utvärdering på skolans område.

Utredningens ambition är att det ska vara möjligt att läsa kapitlen 7–11 för sig, varför texten innehåller vissa smärre upprepningar.

Slutligen görs i *kapitel 12* vissa konsekvensbedömningar.

2 Om evidensinformerad politik och praktik

Intentionen att offentlig verksamhet i högre grad ska utformas med utgångspunkt i evidens är en central utgångspunkt för utredningens arbete. I detta kapitel redovisas fortsättningsvis vissa utgångspunkter och begrepp som har betydelse för förståelsen av betänkandets fortsatta analys.

2.1 Evidens – ett begrepp med flera betydelser

Evidens är ett begrepp som under de senaste tjugo åren fått en allt större spridning. Begreppet kan syfta på tre olika företeelser:

Evidensinformerad politik innebär att beslutsfattare, på olika nivåer och inom olika samhällsområden, så långt möjligt bör fatta beslut som baseras på tillförlitlig kunskap. De förväntas ha informerats om den samlade kunskap som finns inom det aktuella området, i form av forskning, utvärderingar, särskilda utredningar, uppföljningsdata m.m. innan de beslutar om en reform eller annan åtgärd. Det varierar hur starkt olika länder har betonat betydelsen av ett evidensbaserat arbetssätt. Exempelvis Storbritannien har sedan 1990-talet mycket målmedvetet försökt införa ett sådant arbetssätt vid all politikutformning. Även i USA har det funnits ett intresse för metoder som tar utgångspunkt i evidens, inte minst inom den federala utbildningspolitiken. Inom EU har Kommissionen strävat efter att införa evidensbaserade arbetsmetoder och OECD har framstått som en kraftfull tillskyndare av, och pådrivare för, sådana metoder.

Evidensinformerad praktik syftar till att den dagliga verksamheten inom ett område ska vila på systematiskt utvecklad kunskap (evidens). Denna tanke utvecklades först inom det medicinska

området, men har senare spritt sig till exempel till det sociala området och utbildningsområdet. En vanlig tanke är att man genom systematiska översikter av forskning inom ett område ska kunna fastställa ”best practice”. Genom att kontinuerligt införa sådana metoder kan verksamheten successivt förbättras, dvs. verksamhetens insatser (behandling, undervisning etc.) blir effektivare och utbytet av insatta resurser bli större.

Evidens kan också användas i en mer metodologisk betydelse där begreppet avser en särskild *utvärderingsmetodik* som bygger på meta-analyser och systematiska kunskapsöversikter som i sin tur bygger på kvantitativa data insamlade genom en experimentell, kontrafaktisk design¹. Inom den så kallade evidensrörelsen menar man att endast resultat som baserar sig på dessa metoder kan anses utgöra evidens. Metodfrågor i anslutning till den så kallade evidensrörelsen behandlas i kapitel 3 om utvärdering.

Utredningen tar upp alla tre ovan beskrivna aspekter av evidens, dvs. evidens som utgångspunkt för politikutveckling, för praktisk verksamhet och som utvärderingsmetodik. Som en följd av utredningsdirektivens fokus har vi dock lagt större vikt vid evidensbaserad politik snarare än vid evidensbaserad praktik.

2.2 Evidens i politik och praktik

Ett evidensbaserat arbetssätt förutsätter att det finns tillgång till systematiskt framtagna kunskap som täcker samtliga väsentliga frågor inom det aktuella verksamhetsområdet. Ett viktigt problem i samband med evidensbaserat arbete – såväl vid politiskt beslutsfattande som vid utformningen av verksamhetens praktik – är att det finns bristfälligt forskningsstöd inom många politikområden. Det gäller både för själva verksamheten och dess resultat som avseende vilka åtgärder som kan förväntas åstadkomma önskat resultat.

Detta förhållande gäller inte minst det politiska beslutsfattandet om reformer och andra åtgärder på skolans område. Enligt direktiven till utredningen har utvärdering av reformer och politiska beslut under lång tid spelat en relativt undanskymd roll på skolans område. Utredningens direktiv kan tolkas som ett uttryck för en ambition att stärka evidensens betydelse vid utformningen av reformer och

¹ Om begreppen närmare betydelse se vidare kapitel 3.

politiska beslut på nationell nivå, något som i sin tur ställer krav på mer kvalificerade underlag för det politiska beslutsfattandet.

När det gäller skolans verksamhet är kravet på evidens fastställt i lag. Utbildningen ska enligt skollagen (2010:800) vila på vetenskaplig grund och beprövad erfarenhet.²

2.2.1 Evidens i politiskt beslutsfattande

Det är inte ovanligt att politiska beslutsfattare menar att det saknas information som är möjlig att använda som utgångspunkt för ett beslutsfattande grundat på evidens. Det kan saknas policyrelevant forskning eller saknas forskning som ger tillräckligt entydiga svar. Man kan säga att forskningsläget ”sprettar”. Mängden forskning kan också vara så stor att det är svårt att identifiera vilken forskning som är relevant och tillförlitlig. Situationen kan emellertid även vara en följd av att beslutsfattare i allt för liten utsträckning faktiskt intresserar sig för hur kunskapsläget kan förbättras inom det aktuella verksamhetsområdet.

Att många politik- och yrkesområden saknar stöd av omfattande och väl utförd forskning påverkar förutsättningarna för att utgå från evidens såväl vid det politiska beslutsfattandet som vid utformningen av själva verksamheten. Davies³ konstaterar att politiker behöver beakta information från många olika håll: enskilda studier, pilotstudier, fallstudier, expertutlåtanden, utvärderingar, beskrivande analyser, attitydundersökningar, statistiska modeller m.m.

Davies⁴ understryker samtidigt att beslutsfattare utöver denna typ av evidens måste lyssna på såväl den egna politiska organisationen som på den allmänna opinionen och olika intresseorganisationers uppfattningar. Evidensbaserad kunskap, dvs. kunskap som är systematiskt insamlad och sammanställd och bygger på vetenskapligt utförda studier av hög kvalitet, bör därför enligt Davies ingå i det beslutsunderlag som en beslutsfattare har att tillgå. Själva beslutsfattandet kan däremot aldrig förväntas vara endast evidensbaserat.

² 1 kap. 5 § skollagen (2010:800).

³ Davies, Philip (2004) *Is Evidence-Based Government Possible?* Jerry Lee Lecture. Annual Campbell Collaboration Colloquium, Washington D.C.

⁴ Davies, Phillip (2004) a.a.

Gray⁵ menar emellertid att evidens över tid kommer att få ökad betydelse i politikutvecklingsprocessen och att beslutsfattande baserat enbart på åsikter och värderingar kommer att minska i betydelse. OECD:s studier⁶ av olika länders arbete med att utveckla mer evidensbaserade arbetsformer utgår från synsättet att evidens bör ges en ökad betydelse vid det politiska beslutsfattandet. Resultaten av OECD:s studier pekar också på att många länder strävar efter att öka evidensens betydelse vid utformningen av sin politik. Detta innebär naturligtvis inte att politiska värderingar kommer att sakna betydelse. Det handlar här mer om en förskjutning där politiker i ett allt mer komplicerat och komplext samhälle kommer att behöva utgå mer från evidens vid formandet av sin politik.

I en amerikansk studie om användningen av vetenskap i politiska sammanhang konstaterar man att det ligger i sakens natur att politikutformning innehåller värderingar och andra närbesläktade överväganden som ligger utanför den vetenskapliga expertisens domän. Även när det handlar om värderingar, är det rimligt, menar man, att förorda användningen av kunskap som beskriver det problem som ska lösas på ett så exakt sätt som möjligt eller som förutsäger sannolika konsekvenser av föreslagna handlingar. Om beslutsfattare beaktar relevant forskning, så kommer de att förbättra chanserna att förverkliga de avsedda följderna av den politik de för fram.⁷

Användningen av vetenskap i politiska sammanhang kan enligt de amerikanska forskarna inte att skapa "bättre politik". Det handlar i stället om vägledning för att förstå ett problem, de förhållanden som ger upphov till dessa problem, och de resultat som sannolikt kan uppkomma när en viss policy riktar sig mot dessa problem. I denna mycket specifika bemärkelse kan, menar de, samhälls- och naturvetenskaperna ge en mer tillförlitlig ("bättre") vägledning än vad som i övrigt är tillgängligt för beslutsfattare när de funderar över olika frågeställningar.

⁵ Gray, J. A. M. (1977) *Evidence-based Healthcare: How to Make Health Policy and management Decisions*. Här återgiven efter Davies, P. (2004) *a.a.*

⁶ Här avses de länderstudier som OECD genomfört inom ramen för projektet OECD Reviews of Evaluation and Assessment in Education.

⁷ Prewitt, Kenneth, Thomas A. Schwandt and Miron L. Straf (2012) *Using Science as Evidence in Public Policy*. Committee on the Use of Social Science Knowledge in Public Policy. National Research Council of the National Academies. The National Academies Press. Washington, D.C.

2.2.2 Evidensinformerad snarare än evidensbaserad

Överväganden av det slag som redovisats ovan har lett till diskussioner om vilka krav som ska ställas för att något ska anses vara evidensbaserat. Inom vissa delar av evidensrörelsen ställs kraven högt. I princip är det endast (kvantitativa) forskningsresultat grundade på RCT-design (randomised controlled trials) som anses kunna komma i fråga, eftersom det är den enda forskningsdesign utifrån vilken en exakt beräkning anses kunna göras av sannolikheten för att det finns ett kausalt samband mellan den insats som studeras och den effekt som observeras. Andra aktörer har en vidare syn på evidens. Såväl OECD⁸ som Europeiska kommissionen⁹ menar att man kan utgå från både experimentell design och från andra metoder. Utredningen ansluter till detta synsätt. Inte heller den centrala administrationen inom den brittiska regeringen, som får anses ligga i frontlinjen inom detta område, gör en så snäv avgränsning att endast RCT-design räknas.¹⁰

Diskussionen om vad som kan anses vara evidensbaserat har också lett till att vissa uttolkare hellre talar om en evidensinformerad än en evidensbaserad politik eller verksamhet.

Evidensinformerad politik

Med evidensinformerad politik avses enligt Gough m.fl.¹¹ ett förhållningssätt som hjälper människor att fatta välinformerade beslut om politik, åtgärder och projekt genom att sätta *bästa tillgängliga kunskap* i centrum för politikutvecklingen och implementeringen. Det finns dock i dag ingen konsensus bland forskarna om vad begreppet evidensinformerad politik innebär i praktiken. En grundläggande skillnad mellan evidensbaserad och evidensinformerad politik är dock *vad som kan betraktas som belägg eller bevis* (evidens).

Vad som kan utgöra belägg vid utformningen av evidensinformerad politik kan utöver forskningsresultat även röra sig om expert-

⁸ Se t.ex. OECD (2007) *Evidence in Education. Linking research and policy* och OECD (2013) *Synergies for Better learning. An international perspective on evaluation and assessment*.

⁹ Se t.ex. www.ec.europa.eu/europeaid/evaluation/methodology/methods/mth_att_en.htm

¹⁰ Se till exempel *The Sage Handbook of Evaluation*, s. 170 samt *Cabinet Office The Magenta Book: guidance notes for policy evaluation and analysis*. Government Social Research Unit. HM Treasury, 2003, updated 2007, och Spencer, Liz m.fl. *Quality i Qualitative Evaluation: A framework for assessing research evidence*. Cabinet Office, Strategy unit, 2003.

¹¹ Gough, David m.fl. (2011) *Evidence Informed Policymaking in Education in Europe*. EIPEE Final Project Report. Rapporten hänvisar i sin tur till Davies P., (1999) *What is evidence-based education?* *Brittish Journal of Educational Studies* 47(2): 108–121.

kunskap, statistik, särskilda utredningar, systematiskt genomförda enkäter eller intervjuer med företrädare för berörda organisationer eller verksamheter, utvärderingar av tidigare politiska insatser, resultat från internetsökningar, indikatorer och data från uppföljningssystemet, kostnadsberäkningar och/eller output från ekonomiska eller statistiska modelleringar, dokumentanalyser, observationsanalyser (platsbesök) och analyser av administrativa data och enkätdata¹².

Med en så vid tolkning av vad som kan utgöra belägg vid ett evidensinformerat beslutsfattande skulle i princip allt svenskt nationellt politiskt beslutsfattande kunna sägas vara evidensinformerat redan genom det beredningskrav som finns inskrivet i regeringsformen. Den ambitionsnivå, som begreppet signalerar, ställer dock högre krav på evidens än vad som, enligt vår bedömning, många gånger i praktiken är fallet vid nationellt politiskt beslutsfattande i Sverige.

När OECD behandlar frågor om evidensbaserad/-informerad politikutveckling är det främst ländernas olika system för uppföljning och utvärdering man uppehåller sig vid. Viktiga delar i dessa system är olika former av nationella standards (kunskapskrav), prov för att mäta i vad mån eleverna uppnår dessa standards samt upprättande av indikatorer på måluppfyllelse, t.ex. andelen elever som fullföljt en utbildning.¹³ OECD:s fokus ligger således på att länderna i första hand ska bygga upp sina uppföljningssystem, i mindre utsträckning på utvärdering.

Evidensinformerad praktik

Systematiska kunskapsöversikter (systematic reviews) och metaanalyser utgör hörnstenar i en evidensbaserad praktik. Kännetecknande för den klassiska formen av metaanalys är att endast en forskningsdesign som bygger på experimentella, kontrafaktiskt inriktade studier med kontrollgrupp (randomiserade kontrollerade studier) anses motsvara kraven på evidens¹⁴. Denna forskningsdesign har fungerat väl inom delar av det medicinska området – som också var det område där formatet med systematic reviews först tillämp-

¹² Se t.ex. Davies, Philip (2004) *Is evidence-Based Government Possibly?* Jerry Lee lecture. 4th Annual Campbell Collaboration Colloquium. Washington D.C. samt Davies, Phillip, Kathryn Newcomer and Haluk Soydan (2006) *Government as structural context for evaluation* i Shaw, Ian F., Jennifer C. Greene and Melvin M. Mark (Ed.) *The Sage Handbook of Evaluation. Policies, programs and Practices*. Sage publications 2007. S. 171 f.

¹³ OECD (2007) *Evidence in Education: Linking Research and Policy*.

¹⁴ Om begreppens närmare betydelse se vidare kapitel 3.

ades – men är inte alltid lika tillämpbar för verksamheter som karaktäriseras av att de är komplexa¹⁵. Utbildningsområdet karaktäriseras av en sådan komplexitet liksom, för den delen, delar av hälso- och sjukvården och socialtjänsten. Begreppet evidensinformerad är därför relevant även för verksamhetsnivån. Hur begreppet avgränsas från begreppet evidensbaserad skiljer sig dock – som vanligt i dessa sammanhang – åt mellan olika forskare och brukare av dessa begrepp.

I en nyligen framlagd avhandling av Levinsson¹⁶ beskrivs t.ex. begreppet evidensinformerad som ett försök att anpassa principerna för evidensbaserad praktik så att de passar bättre till utbildnings- och skolsammanhanget. En påtaglig skillnad i förhållande till den klassiska modellen är att betydelsen av lärares professionella omdöme och praktiska kunnande betonas mer bland dem som föredrar begreppet evidensinformerad praktik. Avhandlingen hänvisar till en forskare¹⁷ som menar att ”om resultaten från systematiska översikter ska göra skillnad måste de på något vis ramas in av lärares professionella omdöme och införlivas i deras befintliga verksamhet i klassrummet”.

Enligt Levinsson menar representanter för evidensinformerad praktik att det finns ett behov av att vidga evidensbegreppet och att systematiska översikter inte enbart bör fokusera på skolans effektivitet, utan även på andra frågor som är av relevans för yrkesverksamma. Det kan i detta sammanhang vara värt att notera att bland annat två tunga aktörer på området, EPPI-center¹⁸ och Campbell Collaboration, arbetar med att utveckla metoder för att göra systematiska kunskapssammanställningar grundade på kvalitativt inriktad forskning.

Utredningen förordar begreppet evidensinformerad

Utredningen har valt att, där vi inte hänvisar till forskare eller andra författare, använda begreppet evidensinformerad framför evidensbaserad. Vi vill genom detta göra tydligt att vi menar att inte endast

¹⁵ Se kapitel 3 om utvärdering för en diskussion om vad som avses med komplexa verksamheter och om följderna av detta för utvärderingsmetodik m.m.

¹⁶ Levinsson, Magnus (2013) *Evidens och existens. Evidensbaserad undervisning i ljuset av lärares erfarenheter*. Göteborgs Studies in Educational Sciences. Göteborgs universitet.

¹⁷ Cordingley, Philippa (2004) *Teachers using evidence: Using what we know about teaching and learning to reconceptualize evidence-based practice*. I G. Thomas & R. Pring (red.) *Evidence-based practice in education*. Buckingham: Open University press.

¹⁸ The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) at the Institute of Education, London University.

forskning, utan även andra källor, kan utgöra underlag för ett beslutsfattande som utgår från evidens. Vi har också övervägt begreppet kunskapsbaserat, men avstått från att använda detta begrepp då det lätt likställs med evidensbaserat.

2.2.3 Styrningsperspektiv och utvecklingsperspektiv

Frågan om evidens i politiken kan ses från två perspektiv – ett styrningsperspektiv och ett utvecklingsperspektiv. Vilket perspektiv som dominerar i en nation har betydelse för det syfte och den roll som utvärderingen primärt kommer att ha, och får därmed avgörande betydelse för vilket slags utvärderingssystem som kommer att byggas upp.

Oavsett vilket av perspektiven som dominerar det politiska beslutsfattandet kommer det att ställa den som från nationell nivå vill stärka inslagen av evidens i beslutsfattandet inför betydande utmaningar. Det handlar för det första om hur man ska få olika parter – på nationell nivå, på lokal nivå och på verksamhetsnivå – att engagera sig i utvärderingsaktiviteter och att använda sig av resultaten från dessa aktiviteter.

För det andra handlar det om ett förändringsarbete som berör den nationella nivåns egna arbetssätt, dvs. hur arbetet inom departement och myndigheter organiseras och utförs. Inte minst ställs krav på en genomgripande analys av vilka roller som politiker och andra aktörer ska ha samt vilken ansvarsfördelning som ska råda mellan olika instanser. Att övergå till arbetssätt som innebär en förstärkning av evidensinformerade arbetsformer ställer därför även krav på både organisationsutveckling och kompetensuppbyggnad på olika nivåer i systemet.

Utvärderingssystemets utformning och innehåll m.m. påverkas av vilket perspektiv som dominerar.

Styrningsperspektivet

En aspekt på evidens i ett styrningsperspektiv är när en nations utvärderingsaktiviteter ingår i ett system för styrning av den aktuella verksamheten. Dokumenterade prestationer, som mått på skolans kunskapsresultat, kan t.ex. användas för att bedöma i vilken mån mål som fastställts för verksamheten har uppnåtts. Upp-

följning och utvärdering kan med andra ord betraktas som en förutsättning för ett mål- och resultatstyrt styrsystem. De operativt ansvariga förutsätts tolka resultaten och omsätta dessa till förändringar i verksamheten, i syfte att förbättra måloppfyllelsen. Utvärdering kan i detta sammanhang ses som ett sätt att utkräva ansvar från den lokala nivån för de nationellt fastställda målen. Detta sätt att se på utvärdering står nära det synsätt som OECD nyligen presenterat i syntesrapporten *Synergies for Better learning*.¹⁹

En annan aspekt på evidens i ett styrningsperspektiv gäller i vad mån själva utformningen av nya politiska initiativ bygger på utvärderingsresultat eller annan evidens. Själva politikformuleringen förväntas med andra ord i ökad utsträckning utgå från evidens. Under en följd av år har i ett antal länder strävan varit att på central nivå skapa arbetsformer där forsknings- och utvärderingsdata i större utsträckning utgör utgångspunkt för framtida politikutveckling. Exempel på länder i Europa som arbetar med evidensbaserad politik är enligt OECD²⁰ Norge, Danmark, Nederländerna och Storbritannien. Dessutom finns liknande exempel från t.ex. Canada och Nya Zeeland.

Intresset för ett beslutsfattande med utgångspunkt i evidens har enligt OECD²¹ ökat på skolans område, både bland politiskt ansvariga och bland yrkesverksamma. Enligt samma källa är bland annat skälen för detta en ökad oro för utvecklingen av skolornas kunskapsresultat, en ökad tillgång på information och en ökad betoning på prov och utvärdering samt förändringar i det sätt på vilket politiska beslut fattas. Det växande finansdepartementala intresset för utbildningssystemets kostnader och prestationer, insikten om utbildningens roll som drivkraft för innovation och ekonomisk tillväxt samt den uppmärksamhet som riktats mot de svaga kopplingarna mellan forskning, policy och innovation har också, enligt OECD, bidragit till att stärka intresset för evidens i utformningen av politik och praktik.²²

¹⁹ OECD (2013) *Synergies for better learning: An international perspective on evaluation and assessment. An international perspective on evaluation and assessment*. OECD reviews of Evaluation and Assessment in Education.

²⁰ Se OECD (2013) a.a.

²¹ OECD (2007) *Evidence in Education: Linking Research and Policy*.

²² OECD (2007) a. a.

Utvecklingsperspektivet

Ur ett utvecklingsperspektiv bör den evidens som genereras genom olika utvärderingsaktiviteter främst användas för att utveckla och förbättra skolans prestationer. Det gäller inte endast lokal utvärdering utan även nationell utvärdering. Utvärderingsaktiviteterna kan ingå i ett genomtänkt och sammanhållet ramverk, men utgör inte delar i ett styrsystem eller i en top-downorienterad styrkedja. Utvärderingen är i stället tänkt att fungera som utgångspunkt för en lokalt utformad skolutvecklingsstrategi där de yrkesmässigt verksamma anförtros ett stort ansvar för åtgärder som leder till resultatförbättring.

Av de länder utredningen har studerat närmare har vi funnit flera utvärderingssystem vilkas huvudsyfte vi bedömer är att leverera empiri som är användbar lokalt och kan ligga till grund för det lokala utvecklingsarbetet. I länder där utvecklingsperspektivet dominerar, som till exempel Finland och Danmark, har därför rapportering och återkoppling till de operativt verksamma en viktig funktion. Detta gäller även om det förekommer ytterligare syften, som att informera föräldrar och allmänhet samt i viss mån försörja den nationella nivån med underlag.

I länder där utvecklingsperspektivet dominerar är det även vanligt med satsningar på att bygga upp den lokala kapaciteten att hantera och utnyttja uppföljnings- och utvärderingsdata samt annan empiri. Vid utredningens besök i Finland underströks samtidigt att man är mycket försiktig med att införa inslag i utvärderingssystemet som man uppfattar skulle störa den lokala nivåns upplevelse av tillit och utrymme för eget ansvarstagande.

Ett utvärderingssystem som är utformat i ett utvecklingsperspektiv skiljer sig med andra ord på flera sätt från ett system där utvärderingsaktiviteterna främst är kopplade till ansvarsutkrävande (accountability) från lokalt skolansvariga, dvs. skolpolitiker, skolledare och lärare. Detta hindrar dock inte att drag från båda perspektiven kan förekomma parallellt i olika länder.

Enligt utredningen handlar det om att finna en rimlig avvägning mellan de båda perspektiven samtidigt som man är medveten om vilka inslag som faktiskt inte är förenliga med varandra eller som kan motverka vad man totalt sett vill uppnå. Vi återkommer till dessa frågor i kapitel 6.

2.3 Forskning och politik – en problematisk relation

Evidensbaserad politik handlar ytterst om relationen mellan politik och vetenskap – en i grunden problematisk relation. Forskning och politik styrs av olika logiker. Forskningen är till sitt innersta väsen *analytisk* och *kritiskt granskande*, dess objekt är något som funnits eller fortfarande pågår och den leder sällan till konkreta rekommendationer om vilken åtgärd som bör vidtas. Politiken är däremot till sitt väsen *framtids-, lösnings- och handlingsinriktad*. Den vill veta vad som kan och bör göras.

Det finns även en *beroende-oberoende-dimension*. Forskning kräver oberoende från olika parter intressen. Politiken är aldrig oberoende i detta avseende. En politiker har alltid sina ideologiska utgångspunkter och dessutom många grupper att ta hänsyn till: partimedlemmar, opinionsläget i väljarkåren, majoritetsställningen i riksdagen etc. Det är därför inte sannolikt att forskningsresultat – även om de i sig är entydiga och pekar mot en tydlig handlingslinje – rakt av läggs till grund för politiskt handlande.

Problemet med vilka frågor som är *möjliga att beforska*, och vilka frågor som inte är det, lägger också hinder i vägen för forskningens användbarhet i politiska sammanhang. Ju bättre avgränsad en forskningsfråga är, desto högre validitet och reliabilitet kan en studie förväntas ha. Dess intresse för politiken blir dock sannolikt låg. Begränsade frågeställningar kan inte ge svar på hur man ska lösa stora och komplexa samhällsfrågor. Sådana stora och komplexa frågor är samtidigt svåra att beforska med vetenskapligt acceptabla metoder.

Hur *relevant* forskningen kan vara för politiken är en svårlöst fråga. Det finns en uppenbar risk för *politisering* av forskningen om den låter sig styras. För att öka relevansen i forskningen kan politiken gå för långt i en önskan om att styra forskningsfrågan. Utredningen har dock funnit exempel på länder som lyckats skapa olika institutionella arrangemang som kunnat utgöra grund för samverkan mellan forskare, politiker och myndigheter.

Även frågan om *timing* är svårhanterlig. Forskning tar tid medan politiken kräver snabbt handlande. Forskningsresultat (eller resultat från utvärderingar) måste komma vid exakt rätt tidpunkt i förhållande till en berednings- och beslutsprocess för att bli en del av denna. "För tidigt" kan innebära att de kommer för tätt inpå en förändring som redan genomförts. Någon ny förändring är då på kort sikt inte tänkbar även om resultaten skulle indikera att det finns ett sådant behov. "För sent" kan innebära att resultaten inte

kan påverka beslutets utformning, eftersom beredningen redan avslutats.

En särskild aspekt av timing rör *hastigheten i det politiska reformarbetet*. Hanberger²³ menar att utvärderingsobjektet vid utvärdering av politiska reformer är ett "rörligt mål" eftersom förändringar numera sker oupphörligt. Samtidigt blir utvärdering i realtid allt vanligare, dvs. utvärderingar görs under fortlöpande reformering och följer och granskar pågående processer. Detta leder till att frågan om hur sådana förändringar i politiken ska hanteras i varje enskilt fall måste diskuteras noga i samband med att utvärderingsdesignen utformas.

Frågan har stor betydelse för vilka utvärderingsmetoder som kan användas. Vissa mer summativt inriktade modeller utgår från att inga förändringar görs av reformen under implementeringen, dvs. förändringar betraktas som avvikelser från den ursprungliga programteorin²⁴. Andra mer formativt inriktade modeller ger utrymme för att hantera förändringar som ett "naturligt" led i implementering och utvärdering.

Ett ökat fokus på evidensinformerade arbetsformer ligger väl i linje med behovet av att stärka utvärderingens roll vid nationell politikutveckling och styrning, liksom vid lokal skolutveckling. Utredningen menar att det är viktigt att försöka finna former för samverkan mellan forskare och politiker. Genom en utökad samverkan får politiker bättre underlag för sina beslut och forskare får ett ökat intresse för sina resultat. Utredningen återkommer till denna fråga i kapitel 11.

²³ Hanberger, Anders (2001) *What is the Policy Problem? Methodological Challenges in Policy Evaluation*. Evaluation Vol. 7, No. 45.

²⁴ Med programteori avses de föreställningar om orsak och verkan samt om de verkansmekanismer som finns bakom en åtgärd eller reform. Se vidare i kapitel 3.

3 Om utvärdering

När evidens blir allt viktigare såväl i samband med politiskt beslutsfattande som vid utformningen av verksamheten, blir också utvärdering allt viktigare. Utvärdering ingår dessutom som ett naturligt inslag i ett system som, liksom det svenska, bygger på mål- och resultatstyrningens principer. När offentliga verksamheter inte längre, som förr, styrs av precisa regler, utan av ramlagar med generella mål, är utvärdering helt grundläggande för att politikerna ska kunna ta ansvar för den förda politiken.

Detta kan förefalla självklart. Över tid och i olika sammanhang har dock begreppet utvärdering getts olika betydelser och tilldelats olika roller. En utgångspunkt för nedanstående genomgång av olika uppfattningar om utvärdering är utredningens övertygelse om att arbetet med att utvärdera skolpolitiska reformer och andra åtgärder kräver en mångfald av teoretiska utgångspunkter, perspektiv, frågeställningar och metoder.

3.1 Vad menar vi med utvärdering?

Utvärdering (evaluation) betecknas ofta som tillämpad samhällsvetenskap och är ett förhållandevis ungt verksamhetsområde, åtminstone som vi tänker om det i dag – som en systematisk undersökning och värdering av en verksamhet, bedriven på ett transparent och kontrollerbart sätt. Utvärdering är inte någon självständig akademisk disciplin utan ett verksamhetsområde som lånar sina teorier och sin metodologi från olika vetenskapliga discipliner.

Utvärdering utövas i en gränslinje mellan forskning och praktik. Det finns forskare som forskar om utvärdering och det finns forskare som arbetar med att genomföra utvärdering. Merparten av utvärderingar genomförs dock, helt eller delvis, utanför den strikt akademiska miljön. I olika länder finns institut som arbetar med

mer eller mindre forskningsbaserade metoder (se vidare bilaga 5). Dessa institut kan vara helt fristående eller delvis sammanlänkade med en akademisk miljö. De kan också vara mer eller mindre nära kopplade till den nationella politiska nivån. De kan även vara helt fristående företag, t.ex. konsult- och revisionsföretag, som koncentrerat sig på att utföra utvärderingar.

3.1.1 Skilda tolkningar av begreppets innebörd

Begreppet utvärdering används på en mängd olika sätt. I den omfattande internationella litteraturen inom utvärderingsområdet används t.ex. utvärdering som ett paraplybegrepp för ett antal sinsemellan olikartade aktiviteter. Gemensamt för dessa aktiviteter är att de på ett eller annat sätt kritiskt granskar en företeelse eller en verksamhet genom att systematiskt samla information om den och att granskningen utmynnar i någon form av jämförelse eller bedömning av värdet (nyttan) av denna¹. Sedan skiljer sig uppfattningarna åt vad gäller de teoretiska utgångspunkterna, vilka metoder som kan användas (kvantitativa och kvalitativa), om vem som ska göra den slutliga bedömningen etc.

Ett flertal olika indelningsgrunder har föreslagits i syfte att skapa en överblick över hur olika utvärderingsmodeller förhåller sig till varandra. Enligt en sådan indelning kan utvärdering ses som (1) en i huvudsak metodisk kompetens för att mäta, (2) en tolkande och kritiskt värderande verksamhet och (3) en verksamhet inriktad på förändring av både organisationers och individers sätt att fungera².

Andra forskare konstaterar att de olika definitionerna hör samman med olika utvärderingstraditioner, t.ex. den amerikanska programutvärderingen respektive den utvärdering som vuxit fram parallellt med den skandinaviska välfärdsstaten och som inkluderar ett organisatoriskt perspektiv³.

I många framställningar framstår utvärdering som något neutralt och värderingsfritt, en systematiserad metod för att få fram kun-

¹ Lindgren, Lena (2008) *Utvärderingsmonstret. Kvalitets- och resultatmätning i den offentliga sektorn*. Studentlitteratur.

² Karlsson, Ove (1999) *Utvärdering – mer än en metod. Tankar och synsätt i utvärderingsforskning. En översikt*. Åjour. En serie kunskapsöversikter från Svenska kommunförbundet. Nr 3.

³ Foss Hansen, Hanne (2012) *Systematic evaluation governance. New logics in the development of organisational fields*. Offentlig förvaltning. Scandinavian Journal of Public Administration.

skap om något och sedan kunna göra kloka bedömningar om detta något. Andra forskare konstaterar att utvärdering alltid genomförs inom ramen för en styrningsstruktur eller styrningsmodell som både styr utvärderingen och begränsar den⁴.

Detta spänningsförhållande mellan utvärdering som *kunskapande* och utvärdering som en del i ett *styrsystem* går som en underström i nästan allt som tänks och skrivs om utvärdering. Denna spänning kan ses som en återspeglning av det spänningsförhållande mellan ett styrningsperspektiv och ett utvecklingsperspektiv som beskrevs i föregående kapitel om evidensbaserad politik och praktik.

Från mätning till instrument för styrning

Utvärderingens roll och funktion i samhället har utvecklats över tid. Olika forskare beskriver och tolkar denna utveckling på olika sätt. Enligt en uttolkning präglades 1950- och 1960-talen av mätning och kontroll, där fokus för utvärderingen inte var att värdera utfallet och där bedömningen överlämnades till beställaren. Under 1970- och 1980-talen riktades intresset mer mot processer och mot hur resultatet kommit till. Utvärdering skulle inte enbart vara beskrivande utan förväntades också göra en kvalitativ bedömning av hur verksamheten fungerade. Under 1980- och 1990-talen utvecklades sedan en interaktiv utvärdering som bland annat betonade deltagarmedverkan för att förbättra utvärderingens relevans för intressenterna.⁵

En annan historik⁶ talar om olika utvärderingsgenerationer. Den första generationen utvärderingar beskrivs som "the Measurement generation", den andra som "the Description generation", den tredje som "the Judgement generation" och den fjärde som "Responsive Constructivist evaluation".

I en tredje historik⁷ identifieras i stället fyra olika utvärderingsvågor sedan 1960-talet: den vetenskapliga vågen, den dialog-orienterade vågen, den nyliberala vågen och den evidensbaserade vågen.

⁴ Hanberger, Anders (2012) *Framework for exploring the interplay of governance and evaluation*. Offentlig förvaltning. Scandinavian Journal of Public Administration. 16(3): 9–27.

⁵ Karlsson, Ove (1999) *Utvärdering – mer än en metod. Tankar och synsätt i utvärderingsforskning*. Åjour. En serie kunskapsöversikter från Svenska kommunförbundet. Nr. 3.

⁶ Guba och Lincoln. Se t.ex. Gröjer a.a.

⁷ Vedung och Krogstrup. Här citerade efter Foss Hansen, Hanne (2012) a.a.

I en fjärde historik⁸ urskiljs slutligen tre skikt som, enligt forskaren, alla kan återfinnas inom dagens utvärderingspraxis. Det första skiktet, ursprungligen inspirerat från USA, fokuserar utvärdering av specifika program, dvs. experiment och reformer inom den offentliga sektorn. Utvärdering inriktas mot implementering av olika politiska beslut och initiativ och organiseras ad hoc i kölvattnet på dessa.

Det andra skiktet är kopplat till system för att leda verksamhet och förbättra verksamhetens prestationer. Sådana system har utvecklats i en bred grupp anglo-amerikanska länder och fokuserar utvärdering av en viss aktör. Den bärande idén är att en organisation eller verksamhet måste granskas och utvärderas av sin huvudman – i syfte att säkra kvalitet, ansvarsutkrävande och resultat. Denna form av utvärdering organiseras mer rutinemässigt, t.ex. i form av årligen återkommande processer som definierar mål och bedömer i vilken grad målen uppfyllts.

Det tredje skiktet fokuserar utvärdering av en mångfald av enheter, organisationer och nationer. Den bärande idén är att jämförelser och transparens i en organisation främjar lärande samt skapar en stimulerande och problemlösande miljö där aktörerna strävar efter att förbättra sin konkurrenskraft. Utvärderingspraxisen organiseras som återkommande rutiner som omfattar flera organisationer.

Det finns, enligt Foss Hansen⁹, i själva verket två sätt att styra organisationer genom utvärdering. Den första metoden är *målstyrning* (performance management eller management by objectives, MBO). Vid målstyrning handlar det om att utvärdera individuella organisationer i förhållande till kriterier som fastställts av huvudmannen (jfr det andra skiktet). Till skillnad från detta syftar *systematisk utvärderingsstyrning* (SEG) till att jämföra och bedöma flera aktörer sinsemellan genom att använda en gemensam uppsättning av utvärderingskriterier (jfr det tredje skiktet).

Det kan konstateras att SEG under de senare decennierna fått en allt större internationell utbredning och ett starkt institutionellt genomslag.

⁸ Foss Hansen, Hanne (2012) a.a.

⁹ Foss Hansen, Hanne (2012) a.a.

3.1.2 Några avgränsningar

En följd av att olika aktörer definierar begreppet utvärdering på olika sätt är att begreppets relation till vissa andra näraliggande begrepp behöver klargöras.

Uppföljning och utvärdering

I Sverige görs ofta en distinktion mellan uppföljning och utvärdering där uppföljning, något förenklat, svarar på frågan *vad* som skett medan utvärdering även svarar på frågan *varför*. Uppföljning beskrivs som en regelmässig och fortlöpande insamling av information för att kontrollera hur en verksamhet genomförs. Uppföljningen är ofta en intern aktivitet som sker löpande enligt fastställda rutiner, medan utvärdering ofta genomförs av extern expertis på uppdrag av någon i den egna organisationen eller utanför denna.¹⁰

Uppföljning sker således kontinuerligt medan utvärdering sker vid särskilda tillfällen och innebär en djupare analys som avslutas med en värdering eller en bedömning. Uppföljningens mål är att genom systematiskt genomförda beskrivningar som bygger på sakkuppgifter, kontrollera vad som sker i verksamheten, vanligen genom nyckeltal, budgetutfall eller liknande, medan målet vid utvärdering i stället är att producera en bedömning av objektet för utvärderingen¹¹.

Det kan dock vara svårt att dra en tydlig skiljelinje mellan uppföljning och utvärdering eftersom uppföljningsinformation också kräver någon form av analys och värdering för att kunna användas¹². Det kan tilläggas att utvärdering på motsvarande sätt är beroende av uppföljning och ofta bygger på data som samlats in inom ramen för ett uppföljningssystem.

Den distinktion som ofta görs i Sverige mellan uppföljning och utvärdering görs inte på samma sätt i internationella sammanhang. När t.ex. OECD talar om "evaluation" är det många gånger det vi i Sverige skulle kalla för uppföljning man avser. Det handlar med andra ord om att i olika sammanhang redovisa och använda statistik och andra uppföljningsdata. Det bör även noteras att "follow up", dvs. den ordagranna översättningen till engelska av den svenska

¹⁰ Se t.ex. Ove Karlsson a.a., Lena Lindgren (2008) a.a. och Statskontoret (2001) a.a.

¹¹ Karlsson, Ove (1999) a.a.

¹² Lindgren, Lena (2008) a.a.

termen uppföljning, inte har någon självklar förankring i den angloamerikanska diskursen. Där talar man hellre om "performance measures" m.fl. liknande begrepp när man avser "uppföljningsdata".

I OECD:s *Reviews of Evaluation and Assessment in Education*¹³ används termen *evaluation* när man relaterar till bedömningar av effektiviteten i skolor, skolsystem, policys och program, medan man använder *assessment* när man relaterar till bedömningar av individuella framsteg och prestationer i förhållande till lärandemål. Det gäller i det sammanhanget såväl klassrumsbaserad bedömning som storskaliga, externa bedömningar (motsvarande de nationella proven i Sverige) och examinationer.

Med *evaluation and assessment framework* avser OECD ett samordnat arrangemang för uppföljning och utvärdering (av skolor/skolsystem/program) och bedömning (av elevers kunskapsprestationer) som ytterst syftar till att förbättra elevernas resultat inom ett skolsystem.

Detta betänkande ansluter till den svenska tradition där uppföljning och utvärdering ges olika betydelse och där utvärdering innebär en djupare analys och bedömning som inte bara syftar till att fastställa ett förhållande, utan att också förklara det. Som motsvarighet till engelskans övergripande paraplybegrepp "evaluation" används antingen samlingsbegreppet "uppföljning och utvärdering" eller begreppet "utvärderingssystem". Se vidare kapitel 7 om ramverk för utvärdering.

Utvärdering och utveckling

Även relationen mellan utvärdering och utvecklingsarbete kan vara otydlig, vilket bl.a. har att göra med att utvärdering både kan utgöras av en produkt (en utvärderingsrapport) och en process (under pågående utvärdering). I en övergripande mening kan man säga att merparten av all utvärdering har en relation till förändring eller utveckling. Även en utvärdering som är av mer summativ eller kontrollerande karaktär leder, om resultaten inte är de som förväntats, ofta till beslut om förändringar. Det gäller på såväl nationell nivå som verksamhetsnivå. Dessutom sker det ofta en

¹³ OECD (2013) *Synergies for Better Learning. An International Perspective on Evaluation and Assessment*.

påverkan på verksamheten av en utvärdering, även i de fall detta inte varit syftet.

På verksamhetsnivå kan utvärdering dessutom ha en mer nära relation till utveckling. Man kan å ena sidan tala om utvärdering *av* utvecklingsarbete, å andra sidan om utveckling *genom* utvärdering¹⁴. I det senare fallet handlar det om att utvärderingen syftar till att bidra till ett pågående förändringsarbete, dvs. att utvärderingen ges en utvecklingsdrivande funktion. Det är samtidigt viktigt att se vad som skiljer dessa aktiviteter åt. Medan utvecklingsarbetets mål är att komma fram till en förändring som är positiv för verksamheten inklusive intressenterna/brukarna är utvärderingens mål alltid att göra en bedömning eller en jämförelse, t.ex. mellan olika utfall.

I vissa forskningssammanhang talar man dessutom om exempelvis *aktionsforskning*. Denna forskningsansats betonar kunskap som utvecklas genom handling och forskningen bedrivs i ett nära samarbete mellan forskaren och praktikern. Vad som utgör skillnaden mellan aktionsforskning och en utvecklingsinriktad utvärdering kan i praktiken vara svårt att avgöra.

Utvärdering och tillsyn

Tillsyn tillämpas företrädesvis i sammanhang där staten har gjort en utfästelse i relation till en enskild medborgare eller till en grupp av medborgare. Tillsyn har det gemensamt med utvärdering att verksamheter granskas och bedöms. Enligt Statskontoret är en viktig skillnad mellan tillsyn och utvärdering att bedömningskriterierna vid tillsyn alltid är relaterade till den relevanta lagstiftningen, medan kriterierna vid utvärdering fritt kan väljas utifrån vad utvärderaren eller uppdragsgivaren är intresserad av¹⁵. Denna definition av tillsyn överensstämmer i princip med den som ges i regeringens tillsynsskrivelse¹⁶ från 2009. Där framgår dock att det i vissa fall kan finnas skäl att göra avsteg från denna strama definition och användning av begreppet tillsyn, bl.a. mot bakgrund av hur detaljerad lagstiftningen är och hur många tillsynsobjekt det finns.

¹⁴ Karlsson, Ove (1999) a.a.

¹⁵ Statskontoret (2011) *Fristående utvärderingsmyndigheter – en förvaltningspolitisk trend?*

¹⁶ Regeringens skrivelse 2009/10:79 *En tydlig, rättssäker och effektiv tillsyn*.

Utvärdering och forskning

Den kanske mest grundläggande skillnaden mellan forskning och utvärdering är att forskning inte är lika starkt knuten till externt givna ramar för sin undersökning som utvärdering. Utvärdering sker inom en mer begränsad ”uppdragsram” som formulerats av beställaren¹⁷ medan forskningen sker med utgångspunkt i en inom-forskningsstyrd ”teoriram” som formulerats av forskaren.¹⁸ Utvärdering är dessutom mer tydligt inriktad på att göra en bedömning av en specifik verksamhet eller insats, som är relevant för beställaren, medan forskning oftare är inriktad mot kunskapsutveckling och kritisk granskning.

Betecknande för forskning är vidare att det rör sig om en kumulativ process som bygger vidare på tidigare forskning och att nya resultat förväntas ge ett bidrag till en eller flera vetenskapliga diskurser (generaliserbarhet). Utvärdering syftar i stället primärt till att komma till praktisk användning.¹⁹

I Storbritannien har utvärdering inom skolområdet av tradition uppfattats som en form av vetenskap där det problem som ska studeras är bestämt av den som beställt utvärderingen, till skillnad från forskning där forskaren själv avgör vilket problem som ska beforskas. Utvärderarens frihet begränsas också, till skillnad från forskarens, genom syftet med utvärderingen – om det är formativt eller summativt.²⁰

I USA likställdes tidigare utvärdering med utvärdering av elever, dvs. bedömning. Denna starkt begränsade tolkning av vad utvärdering är – eller kan vara – i skolsammanhang, utmanades dock av tillkomsten av så kallad formativ utvärdering, dvs. utvärdering som görs innan en aktivitet har avslutats i syfte att kunna bidra till att anpassa processen och därmed förbättra utfallet av denna. Vid sådan utvärdering läggs fokus lika mycket på hur utbildningsprogrammet fungerar i verksamheten, som på om det leder till de avsedda målen eller inte.²¹

¹⁷ Karlsson, Ove a.a.

¹⁸ Även begreppet *problemram* förekommer. Med detta begrepp avses att man ringar in det som olika parter ser som gemensamma frågor att utvärdera. Karlsson menar att man genom att använda problemram kan ”komma förbi den låsning som riskerar att uppstå i kampen mellan den forskningsinriktade utvärderarens ambition att kvalificera utvärderingen utifrån teoretiska krav och beställarens önskan att styra utvärderingen genom en preciserad uppdragsbeskrivning”.

¹⁹ Statskontoret (2011) a.a.

²⁰ Harlen, Wynne (2009) *Reflections on a personal journey in research*. The Association for Science Education. (www.ase.org.uk/resources/scitutors/research/research-wynne-harlen/)

²¹ Harlen, Wynne (2009) a.a.

Forskningsbaserad utvärdering och utvärderingsforskning

Att det finns en närhet mellan utvärdering och forskning framgår av att man i utvärderingssammanhang ibland talar om *forskningsbaserad utvärdering*, dvs. utvärdering som går på djupet vad gäller teoretiska utgångspunkter, begrepp, metodik m.m. och som oftast genomförs av utvärderare med forskarkompetens.

Begreppet bör inte sammanblandas med begreppet *utvärderingsforskning*. Detta begrepp kan visserligen ges samma innebörd som forskningsbaserad utvärdering, men avser oftast forskning om utvärdering.

3.2 Olika metodologiska ansatser

Fokus i detta avsnitt ligger på utvärdering som sker på nationell nivå. Utvärdering kan givetvis även ske på lokal nivå samt inom enskilda verksamheter och det finns särskilda utvärderingsdesigner som utvecklats för att passa olika sådana syften. Exempel på sådana former av utvärdering är bl.a. följeforskning och brukarutvärdering. Dessa med flera andra liknande ansatser kommer inte att tas upp här.

Även när denna avgränsning gjorts kan det konstateras att antalet teoretiska utgångspunkter, modeller och metoder inom utvärderingsområdet i dag är mycket stort och att nya designer utvecklas ständigt. Området är alltså svåröverskådligt. Så sent som 2012 konstaterade forskare att det saknas en allmänt erkänd terminologi och typologi inom området²². Även en välkänd handbok²³ inom området avstår från att försöka ge en samlad översikt över hur olika teorier och metoder förhåller sig till varandra. Snarare än att slå fast vilken/vilka metoder som är den/de enda rätta har handboken som utgångspunkt att diskutera vad som är god utvärdering. En annan utgångspunkt är att utvärderingsmetod bör väljas utifrån syftet med utvärderingen.

Utredningen anser detta vara insiktsfulla utgångspunkter som det går att finna stöd för såväl inom forskningen som bland dem som praktiskt utför utvärderingar.

²² Gough, David, James Thomas and Sandy Oliver (2012) *Clarifying differences between review designs and methods*. BioMed Central. Systematic reviews.

²³ Ian F. Shaw, Jennifer C. Greene and Melvin M. Mark (2006) *The Sage Handbook of Evaluation*. Sage Publications London, Thousand Oaks, New Delhi.

3.2.1 Effektutvärdering

Den tydligaste skiljelinjen inom utvärderingsområdet när det gäller metodologi är, historiskt sett, den som handlar om kvantitativa och kvalitativa metoder. Denna och andra diskussioner är dock inte specifika för forskningen om utvärdering utan bör ses som återspeglningar av centrala kunskaps- och vetenskapsteoretiska diskussioner som ständigt förs i vetenskapssamhället – diskussioner som både format den vetenskapliga utvecklingen och formats av denna.

Effektutvärdering (effektmätning) uppfattas av många som starkt kopplad till kvantitativ analys och kvantitativ presentation av resultat. Effektutvärdering har också traditionellt sett varit så starkt förknippad med utvärdering att begreppen för vissa närmast framstått som synonyma²⁴.

Den kontrafaktiska utgångspunkten

Med ”effekt” menas per definition alltid en *jämförelse* mellan en faktisk uppmätt (skattad) utveckling och den utveckling som skulle ha inträffat om den gjorda insatsen inte hade gjorts, dvs. den *kontrafaktiska utvecklingen*.

Det randomiserade kontrollerade experimentet (*randomized controlled trials, RCT*) är en kraftfull teknik för effektmätning, som bygger på jämförelser mellan slumpmässigt konstruerade försöks- och kontrollgrupper. En RCT kan med hög validitet och reliabilitet ge svar på frågor om korrelation (samband) och kausalitet (orsak-verkan). Modellen omnämns därför ofta som en ”gyllene standard” för utvärdering.

När det inte finns förutsättningar för att konstruera ett experiment med en slumpmässigt konstruerad kontrollgrupp kan man använda en så kallad *kvasiexperimentell* design. Vid en sådan design åstadkoms en jämförelsegrupp, som i allt väsentligt antas motsvara experimentgruppen, men som inte är slumpmässigt konstruerad. Detta kan ske i efterhand. Ibland kan s.k. naturliga experiment uppstå, som gör att en viss population kan delas in i en experiment- respektive en jämförelsegrupp.

²⁴ Den första generationen av utvärderare har t.ex. benämnts ”the measurement generation”.

Effektutvärderingens möjligheter och begränsningar

Samtidigt som effektutvärdering i många sammanhang är en stark utvärderingsmetod står det klart att effektstudier – precis som varje annan form av utvärdering – också kan förknippas med olika svårigheter och begränsningar. De kan vara tids- och resurskrävande och de kräver också att reform- och implementeringsstrategier utformas på ett sådant sätt att de blir möjliga att genomföra.

Enligt betänkandet *Att nå ut och nå ända fram* (SOU 2009:94) är det ofta svårt att fastställa en sluteffekt om effektkedjan (eller målmedelkedjan) är lång och sker i flera steg. Liknande problem uppstår om flera förändringar sker samtidigt. Det blir då svårt att särskilja de olika effekterna från varandra. Utredningen konstaterar att skolan är ett område där svårigheter med långa effektkedjor och med att särskilja effekter från varandra är vanligt, något som pekar på behovet av kompletterande ansatser.

Synsättet att så många olika slags fakta oftast är involverade i en förändring att det inte går, eller är mycket svårt, att hänföra en viss förändring till en specifik intervention är även företrätt inom forskningen. White konstaterar till exempel ”att mäta är inte att utvärdera”²⁵. Han menar att det i många fall är rimligare att titta efter bidrag (contribution) än efter effekter (attribution). Det finns ofta fler än en händelse som påverkat en viss förändring och många resultat är inte heller dikotoma, dvs. det är inte fråga om antingen eller. Frågan är snarare hur stor andel av den samlade effekten som kan tillskrivas en viss intervention.

En annan invändning som riktats mot effektutvärdering handlar om att denna metod anses utesluta olika former av kvalitativa data, till exempel i form av enkätsvar eller genom observationer. Det har hävdats att det finns angelägna varför-frågor som effektutvärdering kan ha svårt att besvara, dvs. en typ av frågor som politiker och beslutsfattare ofta vill ha svar på.

Det finns forskare som menar att utvärdering av politiska interventioner måste gå längre än vad som vanligen sker vid studier av hur politik och politiskt beslutsfattande bidrar till utveckling av demokrati och politisk legitimitet.²⁶ Frågor måste också ställas om huruvida politiken är väl utformad för den särskilda problemsituationen och vilka värderingar och vilken ordning den främjar. Frågor

²⁵ White, Howard (2010) *A contribution to Current Debates in Impact Evaluation*. Evaluation, No. 16:53.

²⁶ Hanberger, Anders (2001) *What is the Policy Problem? Methodological Challenges in Policy Evaluation*. Evaluation, Vol. 7, No. 45.

av detta slag kan, menar man, inte besvaras inom ramen för en strikt RCT-studie.

Danmarks Evalueringsinstitut (EVA) är inne på liknande tankegångar. Vill man undersöka om insatsen har haft effekt eller inte bör man enligt EVA undersöka möjligheterna att genomföra en effektutvärdering med en experimentell ansats. Vill man i stället förstå hur och på vilket sätt en insats påverkar eller inte påverkar ett visst resultat, därför att man kanske vill förbättra eller justera insatsen, bör man använda en processinriktad, dvs. en programteoretisk, ansats. Detta behöver enligt EVA inte uppfattas som att det finns en hierarki mellan olika metoder. De svarar helt enkelt mot olika frågeställningar och levererar olika slags kunskap. Den experimentella ansatsen kan t.ex. ge ett kvantitativt mått på styrkan i ett samband, medan den processinriktade ansatsen kan bidra till en förståelse och förklaring av ett sådant samband (eller till avsaknaden av ett sådant). Det kan därför finnas goda skäl att kombinera dessa båda ansatser i stället för att se dem som uteslutande varandra.²⁷

Utredningen konstaterar att det hela tiden pågår en utveckling av metoder för effektutvärdering, inte minst till följd av att den statistiska metodiken förfinas alltmer. Inte sällan handlar de metodologiska nyheterna om mer avancerade modeller för att hantera bias²⁸ av olika slag. Den experimentella ansatsen utvecklas också på så sätt att man inte endast intresserar sig för effekter i form av resultat utan även för att undersöka processer och mekanismer.²⁹

3.2.2 Teoribaserad utvärdering

Teoribaserad eller teoridriven utvärdering kan vid sidan av effektutvärdering ses som en andra huvudinriktning inom utvärdering.

De teoribaserade utvärderingsmodellerna utgår från en teoretisk referensram som beskriver utvärderingsobjektets tillkomst, organisation, mål och verksamhet, samt kopplingar mellan dessa och mellan organisationen och omvärlden.

²⁷ Danmarks Evalueringsinstitut *Hvordan kan man evaluere effekt?* (www.eva.dk/udgivelser/2012/effektevaluering-notat)

²⁸ Snedvridning. Allmänt: felaktigheter i vetenskapliga resultat p.g.a. något systematiskt fel i forskningsprocessen vad gäller t.ex. insamling av data, bearbetning eller analys av resultat. Källa: NE.

²⁹ Se t.ex. Jens Ludwig, Jeffrey R. Kling and Sendhil Mullainathan (2011) *Mechanism Experiments and Policy Evaluations*. Journal of Economic Perspectives, Vol. 25, Nr. 3.

EU-kommissionen³⁰ anger att teoribaserad utvärdering fokuserar på de teorier som ligger inbäddade i en policy eller ett program, dvs. de antaganden som beslutsfattare och intressenter gör avseende de förutsättningar, mekanismer och ramar som krävs för att en intervention ska fungera. Teoribaserade effektutvärderingar testar dessa antaganden mot de resultat som observeras (empiri) i samband med de olika stegen i programteorin och undersöker även andra faktorer som kunnat påverka utfallet. Därigenom blir det möjligt att förklara varför och hur resultaten har inträffat och att bedöma i vilken mån det specifika programmet bidragit till resultaten och i vad mån andra faktorer påverkat dessa.

Kommissionen menar att det är av största betydelse för beslutsfattandet att förstå varför och under vilka förutsättningar en intervention av något slag påverkar och ger upphov till effekter. Teoribaserade utvärderingar gör det möjligt att bedöma om ett program fungerar, hur det fungerar och i vilket sammanhang. Samtidigt skapar de en ram för att bedöma effekten av komplexa program som integrerar olika interventioner och kan användas för verksamhet som inte lämpar sig för utvärderingar av kontrafaktiska effekter.

Programteori

En utgångspunkt för teoribaserad utvärdering är att det går att bilda sig en uppfattning om vilka antaganden som beslutsfattaren gjort om hur reformen eller åtgärden är tänkt att verka för att uppnå avsett syfte, en s.k. programteori.

Programteori kan beskrivas som välgrundade och väl motiverade uppfattningar om hur och varför en intervention antas påverka något³¹. Genom att redogöra för de olika sätt som en intervention kan tänkas påverka en verksamhet, en situation eller ett resultat kan man klargöra vilka mekanismer som fungerar och hur de gör det. Programteorin ger stöd för vad som bör undersökas i utvärderingen och i vilken ordning samt är ett viktigt led i bedömningen av en eller flera effekter. Programteorin bör därför utformas samtidigt som själva utvärderingsfrågan formuleras.

Programteorin uttrycks sällan explicit i ett visst beslut eller i de underlag som tagits fram inför beslutet. Inom forskningen om

³⁰ www.ec.europa.eu/regional_policy/information/evaluations/impact_faq_theor_en.cfm

³¹ Dahler-Larsen, Peter (2005) *Vaikuttavuuden arviointi. (Utvärdering av effektivitet)*. Stakes.

utvärdering har det utvecklats olika metoder för hur en programteori kan konstrueras i efterhand. Programteorin bör utarbetas med stöd av olika källor som kan ge en uppfattning om t.ex. en regerings antaganden om relationen mellan åtgärd och utfall för en tänkt reform. Källor kan i detta fall vara t.ex. bakomliggande dokument eller information från intervjuer med personer som haft insyn i förberedelserna för reformen. Svaret ges sällan i en proposition.

Programteorin utformas ofta med hjälp av formen ”om – så”³². För detta finns olika metoder, t.ex. kan man ta fram s.k. CMO-konfigurationer (context-mechanism-outcomes) som visar hur en viss mekanism producerar ett visst resultat under vissa förhållanden. Försök har även gjorts att fastställa effekten av sådana mekanismer med hjälp av experiment. De metoder som utvecklats för att göra effektutvärderingar har på så sätt visat sig användbara för att bekräfta eller förkasta de resultat som framkommit i studier med en programteoretisk ansats³³.

3.2.3 Utvärdering av komplexa system

Utgångspunkten för den typ av programteori som den teori-baserade utvärderingsansatsen bygger på, är att det föreligger linjära orsakssamband (kausalitet), dvs. att saker inträffar i en viss följd. ”Om A görs, så inträffar B och resultatet blir C ...”. Många har dock konstaterat att offentlig verksamhet sällan består av sådana enkla situationer. Snarare präglas offentlig verksamhet, som till exempel, sjuk- och hälsovård, socialt arbete och utbildning, av att verksamheten är komplex.

Man skiljer ofta på enkla, komplicerade och komplexa situationer eller verksamheter. En *intervention i en enkel situation* karaktäriseras av det är enkelt att utforma åtgärden och att effekten av en denna enkelt kan förutses och upprepas med gott resultat. Inom sjukvården brukar man exemplifiera med att ta en huvudvärkstablett. En *intervention i en komplicerad situation* karaktäriseras av att åtgärden visserligen är komplicerad, men att effekten kan förutses och upprepas. Inom sjukvården brukar knäledsoperationen ges som exempel. Ett exempel från ett annat område gäller rymdfärder, som visserligen bygger på en komplicerad teknik, men när denna väl har

³² Leeuw, Frans L. (2003) a.a.

³³ Ludwig, Jens & Jeffrey R. Kling & Sendhil Mullainathan (2011) *Mechanism Experiments and Policy Evaluations*, Journal of Economic Perspectives, Vol 25 No 3.

utvecklats kan färden upprepas med förväntat resultat. En *intervention i en komplex situation* kännetecknas däremot av att en åtgärd inte självklart får samma effekt när den upprepas. Barnuppfostran har angivits som ett exempel på en komplex situation.³⁴

Komplexa situationer kännetecknas av *icke-linjär kausalitet*, dvs. att sambandet mellan orsak och verkan inte är tydligt, direkt eller proportionerligt. Komplexa situationer präglas även av *oförutsägbarhet* eftersom det finns många aktörer som påverkar och påverkas av varandras beteenden, och som i sin tur påverkar och påverkas av systemet samtidigt som externa faktorer ständigt påverkar situationen.

Vidare kan det finnas *flera samtidiga eller flera alternativa kausala vägar*, dvs. det finns flera olika sätt som verksamheten kan genomföras på för att de slutgiltiga effekterna ska uppstå. Forskare³⁵ talar till exempel om *framväxande resultat* (emergent outcomes), dvs. att de specifika resultaten uppstår under hand under implementeringen av en intervention. Fenomenet med framväxande resultat är kopplat till att människors beteendemönster inte är konstanta i komplexa, föränderliga system. När systemets omgivning förändras så förändras även beteendet hos de som agerar i systemet, och, som en följd av detta, förändras systemet som helhet. Sådana framväxande resultat är vanliga i samband med omfattande samhälleliga interventioner.

Det är lätt att konstatera att utbildning är en komplex verksamhet, där samma åtgärd långt ifrån alltid leder till samma effekt inom alla skolenheter. Ett skäl till detta är att interventioner i skolan i hög grad är beroende av mänskligt beteende. Skolan är därtill ett omfattande socialt system med flera olika ansvarsnivåer. Insatser av olika slag beslutas på en nivå i systemet men implementeras i en lokal verksamhet där varje miljö är unik. Som en följd av detta uppstår varierande (eller heterogena) effekter.

Eftersom mänskligt beteende varierar under påverkan av olika kontexter är det nödvändigt att intressera sig för teorier som kan bidra till att förklara och "förutsäga" hur särskilda kontexter påverkar individer att agera på ett visst sätt och därigenom åstadkomma ett visst specifikt utfall eller resultat.

³⁴ Glouberman, Sholom & Brenda Zimmerman (2002) Complicated and Complex Systems: What Would Successful Reform of Medicare look like? Commission on the Future of Health Care in Canada. Discussion paper, no. 8.

³⁵ Rogers, Patricia (2008) *Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions*. Evaluation, 14:29.

Forskningen har gjort vissa framsteg i fråga om möjligheten att formulera detta slags förklaringar och sådan "förutsägbarhet". *Realist evaluation*³⁶ är en ansats som bland annat fokuserar vilka grupper som gynnats respektive inte gynnats av ett program eller en åtgärd. Bedömningen av en åtgärds effektivitet handlar då inte endast om att bedöma utfallet i termer av orsak och verkan, t.ex. genom att ställa en grupp som berörts av åtgärden mot en grupp som inte berörts, utan också om att bedöma variationen inom den grupp eller mellan de grupper som berörts av åtgärden.

Slutsatsen är således inte att det är omöjligt att utvärdera interventioner inom en komplex verksamhet som skolan. Däremot krävs en god utvärderingsdesign, som i många fall bygger på en variation av olika metoder och är utformad med utgångspunkt i en väl genomtänkt reformstrategi (se vidare kapitel 6).

3.2.4 Systematiska översikter

En systematisk översikt (systematic review) är inte en utvärderingsdesign utan en sammanställning av "bästa möjliga kunskap" inom ett visst område. Arbetet med att ta fram en systematisk översikt följer ett rigoröst protokoll som syftar till att säkerställa att den slutliga bedömningen bygger på ett urval av relevanta studier av högsta vetenskapliga kvalitet. Av protokollet framgår vad som krävs i form av sökstrategi vid sökning i databaser, transparens vad gäller urvalsmetoder och former för att säkerställa vetenskaplig kvalitet. En vanlig metod är peer review, dvs. att kvalificerade oberoende forskare granskar metoder och resultat.

Systematiska översikter har kommit att bli en viktig beståndsdel i den strömning för evidensbaserad politik och praktik som vuxit sig stark i många länder.³⁷ Strömningen går tillbaka till den brittiske läkaren Archibald Cochrane som i början av 1970-talet hävdade att medicinsk behandling skulle kunna bli mycket effektivare om man i större utsträckning arbetade med vetenskapligt kvalitetssäkrade metoder³⁸.

³⁶ Pawson, Ray & Nick Tilley (1997) *Realistic Evaluation*. Sage.

³⁷ Framställningen utgår från Bohlin, Ingemar (2010) *Systematiska översikter, vetenskaplig kumulativitet och evidensbaserad pedagogik*. Pedagogisk forskning i Sverige, årg. 15, nr. 2/3.

³⁸ Cochranes, Archibald (1972) *Effectiveness and Efficiency: Random Reflections on Health Services*. London: Nuffield Provincial Hospitals Trust, 1972. (Reprinted in 1989 in association with the BMJ, Reprinted in 1999 for Nuffield Trust by the Royal Society of Medicine Press, London.

Det stora genomslaget för Cochranes idéer kom sedan andra forskare³⁹ utvecklat en metod för att jämföra effekter av olika behandlingar och att genom ett enhetligt mått, *effektstorlek*, göra individuella undersökningar jämförbara. En *metaanalys* är den term som används när en sammanvägning görs av ett stort antal studiers effektstorlek till ett sammanvägt mått.

Metoden att fastställa effektstorlekar medförde att de systematiska översikterna till en början helt kom att bygga på kvantitativa metoder. Tidigt började dock en kritik växa fram mot att metoden uteslöt andra, mer kvalitativt inriktade forskningsdesigner. Forskare med samhällsvetenskaplig och social inriktning menade att det är både möjligt och nödvändigt att göra systematiska översikter byggda även på kvalitativt inriktade studier. Annars kommer helt enkelt inte ett stort antal frågor att kunna besvaras. Successivt har ståndpunkten vunnit mark att även kvalitativa studier, som bygger på t.ex. observationer och intervjuer, kan ingå i underlaget för systematiska översikter. *Metasyntes* har kommit att bli en samlingsbeteckning för metoder med vars hjälp resultaten av kvalitativa studier kan integreras.

Inom vetenskapssamhället har en intensiv diskussion förts om den evidensbaserade ansatsens anspråk och begränsningar. Å ena sidan finns ett brett stöd för behovet av kunskapsöversikter och synteser av forskningsläget inom olika områden – dels som grund för den fortsatta inomvetenskapliga utvecklingen, dels som grund för professionell praktik och beslutsfattande på policynivå. Å andra sidan har det funnits en diskussion om vad som kan utgöra evidens och därmed vilken slags forskning som kan ingå i en systematisk översikt. Diskussionen har även rört hur forskningsresultat kan omsättas i god praktik och vilka förutsättningar som krävs för att detta ska kunna ske.

Systematiska översikter syftar till att ta fram bästa möjliga kunskap som grund för politiska beslut eller för praktiskt verksamhetsbeslut om handlingar, aktiviteter, processer m.m. De har alltså en viktig roll att spela i den evidensinformerade politik och verksamhet som diskuterades i föregående kapitel.

Systematiska översikter syntetiserar redan befintlig kunskap. Om det saknas studier (nationella eller internationella), eller om

³⁹ Gene Glass tillsammans med Mary Lee Smith.

dessas inte håller tillräcklig vetenskaplig kvalitet, kan inte heller några systematiska översikter produceras.

3.3 Sammanfattande bedömning

Genomgången av utvärderingslitteraturen tydliggör att det kan vara svårt att orientera sig rätt bland alla olika teorier och utgångspunkter. Det framgår samtidigt att det också finns goda möjligheter att genomföra utvärdering, även i de fall förutsättningarna i förstone inte verkar finnas.

Gemensamt för god utvärdering av en reform eller insats är sammanfattningsvis att det finns

- en systematisk insamling av data (kvantitativ eller kvalitativ) till grund för analys, tolkningar och bedömningar,
- en metodik vid behandlingen av data (kvantitativ eller kvalitativ) som uppfyller rimliga krav på validitet och reliabilitet, samt
- en värdering (av utvärderingsobjektets nytta, effekter, påverkan eller motsvarande) i förhållande till något slags måttstock (riktvärden, kriterier eller referenspunkter).

Utöver dessa krav kan utvärderingar vara av en mängd olika slag, för att möta olika krav och förutsättningar: effektutvärdering, programteoretisk utvärdering, processutvärdering, etc.

Det behövs en mångfald av metoder

Genomgången av forskning inom området visar att flertalet utvärderingsexperter och utvärderingsforskare är överens om att metoder för utvärdering bör väljas med utgångspunkt i själva objektet för utvärderingen och i de frågeställningar som ska utvärderas. Flerparten forskare menar också att kombinationer av olika metoder kan vara en god strategi för att fånga olika aspekter av en mer eller mindre sammansatt verklighet.

I likhet med flertalet forskare och utvärderare ansluter vi oss till ett utvärderingsbegrepp som kan omfatta såväl att mäta effekter och skatta samband som att bedöma eller värdera ett utvärderingsobjekt, att söka efter förklaringar till varför vissa effekter eller

resultat uppstått (eller inte uppstått) samt att försöka förstå olika processer och variationer i utfall.

Valet av metod för en specifik utvärdering bör vidare göras utifrån den eller de aktuella utvärderingsfrågorna och kvantitativt och kvalitativt inriktade metoder bör inte ställas mot varandra. Det behövs utvärdering av båda slagen.

Utvärdering av reformer ställer krav på att beställaren har en hög metodisk medvetenhet. En sådan medvetenhet innebär bland annat att vara uppmärksam på att det kan vara riskabelt, många gånger direkt olämpligt, att slå fast vilken slags utvärderingsmetod som ska användas – metodfrågor hanteras bäst av utvärderaren. För att åstadkomma relevanta utvärderingar av god kvalitet bör beställaren av utvärderingen däremot noga överväga vad det är för kunskaper man vill få ut som resultat av utvärderingsinsatsen, dvs. i vilket syfte utvärderingen ska genomföras, och hur man vill att utvärderingens resultat ska användas.

Utvärderingssystemet bör kunna hantera komplexitet

Forskningen har tagit fasta på att många offentliga verksamheter är sammansatta till sin karaktär. Man har använt olika begrepp för att beteckna denna karaktär – enkla respektive komplexa situationer, tama respektive svårlösta samhällsproblem⁴⁰ – men den verklighet man velat komma åt uppfattar vi är densamma.

Distinktionen mellan enkla och komplexa situationer är av stor betydelse för hur en utvärdering kan utformas och vilken slags vetande en utvärdering kan resultera i. Skolans verksamhet uppvisar alla kännetecken på en komplex och svårlöst verklighet. Detta får till följd att även interventioner i skolan i de flesta fall har en komplex karaktär, vilket både den övergripande reformstrategin och valet av metoder för utvärdering bör förhålla sig till.

Effektstudier med RCT-metodik kan i många sammanhang vara lämpliga. I andra sammanhang kan de leda till att regeringen endast får svar på mer begränsade frågeställningar och kommer att bli utan svar på angelägna frågor, som den kunde fått svar på om även andra utvärderingsmodeller använts. För att regeringen ska få maximalt utbyte av sina utvärderingsuppdrag ställs därför krav på beställar-kompetens och på att metodval inte låses i onödan.

⁴⁰ Begreppen tama respektive svårlösta samhällsproblem diskuteras t.ex. i betänkandet *Att tänka nytt för att göra nytta – om perspektivskiften i offentlig verksamhet* (SOU 2013:40).

Mångfald och komplexitet påverkar organiseringen

Utvärdering har en roll i olika skeden av en reform (före, under och efter) och kan genomföras med skilda strategier och metoder. Det medför att utvärderingen kan behöva organiseras på olika sätt. Det kan ske helt eller delvis inom ramen för ett existerande utvärderingssystem, men det kan också ske genom att det skapas särskilda projekt eller uppdrag kopplade till en viss reform. Här väcks frågor om hur ett utvärderingssystem bör vara utformat för att kunna bidra med den kunskap som behövs, inte minst för att kunna identifiera behov av framtida politiska reformer eller andra åtgärder. Vilka beståndsdelar behöver det bestå av, hur hör de ihop, vem ansvarar för vad, vilken information/data behövs, etc? Här aktualiseras även frågor om när det behövs mer oberoende särskilda utvärderingar av reformer och hur dessa i så fall ska organiseras och förhålla sig till det befintliga utvärderingssystemet. Utredningen återkommer till dessa frågor i det följande.

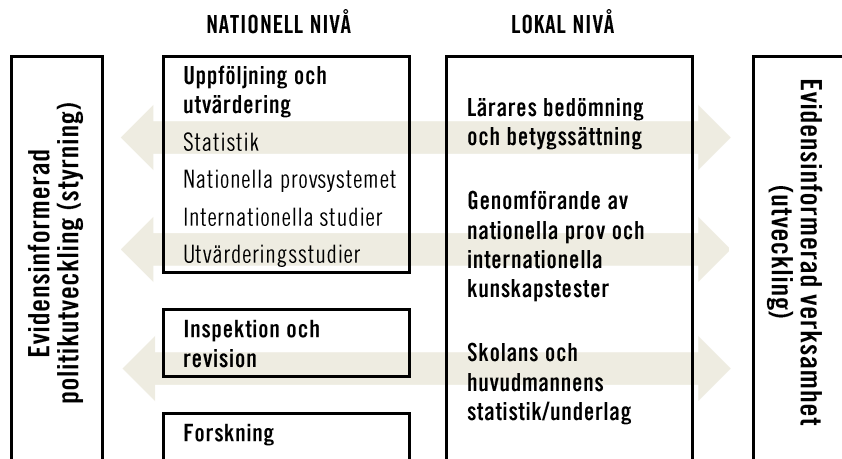
4 Dagens utvärderingssystem

För den fortsatta diskussionen om utvärdering av reformer, och behovet av data för sådan utvärdering, behöver utgångspunkt tas i dagens nationella system för uppföljning och utvärdering av skolan. Utöver uppföljning och utvärdering består systemet av viss tillsyn och granskning samt viss forskning. Systemet bidrar fortlöpande till systemkunskap om skolan, och utgör underlag för utvärdering av reformer. Här redogörs därför för dagens nationella uppföljnings- och utvärderingssystem. Redovisningen avser de väsentligaste inslagen i utvärderingssystemet och de aktörer som ansvarar för utvärderingen.

4.1 Det nationella uppföljnings- och utvärderingssystemet

Huvudobjektet för detta avsnitt är det nationella systemet för uppföljning och utvärdering av skolan. Figur 4.1 är ett försök att schematiskt åskådliggöra den nationella uppföljningen och utvärderingens olika beståndsdelar och dess placering i gränslinjen mellan den nationella politikens styrning och skolans praktik. Modellen är extensiv i den meningen att den utöver vad som självklart kan räknas in som uppföljning och utvärdering även omfattar viss inspektion och granskning samt viss forskning. Gränserna mellan de olika aktiviteterna är inte skarpa. Till exempel kan en utvärdering genomföras med underlag från statistiksystemet eller med utgångspunkt i data från t.ex. de internationella studierna. En utvärdering kan också genomföras som ett forskningsprojekt, etc. Modellen syftar till att skapa en översiktlig struktur för den fortsatta framställningen.

Figur 4.1 Det nationella uppföljnings- och utvärderingssystemet för skolan



Modellen syftar till att tydliggöra att det nationella systemet å ena sidan vilar på lärares bedömning och betygssättning och på skolans och huvudmannens uppgiftslämning, dvs. på uppgifter från verksamheten, och å andra sidan kan bidra till lokal skolutveckling. Vad som inte framgår av bilden är bland annat det uppgiftslämnande från den lokala nivån som består i att elever, lärare och skolledare svarar på enkäter eller på annat sätt bidrar med underlag till olika regelbundna studier, liksom det uppgiftslämnande som är förknippat med sådana unika utvärderingsstudier som inte ingår i den regelbundna uppgiftsinsamlingen.

Vad som dessutom inte framgår, men som vi återkommer till i det följande, är vilka de centrala aktörerna är och ansvarsfördelningen dem emellan. Eftersom bilden ovan fokuserar det nationella uppföljnings- och utvärderingssystemet framgår inte heller de viktiga utvärderingsaktiviteter som bör genomföras av skolan och huvudmannen för att svara på lokala behov.

4.1.1 Uppföljning och utvärdering

Skolverket har det huvudsakliga ansvaret för det nationella uppföljningssystemet. Nationella utvärderingsstudier genomförs av Skolverket och av även andra aktörer.

Det nationella uppföljningssystemet

Skolverkets nationella uppföljning

Skolverket är statistikansvarig myndighet för den *officiella statistiken* om skolväsendet¹. Det är dock Statistiska Centralbyrån (SCB) som på Skolverkets uppdrag svarar för att samla in uppgifterna från skolhuvudmännen.

Syftet är att den officiella statistiken ska ge en samlad bild av skolans verksamhet. Varje år samlas uppgifter in om barn, elever, personal, kostnader och utbildningsresultat. Uppgifterna sammanställs till statistik per verksamhetsform och skolform. Statistiken presenteras på såväl riks-, läns-, kommun- som skolnivå.

För grundskolan finns *strukturella uppgifter* om t.ex.

- antal elever och grundskolor, fördelat på huvudman, årskurs samt kön,
- moderna språk, svenska som andraspråk och modersmål,
- antal lärare och annan pedagogisk personal som arbetar i skolorna, lärartäthet och andelen lärare med pedagogisk högskoleexamen samt vilka ämnen lärarna undervisar i, samt
- totalkostnader och kostnader per elev samt kostnader för undervisning, lokaler, läromedel, elevvård, skolmåltider m.m.

Dessutom finns statistik över studieresultaten. Den består av *slutbetyg för eleverna i årskurs 9* bland annat i form av genomsnittligt meritvärde, behörighet till nationella program i gymnasieskolan och betyg per ämne. Från och med 2013 finns även *betygsresultat för eleverna i årskurs 6*. Betygsuppgifterna fördelas på huvudman, årskurs, kön, svensk respektive utländsk bakgrund och i vissa fall även efter invandringsår samt på föräldrarnas utbildningsnivå.

Vidare finns statistik över *resultat från de nationella proven* för eleverna i årskurs 9. För vissa år finns även resultat från årskurs 5 och årskurs 3. Från och med 2012 ersätter resultat från årskurs 6 de tidigare resultaten från årskurs 5. Uppgifter fördelas på andel elever totalt, per kön, utländsk bakgrund och i vissa fall efter föräldrarnas högsta utbildningsnivå. (Se vidare nedan.)

Karaktäristiskt för den officiella statistiken är att den är stabil över tid. Justeringar med hänsyn till förändrade strukturer och

¹ Enligt förordningen (2001:100) om den officiella statistiken.

bestämmelser måste naturligtvis genomföras, men stor vikt läggs vid att det ska vara möjligt att följa utveckling över tid. En hel del uppgifter lämpar sig dock inte särskilt väl för jämförelser över tid. När man vill följa reformer är det t.ex. ett problem att statistiken över betyg och resultat på nationella prov knappast lämpar sig för trendanalyser. Det gäller särskilt i samband med en reformperiod av det slag som vi nu befinner oss i, då förändringar i både betygs- och provsystemet ingår i det samlade reformpaketet. Vi har belyst detta i vårt delbetänkande² och återkommer till frågan i kapitel 9.

Utöver den officiella statistiken samlar Skolverket regelbundet in uppgifter om attityder till skolan, som publiceras i rapporten *Attityder till skolan*, och som måste anses ingå i det nationella uppföljningssystemet. Samma sak gäller t.ex. Skolverkets uppföljning om it-användning respektive om simkunnighet.

Andra aktörers insamling av longitudinella uppföljningsdata

Sedan början av 1960-talet har Statistiska centralbyrån (SCB) i samverkan med olika forskningsinstitutioner genomfört uppföljningsundersökningar inom skolväsendet. Dessa undersökningar är sedan 1990 sammanslagna till ett forskningsprojekt vid Göteborgs universitet benämnt "Utvärdering Genom Uppföljning" (UGU), fortfarande i samarbete med SCB.

Uppföljningsundersökningarna är ett led i den nationella utvärderingen av skolan och utgör underlag för forskning inom framförallt det samhällsvetenskapliga forskningsområdet. Hittills har nio uppföljningsundersökningar (s.k. elevpaneler), som baseras på stora och riksrepresentativa stickprov från olika årskullar av elever, startat. Det första urvalet gjordes bland födda 1948. För varje stickprov insamlas administrativa data och enkätuppgifter. Vidare ingår resultat från olika kognitiva test och kunskapsprov. Dessa görs av Pedagogiska institutionen vid Göteborgs universitet. De har varit de samma ända sedan den första kohorten, vilket gör databasen unik även i ett internationellt perspektiv. Panelerna följs sedan de slutat grundskolan med hjälp av registerdata samt genom enkätundersökningar som vänder sig till elever, lärare och föräldrar.

Avsikten med de omfattande datainsamlingarna är att möjliggöra såväl längdsnitts- som tvärsnittsundersökningar av stora och representativa elevurval. Härigenom har man bl.a. kunnat belysa

² SOU 2013:30 *Det tar tid – om effekter av skolpolitiska reformer*.

hur olika faktorer i uppväxtmiljön successivt påverkat studieval och studieprestationer, i vilken grad denna påverkan förändrats mellan olika årskullar samt vilken betydelse olika utbildningspolitiska insatser haft i detta sammanhang. Insamlade data har också gett underlag till ett antal studier inom andra discipliner inom det samhällsvetenskapliga forskningsområdet samt ett antal utvecklingspsykologiska studier.

Sedan 1980 svarar SCB för de s.k. elevpanelerna för longitudinella studier, vilka ingår i den officiella statistiken. SCB:s insamling utgör även ett av underlagen för UGU, som dock kompletterar underlaget med egna tester och enkäter. SCB följer ett urval av elever genom grundskolan från och med skolår 3. Eleverna följs sedan vidare genom gymnasium, högskola och ut i arbetslivet. Sedan 1980 har paneler hittills startats åren 1980 för elever i årskurs 6 samt 1982, 1987, 1992, 1997, 2002 respektive 2008 för elever i årskurs 3. Frågor ställs varje år till skolan så länge som eleven går i grundskolan. Datainsamlingen sker genom en enkät som ställs till skolan om bl.a. stödundervisning, hemspråksundervisning, språkval och betyg. Dessa data kompletteras sedan med uppgifter ur register för att följa eleven efter grundskolans slut. Elevpanelerna är i första hand ett samarbete mellan SCB och UGU vid Göteborgs universitet.

Nationella provsystemet

Skolverket har regeringens uppdrag att svara för det nationella provsystemet, som består av nationella prov och andra bedömningsstöd. De nationella proven utformas dock inom universitetsinstitutioner, på uppdrag av Skolverket.

Det är regeringen som beslutar om i vilka ämnen, årskurser och skolformer som nationella prov ska genomföras. För årskurs 3 i grundskolan ges i dag obligatoriska nationella ämnesprov i matematik samt svenska och svenska som andraspråk. Ämnesproven i årskurs 6 respektive i årskurs 9 ges i engelska, matematik, svenska och svenska som andraspråk samt i biologi, fysik eller kemi och geografi, historia, religionskunskap eller samhällskunskap. Vid samma provtillfälle genomför skolan ett av ämnesproven i NO-ämnen respektive ett av ämnesproven i SO-ämnen. I specialskolan genomför eleverna motsvarande ämnesprov i årskurserna 4, 7 och 10. I sameskolan, som är en sexårig skolform, genomförs proven i årskurs 3 och 6.

Enligt de olika uppdrag³ utifrån vilket dagens nationella provsystem ursprungligen utformats har systemet flera olika syften. Systemet ska

- bidra till ökad måluppfyllelse för eleverna,
- förtydliga målen och visa på elevers starka och svaga sidor,
- konkretisera kursmål och betygskriterier,
- stödja en likvärdig och rättvis bedömning och betygssättning och
- ge underlag för en analys av i vilken utsträckning kunskapsmålen nås på skolnivå, på huvudmannanivå och på nationell nivå.

De nationella proven är inte examensprov, utan ska vara en del av lärarens samlade information om en elevs kunskaper. De har framför allt en summativ funktion, vilket betyder att de ska fungera som en avstämningspunkt i slutet av vårterminen i respektive årskurs/kurs och visa vilka kvaliteter eleven har i sina kunskaper i de ämnen där proven genomförs. De nationella proven kan emellertid även i viss utsträckning användas formativt, som en del i den bedömning för lärande som är en del av undervisningen. Provresultaten ger information om vilka kunskaper som utgör styrkor och vilka kunskaper som eleven genom undervisningen behöver utveckla mer. Proven ger även en bild av hur undervisningen har fungerat, vilket i sin tur kan signalera om hur undervisningen kan utvecklas.

Som komplement till de nationella ämnesproven utvecklar Skolverket kontinuerligt bedömningsstöd i form av kartläggnings- och diagnosmaterial, prov och enskilda uppgifter med tillhörande bedömningsanvisningar och diskussionsmaterial.

Skolverket konstaterar i en rapport⁴ till regeringen med förslag till hur systemet bör utvecklas att det är svårt att med ett enskilt instrument möta alla dessa syften och att provsystemet därför består av olika typer av material. Utredningen konstaterar att även om de olika syftena gäller för det nationella provsystemet som helhet, dvs. inklusive de bedömningsstödjande insatserna, förväntas ändå de nationella proven ensamma svara upp mot merparten av syftena. Regeringen konstaterar i november 2011 i ett uppdrag⁵ till

³ U2004/5293/S, U2005/8381/G, U2006/8951/S, U2008/1045/SV och U2008/2645/S.

⁴ Skolverket (dnr. 2009:337) *Redovisning av regeringsuppdrag att ge förslag på hur det nationella provsystemet bör utvecklas och utformas.*

⁵ Skolverket (dnr. 2011/6543/S) *Uppdrag om nationella prov.*

Skolverket att de nationella proven inte i nuläget är utformade för att mäta kunskapsutvecklingen över tid. Det framgår av uppdraget att det förväntas att detta ska vara möjligt. Vi återkommer i kapitel 9 till frågan om förutsättningarna att knyta så många olika förväntningar till ett och samma instrument som det görs i dag vad gäller de nationella proven.

Internationella studier

För att hålla för trendanalys krävs i de flesta fall att statistik eller studier utformas med detta syfte för ögonen. De enda studier som Sverige för närvarande deltar i med någon regelbundenhet, och som svarar mot de krav som kan ställas på trendanalyser, är de internationella kunskapsmätningarna PIRLS⁶, TIMSS⁷ och PISA⁸. Förutom kunskapsprov innehåller studierna enkäter till elever och rektorer samt – vad gäller PIRLS och TIMSS – även till lärare och ibland till föräldrar. Detta gör det möjligt att också följa de olika gruppernas attityder och inställning till skolan och skolarbetet.

För närvarande är strategin att Sverige ska delta i samtliga tre undersökningar och att detta ska ske för både yngre och äldre elever i grundskolan. Bakåt i tiden har dock deltagandet varierat något. Sverige har deltagit i samtliga IEA⁹:s PIRLS-undersökningar av läsförmågan för elever i årkurs 4, dvs. sedan 2001. Undersökningen genomförs vart femte år. Sverige deltog även 1971 och 1991 i IEA:s tidigare undersökningar om läsning.

Sverige har även deltagit i samtliga OECD:s PISA-undersökningar av 15-åringars kunskaper i matematik, naturvetenskap och läsförståelse, varav den första genomfördes 2000. Studien genomförs vart tredje år med olika huvudområden vid de olika undersökningstillfällena. Den trendutveckling som kan identifieras mäts i

⁶ PIRLS (Progress in Reading Literacy Study) är en internationell studie som undersöker elevers läsförmåga i årkurs 4. Studien organiseras av The International Association for the Evaluation of Educational Achievement (IEA).

⁷ TIMSS (Trends in International Mathematics and Science Study) undersöker kunskaper i matematik och NO hos elever i årskurs 4 och årskurs 8. Studien organiseras av The International Association for the Evaluation of Educational Achievement (IEA) och genomförs vart fjärde år.

⁸ PISA (Programme for International Student Assessment) är ett OECD-projekt som syftar till att undersöka i vilken grad respektive lands utbildningssystem bidrar till att femtonåriga elever, som snart kommer att ha avslutat den obligatoriska skolan, är rustade att möta framtiden. Studien genomförs för kunskapsområdena matematik, naturvetenskap och läsförståelse.

⁹ The International Association for the Evaluation of Educational Achievement.

första hand i förhållande till det år när området i fråga senast var huvudområde.

Första gången Sverige deltog i en IEA-studie om matematik var 1964 och om de naturvetenskapliga ämnena 1970. Därefter har elevers kunskaper i matematik och/eller de naturvetenskapliga ämnena undersökts i IEA-studier 1980 och 1983. Sedan 1995 då TIMSS introducerades har ambitionen varit att mäta förändring i resultat över tid, dvs. att mäta trend. Detta har skett genom att delar av uppgifterna i tidigare undersökningar, så kallade ankaruppgifter, återanvänds. Undersökningen 1995 var den första där kunskaps-testen i matematik och de naturvetenskapliga ämnena samordnades.

Sveriges deltagande i TIMSS har varierat något över tid. Den första undersökningen 1995 genomfördes för ett stort antal åldersgrupper. Sverige valde att för grundskolan delta med elever som var 13 år, dvs. elever i årskurs 6 eller 7. Följande undersökning 1999 riktades endast till elever i årskurs 8. Sverige valde dock att inte delta. År 2003 deltog Sverige i undersökningen med elever från årskurs 8, men inte med elever från årskurs 4, vilket var ett erbjudande. I undersökningarna 2007 respektive 2011 har Sverige deltagit med elever från både årskurs 4 och årskurs 8.

TIMSS Advanced¹⁰ undersöker kunskaper i avancerad matematik och fysik hos elever som går sista året på gymnasiet. Studien har genomförts vid två tillfällen, 1995 och 2008. Sverige har deltagit vid båda tillfällena. Nästa gång studien genomförs blir 2015.

Sverige deltar dessutom i flera andra internationella studier, som än så länge inte kan ligga till grund för trendanalyser, helt enkelt för att de inte har återupprepats med given design tillräckligt många gånger.

År 2009 genomförde IEA en studie om medborgar- och samhällsfrågor benämnd ICCS¹¹. Syftet med ICCS är att undersöka hur unga människor i olika länder förbereds för att axla sin roll som medborgare. Studien bygger på tidigare internationella undersökningar om samhällsfrågor men är anpassad till de nya utmaningar som finns när det gäller att utbilda unga människor till medborgare. Tidigare studier är den så kallade sexämnesstudien 1971 och CIVED¹² som i grundskolan genomfördes 1999 bland 14- och 15-

¹⁰ TIMSS Advanced (Trends in International Mathematics and Science Study (advanced)). Studien organiseras av The International Association for the Evaluation of Educational Achievement (IEA).

¹¹ International Civic and Citizenship Education Study (ICCS).

¹² Civic Education Study (CIVED).

åringar. Även om ICCS 2009 till stor del bygger på CIVED, som inte kommer att upprepas, är det för tidigt att tala om trendmätning. För närvarande planeras ICCS 2016 för årskurs 8 i grundskolan.

Internationella språkstudien ESLC¹³ är ett projekt som är initierat av EU-kommissionen. Studien genomfördes för första gången 2011. Syftet med studien var att ge EU:s medlemsländer jämförande data om elevers kunskaper i främmande språk och att ge insikter om hur språkundervisningen kan leda till förbättrade kunskaper i främmande språk. I den studie som genomförts prövades de fem största språken som lärs ut inom EU, dvs. engelska, tyska, franska, spanska och italienska. Varje land som var med i studien deltog med sina två största främmande språk av dessa fem språk. I studien prövades de tre förmågorna läsa, lyssna och skriva. Diskussioner förs för närvarande om hur man ska gå vidare. Inriktningen är troligen att man nu vill begränsa sig till att testa kunskaperna i engelska. Dessutom övervägs om detta ska genomföras i form av en återupprepning av ESLC eller om man i stället ska ansluta till en av IEA planerad studie ETLS (English Teaching and Learning Study) som planeras för 2017.

TALIS¹⁴ slutligen är en studie som fokuserar på lärares yrkesvardag. Syftet med studien är att den ska fylla de kunskapsluckor som finns inom internationella jämförelser om utbildningssystem, om undervisning och om lärares och rektorers arbetsvillkor, dvs. aspekter som har betydelse för elevers lärande och för lärares och rektorers tillfredsställelse i sitt arbete. I studien frågar man bland annat efter vilka möjligheter lärare och rektorer har till kompetensutveckling, hur de fördelar arbetstiden mellan olika sysslor och vilken skolledarstil rektorn har. Man frågar också om trivsel, klassrums- och skolklimat och om lärarna anser att de har tillräcklig kunskap för den undervisning som de bedriver. TALIS omfattar i första hand lärare och rektorer i grundskolans årskurs 7–9.

TALIS är initierad av OECD och är ett samarbetsprojekt mellan OECD och deltagarländernas myndigheter med ansvar för utbildning. Projektet stöds av OECD:s fackliga rådgivarkommitté¹⁵ och EU-kommissionen. TALIS genomfördes första gången 2008 och upprepades 2013. Sverige deltog först 2013. Det finns ambitioner att studien ska upprepas framöver.

¹³ European Survey of Language Competences (ESLC).

¹⁴ The Teaching and Learning International Survey (TALIS).

¹⁵ TUAC.

De internationella studierna är tänkta för jämförelser mellan olika länder och baseras på små urval av elever i varje land. Det går därför inte att med försvarbar statistisk säkerhet bryta ner underlaget för att analysera skillnader t.ex. mellan olika typer av kommuner eller mellan skolor med olika socioekonomisk sammansättning etc. De internationella studierna kan därför endast spela en begränsad roll vid nationell reformutvärdering, om man inte väljer att nationellt utöka urvalet av elever för att möjliggöra en sådan analys.

Nationella utvärderingsstudier

Nationella utvärderingsstudier genomförs som nämnts ovan av olika myndigheter och andra aktörer. I många fall kan regeringens val av myndighet för ett uppdrag förklaras av myndigheternas skilda uppdrag.

Skolverkets utvärdering av skolan

Skolverket genomför utvärderingar både på regeringens uppdrag och på eget initiativ. Utrymmet för egeninitierade studier har dock minskat över tid och är nu starkt begränsat på grund av att verkets resurser i hög grad upptas av uppdrag givna av regeringen.

Skolverkets utvärderingsstudier varierar till sin karaktär. Fler-talet genomförs på uppdrag av regeringen och kan handla om att utvärdera satsningar av olika slag, t.ex. kompetensutvecklings- eller skolutvecklingssatsningar. Skolverkets utvärdering bidrar till en generaliserbar bild av läget i den svenska skolan, exempelvis genom registerstudier. Skolverket genomför emellertid även kvalitativa utvärderingsstudier med små urval av skolor, som kan bidra med fördjupad kunskap om exempelvis skolans arbetssätt och processer. Liksom för Skolinspektionens kvalitetsgranskningar (se nedan) är då generaliserbarheten mer begränsad.

Över tid finns också exempel på nationella utvärderingsstudier inriktade på att mäta kunskaperna i den svenska skolan, i vissa fall med ambitionen att det ska vara möjligt att följa trender, och/eller att bidra med systemkunskap. Senaste exemplet på en mer omfattande sådan satsning gäller den nationella utvärderingen av grundskolan från 2003 (NU 03). Denna typ av utvärderingar har hittills genomförts inom ramen för myndighetens uppdrag som det är

formulerat i Skolverkets instruktion¹⁶ och inte till följd av särskilda uppdrag från regeringen. Dessa utvärderingar bör karaktäriseras som utvärdering av skolsystemets funktionssätt (systemutvärdering), snarare än som reformutvärdering.

Skolverket upphandlar både hela utvärderingar och mer avgränsade insatser av externa utvärderare, främst universitet och högskolor. Det kan ske på initiativ av regeringen eller på eget initiativ. Orsaken kan vara att kompetensen inte finns inom verket. I många fall handlar det dock om att utvärderingarna eller uppföljningarna avser verkets egna insatser, t.ex. i samband med implementeringen av en reform. Implementeringen av mål och nationella prov i årskurs 3 samt implementeringen av Lgr11 är sådana exempel. Myndigheten bör inte utvärdera den egna verksamheten varför det i sådana sammanhang är lämpligt med extern utvärdering.

IFAU:s utvärdering av skolan

Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU) har sedan februari 2012 enligt sin instruktion¹⁷ i uppdrag att bland annat utvärdera effekter av olika reformer och åtgärder inom utbildningsväsendet. Med utvärdering avses enligt instruktionen både effektutvärdering och processutvärdering. Före 2012 var verksamheten mer renodlat fokuserad på arbetsmarknadspolitiska frågor och kopplingen till utbildningsväsendet handlade om utvärdering av effekterna på arbetsmarknaden av åtgärder inom utbildningsväsendet. IFAU:s analyser kännetecknas av en hög metodkompetens inom det kvantitativa området.

Sedan IFAU:s uppdrag utvidgats till att också omfatta effekter av reformer och åtgärder inom utbildningsväsendet har institutet fått ett särskilt uppdrag¹⁸ av regeringen att utvärdera de senaste årens stora reformer på skolområdet, mer specifikt för grundskolan och gymnasieskolan. Det framgår av direktiven att uppdraget avser de reformer som genomfördes under 1990-talet. Där framgår att det finns ”anledning att utvärdera vilka effekter de stora reformerna, som genomfördes under 1990-talet, fick för skolväsendets utveckling”. Samtidigt slår regeringen fast att genom detta uppdrag har Riksdagens tillkännagivande från 2011 till regeringen om att

¹⁶ Förordning (2011:555) med instruktion för Statens skolverk.

¹⁷ Förordning (2012:16) med instruktion för Institutet för arbetsmarknads- och utbildningspolitisk utvärdering.

¹⁸ U2012/7283/S, U2011/2269/S.

tillsätta en utredning med uppdrag att analysera de senaste årens skolreformer – kommunaliseringen, det fria skolvalet och fristående skolors etablering – slutligen omhändertagits. De stora reformer som genomförts sedan 2007, t.ex. Lgr11 och betyg från årskurs 6, omfattas dock inte av detta särskilda uppdrag. Myndigheten kan naturligtvis i mån av resurser själv ta initiativ till sådana utvärderingar inom ramen för sin instruktion.

Under hösten 2013 publicerade IFAU rapporten *Effekter av 1992 års skolvalsreform*. En fördjupning av denna analys pågår för närvarande inom projektet *Skolvalsreformerna och likvärdigheten*. Andra exempel på pågående projekt är *Reformer, resurser och resultat i den svenska skolan: Om decentralisering, skolval och fri-skolor* samt projektet *Hur fördelar kommuner resurser till skolan?* För närvarande pågår, enligt uppgift, ett fyrtiotal projekt som berör skolan.

Statskontorets utvärdering av skolan

Statskontoret ska t.ex. enligt sin instruktion¹⁹ på regeringens uppdrag bistå med utredningar, utvärderingar och uppföljningar av statlig och statligt finansierad verksamhet samt av övergripande frågor om den offentliga förvaltningens funktionssätt. I detta ingår att analysera verksamheter och myndigheter ur ett effektivitetsperspektiv, redovisa effekter av statliga åtgärder samt ge underlag för omprövning och effektivisering.

Myndighetens verksamhet styrs av de uppdrag som regeringen lägger på Statskontoret. Departementen och kommittéerna kan också efter överenskommelse anlita Statskontoret för utredningsinsatser. Bredden på de olika utredningar som Statskontoret genomför är stor.

Antalet uppdrag från regeringen till Statskontoret att utvärdera frågor som berör grundskolan eller motsvarande skolformer är begränsat. Sedan 2000 handlar det om en handfull uppdrag. I de flesta fall kan Statskontorets roll motiveras med att utvärderingarna avser insatser där Skolverket haft en roll vid genomförandet, och således inte kan anses vara oberoende som utvärderare. Det gäller t.ex. uppdraget att utvärdera Lärarlyftet och nyligen ett uppdrag att utvärdera jämställdhetsinsatser inom skolväsendet. För några år sedan genomförde Statskontoret ett uppdrag från regeringen att

¹⁹ Förordning (2007:827) med instruktion för Statskontoret.

utreda frågor om ersättning för elever i specialskolan och gymnasieskolan för elever med funktionsnedsättning. Även i början av 2000-talet gavs myndigheten uppdrag som mer utgick från dess kompetensprofil, t.ex. att utvärdera försöksverksamheten med lokala styrelser eller satsningen på IT i skolan. Längre tillbaka har Statskontoret t.ex. biträtt med analys av styrningsfrågor inom skolområdet.

Det offentliga utredningsväsendet

I ett historiskt perspektiv har mycket av utvärdering/motsvarande skett inom ramen för det offentliga utredningsväsendet. Det finns fortfarande exempel där en offentlig utredning genomför egna utvärderingar eller lägger ut uppdrag på andra aktörer, t.ex. forskare, myndigheter, konsulter, etc. Som olika källor²⁰ visar har det dock blivit successivt allt mer ovanligt att offentliga utredningar ges uppdrag som förutsätter att utvärdering genomförs. Uppdragen är ofta avgränsade och förhållandevis styrda. Ofta saknas också tid och resurser att genomföra mer omfattande utvärdering.

4.1.2 Inspektion och revision

Det internationella utvärderingsbegrepp (evaluation) som OECD använder är som tidigare nämnts brett och omfattar även inspektion och revision. Det är emellertid inte självklart att räkna in inspektion och revision i ett nationellt uppföljnings- och utvärderingssystem. Vi har tagit med detta i vår modell, men menar att långt ifrån all nationell inspektion och revision hör hemma inom uppföljnings- och utvärderingssystemet. Det kan dock ha ett värde att här redovisa hur inspektion och granskning förhåller sig till uppföljnings- och utvärderingssystemet. Inte minst gäller detta för den del av inspektionsverksamheten som gränsar till utvärdering.

²⁰ Se t.ex. Statskontoret (2000) Redovisning av uppdrag att genomföra en kartläggning av det samlade utvärderingsarbetet inom de olika utgiftsområdena och Amnå, Erik (2013) "Speak Truth to Power" *Statsvetarna och kommittéväsendet*. Statsvetenskaplig Tidskrift 112 (5).

Inspektion

Statens skolinspektions ansvar för tillsyn och kvalitetsgranskning är reglerat i skollagen²¹. Med *tillsyn* avses enligt denna bestämmelse en självständig granskning som har till syfte att kontrollera om den verksamhet som granskas uppfyller de krav som följer av lagar och andra föreskrifter. I tillsynen ingår att fatta de beslut om åtgärder som kan behövas för att den huvudman som bedriver verksamheten ska rätta fel som upptäckts vid granskningen.

Skolinspektionen ska även genom *kvalitetsgranskning* granska kvaliteten i sådan utbildning och annan verksamhet som står under dess tillsyn eller under tillsyn av en kommun. Granskningen ska avse den granskade utbildningens eller verksamhetens kvalitet i förhållande till mål och andra riktlinjer.

Skolinspektionen har således inte i uppdrag att följa upp och utvärdera utan att, genom tillsyn och kvalitetsgranskning, granska de verksamheter som står under myndighetens tillsyn. Som underlag för den regelbundna tillsynen genomför dock inspektionen varje år en enkätundersökning som riktar sig till elever, föräldrar och pedagogisk personal på de skolor som ska inspekteras.

Myndigheten ska också sammanfatta och publicera resultaten av sitt arbete med tillsyn och kvalitetsgranskning. Det sker dels i en årlig rapport till regeringen samt i övergripande rapporter från t.ex. genomförda kvalitetsgranskningar.

Analysen av Skolinspektionens tillsynsbeslut resulterar i statistik över vanliga bristområden, uppgifter som kan betraktas som en form av uppföljningsdata. Det bör dock uppmärksammas att det inte handlar om statistiskt säkerställda urval, utan om sammanställningar av de tillsyner som genomförts ett visst år. Eftersom det handlar om tillsyn av cirka tjugo procent av alla skolor per år bör dock underlagen kunna tillskrivas viss trovärdighet vad gäller de frågor tillsynen inriktats mot. Det bör dock noteras att tillsynen aldrig kan vara fullständig, utan att den endast tar upp områden som Skolinspektionen granskat.

När det gäller kvalitetsgranskningarna är gränsdragningen till bland annat Skolverkets utvärdering inte helt tydlig. Utvärdering kan användas som metod för att genomföra kvalitetsgranskning. Litteraturstudier ingår alltid som utgångspunkt för granskningarna. Syftet skiljer sig dock åt. Skolinspektionens kvalitetsgranskning syftar i första hand till att genom återföring påverka verksamheten

²¹ 26 kap. skollagen (2010:800).

vid de skolor som granskas. De sammanfattande rapporterna bidrar samtidigt till bilden av läget i den svenska skolan, utan att de för den skull har kravet på sig att vara generaliserbara.

Skolinspektionen har inte något explicit uppdrag i instruktionen att granska reformer på skolans område. Myndigheten har inte heller getts några sådana uppdrag av regeringen i samband med de reformer som genomförts sedan 2007, varken i regleringsbrev eller som särskilda uppdrag. I vissa fall inriktas dock tillsynen särskilt mot att kontrollera att nya eller ändrade bestämmelser följs. Det är dock inte samma sak som att exempelvis granska hur skolhuvudmännen tar ansvar för genomförandet av en större reform eller hur inslagen i reformen påverkat undervisningen.

Revision

Riksrevisionen har ett kraftfullt granskningsmandat som möjliggör en sammanhållen och heltäckande granskning av den statliga verksamheten. Riksrevisionen kan som enda myndighet granska hela den statliga beslutskedjan – riksdag, regering och myndigheter.

Effektivitetsrevisionen är den av Riksrevisionens verksamheter som metodmässigt ligger närmast utvärdering. Den innebär en granskning av hur effektiv den statliga verksamheten är. Riksrevisionen beslutar självständigt vad som ska granskas genom effektivitetsrevision och hur granskningen ska genomföras. Granskningarna utgår från riksdagens beslut och omfattar regeringens genomförande, styrning, uppföljning och rapportering till riksdagen samt myndigheternas genomförande och rapportering till regeringen. När det gäller reformer på skolområdet kan revisionen, med sitt fokus på statens agerande, i första hand granska hur effektiva regeringens och skolmyndigheternas insatser är för att genomföra förändringarna.

Inom utbildningsområdet har Riksrevisionen sedan sin tillkomst 2003 lagt tyngdpunkten i sina granskningar på styrningen av den högre utbildningen. Endast ett fåtal granskningar inom effektivitetsrevisionen har genomförts med inriktning på statens styrning av skolan. Statens styrning av betyg och betygssättning har dock granskats vid två tillfällen. Under de senaste åren har även Skolinspektionens verksamhet satts i fokus, med två granskningar. Under 2013 publicerades dessutom en granskning av statens främjande av kunskapsspridning inom skolan.

Riksrevisionens verksamhet har, genom sin uppgift och sitt oberoende från regeringen, inte någon roll i det som vi valt att kalla uppföljnings- och utvärderingssystemet. Det hindrar inte att revisionens rapporter kan bidra med kunskap som i sin tur påverkar beslutsfattandet genom att peka på behov av förändring.

4.1.3 Forskning

Utöver ovan nämnda myndigheter kan även forskare vid universitet och högskolor genomföra vad vi i kapitel 3 benämnt forskningsbaserade utvärderingar av olika reformer och satsningar inom skolområdet. Detta kan ske på uppdrag av regeringen. På skolans område sker det dock i de flesta fall på uppdrag av en myndighet, i första hand Skolverket, eller en offentlig utredning som i sin tur har fått uppdraget från regeringen. Till skillnad från den oberoende forskningen genomför forskarna i detta fall uppdrag där beställaren styr vad forskaren ska utvärdera. I vilken utsträckning beställningen också styr hur utvärderingen ska genomföras varierar, men det är inte ovanligt att metodval ingår i de avtal som upprättas mellan beställare och utförare.

Det finns exempel på uppdrag till forskare inom universitet eller högskolor att utvärdera åtgärder eller insatser som regeringen har initierat. Flertalet har Skolverket som beställare och ingår som del av större uppdrag från regeringen till Skolverket att svara för insatser och samtidigt utvärdera dem. Att utvärderingen inte genomförs av Skolverket kan, som nämnts ovan, bero på att Skolverket i så fall skulle komma att utvärdera sina egna insatser. Det kan dock även vara en resurs- eller kompetensfråga. Uppdragen kan exempelvis handla om att utvärdera Skolverkets insatser för att stödja implementeringen av en viss reform eller att utvärdera ett program för att stödja utvecklingen av undervisningen i ett visst ämne och där Skolverket har en roll. Att denna typ av utvärderingar ska genomföras av en extern aktör fastställs ofta redan i regeringens uppdrag till Skolverket, men beslutet kan också vara verkets eget.

På motsvarande sätt som uppdrag kan läggas på forskare inom universitet och högskolor kan uppdrag också läggas ut på andra externa utvärderare, t.ex. på konsultföretag som inriktat sig på utredning och utvärdering inom offentlig sektor. Utvärderarna kan även i dessa fall ha en forskarutbildning, men behöver inte ha det. Konsultföretag har dessutom ofta specifika modeller för hur de

lägger upp och genomför utvärdering och analys, modeller som i skiftande grad ansluter till forskningsbaserade modeller och metoder.

Den oberoende forskningen har inte som sin primära uppgift att vara policyrelevant på det sätt som det förutsätts att uppdragsforskning eller rapporteringen från ett nationellt utvärderingssystem ska vara, och är det oftast inte heller. Genom att systematiskt sammanställa resultat från forskning kan dock även den oberoende forskningen bidra till ett evidensinformerat beslutsfattande och, inte minst, till en evidensinformerad praktik.

Det finns också exempel på oberoende forskning som i hög grad är policyrelevant. Exempel på detta är den forskning som bedrivits, bland annat vid IFAU, om friskolereformens effekter på studieresultaten²², om betygssättningens effekter på studieresultaten²³ samt om långsiktiga effekter av ökad lärartäthet²⁴. Merparten av denna forskning utgörs av kvantitativt inriktade effektstudier. Dessa forskningsinsatser kan inte på egen hand fylla det behov av utvärdering av skolan som utredningen pekat på, men de visar att det finns både kapacitet och kompetens att genomföra forskningsbaserad utvärdering av reformer på skolans område i Sverige, även om de inte ingår i något nationellt utvärderingsprogram.

Vetenskapsrådet (VR) har som huvuduppgift att fördela medel till forskning. VR stödjer grundforskning inom utbildningsvetenskap, dvs. forskning om bildning, utbildning, undervisning och lärande. Bland de projekt som får stöd från VR finns forskning som avser effekter av reformer. Det handlar dock inte om någon fullständig reformutvärdering utan om utvärdering av avgränsade inslag i reformer. VR arbetar även med att göra forskningsresultat tillgängliga. Som ett led i detta arbete sammanställer rådet årligen under namnet *Resultatdialog* kunskapsöversikter över olika områden och redogör för den senaste forskningen inom utbildnings-

²² Se t.ex. Hartman, Laura (red.), Anders Anell, Eva Mörk, Jonas Vlachos, m.fl. (2011) *Konkurrensens konsekvenser – vad händer med svenskt välfärd?* SNS Förlag. Anders Böhlmark och Helena Holmlund (2012) *Lika möjligheter? Familjebakgrund och skolprestationer 1998-2010*. IFAU, rapport 2012:14. Anders Böhlmark och Mikael Lindahl (2012) *Har den växande friskolesektorn varit bra för elevernas utbildningsresultat på kort och lång sikt?* IFAU, rapport 2012:17.

²³ Se t.ex. Sjögren, Anna (2010) *Betygssatta barn – spelar det någon roll i längden?* IFAU, rapport 2010:8 och Christina Cliffordson (2004) *De målrelaterade gymnasiebetygens prognosförmåga*. Pedagogisk Forskning i Sverige, 9(2) samt Christina Cliffordson (2004). *Betygsinflation i de målrelaterade gymnasiebetygen*. Pedagogisk Forskning i Sverige, 9(1).

²⁴ Fredriksson, Peter, Björn Öckert och Hester Oosterbeek (2012) *Långsiktiga effekter av mindre klasser*. IFAU, rapport 2012:4.

vetenskap. Det gemensamma för de redovisade projekten är att de alla har finansierats av Vetenskapsrådet.²⁵

4.1.4 Hur kommer reformutvärderingen in?

Som nämnts inledningsvis kan reformer utvärderas inom ramen för det fortlöpande uppföljnings- och utvärderingssystemet. Det bygger då på att systemet är så heltäckande och har sådana kvaliteteter att svaren på de frågor man har anledning att ställa så att säga faller ut av sig själva, dvs. att de utvecklingstendenser man kan se kan härledas till en specifik reform. Det är dock sannolikt ett önsketänkande att systemet skulle ge en så kvalificerad kunskap att det inte finns behov av att genomföra särskilda utvärderingar av en reform eller ett reformskede.

I de flesta fall behövs också särskilda utvärderingar. Utvärderingen av en reform handlar ju inte bara om huruvida utvecklingen av resultaten på de internationella studierna går i rätt riktning eller inte, utan även om denna utveckling faktiskt kan härledas till själva reformen. Utvärderingen av en reform kan behöva fokusera på om de mekanismer som reformen bygger på verkligen fungerar som förväntat. Om de inte gör det kan utvärderingen behöva svara på varför de inte gör det och vad som kan göras för att få detta att inträffa, etc. Det vanligaste är kanske att något fungerar, men inte fullt ut och/eller med stora variationer. Det finns alltså ofta utrymme för förbättringar. För att i sådana fall kunna bidra till förbättringar av systemets prestationer räcker det inte med att känna till utfallet, det krävs även kunskap och förståelse för hur systemet fungerar och varför resultaten varierar.

Även med ett väl fungerande fortlöpande uppföljnings- och utvärderingssystem kommer därför särskilda reformutvärderingar att behöva genomföras. Sådana särskilda utvärderingsprojekt är att betrakta som utvärderingsstudier i modellen ovan. Sådana särskilda utvärderingsstudier kan hämta empiriskt underlag från hela eller delar av uppföljnings- och utvärderingssystemet, men också ta fram egna underlag. De kan genomföras av en myndighet eller av forskare på uppdrag av regeringen eller av en myndighet.

²⁵ Vetenskapsrådets webbplats.

4.2 Sammanfattande bedömning

Det utvidgade uppföljnings- och utvärderingssystem som redogjorts för ovan har, med dess kvaliteter och utvecklingsbehov, stor betydelse för möjligheterna att utvärdera reformer. Det består i dag av ett stort antal olika aktiviteter som i vissa fall kompletterar varandra. I andra fall är gränsdragningen mellan olika aktiviteter inte helt tydlig. Det gäller t.ex. mellan Skolverkets utvärdering och Skolinspektionens kvalitetsgranskning. Det finns också exempel där de syften som tillskrivs ett specifikt instrument knappast är möjliga att infria med ett och samma instrument. Det gäller för de nationella proven. Slutligen har systemet luckor, dvs. det finns frågor som det inte kan besvara på ett tillfredsställande sätt. Det gäller t.ex. hur skolans resultat och måluppfyllelse utvecklas över tid. De internationella studiernas trendanalyser behöver enligt vår uppfattning kompletteras med nationella trendanalyser. Vi återkommer till detta i kapitel 9.

Vid sidan om det nationella utvärderingssystemet finns ett stort antal andra aktörer som genomför studier och utvärderingar som avser skolan, t.ex. organisationer som t.ex. Sveriges Kommuner och Landsting (SKL), lärar- och skollädaorganisationer, Studieförbundet Näringsliv och Samhälle (SNS), etc. Även Expertgruppen för Studier i Offentlig ekonomi (ESO) genomför i vissa fall studier med koppling till skolan.

5 Reformutvärdering – historik och nuläge

I detta kapitel görs inledningsvis en historisk tillbakablick över reformutvärdering i Sverige på skolans område och framför allt över den diskussion som förts om reformutvärdering. Därefter redogör vi för dagens situation vad gäller reformutvärdering och för vissa övriga iakttagelser vi gjort avseende dagens nationella utvärdering av skolan.

5.1 En historisk tillbakablick

Av bilaga 3¹ framgår att utvärdering, initierad av regering och riksdag, av olika reformers utfall historiskt har varit mycket ovanlig i Sverige. Vi har samtidigt kunnat konstatera att frågan om utvärderingens roll och utvärdering av reformer på skolområdet återkommande har diskuterats i olika sammanhang.

5.1.1 1900-talets mitt – en period av stora utredningar²

1962 års grundskolereform var slutet på en lång process som startade med 1940 års skolutredning, följd av 1946 års skolkommision. Ett viktigt steg var 1950 års beslut om ett försök med en nioårig enhetsskola, ett försök som under 1950-talet följdes av Skolöverstyrelsen (SÖ). Slutligen utvärderades försöket politiskt av 1957 års skolberedning, som också lämnade förslag till inrättande av grundskolan. Grundskolan infördes rullande och SÖ fick återigen uppdraget att följa reformen.

¹ Björnsson, Mats (2013) *Kunskap och skolreform. En studie om underlag för och utvärdering av reformer på skolområdet*.

² Avsnittet baseras på bilaga 3.

Karaktäristiskt för 1900-talets mitt är att reformer förbereddes noggrant med offentliga utredningar som var omfattande i både tid och betänkandenas omfång. Till flera av betänkandena hörde ett stort antal bilagor av bl.a. forskare och andra experter. Försöksverksamhet genomfördes och utvärderades. Hela processen präglades av ett formativt förhållningssätt och den tidens mest avancerade pedagogiska forskning användes i reformarbetet.

Utformningen av reformerna måste således anses vara baserade på evidens. Det formativa inslaget var också påtagligt under den successiva utformningen av det som kom att bli grundskolan. Man kan däremot inte påstå att det togs några initiativ från regering eller riksdag till att utvärdera utfallet av grundskolans införande. Fokus låg på att så långt som möjligt i förväg säkra de beslut som fattades och att justera under väg, inte att utvärdera i efterhand. Långt senare har forskare visat att reformen bidrog till en högre utbildningsnivå samt att den hade en klassutjämnande effekt³.

5.1.2 En bild från 1970-talet

Frånvaron av reformutvärdering har sedan uppmärksammats återkommande. I början av 1970-talet konstaterade t.ex. professorn i pedagogik Urban Dahllöf att:

Systematiska undersökningar som syftar till att följa upp och utvärdera skolreformer lyser ännu med sin frånvaro. Det går i dagens läge inte att få någon samlad bild av läget vare sig på kunskapssidan eller i andra avseenden.⁴

Dahllöf argumenterade samtidigt för att utvärderingen av skolan inte fick utformas alltför snävt. Han underströk att utvärderingen påverkas av att skolan har både ett kunskapsmål och ett demokrati-mål, och att den därtill ska fungera socialt utjämnande. Utvärderingen av skolan måste därför, utöver elevernas resultat, även omfatta de organisatoriska *ramar* som omger skolverksamheten (studievägar, timplaner, klassindelningar, lokaler, läromedel m.m.) samt de *processer* som försiggår där (undervisnings- och arbetsformer, den tid som i praktiken tas i anspråk för olika kunskaps- och färdighetsmoment m.m.).

³ Meghir, Costa och Mårten Palme (2004) *Education reform, ability and family background*. IFS Working Papers, W04/10, Institute for Fiscal Studies.

⁴ Dahllöf, Urban (1971) *Svensk utbildningsplanering under 25 år*. Studentlitteratur. Lund.

Dahllöf konstaterade att ett och samma system beroende på olika faktorer kan ha olika *verkningsgrad* och att det därför är viktigt att försöka förstå samspelet mellan mål, ramar, processer och resultat. Han menade att sådana analyser bör användas vid uppföljning och utvärdering av reformer i skolan och genomföras relativt snart efter en reform. Detta för att kunna tjäna som vägledning när det gäller beslut om vilka (kompletterande) åtgärder som i första hand måste sättas in för att få det nya systemet att fungera enligt intentionerna. Dahllöf menade att inriktningen bör vara att studera undervisningsprocessen och dess relation till de yttre organisatoriska ramarna och slutresultaten.

Enligt Dahllöf behövdes en breddning av utvärderingsprogrammet från resultattestningar – redan de sällsynta – till studier i syfte att förstå hela sambandet mellan utgångsläge, undervisningsprocess och resultat i relation till uppställda mål.

Utvärderingen borde enligt Dahllöf vara allsidig och täcka olika ämnen såväl som olika aspekter av skolans mål. Den borde läggas upp för att ge underlag för beslut om vilken åtgärd, eller vilken kombination av åtgärder som är de mest adekvata i varje enskilt fall.

Vi (behöver) inte bara veta att skolan fungerar mer eller mindre väl i olika sammanhang utan också varför. Vi behöver få ett betydligt bättre grepp om de olika mekanismerna i undervisningsprocessen, vilken roll som olika metoder och yttre ramar spelar för vad som faktiskt händer i skolarbetet både från den effektiva inlärnings och den sociala fostrans synpunkt. Här gäller om någonsin satsen att ingenting är så praktiskt som en god teori.

En bättre förståelse av hur skolsystemet faktiskt fungerar i sitt inre arbete gentemot de givna målen blir också av stor betydelse då det är dags att göra en storrevision som innefattar en omprövning av målen.⁵

5.1.3 Den fortsatta debatten om den nationella utvärderingen

Under 1980-talet pågick således ett arbete med att forma en nationell utvärdering av skolan inom SÖ. 1989 genomfördes en första studie, som omfattade elever i årskurs 2 och 5 i grundskolan. Resultaten publicerades 1991 en kort tid innan myndigheten lades ned. Skolverket har sedan genomfört resultatmätningar med varierande inriktning och omfattning 1992, 1995, 1998 och 2003. Se vidare kapitel 4 samt bilaga 3.

⁵ Dahllöf, Urban (1971) a.a.

Den form som SÖ gav utvärderingen kritiserades i en skrift av fem professorer i pedagogik.⁶ Man menade att resultatmätningarna på elevnivå tilldelats alltför stor tyngd i förhållande till kunskaper om undervisningens ramar och processer och att det i SÖ:s program knappast klargjordes hur data skulle analyseras och tolkas, samt hur och till vad de skulle användas. Man menade att de redovisade syftena var otydliga och att intresset för hur resultaten skulle återkopplas till lokal nivå var ytterst svagt.

Till exempel hänvisade Dahllöf till att en rad pedagogikforskare ”ställt sig öppet kritiska inte mot kunskapsprövningar som sådana” utan till att ”testprogrammet blivit överdimensionerat och felbalanserat”. Man menade att en bättre balans mellan resultatinformation samt information om vad undervisningen innehållit och med vilka metoder den bedrivits i olika skolmiljöer och för skilda elevgrupper skulle göra det betydligt lättare att förklara bakgrunden till eventuella brister. ”Ju bättre förklaringar man kan ge till starka och svaga sidor i undervisningen, desto lättare blir det att genomföra ett åtgärdsprogram i syfte att förbättra den på sikt.”

Dahllöf menade vidare att SÖ aldrig hade tagit något samlat grepp över utvärderingsfrågan. Han hävdade att en nationell utvärdering inte endast kan begränsa sig till att ange vad som är bra och dåligt. Den måste även kunna svara på frågor om *varför* och om *vilka åtgärder som bör vidtas* och att detta ställer krav på olika former av utvärdering.

Sigbrit Franke menade att man måste fråga sig vilket huvudsyftet med det nationella utvärderingsprogrammet var och borde vara. Var det kontroll/styrning eller förändring/förbättring?

Ulf P. Lundgren menade att förändringen mot ett alltmer decentraliserat skolsystem ställde specifika krav på hur utvärderingen skulle utformas centralt, regionalt och lokalt. De metoder och modeller som används påverkar styrningen av skolan, såväl vad gäller balansen mellan politisk och professionell styrning som vad gäller balansen mellan central och lokal styrning, och möjliggör utvecklandet av kunskap om skolan.

Kritiken mot skolan byggde dock enligt Lundgren sällan på kunskap om vad den producerade och till vilket pris. Utvärderingen av skolan måste därför bl.a. ske utifrån ett brett samhällsteoretiskt perspektiv. Det innebär att tolkningarna av utvärderingen enligt

⁶ Franke-Wikberg, Sigbrit (red.), Urban Dahllöf, Daniel Kallos, Ulf P. Lundgren och Erik Wallin (1989) *Skolan och utvärderingen. Fem professorer tar ordet*. HLS Förlag.

Lundgren borde ske med hänsyn till övergripande kunskap om relationer mellan utbildning, produktion och samhällsliv.

Lundgren menade vidare att förändringar av skolan krävde fortbildning såväl av lärare som av den centrala skoladministrationen. I en vidgad kompetens ligger en vidgad förmåga att ta tillvara kunskap och erfarenheter, att föra en öppen diskussion, att lyssna till kritik och att bygga upp ett eget minne.

Det är inte uppgiften att genomföra ett nationellt utvärderingsprogram jag sett som problemet, utan förmågan att se detta program i ett sammanhang och att kunna bygga det stegvis, så att det skapar en användbar och tillgänglig kunskap. Detta är en utmaning som ett nationellt utvärderingsprogram ställer i en administrativ struktur med rötter i central kontroll och där makt i bland förväxlas med kunskap.⁷

5.1.4 Övergången till mål- och resultatstyrning i början av 1990-talet

Den övergång till mål- och resultatstyrning av den svenska skolan som skedde i början av 1990-talet ingick i en mer omfattande förändring av den statliga styrningen generellt. För skolans del skedde övergången som en följd av riksdagens beslut om två propositioner – propositionen om skolans utveckling och styrning samt propositionen om ansvaret för skolan.

Propositionen om skolans utveckling och styrning

I propositionen om skolans utveckling och styrning⁸ konstaterades att traditionen att utvärdera verksamheten varit svag i svensk skola "om man med utvärdering menar att verksamhetens resultat ställs i relation till målen". Vidare konstaterades att målstyrning förutsätter att läroplanens mål närmare preciseras i det lokala planeringsarbetet. Ett styrsystem som i större utsträckning skulle bygga på målstyrning förutsatte, enligt regeringen, kraftigt ökade insatser för utvärdering på alla nivåer av skolväsendet.

Enligt regeringen borde uppgiften att välja undervisningsmetod och att prioritera vid urval av stoff, både av principiella och praktiska skäl, alltjämt förbehållas lärarna och eleverna, så som läropla-

⁷ Franke-Wikberg, Sigbrit (red.), Urban Dahllöf, Daniel Kallos, Ulf P. Lundgren och Erik Wallin (1989) *Skolan och utvärderingen. Fem professorer tar ordet*. HLS Förlag.

⁸ Prop. 1988/89:4 *Skolans utveckling och styrning*.

nerna föreskrev. Den centrala skolmyndigheten förväntades dock noga följa och låta utvärdera undervisningens resultat och vid behov kräva förändringar och förbättringar.

De ökade kraven på utvärdering gällde såväl undervisningens resultat i relation till uppsatta mål och riktlinjer som hela den genomförandeprocess, varigenom målen och intentionerna skulle förverkligas. Regeringen menade att de skolpolitiskt ansvariga måste få kunskap om hur deras intentioner förverkligas, och i vilken grad de mål som de angett verkligen uppnås, för att med detta som utgångspunkt effektivt kunna styra skolans utveckling i önskad riktning. På alla nivåer torde särskilt den mer *kvalitativt* betonade utvärderingen behöva förbättras.

Regeringen omnämnde i propositionen det arbete som vid denna tidpunkt pågick inom SÖ med att utveckla innehåll och former för en nationell utvärdering av skola och vuxenutbildning. Detta var enligt regeringen en naturlig arbetsuppgift för SÖ, men den allmänna inriktningen av en sådan utvärdering borde dock avgöras av riksdag och regering. Att SÖ tilldelades ett ansvar för utvärderingen hindrade inte, enligt regeringen, att också vetenskapliga institutioner eller andra organ kunde komma ifråga för särskilda utvärderingsuppdrag initierade av riksdagen eller regeringen.

Regeringen menade att det var angeläget att utvärderingskompetensen hos SÖ förstärktes. Som ett led i denna förstärkning och för att få en bredare vetenskaplig förankring föreslogs att SÖ skulle knyta en *referensgrupp* eller *rådgivargrupp* av *vetenskaplig expertkaraktär* till det särskilda nationella utvärderingsprogram, som SÖ efter principbeslut av riksdagen och regeringen vid denna tidpunkt arbetade med att utveckla. Gruppens uppgift skulle vara att biträda vid utformningen av programmet, vid värderingen av de data som insamlats för utvärderingen samt vid bedömningen av vilka delar av programmet som borde ägnas särskild uppmärksamhet. SÖ och länskolnämnderna föreslogs regelbundet publicera sammanställningar av de resultat som de centrala och regionala uppföljningarna och utvärderingarna gav.

Med utvecklingen av informationsteknologin skapades vid denna tidpunkt nya möjligheter att få tillgång till data om skolan och att länka olika data och utvärderingar till varandra. Regeringen menade därför att en utredning borde genomföras om behovet och utformningen av ett bättre informationssystem för skolan och vuxenutbildningen. Utöver en genomgång av tillgången till information genom den befintliga forskningen och utvärderingen, och av

den existerande informationen i övrigt, skulle analyseras vilket behov av information som övergången till en ökad målstyrning medförde.

I det föreslagna utredningsarbetet föreslogs också ingå att utforma bättre modeller för uppföljning och utvärdering på olika nivåer. Därvid skulle möjligheterna att effektivisera den mer kvalitativt betonade utvärderingen särskilt uppmärksammas. Även möjligheterna att bedöma effekterna av den förda utbildningspolitiken och den förändrade styrningen och ansvarsfördelningen skulle uppmärksammas, liksom effekterna av att använda utvärdering som kontroll- och förändringsinstrument.

Någon utredning av det slag som föreslogs blev aldrig tillsatt. Regeringen återkom i stället till frågan i den s.k. ansvarspropositionen och i det interna beredningsarbete som följde på denna.

Ansvarspropositionen

I den s.k. ansvarspropositionen⁹ konstaterade regeringen återigen att uppföljning och utvärdering skulle bedrivas såväl i den enskilda skolan som på kommunal och nationell nivå. De av riksdagen och regeringen fastställda målen måste kunna följas upp och utvärderas nationellt. Avsikten var att riksdagen och regeringen regelbundet skulle få en bild av tillståndet i skolan genom rapporter och utvärderingar. Den då befintliga uppföljningen och utvärderingen av verksamheten i skolan betraktades som osystematisk och fragmentarisk. Det var svårt att få en helhetsbild, samtidigt som det saknades uppgifter om hur olika kommuner drev sin verksamhet och vilka resultat som uppnåddes.

I anslutning till arbetet med att utforma riktlinjerna för det kommande läroplansarbetet pågick också ett arbete med att ta fram riktlinjer för den nationella utvärderingen av undervisningsresultat. Regeringen menade att denna utvärdering måste kompletteras av en nationell uppföljning som ger möjligheter till att mer systematiskt mellan kommunerna jämföra prestations- och kostnadsutvecklingen i de olika skolformerna, för de samlade statliga och kommunala resurserna.

Det nationella utvärderingsprogram som SÖ påbörjat skulle enligt regeringen komma att utvecklas successivt. En viktig uppgift förväntades bli att planera och samordna arbetet så att en samlad bild av skolans verksamhet skulle kunna presenteras, som underlag

⁹ Prop. 1990/91:18 *Ansaret för skolan*.

för ställningstaganden i fråga om skolans framtida styrning och utveckling. Det borde, enligt propositionen, ankomma på regeringen att ge riktlinjer för det fortsatta arbetet.

Vad hände?

Som nämnts ovan tillsattes aldrig den utredning om ett informationssystem för skolan och vuxenutbildningen som regeringen aviserade i styrpropositionen. Inte heller fullföljdes SÖ:s nationella utvärderingsprogram. Ansvar för utvärderingen kom i stället att föras till Skolverket när detta inrättades 1991 samtidigt som SÖ lades ned. Skolverket fick ansvaret för uppföljning och utvärdering. Inför inrättandet av Skolverket genomfördes ett internt arbete med att utforma uppföljningssystemet inom Utbildningsdepartementet.

Skolverket har sedan sin tillkomst på eget initiativ genomfört flera nationella utvärderingar av grundskolan. Den senaste stora utvärderingen genomfördes 2003 (NU 03)¹⁰. Några riktlinjer för det fortsatta arbetet med att utforma ett nationellt utvärderingsprogram för skolan, på det sätt ansvarspropositionen aviserar, har inte getts av regeringen.

5.1.5 Sammantaget – ingen reformutvärdering att tala om

Samtantaget kan konstateras att det från 1900-talets mitt fram till 1990-talets början inte finns några exempel på summativa reformutvärderingar på skolans område som initierats av regering eller riksdag. Införandet av grundskolan föregicks dock av omfattande utredningar och försöksverksamheter som utvärderades. Det fanns således i det fallet ett starkt formativt inslag, där den bästa kunskapen togs tillvara.

Utvärdering diskuterades dock livligt under perioden, av såväl regering som forskare. Endast i begränsad omfattning har denna debatt gällt explicit reformutvärdering. I högre grad har de ambitioner som uttryckts gällt uppföljnings- och utvärderingssystemet generellt, dvs. handlat om hur en ökad kunskap om skolan och skolans prestationer ska kunna etableras.

¹⁰ Om de nationella utvärderingarna av grundskolan se bl.a. bilaga 3.

5.2 Var står vi i dag?

Ser man tillbaka på perioden fram till början av 1990-talet och jämför med perioden därefter kan man konstatera att mycket känns igen och att inte mycket har hänt, åtminstone inte när det gäller reformutvärdering. När det gäller uppföljnings- och utvärderings-systemet har en utveckling skett, men det kvarstår också ambitioner att infria.

5.2.1 Reformutvärdering fortfarande ovanlig

Om vi med reformutvärdering menar utvärdering i anslutning till, eller av, större reformer går det snabbt att sammanfatta de initiativ som under de senaste 20-25 åren tagits från central politisk nivå för att utvärdera reformer på skolans område: Tills helt nyligen har de varit helt obefintliga. Med reform avser vi här förändring av bärande delar av skolsystemet som är reglerade i lag eller förordning. Den nationella utvärderingen av grundskolan (NU 03), som Skolverket genomförde på eget initiativ, fick funktionen av en reformutvärdering av bl.a. införandet av Lpo 94, trots att den i grunden när en systemutvärdering.

Till skillnad från vad som gällde vid mitten av 1900-talet har större reformer som t.ex. kommunaliseringen av skolan respektive införandet av mål- och resultatstyrning av skolan genomförts utan föregående utredning eller försök. Införande av nya läroplaner har föregåtts av utredningar, däremot inte av några försök. Försöksverksamheter har dock genomförts inom vissa andra mer avgränsade områden (se vidare kapitel 8).

Några politiskt initierade beslut att utvärdera 1990-talets reformer togs inte inför eller i samband med att de genomfördes. Tjugo år senare har regeringen, som nämnts i föregående kapitel, sedan februari 2012 gett Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU) som en instruktionsenlig uppgift att bl.a. utvärdera effekter av olika reformer och åtgärder inom utbildningsväsendet¹¹. Efter riksdagens tillkännagivande 2011 om att tillsätta en utredning med uppdrag att analysera de senaste årens skolreformer¹² – kommunaliseringen, det fria skolvalet och fristå-

¹¹ Förordning (2012:16) med instruktion för Institutet för arbetsmarknads- och utbildningspolitisk utvärdering.

¹² Observera att med de senaste årens skolreformer avses här de stora reformer som genomfördes på 1990-talet, inte de reformer som genomförts sedan 2006.

ende skolors etablering – har regeringen dessutom tagit initiativ till två utredningar avseende dessa 1990-talsreformer:

- I augusti 2012 tillkallade regeringen en särskild utredare med uppdrag¹³ att utvärdera eventuella kausala effekter av kommunaliseringen på elevernas studieresultat, läraryrkets status, lärarnas och rektorernas arbetsuppgifter, samt likvärdigheten i skolväsendet. Utredningen redovisade sina resultat i början av februari 2014¹⁴.
- I december 2012 fick IFAU ett särskilt uppdrag¹⁵ av regeringen att utvärdera de stora reformer på skolområdet som genomfördes under 1990-talet. Uppdraget är begränsat till grundskolan och gymnasieskolan. IFAU ska särskilt beakta aspekter som betygsresultat och fullföljande av utbildning, liksom kön, utländsk bakgrund och tid i Sverige. Uppdraget ska redovisas senast den 15 december 2014.

Än så länge finns inga beslut om utvärdering av de reformer som genomförts under perioden från 2006. Det handlar om så viktiga reformer som t.ex. införandet av Lgr11 med nya kursplaner kompletterade med kunskapskrav, införandet av skriftliga omdömen som får vara betygsliknande från årskurs 1, betyg från årskurs 6 samt nationella prov i fler ämnen. Däremot har regeringen initierat mer avgränsade utvärderingar (se nedan).

5.2.2 Skolverkets utvärderingsuppdrag

Som redovisats ovan har Skolverket som ett centralt uppdrag att följa upp och utvärdera skolan. Detta sker genom den fortlöpande uppföljningen och utvärderingen samt genom särskilda regeringsuppdrag.

Den fortlöpande uppföljningen och utvärderingen som har redovisats i föregående kapitel utgör en viktig grund för möjligheterna att utvärdera reformer. Det bygger på det system som initierades när Skolverket bildades i början på 1990-talet och består av bl.a. statistik och de internationella kunskapsmätningarna. Systemet har naturligtvis utvecklats och anpassats över tid.

¹³ Dir. 2012:84.

¹⁴ SOU 2014:5 *Staten får inte abdikera – om kommunaliseringen av den svenska skolan*.

¹⁵ U2012/7283/S, U2011/2269/S.

Utredningen har valt att gå igenom regeringens särskilda utvärderingsuppdrag till Skolverket och Skolverkets återrapporteringar av dessa för perioden 2002–2013¹⁶. Vi har fokuserat bl.a. på vilken typ av utvärderingsuppdrag Skolverket har fått av regeringen. Vi har vidare bedömt om uppdragen avsett utvärdering av reformer och hur vanligt det är med sådana särskilda utvärderingsuppdrag.

Vi har granskat uppdrag givna såväl i Skolverkets regleringsbrev som i särskild ordning samt i viss mån även krav på utvärdering som formuleras i förordningar. Vi bedömer att granskningen omfattar merparten av de uppdrag som under perioden innehållit krav på utvärdering, eller något som motsvarar utvärdering. Sammantaget rör det sig om drygt 25 uppdrag.

Uppdragen handlar inte om reformutvärdering

Det enkla svaret på frågan om vilken typ av utvärderingsuppdrag Skolverket får är att regeringens uppdrag inte avser utvärdering av reformer. De avser i stället ofta mer avgränsade frågeställningar eller är kopplade till olika satsningar, t.ex. statsbidrag.

Exempel på uppdrag om utvärdering av *särskilda satsningar* är statsbidrag för åtgärder som syftar till att stärka arbetet med basfärdigheterna läsa, skriva och räkna (2009), utbildningsinsatser för lärare, utvecklingsinsatser inom områdena naturvetenskap och teknik (2012) samt statsbidrag till personalförstärkningar i skola och fritidshem (2003). I dessa uppdrag anger regeringen ofta att man vill att verket ska visa effekter av åtgärderna. Av redovisningen av uppdraget om personalförstärkningar framgår att Skolverket låtit forskare vid Uppsala universitet göra en kvantitativ effektstudie. Denna visar inga signifikanta effekter av lärartäthet på den genomsnittliga elevens studieresultat, men ökade resurser har förbättrat studieresultaten för elever med lågutbildade föräldrar. I de båda övriga utvärderingarna diskuteras effekter mer kvalitativt med utgångspunkt i lärares bedömningar i enkätsvar.

Regeringens uppdrag om utvärdering kopplas även till pågående *försöksverksamheter*, t.ex. försöksverksamhet med riksrekryterande gymnasial spetsutbildning (2013), försöksverksamhet med ökad undervisningstid i svenska eller svenska som andraspråk för nyanlända elever i grundskolan (2013). Om de utvärderingar som rege-

¹⁶ Ett urval av dessa ingår i bilaga 3 *Exempel på uppföljningar och utvärderingar genomförda av Skolverket på regeringens uppdrag åren 2002 – 2013*.

ringen efterfrågat i anknytning till dessa båda försöksverksamheter läggs till grund för beslut om det fortsatta reformarbetet kan dessa processer sägas utgöra exempel på ett evidensinformerat arbetssätt.

De uppdrag om utvärdering som berör reformer begränsar sig till utvärdering av Skolverkets *insatser* och handlar inte om reformerna *i sig*. Skolverket har t.ex. haft i uppdrag¹⁷ att initiera extern utvärdering av verkets insatser i samband med implementeringen av Lgr11. Ett annat exempel på detta är uppdraget¹⁸ att stödja införandet av en förändrad individuell utvecklingsplan med skriftliga omdömen. Här efterfrågade regeringen en redogörelse för genomförda åtgärder och kostnader samt en utvärdering av insatserna, däremot inte av själva reformen.

Oklara uppdrag påverkar genomförande och rapportering

Utöver att visa att de särskilda uppföljnings- och utvärderingsuppdragen inte avser reformutvärdering visar utredningens genomgång även att det finns ett behov att utveckla det sätt på vilket uppdrag till Skolverket formuleras.

Oklar begreppsanvändning i uppdragen ...

Utredningen har kunnat konstatera att det inte är alldeles oproblemiskt att avgöra när regeringen förväntar sig en utvärdering och när den inte gör det. En enkel utgångspunkt är att alltid förutsätta att regeringen beställer en utvärdering när den explicit begagnar sig av detta begrepp – annars inte. Det är dock inte självklart att så är fallet. Förutom uttrycken ”utvärdering” och ”utvärdera” använder regeringen en rad andra uttryckssätt¹⁹.

¹⁷ U2010/259/S Uppdrag till Statens skolverk om implementering av skolreformer.

¹⁸ 2008/4180/S Uppdrag om implementeringsinsatser rörande en förändrad individuell utvecklingsplan med skriftliga omdömen.

¹⁹ Bedömningen ska innehålla en fördjupad analys...; Genomföra en fördjupad studie...; Pekka på utvecklingsområden...; Följa upp...; Av redovisningen av verksamheten ska framgå resultatet, samt hur (xxx) bidragit till uppföljningen och utvecklingen av kvaliteten i...; Löpande följa upp och analysera...; Undersöka hur...; Av redovisningen ska bland annat framgå hur arbetet med satsningen fortskrider...; Kartlägga omfattningen av problemet samt föreslå...; Erfarenheterna av genomförandet av uppdraget ska redovisas...; Återrapportera utvecklingen av...; Informera RK (Utbildningsdepartementet) om uppdragets genomförande...; Redovisa och analysera utvecklingen av verksamheten...; Göra en registerstudie...; Kartlägga och analysera...; Genom fältstudier undersöka hur...; Utvärderingen ska ske i formen av en uppföljande analys...

Om det endast står att verket ska *redovisa* något, *följa upp* något eller genomföra en *kartläggning* bör man kunna utgå från att det inte förväntas en utvärdering genomförd med utvärderingsmetodik. När Skolverket ska *följa upp och analysera* eller *kartlägga och analysera* är detta mer oklart, liksom när verket ska genomföra en *fördjupad studie* eller *undersöka hur*, eller slutligen när verket ska genomföra en *utvärdering i formen av en uppföljande analys*. Varken uppdrag givna i regleringsbrev eller uppdrag givna genom separata regeringsbeslut är således alltid tydliga med vad regeringen förväntar sig. Handlar det om en utvärdering som kräver en utvecklad metod och som förväntas utmynna i en värdering av utvärderingsobjektet och kanske även innehålla en rekommendation? Eller handlar det om en enklare kartläggning utan förväntan på analys och bedömning?

... bidrar till svag styrning av genomförande och rapportering

Enligt vår bedömning har det sätt som uppdragen formulerats över tid fått återverkningar på det sätt som Skolverket redovisar uppdragen – mer eller mindre likartat oavsett hur beställningen formuleras. Redovisningarna bygger oftast på kvantitativa data hämtade ur uppföljningssystemet och kvalitativa data från enkäter, intervjuer och fallstudier. Redovisningarna kan även innehålla olika former av slutsatser, bedömningar, rekommendationer eller åtgärdsförslag. Det gäller oavsett om beställningen formulerats som en uppföljning, en utvärdering, en kartläggning eller en fördjupad studie.

I flera fall benämner myndigheten sin återslagrapportering för utvärdering trots att regeringen i sitt uppdrag inte uttryckligen efterfrågat en sådan. Detta kan tolkas som att Skolverket inom ramen för sitt instruktionsenliga uppdrag valt en ambitiösare ansats än vad det aktuella uppdraget kräver, som ett led i behovet att bygga upp kunskap om den aktuella frågeställningen för egna och omgivningens behov (för normering, tillsyn och utvecklingsinsatser). Det kan emellertid även vara uttryck för en osäkerhet om vad regeringen egentligen efterfrågat. I vissa fall kan det även ha förekommit en uppdragsdialog mellan Skolverket och departementet som klargjort detta.

Det är även möjligt att uppdrag där regeringen efterfrågat en utvärdering visat sig ha en sådan begränsad karaktär att det framstått som oförsvärligt att genomföra en resurskrävande och metodisk avancerad utvärdering, och att myndigheten därför inte

redovisat en sådan. Vidare kan det finnas exempel där frågan varit av den arten att den inte låtit sig utvärderas med en acceptabel metod. Även i dessa fall kan det ha förekommit en uppdragsdialog.

Att Skolverkets metodval är så likartat oberoende av hur uppdraget har formulerats kan således, enligt vår uppfattning, i viss mån härledas till att uppdragen är otydligt formulerade. Bland annat därför skapas också en osäkerhet om huruvida beställaren faktiskt använder ett begrepp i den betydelse det ges inom utvärderingsteorin eller om begreppet används i en allmänare mening. Bland annat begreppet effekter kan ibland användas på ett sätt så att det är oklart om det används i en mer vardaglig mening eller om det är en effektstudie som avses.

Vi har dessutom kunnat konstatera att de utvärderingar vi gått igenom inte innehåller någon egentlig diskussion om utvärderingsdesign och möjliga utvärderingsmetoder. Det man oftast avser när man talar om metod är i själva verket olika tekniker för datainsamling (enkät, data från uppföljningssystemet, dokumentstudie, observation, fallstudie, etc.), dvs. myndigheten redovisar/förklarar det sätt som den aktuella studien faktiskt genomförts på. Enligt Skolverket utesluter inte detta att t.ex. översikter har gjorts och att programteoretiska resonemang eller genomgångar av verksamhetslogiker har genomförts och att dessa redovisats i projektens arbetsplaner. Att de inte redovisas i verkets externa rapporter beror, enligt verket, på att dessa ska vara målgruppsanpassade.

Utredningen har inte, bland de utvärderingar Skolverket genomfört eller låtit genomföra på regeringens uppdrag, funnit något exempel där en effektstudie med randomiserad kontrollgrupp genomförts, eller någon studie med en programteoretisk ansats. I ett fall, där ett konsultföretag genomfört en utvärdering på Skolverkets uppdrag, förs ett inledande resonemang med programteoretiska utgångspunkter. I några fall har myndigheten låtit forskare ta fram underlag för sin rapportering. Dessa har huvudsakligen haft en nationalekonomisk utgångspunkt för metodval och val av frågeställningar. I flera av rapporterna återkommer verket till svårigheterna med att mäta effekter.

5.3 En internationell utblick

Utredningen har studerat ett antal OECD-länders arbete med reformer och framför allt utvärdering. Genomgången visar att länderna har delvis olika strategier för utvärderingar av reformer, delvis som en konsekvens av att länderna har olika reformstrategier.

I flera länder har t.ex. fokus flyttat från utvärdering av reformer i form av särskilda forskningsbaserade studier till att i högre grad söka utveckla utvärderingssystem som löpande tillhandahåller data som grund för bedömningar av hur skolan utvecklas. Denna förskjutning har i vissa fall varit kopplad till att man har lämnat strategin att genomföra stora strukturreformer till förmån för mer partiella reformer och satsningar på att förstärka den lokala förmågan till skolutveckling (capacity building). Detta innebär dock inte att länderna helt har slutat att vid behov genomföra utvärderingar av skolpolitiska reformer. Det finns också exempel på länder som har genomfört större skolreformer och där särskilda forskningsbaserade studier har getts en mer framskjuten roll vid utvärderingen av dessa.

Vi har även funnit att de system för kunskapsförsörjning som byggs upp i andra länder oftast sätter utvecklingsperspektivet främst. System för utvärdering och andra former av kunskapsinhämtande syftar främst till att stödja den lokala utvecklingen. Det innebär att även de beslut som fattas på nationell politisk nivå har denna inriktning, dvs. utvecklingen förefaller gå mot en utvecklingsstrategi som lägger störst vikt vid att bygga upp den lokala kapaciteten att själv ta ansvar för sin utveckling och sina prestationer. Centralt planerade reformer utformas med detta för ögonen.

Denna trend är inte rätlinjig och tar sig inte samma väg i alla länder, men vi anser ändå att en sådan trend kan urskiljas. En sådan trend kan under senare år även avläsas i England, dvs. ett av de länder som tidigare haft det tydligaste kontroll- och styrningsperspektivet i synen på utvärderingens roll och syfte inom utbildningssystemet. Under senare år har till exempel skolornas förmåga till egen utveckling betonats allt kraftigare i de engelska regeringarnas policydokument.²⁰

²⁰ Se till exempel dokumenten *A new Relationship with Schools*. Department for education and skills tillsammans med Office for standards in education (2004) och *The Importance of Teaching. The schools White Paper 2010*. Department for Education (2010).

Det engelska utbildningsministeriets utvärderingar är av olika omfattning och inriktning – särskilda satsningar och utbildningsprogram hör till det som utvärderas oftast – och genomförs av olika forskargrupper eller av enskilda konsultföretag²¹. Man har även avtal med olika forskningsinstitutioner som bidrar med vetenskapliga underlag inom sina respektive kunskapsområden²².

För att utvärdera nationella reformer och strategier beställer utbildningsministeriet i Nya Zeeland oberoende utvärderingar från externa akademiker och forskare och/eller samlar information från ett stort antal källor (t.ex. granskningen och utvärderingen av den nationella standarden). Sådana utvärderingar används både för summativa syften (bedömningar av effektiviteten och värdet av policies och program) och för formativa syften (för anpassning, modifiering och förbättring av pågående reformer och program). Ett exempel på detta slags utvärderingar är den utvärdering av den nya zeeländska läroplanen och de nationella standarderna (ungefär motsvarande de svenska kunskapskraven) som pågår för närvarande inom projektet The National Standards: School Sample Monitoring and Evaluation Project. Projektet genomförs av ett kontrakterat utvärderingsteam och samlar in data från ett urval av skolor under perioden 2009-2013²³.

Danmark har under senare år arbetat intensivt med att stärka utvärderingskulturen i skolan. Ett viktigt inslag i detta arbete har varit beslutet om att införa ett nationellt provsystem i folkeskolan²⁴. I samband med detta beslutades även om att det nya systemet skulle utvärderas. Det nationella provsystemet började tillämpas i full skala 2010 och en första utvärdering av systemet genomfördes av det fristående konsultföretaget Ramböll under perioden november 2012 till september 2013²⁵.

I Norge har under de senaste decennierna stora forskningsbaserade utvärderingsstudier genomförts i samband med att de stora skolreformerna Reform 94, Reform 97 och Kunskapslyftet (2006) genomförts. Utvärderingen av Kunskapslyftet var omfattande. Enligt uppgift satsades totalt 60 miljoner norska kronor. Under 2009 mottog Utdanningsdirektoratet, utöver ett stort antal delrap-

²¹ Se www.gov.uk/contracts-finder

²² Centre for Analysis of Youth Transitions, Centre for Understanding Behaviour Change och The Childhood Wellbeing Research Centre. Se DfE:s hemsida.

²³ Uppgifterna om Nya Zeeland är hämtade från det underlag som lämnats av Nya Zeeland i New Zealand Country Background Report (2010) Denna rapport utgör ett av underlagen för OECD:s Review of Evaluation and Assessment in education. New Zealand (2012).

²⁴ En mer utförlig redovisning av de initiativ som tagits i Danmark finns i bilaga 5.

²⁵ Ramböll (2013) *Evaluerings af de nationelle test i folkeskolen*.

porter, femton slutrapporter från olika forsknings- och utvärderingsprojekt. Projektet resulterade i en nära koppling mellan forskning och policyutformning. Forskningen har haft betydelse dels för att stärka styrningen av skolsystemet (och den har påverkat dess struktur och kultur), dels för att stärka lärarnas och elevernas arbete och lärande, dvs. skolans praxis. Det senare har dock skett i mindre utsträckning än det förra.²⁶

5.4 Sammanfattande bedömning

Karaktäristiskt för reformarbetet i Sverige vid 1900-talets mitt var på skolans område att större reformer förbereddes genom omfattande utredningar. Reformerna som genomförts under de senaste decennierna har ofta föregåtts av utredningar med ett mer begränsat utrymme för att på ett mer djuplodande sätt utreda den aktuella frågeställningen, t.ex. genom att initiera utvärdering och forskning eller genom att i förväg pröva tänkta lösningar genom försök eller stegvis införande. Detta innebär inte att det helt saknats utredningsinsatser inför reformer eller att det inte genomförts några försöksverksamheter alls. Det finns emellertid också exempel på större reformer från början på 1990-talet som genomfördes utan föregående utredning (se bilaga 3).

Under de senaste decennierna har det offentliga utredningsväsendet generellt sett getts en delvis annan roll än det hade tidigare. Utredningstiderna har förkortats och ger inte i samma utsträckning som förr utrymme för att initiera egna utvärderingar eller separata studier av t.ex. forskare. Det är i dag vanligt med direktiv som tydligt anger inriktningen av arbetet. Dessa förhållanden får ses som en konsekvens av en allt snabbare förändringstakt och krav på snabbare beredningsprocesser²⁷.

Samtidigt har ambitionen om att de politiska besluten ska vara evidensinformerade förstärkts. Utredningen menar att skolpolitiska reformer har mycket att vinna på att utformas i processer där evidens och utvärdering ges en mer central roll än under de senaste decennierna. För att en sådan ambition ska kunna infrias behöver förutsättningarna för att ta tillvara systematiskt uppbyggd kunskap

²⁶ En mer utförlig redovisning av de initiativ som tagits i Norge finns i bilaga 5.

²⁷ Se t.ex. Amnå, Erik (2013) *Speak truth to Power. Statsvetarna och kommittéväsendet*. Statsvetenskaplig Tidskrift 112 (5).

i beslutsprocesserna förbättras. Det kan innebära att dessa behöver ges mer tid än i dag.

Det är dock knappast realistiskt, eller ens önskvärt, att återvända till situationen vid 1900-talets mitt där beredningsprocesserna kunde sträcka sig över mer än ett decennium. Större krav måste därför också ställas på den kunskap som kan fås ur såväl det löpande uppföljnings- och utvärderingssystemet som från annan utvärdering, t.ex. av redan genomförda reformer. Frågan om hur politikens behov av evidens kan matchas med produktionen av forsknings- och utvärderingsresultat har blivit kritisk. Kunskapen måste i hög utsträckning redan finnas på plats när de politiska besluten ska fattas.

Det är mot denna bakgrund slående hur litet som hänt över tid när det gäller utvärdering av reformer på skolans område i Sverige. I hög grad gäller fortfarande Urban Dahllöfs konstaterande från början av 1970-talet, att det saknas systematiska undersökningar som syftar till att följa upp och utvärdera skolreformer. Förutom de båda utredningar som nyligen tillsatts, och som avser att utvärdera förändringar som skedde för 20 år sedan, har regering och riksdag inte tagit initiativ till någon systematisk utvärdering av de reformer som genomförts under senare år. Stora reformer i Sverige har således genomförts, och genomförs fortfarande, utan att bli föremål för en från början planerad utvärdering, utöver den som sker av informationsinsatserna i implementeringens första fas.

Behovet av ett väl fungerande uppföljnings- och utvärderingssystem har diskuterats återkommande under åren från 1990-talets mitt och framåt, av både forskare och av regeringen. Frågor som aktualiserats handlar om vad systemet ska kunna ge svar på, balansen mellan kvantitativt och kvalitativt inriktad utvärdering, olika nivåers behov av utvärderingsinformation, etc. Det nationella uppföljningssystem som Skolverket svarar för innebär mot denna bakgrund en viktig utveckling jämfört med tidigare, även om det fortfarande finns utrymme för att förstärka systemet, något vi återkommer till i flera av de följande kapitlen.

Det uppdrag som IFAU fick 2012 att utvärdera effekter av olika reformer och åtgärder inom utbildningsväsendet innebär att det tillkommit en aktör med en specialiserad kompetens för just denna typ av utvärdering. Det är värdefullt, men innebär enligt vår bedömning inte att behovet av reformutvärdering är täckt. För utvärdering av reformer behövs enligt vår uppfattning en bredd

både vad gäller utvärderingsdesign och vad gäller de forskningsdiscipliner som företräds. Det handlar om både kvantitativt och kvalitativt inriktad utvärdering.

När det gäller utvärderingen av skolan och av reformer på skolans område finns således, enligt vår uppfattning, fortfarande ett stort utrymme för utveckling som främst har att göra med vilken typ av utvärdering som efterfrågas, men också på hur beställningarna av utvärdering utformas.

Genomgången av regeringens utvärderingsuppdrag till Skolverket visar både på en frånvaro av mer omfattande utvärderingsuppdrag från regeringens sida och på att verket i begränsad omfattning har förutsättningar att ta egna initiativ till större utvärderingar, t.ex. utvärderingar av reformer²⁸. Genomgången visar även att det finns ett behov av ökad tydlighet från regeringens sida när uppdrag ges. Formuleringarna bör vara mer konsistenta och metodiskt medvetna, så att utföraren inte behöver tveka om huruvida det är en utvärdering som efterfrågas, med de krav det ställer på metod, eller om det är en enklare kartläggning. Att formulera beställningar kräver således en grundläggande metodisk medvetenhet. Det gäller bland annat att formulera utvärderingsbara frågeställningar, utan att för den skull detaljstyra genomförandet.

I en internationell belysning kan vi konstatera att Sverige i flera avseenden befinner sig väl framme. Vi har till exempel länge haft ett omfattande uppföljningssystem som försett såväl den centrala nivån som skolhuvudmännen med grundläggande data om kunskapsresultat och andra prestationer, kostnader samt socioekonomiska bakgrundsdata.

I andra avseenden riskerar dock Sverige att hamna på efterkälken, till exempel vad gäller i vilken utsträckning som olika reformer och andra insatser utvärderas och i vilken utsträckning evidens läggs till grund för beslut om reformer och andra förändringar. Medan utvecklingen i många länder har gått mot att sätta utvecklingsperspektivet främst har tendensen i Sverige under senare år varit att i allt högre utsträckning stärka kontroll- och styrningsperspektivet.

²⁸ Ett undantag från detta gäller utvärderingen av förskolereformen från 1998 som genomfördes på Skolverkets initiativ. Utvärderingen redovisades i två rapporter: Skolverkets Rapport 239 (2004) *Förskola i brytningstid. Nationell utvärdering av förskolan* och Skolverkets rapport 318 (2008) *Tio år efter förskolereformen. Nationell utvärdering av förskolan*.

6 Reformutvärdering – strategi för framtiden

I detta kapitel diskuteras mer ingående en av de centrala frågorna i utredningens uppdrag, dvs. den om hur reformer på skolområdet bör genomföras för att underlätta utvärdering. En utgångspunkt är iakttagelserna från föregående kapitel, dels att det i dag ofta ges ett begränsat utrymme för ett evidensinformerat beslutsfattande inför att beslut om reformer fattas, dels att det, med några få undantag, inte under överskådlig tid getts några uppdrag att utvärdera de reformer som beslutats. En annan utgångspunkt är att förutsättningarna för, och genomförandet av, reformutvärdering behöver förbättras.

6.1 Reformutvärdering och implementering

Utredningen konstaterade i kapitel 3 att utvärdering är ett mångfacetterat begrepp. Om man dessutom kopplar utvärdering till reformer, på det sätt som görs i vårt uppdrag, finns anledning att betrakta utvärderingen i relation till implementeringen av reformer. I vårt föregående betänkande¹ presenterades en modell för implementering av reformer, som i sin tur bygger på en rad forskares olika modeller.

Modellen består sammanfattningsvis av följande faser: (1) *Politikformuleringsfasen* som bl.a. omfattar problemformulering och beslut om reform-, implementerings- och utvärderingsstrategi, (2) *Processimplementeringsfasen* som bl.a. omfattar initiala implementeringsinsatser och formativ utvärdering, och (3) *Institutionaliseringsfasen*, dvs. den fas då reformen implementerats fullständigt och kommer att utgöra en integrerad del i verksamheterna.

¹ SOU 2013:30 *Det tar tid – om effekter av skolpolitiska reformer*.

Med detta synsätt inleds således implementeringen av reformer redan när problem identifieras och politiken utformas. Implementeringen fortgår sedan under hela processimplementeringsfasen och fullföljs när reformerna institutionaliseras. Sammantaget handlar detta om en långvarig process². När det gäller större skolreformer som till exempel införandet av en ny läroplan kan det handla om från fem till mer än femton år innan en reform kan sägas vara institutionaliserad³.

Detta sätt att beskriva en implementering i olika faser kan uppfattas som självklar när det handlar om politiska reformer beslutade på nationell nivå. Det är dock lätt att förledas att tro att reformer som beslutas i ett "top-down-perspektiv" enkelt kan "hållas genom systemet" och kommer att genomföras i verksamheten på det sätt som det var avsett. Detta är naturligtvis inte fallet. Reformers "avdrift" är ett välkänt fenomen, dvs. varje realiserad reform kommer att avvika från intentionerna i större eller mindre utsträckning. Denna avdrift är i själva verket ofta betydande. Att genomföra politiskt beslutade intentioner i en mångfacetterad och komplex verksamhet, som står under korstryck från flera olika håll, är således inte någon enkel och fullt rationell process. Den stora utmaningen för politiska beslutsfattare ligger därför i att försöka påverka det man ytterst vill påverka.

Reformutvärdering – före, under och efter

Reformutvärdering kan ses i såväl ett före-, under- som efterperspektiv. Inför beslut om genomförande av en reform krävs en ordentlig problemanalys och en kvalificerad analys av hur de olika åtgärderna kan förväntas påverka de problem som ska lösas, dvs. en policyanalys eller programteoretisk analys. Det är genom en sådan analys som beslutet om en reform kan sägas bli evidensinformerat.

Före ett reformbeslut handlar det således om att utgå från utvärderingar som gjorts av tidigare genomförda reformer, från forskning och särskilda utvärderingsstudier samt från den kunskap som hämtas ur det befintliga utvärderingssystemet. Dessa olika underlag kan ge utgångspunkter för att formulera målsättningar för reformen samt vägledning om vilka åtgärder som kan förväntas fungera eller inte, vilka oavsedda effekter som kan förväntas uppkomma

² Se vidare SOU 2013:30.

³ Åsén, Gunnar Bilaga 2 till SOU 2013:30 *Det tar tid – om effekter av skolpolitiska reformer*.

etc. Tillsammans med andra underlag kan de bidra till ett evidensinformerat beslutsfattande. Till detta kan även analyser som bygger på olika metoder för att i förväg göra prognoser om kommande utfall bidra, liksom att reformer och reformstrategier testas i förväg genom försöksverksamheter. Utredningen återkommer till frågan om försöksverksamheter i kapitel 8.

Under implementeringsprocessen kan olika former av formativ utvärdering genomföras. Resultatet av denna utvärdering kan användas för att vid behov genomföra justeringar i reformen eller för att påverka processer som inte fungerar som det varit tänkt. Beroende på reformstrategi kan det under denna fas också vara aktuellt att, på lokal verksamhetsnivå, arbeta med utvecklingsstödande utvärdering. Under implementeringsprocessen kan också datainsamling genomföras för olika former av summativ utvärdering.

Efter det att en reform implementerats, dvs. när reformen slutligen har institutionaliserats, kan en summativ utvärdering ge svar på om reformen gett de resultat som den förväntades att ge. Den kan även bidra med kunskaper om hur verksamheten påverkats, varför det uppstått variationer och avvikelser etc. Detta är kunskaper som behövs inför utformningen av en eventuell ny reform. Summativ utvärdering är med andra ord på nationell nivå ett betydelsefullt sätt att få underlag eller "ingångsvärden" inför framtida reformer.

Vid stora reformer som tar lång tid att institutionalisera, t.ex. läroplansreformer, kan det många gånger vara svårt att genom en summativ utvärdering isolera effekterna av reformen från andra förändringar som skett över tid. Ju längre effektkedjan, eller målmedelkedjan är, desto svårare är det att fastställa en slutlig effekt eller att mer precist svara på varför vissa effekter har uppstått.

Utvärderingen bör planeras i tid

Oavsett vilken form av utvärdering det handlar om bör utvärderingsinsatserna planeras i tid för att kunna genomföras med god kvalitet för att ge den information som efterfrågas. Enligt betänkandet *Att nå ut och nå ända fram* (SOU 2009:94) är en helt avgörande förutsättning för effektmätning att det finns en planering redan när en reform sjösätts. Utredningen delar i princip denna uppfattning även om det numera finns avancerade metoder som i vissa fall kan göra det möjligt att initiera en utvärdering i efterhand.

Att avstå från att från början planera för utvärderingen är dock inte att rekommendera. Dessutom har det sätt på vilket en reform genomförs stor betydelse för hur den kan utvärderas, vilket vi återkommer till längre fram.

Det finns dessutom andra typer av utvärderingar som definitivt inte kan initieras i efterhand. Det gäller exempelvis formativ utvärdering som ska genomföras under implementeringen. Vill man basera utvärdering på t.ex. observationer av hur arbetet i klassrummet påverkas av en reform kan detta inte heller göras i efterhand. Det kan behövas observationer såväl inför, under som efter reformen. För att veta hur en viss praxis påverkats av en reform behöver man veta hur denna praxis såg ut innan reformen trädde i kraft.

Det behövs ett systemperspektiv

I kapitel 3 bedömde utredningen att skolans utvärderingssystem behöver kunna hantera komplexitet och därför bör utvärderas med en mångfald av metoder. Utredningen konstaterar att utvärdering av reformer även förutsätter ett systemperspektiv, dvs. en helhets-syn på systemets olika delar och hur de samverkar.

I en amerikansk rapport konstateras att politiska interventioner äger rum i stora, komplexa, dynamiska sociala system. Ett systemperspektiv hjälper beslutsfattare och forskare att tänka brett kring de många effekter som en policy kan skapa och om de sätt som en planerad social intervention interagerar med andra redan existerande interventioner och institutionella praktiker. Det är sällan, menar forskarna, som studiet av enskilda komponenter i ett system kan leda till en fullständig förståelse av systemet.”⁴

Innovationsrådet är inne på likartade tankar och menar i sitt slutbetänkande⁵ att det krävs en *systemansats* för att hantera komplexa samhällsproblem. En sådan ansats kännetecknas enligt rådet av följande:

- Kontinuerligt lärande av vad som fungerar (snarare än att ange mål för vad som ska uppnås)

⁴ Prewitt, Kenneth, Thomas A. Schwandt and Miron L. Straf. (2012) *Using Science as Evidence in Public Policy*. Committee on the Use of Social Science Knowledge in Public Policy. National Research Council of the National Academies. The National Academies Press. Washington, D.C.

⁵ SOU 2013:40 *Att tänka nytt för att göra nytta*.

- Förbättra hela systemets resultat utifrån vad slutanvändarna av systemet anser om utfallet (inte bara vad uppdragsgivare och de som arbetar i systemet anser om resultatet)
- En politisk beslutsprocess som fokuserar på att förbättra själva systemet och dess verksamhetsprocess (snarare än kontroll av enskilda myndigheter och funktioner)
- Ett lyssnande förhållningssätt och involvering av brukare och andra intressenter (i stället för att instruera)
- Att hitta innovativa lösningar för att lösa grundproblemen (snarare än att agera på symptomen)

Utredningen menar mot denna bakgrund att ett utvärderingsprogram som utformas ur ett brett systemperspektiv kan bidra till att begränsa det slags stuprörseende som gärna blir följden när endast en fråga eller delkomponent belyses i taget. En systemansats ökar möjligheterna att belysa relationerna mellan de olika komponenterna, identifiera motsägelsefullheter och konflikter, förklara hur olika aktörers ageranden samspelar och vilka konsekvenser det kan få, samt bidrar till förståelse för varför beslut som kan förefalla rimliga i stället kan motverka syftet.

Detta synsätt ligger nära den breda utvärdering av skolan som Dahllöf m.fl. förespråkade redan på 1980-talet och som utredningen redovisat i kapitel 5. Han pekade bl.a. på att en alltför ensidigt inriktad resultatutvärdering innebär att alltför många viktiga frågor kommer att falla utanför det som studeras och värderas.

Dahllöf menade att det inte räcker med att studera resultaten (elevernas kunskaper) eller det som äger rum i undervisningen ("den svarta lådan") var för sig. Utvärderingen behöver också inriktas på att granska relationerna mellan det som händer i klassrummet och de olika slags ramar som finns för verksamheten. Genom att utvärdera reformer i ett sådant systemperspektiv kan en sådan kunskap om systemet och dess funktionssätt uppnås som gör att de ansvariga på olika nivåer med ökad säkerhet kan avgöra vilka förändringar som bör göras i vilken situation och vilka avsedda och oavsedda effekter som kan förväntas av den ena eller andra åtgärden.

6.2 Reformstrategier och utvärdering

Vid reformering bör således utvärderingen av en reform designas redan när strategin för själva reformen utformas. I kapitel 2 konstateras att evidensinformerad politik kan bedrivas ur såväl ett styrningsperspektiv som ett utvecklingsperspektiv. I kapitel 3 har redogjorts för olika utvärderingsansatser. Här söker vi knyta ihop detta och närma oss frågan om sambandet mellan reformstrategi och utvärderingsansats.

Om reformstrategier

En reformstrategi handlar om vilka slags åtgärder man tror kommer att påverka en verksamhet i önskad riktning och om de mekanismer man tror kommer att vara verksamma för att åstadkomma denna påverkan. Sådana strategier växer fram utifrån olika föreställningar om vad som gynnar förändring och utveckling och om vilka metoder för statlig styrning som är effektiva.

En typ av reformstrategi utgår från uppfattningen att staten genom reformer direkt kan påverka den lokala verksamheten. En förändring ses då som något som sker rationellt, närmast med automatik, som en följd av att ett beslut har fattats. Detta är – eller har åtminstone varit – en vanlig utgångspunkt vid statlig styrning. Denna strategi sammanfaller i hög grad med klassisk top-down-inriktad implementering, där utvärderingens roll i första hand blir att bekräfta eller motsäga att den förväntade effekten uppstått samt eventuellt också förklara varför detta skett, så att justeringar av reformen kan genomföras. Utvärderingen förväntas däremot inte ha någon direkt utvecklingsstödande funktion. Ett stort antal forskningsstudier och utvärderingar runt om i världen har funnit att reformstrategier som vilar på sådana rationella föreställningar har haft begränsade framgångar, särskilt inom komplexa verksamheter som t.ex. utbildning.

En annan typ av reformstrategi utgår från uppfattningen att det är svårt att påverka den operativa verksamheten enbart genom t.ex. förändrade styrdokument, dvs. med lagar, förordningar m.m. Förändring ses här som något som präglas av irrationalitet, intresse-motsättningar, framsteg och misslyckanden i en mer eller mindre oförutsägbar blandning. Staten bör i denna situation inrikta sina initiativ och insatser på åtgärder som har potential att stödja lokala

utvecklingsprocesser. Styrningen blir då mer indirekt och inriktad på att försöka påverka de grundförutsättningar och villkor som gäller där verksamheten faktiskt bedrivs. Med denna reformstrategi kommer utvärderingens formativa funktion att betonas och, åtminstone på verksamhetsnivå, utvärderingen också ges en mer utvecklingsstödande funktion.

Olika reformstrategier – i Sverige och i omvärlden

OECD:s olika granskningar och översikter⁶ visar att regeringarna i flera länder – däribland Nederländerna och Storbritannien – under senare decennier mer och mer övergivit strategin att försöka påverka skolans inre arbete med storskaliga, nationella reformer. I stället arbetar man allt mer med att försöka förstärka den lokala nivåns förmåga att själv ta ansvar för sin egen utveckling och prestationsförbättring.

Detta byte av reformstrategi kan ses som ett led i övergången från *government* till *governance*, en övergång som innebär att politikutvecklingen inte längre enbart utformas av regeringen och riksdagen – även om det givetvis fortfarande är där som de nationella politiska besluten fattas. Enligt teorin om *governance* utformas den faktiska politiken i stället i olika nätverk, dvs. av fler aktörer (företagare, konsument, brukare, verksamhetsföreträdare, etc.). Inslaget av *governance* anses vara särskilt framträdande i de lokala beslutsprocesserna.

Även Innovationsrådet⁷ konstaterar att mycket tyder på att en hierarkisk och auktoritativ styrning håller på att ersättas av horisontella, nätverksliknande och mer egalitära styrmekanismer.

Det tycks råda enighet om att samhällsstyrningen har blivit allt mer komplicerad, vilket borde medföra ändrade förutsättningar för, och krav på politikens roll i styrningen.⁸

Utredningen kan konstatera att den statliga styrningen av den svenska skolan har förändrats drastiskt under senare decennier. Från att ha haft ett centraliserat skolväsende med ett starkt inslag av *regelstyrning* övergick staten i början av 1990-talet till en decentraliserad styrning som skulle bygga på en *mål- och resultatstyrning*.

⁶ Se t.ex. OECD (2013) a.a.

⁷ SOU 2013:40 *Att tänka nytt för att göra nytta – om perspektivskiften i offentlig verksamhet*.

⁸ SOU 2013:40 a.a.

Man kan i efterhand konstatera att denna mål- och resultatstyrning inledningsvis handlade mest om målstyrning för att under senare år få ett ökat fokus på resultatstyrning. Det senare har skapat tendenser till en recentralisering som skett parallellt med en viss återreglering av skolan. Samtidigt har konkurrens mellan skolor införts vilket skapat inslag av en *marknadsstyrning*. De yrkesverksammas friutrymme genom *professionsstyrning* har enligt många bedömare minskat. Förändrade styrningsstrategier får givetvis följder för vilken slags utvärderingsinsatser som behövs.

Utvecklingen har lett till att styrningen av skolan i dag karaktäriseras av att många av dessa styrningslogiker lever ett parallellt och i vissa fall kanske ett konfliktfyllt liv. Det ligger dock inte inom vårt uppdrag att utreda konsekvenserna av denna situation annat än utifrån vilka konsekvenser detta har för utvärderingen av skolan.

Olika slags reformstrategier samverkar med olika slags reformutvärdering och utvärderingssystem

Vilken slags utvärdering som framstår som meningsfull att genomföra beror givetvis på det konkreta innehållet i den reform som genomförs. Den valda reformstrategin har konsekvenser för vad man behöver få veta om de förändringar som genomförs, vilka data och processer man behöver ha kännedom om, vilka utvärderingsmetoder som bör väljas osv. Reformstrategin har även betydelse för valet av implementeringsstrategi. Det är till exempel stor skillnad på implementeringen av en strukturreform och implementeringen av en insats för att stödja lokal skolutveckling.

Valet av reformstrategi har dessutom betydelse för vilka slags frågor som kommer att ställas till uppföljnings- och utvärderingssystemet och som detta kommer att förväntas kunna besvara. Utvärderingssystemet bör därför å ena sidan vara anpassat till aktuella styrningsmodeller och rådande övergripande nationella reformstrategier, å andra sidan bör det inte vara konserverande utan kunna följa, och till och med bidra till, utvecklingen av dessa modeller och strategier. Utvärderingssystemets utformning är också beroende av antalet nivåer och aktörer som finns inom skolsystemet.

Sammantaget kännetecknas det svenska skolsystemet i dag av många olika typer av aktörer på olika nivåer – offentliga och privata

– och av att det förekommer sinsemellan olika och delvis motstridiga styrningslogiker (se ovan). Detta förhållande har betydelse både för vilken slags utvärderingssystem som behöver utvecklas och för vilket slags frågor som utvärderingarna bör kunna besvara. Utvärderingssystemet måste t.ex. ha kapacitet att samla in data, granska, analysera och värdera/bedöma den variation som ett sådant decentraliserat och pluralistiskt system skapar vad gäller ekonomiska och rättsliga förutsättningar, olikheter i ansvarsförhållanden och samverkansformer etc.

De huvudinriktningar inom utvärdering som redovisats i kapitel 3, dvs. effektutvärdering och teoribaserad utvärdering, men även de ansatser som utvecklats särskilt för utvärdering av komplexa system, lämpar sig olika väl för att ge svar på olika slags frågor. För att förstå vad som händer i en komplex situation behöver, som tidigare understrukits, utvärderingssystemet ge utrymme för en mångfald av perspektiv och utvärderingsmetoder. Inte minst bör så kallade *mixed methods*, där både kvantitativa och kvalitativa metoder utnyttjas, komma till användning.

Utvärdering kan främja utveckling

Utvärdering kan även användas för att mer explicit främja implementering. Genomförande av en reform gynnas av att det finns en genomtänkt och långsiktig implementeringsstrategi som sträcker sig längre än till den centrala skolmyndighetens initiala implementeringsinsatser och som kännetecknas av dialog med den lokala nivån. Ett sätt att skapa dialog är att besluta om en tidig utvärderingsinsats som utformas så att den pågår parallellt med den initiala implementeringen och ger de operativt verksamma en möjlighet att bidra med sina erfarenheter dels med avseende på vad som bör beaktas vid utvärderingen, dels i samband med den empiri som samlas in.

Ett annat sätt att stärka utvärdering och användning av utvärderingsresultat är att utforma en implementeringsstrategi som är inriktad på kapacitetsbyggande insatser såväl på huvudmannanivå som på verksamhetsnivå (skolan) vad gäller att följa, dokumentera och utvärdera sitt eget arbete med implementeringen av reformen. En sådan strategi kan bidra till att öka huvudmännens och de operativt ansvarigas intresse för den nationella utvärderingen. Detta ökar i sin tur förutsättningarna för att resultaten av utvärderingen

ska tas om hand lokalt. David Gough⁹, ledande forskare vid EPPI-Centre vid London University, menar t.ex. att det behövs fler efterfrågestyrda utvärderingar, dvs. utvärderingar som svarar mot verksamhetens behov, för att öka sannolikheten för att utvärderingarna ska komma till användning.

Ytterligare sätt att stödja utveckling genom utvärdering är att initiera utvärderingar som är utformade som följeutvärdering eller aktionsforskning.

6.3 Regeringens och Regeringskansliets arbete

För att stärka ett evidensinformerat arbetssätt i politikutvecklingen och i styrningen av hur reformer utvärderas behöver arbetsformerna inom Regeringskansliet utvecklas. Regeringen har det yttersta ansvaret för att efterfråga såväl evidens inför beslut om reformer som utvärdering av genomförda reformer. Regeringen har även ansvaret för att arbetsformer etableras inom Regeringskansliet som bidrar till ett sådant evidensinformerat beslutsfattande och till att uppdrag om kvalificerad utvärdering kommer till stånd.

6.3.1 OECD om det centrala regeringsarbetet i Sverige

OECD har inom ramen för projektet *Value for Money in Government* genomfört en studie av det centrala regeringsarbetet i Sverige¹⁰. Rapporten innehåller bland annat en analys av den roll som utvärdering spelar i det svenska regeringsarbetet.

Det finns enligt OECD en generell tendens att man på tjänstemannanivå försvarar existerande program, varför det är viktigt att ministrar själva använder sig av möjligheten att genomföra kritiska utvärderingar. Det primära ansvaret för politiska reformer och program bör ligga på ansvarig minister snarare än på de verkställande myndigheterna, även om dessa bör vara delaktiga under hela processen. OECD noterar vidare att Sverige har många utvärderingsmyndigheter.

OECD konstaterar att såväl Kanada som Storbritannien infört utvärdering som en central aktivitet för hela det departementala

⁹ Gough, David (2007) Artikeln *The Evidence for Policy and Practice Information and Coordination (EPPI) Centre, United Kingdom* i OECD & CERI (2007) *Knowledge Management. Evidence in Education: Linking Research and Policy*.

¹⁰ OECD (2013) *Value for Money in Government. Sweden*.

regeringsarbetet i syfte att stärka kvaliteten i besluten och att skapa incitament till att genomföra policyutvärdering som är relevant och rigorös. OECD rekommenderar att Sverige låter sig inspireras av dessa båda länder. Bland annat bör en utvärderingsmetodologisk standard etableras.

För svenskt vidkommande har OECD gjort iakttagelsen att det svenska regeringskansliets kapacitet vad gäller policyutveckling är svagare än i flera andra länder. Som en följd av den svenska ansvarsfördelningen mellan departement och myndigheter har det departementala arbetet i Sverige enligt OECD kommit att fokusera på styrningens juridiska och ekonomiska aspekter.¹¹

6.3.2 Kopplingen mellan reformstrategier och reformutvärdering i Regeringskansliets arbete

De svenska fackdepartementen ingår i ett gemensamt regeringskansli som förbereder regeringens beslut om såväl reformer som om utvärdering av desamma. Många skäl talar för att arbetsformer och rutiner i stor utsträckning ska vara gemensamma inom Regeringskansliet. Regeringens och Regeringskansliets arbetsformer måste också ses ur ett övergripande demokratiskt perspektiv. Beredningen av en enskild reform eller åtgärd kan många gånger kompliceras av olika faktorer, som t.ex. pressade tidsplaner, kopplingar till andra beslut, finansiering, meningsskiljaktigheter av mer politisk karaktär, etc. Att ytterligare komplicera beredningen genom att tydligare föra in frågan om reformutvärdering kan uppfattas som försvårande.

Enligt vår uppfattning är det ändå nödvändigt, inte minst mot bakgrund av den frågeställning som direktiven lyfter om hur reformer bör genomföras för att bättre kunna utvärderas. Regeringskansliet (Utbildningsdepartement) kan genom sina arbetsformer och beslut bidra till att underlätta utvärdering av reformer. Det handlar bland annat om hur utvärderingsfrågorna prioriteras i det departementala arbetet och om hur inslagen av evidensinformerad politikutveckling – som ju hör nära samman med utvärdering – kan förstärkas i de interna arbetsprocesserna.

¹¹ OECD (2013) *Value for Money in Government: Sweden*.

Internationella erfarenheter visar att inslaget av evidens i politikutvecklingen kan stärkas

Utredningen konstaterar att intresset för evidensinformerad policy ökat inom många OECD-länder.¹² Storbritannien hör, som nämnts ovan, till de länder som arbetat mest intensivt med detta. Man har t.ex. infört en utvärderingsstandard för hela det centrala, departementala regeringsarbetet i syfte att stärka kvaliteten och att skapa incitament för att genomföra policyutvärdering. Redan 2003 utgavs en s.k. grönbok om hur utvärdering kan användas inom den centrala styrningen.¹³ Syftet med grönboken är att säkra att ingen policy, inget program eller projekt ska antas utan att frågor ställs om huruvida det finns ett bättre sätt att nå ett utpekat mål eller om det finns en bättre användning av dessa resurser. Men syftet är inte enbart fiskalt. Det handlar också om att bedöma den strategiska påverkan som beslutet (om en intervention eller ett stöd eller något annat dylikt) kan antas få och om realismen i beslutet, dvs. möjligheterna att uppnå målen.

Vidare har en omfattande vägledning eller handbok tagits fram, *The Magenta Book*, som på ett både brett och fördjupat sätt redogör för vilken slags utvärdering som bör genomföras.¹⁴

Inom det brittiska utbildningsdepartementet (DfE) har också olika insatser gjorts för att stärka departementet som kunskapsorganisation. Det har bl.a. handlat om insatser för att förstärka departementets analytiska förmåga. Insatserna har gjorts med utgångspunkt i en övertygelse om att om kunskapsstyrning ska få genomslag på alla nivåer inom verksamhetsområdet så måste departementet självt visa ledarskap och gå före och visa vägen.¹⁵

¹² OECD (2007) Knowledge Management. Evidence in Education. Linking research and policy.

¹³ HM Treasury (2011) *The Green Book. Appraisal and Evaluation in Central Government*. Treasury Guidance. London:TSO. Reviderad 2011.

¹⁴ HM Treasury (2011) *The Magenta Book. Guidance for evaluation*. Utöver huvuddokumentet har ett antal bakgrundsdokument publicerats i *The Magenta Book: guidance notes for policy evaluation and analysis. Magenta Book Background Papers*. Government Social Research Unit. HM Treasury (2003, updated 2007). Ytterligare två dokument har publicerats som behandlar olika sidor av kvantitativa och kvalitativa utvärderingsmetoder: *Quality in policy impact evaluation: understanding the effects of policy from other influences* (supplementary Magenta Book guidance). HM Treasury (2012) samt *Quality in qualitative evaluation: a framework for assessing research evidence* (supplementary Magenta Book guidance). HM Treasury (2012).

¹⁵ Department for Education (2013) *Analytical Review*. The Department (By Steering Group).

Ett annat exempel gäller Norge där intresset för evidens-informerade arbetssätt varit växande under de senaste regeringskonstellationerna.¹⁶

Utvärdering av reformer

Frågan om hur reformer bör genomföras för att underlätta utvärdering handlar i stor utsträckning om att utvärderingsfrågorna behöver få en mer framskjuten roll i det departementala arbetet. Det handlar t.ex. om att öka användningen av utvärderingar av tidigare reformer, att klargöra vilken programteori som reformer och åtgärder bygger på, dvs. att explicit beskriva hur en viss åtgärd förväntas bidra till måluppfyllelsen och vilka mekanismer som är tänkta att verka för att de avsedda resultaten ska uppstå.

Förbättrade förutsättningar för utvärdering av reformer handlar inte minst om att faktiskt prioritera utvärderingsfrågorna i samband med *planeringen* av reformer och andra mer genomgripande åtgärder. Som vi visat i föregående kapitel ingår det inte i den svenska traditionen att på skolans område planera för utvärdering inför att reformer beslutas.

Strategiska frågor som regeringen har att ta ställning till

Vi har redan understrukt betydelsen av att reformutvärdering planeras parallellt med att reformstrategin formas. Detta har dock inte endast att göra med det tämligen självklara att förutsättningarna för att genomföra vissa typer av utvärdering, t.ex. effektstudier, är bättre om de planeras från början. Framför allt finns ett starkt samband mellan reformstrategi och utvärderingsstrategi. Valet av reformstrategi styr valet av utvärderingsansats. Omvänt kan utvärdering bidra till genomförandet av en reformstrategi dels genom att fungera formativt i förhållande till den nationella beslutsnivån, dels genom att fungera utvecklingsstödande på den lokala nivån, i verksamheten.

Att utforma reformstrategin och utvärderingsstrategin så att de samverkar på detta sätt måste, som vi ser det, vara ett ansvar för regeringen och Regeringskansliet. Detta är emellertid inte någon

¹⁶ Uppgifterna har lämnats av företrädare för Kunskapsdepartementet som utredningen varit i kontakt med och bedömningen gäller för tiden fram till och med våren 2013.

enkel uppgift, varför utformning och utvärdering av reformer kan vinna på att planeras efter samråd med forskare eller andra utvärderare. Detta bör i så fall ske i ett tidigt stadium av en beredningsprocess. Forskare kan t.ex. bidra med synpunkter på hur en reform eller annan åtgärd bör utformas och implementeras för att den ska vara möjlig att utvärdera med en forskningsbaserad design. Utvecklade och tillitsfulla relationer mellan Regeringskansliet och de institutioner som äger utvärderingskompetens, dvs. myndigheter, universitet, högskolor, särskilda statliga eller privata institut, underlättar givetvis ett sådant samrådsförfarande. Vi återkommer till hur en samverkan kring utvärderingsfrågorna kan organiseras i kapitel 11.

En form av reformstrategi som måste beslutas initialt och som skapar sina speciella förutsättningar för utvärdering är att genomföra försöksverksamheter eller använda sig av stegvis införande. I t.ex. Storbritannien ingår pilotstudier och försöksverksamheter som ett led i de initiativ som tagits för att stärka inslaget av evidensinformerad politikutveckling. Man menar där att inga reformer bör införas i full skala innan man på olika sätt övertygat sig om att reformen fungerar så som det varit tänkt.¹⁷ Även om det ur ett utvärderingsperspektiv är fördelaktigt om det går att genomföra en försöksverksamhet eller att stegvist införa en reform, finns det också situationer där detta inte låter sig göras. Vi återkommer till detta i kapitel 8.

Ytterligare en aspekt av planering handlar om antalet reformer som genomförs samtidigt. Självfallet beror den slutliga utformningen av ett reformpaket på en kombination av olika sakförhållanden och politiska bedömningar. Det måste dock framhållas att antalet reformer som implementeras samtidigt påverkar förutsättningarna för utvärdering. Ju fler reformer som genomförs sam, desto svårare är det att fastställa vilka åtgärder som bidragit till vilka förändringar och med vilken styrka. Detta konstaterades även i samband med Kunskapslyftet, dvs. det senaste stora skolpolitiska reformpaketet i Norge.

Att efterhand göra successiva förändringar i beslutade reformer, t.ex. i syfte att förbättra reformens funktionssätt, är en vanlig implementeringsstrategi. För en utvärdering som nära följer genomförandet av en reform är det därför ofta ett huvudsyfte att tidigt identifiera sådana delar av en reform som kan behöva justeras (real-

¹⁷ Se t.ex. Cabinet Office (2003) *Trying It Out. The Role of 'Pilots' in Policy-Making*. Report of a Review of Government Pilots. Government Chief Social Researcher's Office.

time evaluation). Samtidigt kan sådana förändringar bli ett problem för en mer summativt inriktad utvärdering. Det kan t.ex. bli svårt att uppskatta vilken av förändringarna som bidragit till vilken av de olika effekter som uppstått och med vilken styrka de gjort detta.

Detta innebär inte att den centrala nivån bör avstå från att besluta om förändringar som man anser är nödvändiga. Ett sådant ”reformstopp” i syfte att förbättra utvärderingsbarheten vore orimligt. Man måste dock vara medveten om att möjligheterna att utvärdera försvåras om reformer förändras med korta intervall. Möjligen är det också ett argument till varför reformer bör stå på en god empirisk grund och vara väl genomtänkta innan de beslutas.

6.4 Reformutvärderingens roll bör stärkas

Utredningens bedömning: Utformningen av reformer och utvärderingen av dessa står i ett ömsesidigt förhållande till varandra. Beslut om reformstrategi påverkar förutsättningarna för utvärdering. Samtidigt kan beslut om utvärdering, om det fattas tidigt inför genomförandet av en reform, påverka såväl utformningen som implementeringen av denna. Frågan om reformutvärdering bör därför prioriteras redan i samband med planeringen av reformer och andra mer genomgripande åtgärder.

Utvärderingen kan bidra till genomförandet av reformstrategin genom att fungera formativt i förhållande till den nationella beslutsnivån och utvecklingsstödande på verksamhetsnivån.

En planerad och systematisk utvärdering av reformer är enligt utredningens uppfattning nödvändig för att säkerställa en både bred och djup systemkunskap om skolan. Inte minst är det viktigt för det politiska ansvarsutkrävandet att det går att få veta om den förda politiken har avsedda effekter och om de allmänna medlen allokeras på ett effektivt sätt. Utvärdering av reformer ger också kunskap som kan bidra till att öka genomslaget för reformen respektive signalera om något håller på att gå fel. Slutligen kan reformutvärdering bidra med kunskap inför kommande reformer.

Reformutvärdering har en nära relation till implementeringen av reformer, och dess olika faser. Till att börja med bör resultat från forskning och utvärdering av rådande förhållanden utgöra underlag

för att beslutet om att genomföra en reform ska kunna sägas vara evidensinformerat. Under processimplementeringen av reformen kan utvärdering fylla en funktion genom att fungera formativt och utvecklingsstödande. När reformen slutligen institutionaliserats kan resultat från den summativa utvärderingen ge kunskap om hur verksamheten påverkats, varför det uppstått variationer etc. På nationell nivå kan sedan denna summativa utvärdering utgöra ”ingångsvärden” inför eventuella framtida reformer.

Reformutvärdering har således centrala och viktiga roller vid reformering av offentlig verksamhet, i det här fallet av skolan. Samtidigt har vi i föregående kapitel visat att det är, och har varit, ovanligt med nationellt initierad reformutvärdering på skolans område. Under de senaste åren har dock vissa initiativ tagits till att utvärdera de reformer som genomfördes under 1990-talet. Dessa initiativ har dock tagits först efter tjugo år.

För att reformutvärderingen ska kunna ges den centrala roll den enligt vår uppfattning bör ha, behöver frågan prioriteras på ett helt annat sätt än i dag redan i samband med planeringen av reformer och andra mer genomgripande åtgärder. Det är först när planeringen av reformer, reformimplementeringen och reformutvärderingen sker integrerat som utvärderingens fulla potential kan tas till vara. De beslut om utvärdering som fattas tidigt, inför införandet av en reform, kan komma att påverka både utformning och genomförande av reformen.

Arbetet med detta betänkande och utredningens delbetänkande¹⁸ har stärkt oss i uppfattningen att det är långt ifrån enkelt att förändra skolan genom strukturella reformer. Det tar tid innan effekter uppstår så att de kan bli mätbara på nationell nivå. Dessutom är de effekter man kan förvänta sig ofta inte så stora som man kanske önskat. Detta innebär inte att man kan avstå från att försöka att förbättra förutsättningarna för skolan. Det understryker dock behovet av att beslut om reformer vilar på en ordentlig kunskapsgrund. Det bör för varje åtgärd som vidtas framgå tydligt på vilket sätt åtgärden är tänkt att påverka skolan, och vad som styrker denna slutsats. Forskning och utvärdering har här en central roll både inför att reformer ska utformas och vid utvärderingen av dem. Utvärdering av reformer kan i sin tur resultera i kunskap som kan

¹⁸ SOU 2013:30 *Det tar tid – om effekter av skolpolitiska reformer*.

ligga till grund för evidensinformerade beslut inför eventuella framtida reformer.

Enligt vår uppfattning finns följaktligen starka argument för att tydligare föra in frågan om reformutvärdering i regeringens och Utbildningsdepartementets arbete med att utforma reformer och andra mer genomgripande förändringar. Vi har i det ovanstående pekat på några frågor som är viktiga att ta ställning till i detta sammanhang samt på behovet av att i högre grad använda kvalificerad utvärderingskompetens vid utformningen av utvärderingsstrategier i samband med beslut om reformer.

Med denna utgångspunkt kommer vi i det följande fördjupa oss i vissa för utvärderingen av reformer viktiga frågor. Det handlar om

- hur en överblick över utvärderingslandskapet kan förbättra förutsättningarna för att styra utvärderingen av skolan (kapitel 7),
- möjligheterna med att använda försöksverksamhet och stegvis införande vid reformeringen av skolan (kapitel 8),
- behovet av bättre underlag för att följa kunskapsresultaten (kapitel 9),
- behovet av data för att utvärdera reformer (kapitel 10), samt
- hur utvärdering av skolan kan organiseras på ett mer ändamålsenligt sätt än i dag (kapitel 11).

7 Ramverk för utvärdering av skolan – för bättre överblick och styrning

Av kapitel 4 framgår att det svenska uppföljnings- och utvärderingssystemet består av ett stort antal komponenter och många olika aktörer är involverade. I detta kapitel närmar vi oss därför den fråga om ett nationellt ramverk för utvärdering som aktualiserats av OECD.

7.1 OECD om behovet av ramverk för utvärdering av utbildningen

I en OECD-rapport¹ om utvärdering och bedömning av utbildning konstateras att många länder nu tar initiativ till att utveckla och stärka kapaciteten för att utvärdera sina utbildningssystem. Utvärdering och uppföljning av elevresultat ses i många länder som en motor för den framtida reformagendan. Den fulla potentialen av detta kommer dock, enligt OECD, inte att förverkligas om inte ländernas ramverk för utvärdering av utbildningen utvecklas till fullt sammanhängande helheter. Synergier mellan olika delkomponenter i systemet bör tas till vara, dubblering av procedurer bör undvikas och motsägelser mellan olika målsättningar förhindras. OECD pläderar starkt för att det arbete som inletts i många länder med att utveckla sådana ramverk bör fortsätta.

¹ OECD (2013) *Synergies for Better Learning: International Perspective on Evaluation and Assessment*. OECD Reviews of Evaluation and Assessment in Education.

Vad menar OECD med ramverk?

Att utarbeta ett strategiskt dokument som beskriver ett ramverk för utvärdering kan enligt OECD vara en god idé. Ett sådant dokument ska tydliggöra syftet med ett utvärderingssystem samt hur det är länkat till utbildningssystemet som helhet, till andra förändringar och till förbättrade elevresultat. Dokumentet bör även ange hur informationen i systemet ska rapporteras och beskriva hur olika komponenter i systemet kan producera resultat som är användbara för klassrumspraktiken och för olika skolutvecklingsaktiviteter.

Uppföljning och utvärdering har enligt OECD inget värde om det inte leder till förbättringar i klassrumspraktiken. OECD lägger därför stor vikt vid att utvärderingssystemet på ett tillförlitligt och relevant sätt fångar upp målen för utbildningen på skol- och elevnivå. Annars förfelas hela idén med systemet, menar man. Inte minst undergrävs dess legitimitet bland de professionella. Särskild uppmärksamhet bör därför riktas mot att identifiera oavsedda effekter eftersom uppföljning och utvärdering har en betydande potential att påverka beteendet hos skolans aktörer.

Ramverket bör även bidra till att klargöra ansvarsfördelningen mellan olika aktörer vad gäller de olika delarna i systemet och främja nätverkande och samverkan mellan de personer som arbetar med olika utvärderings- och uppföljningsaktiviteter.

OECD om Sverige

OECD:s rapport baseras på granskningsrapporter från ett stort antal länder². I Sverigerapporten³ konstateras att utvärdering och bedömning har stor betydelse i vårt land på grund av att vi har ett starkt decentraliserat utbildningssystem. De yrkesverksamma behöver få information och feedback i syfte att utveckla kvaliteten på sitt arbete samtidigt som regeringen behöver kunna följa upp om nationella mål för kvalitet och likvärdighet i utbildningen uppnås. Det svenska förhållningssättet kombinerar enligt OECD nationellt uppsatta mål och standarder samt central provutveckling med en hög

² De länder som ingått i studien är Australien, Belgien (den flamländska delen), Chile, Danmark, Luxemburg, Mexiko, Nederländerna, Norge, Nya Zeeland, Portugal, Slovakien, Sverige, Tjeckien och Storbritannien (Nordirland).

³ OECD (2011) *Reviews of Evaluation and Assessment in Education. Sweden*.

grad av förtroende för skolans personal att själva utföra utvärdering och bedömning.

Enligt OECD har Sverige många delar i ett uppföljningssystem på plats. De olika delarna är väl utformade, men man menar att det saknas en strategisk plan eller ett nationellt ramverk för utvärdering och bedömning. En sådan plan eller ett sådant ramverk skulle enligt OECD kunna säkerställa konsekvens och helhetssyn samt bidra till att optimera anpassningen mellan de olika komponenterna, ge en överblickbarhet och utgöra en referens för aktörer som arbetar med utvärdering och bedömning inom utbildningssektorn, reda ut kraven på och villkoren för utvärdering och bedömning på olika nivåer i systemet, förtydliga ansvarsfördelningen och kartlägga utbudet av de verktyg som finns för att årstadkomma optimal praxis. Ramverket bör enligt OECD bland annat kompletteras med beskrivningar av den kompetens som de som ansvarar för utvärdering bör ha.

OECD pekar vidare på olika möjligheter för att få fram mer pålitliga verktyg för systemuppföljning. Frågor som enligt rapporten bör utredas gäller bland annat de nationella proven, stickprovsbaserad kartläggning av studieresultat över tid samt användningen av de internationella studierna. Utredningen återkommer till dessa frågor i det följande. Enligt OECD bör Sverige dessutom främja en ökad användning av de data och den information som redan finns. Man menar vidare att tillförlitligheten vid rättningen av nationella prov behöver förbättras samt att den externa utvärderingen i större utsträckning bör grundas på intern självutvärdering.

Hur ser det ut i andra länder?

I de flesta länderrapporter som ingår i OECD:s studie återkommer man till behovet av att utveckla utvärderingssystemen och stärka de enskilda delarna i dessa samt till behovet av att skapa ett tydligare ramverk för utvärdering och/eller att tydliggöra kopplingen mellan olika komponenter i systemet. Det gäller även för länder som, liksom Sverige, anses ha kommit en god bit på väg i att etablera ett fungerande utvärderingssystem.

OECD menar t.ex. att Danmark har vidtagit flera åtgärder för att stärka utvärderingskulturen, men förordar ändå att landet ska ta fram en strategisk utvecklingsplan för att slutföra arbetet med att skapa ett utvärderingsramverk. Norges utvärderingsramverk behöver

enligt OECD kompletteras och kopplingen mellan de olika komponenterna behöver förstärkas. Etc.

Genom kompletterande litteraturstudier av bland annat forskningslitteratur, och i några fall genom besök i länderna, har utredningen sökt fördjupa bilden av uppföljnings- och utvärderingsystemen i några länder och systemens roll i samband med genomförandet av reformer⁴. Det närmaste vi kommit ett dokumenterat ramverk för utvärdering är i Finland, dvs. ett land som inte deltog i OECD:s studie. I januari 2012 beslutade det dåvarande Utbildnings- och kulturministeriet om en nationell utvärderingsplan för 2012-2015⁵. Denna innehåller vissa utgångspunkter för utvärderingen, organiseringen av utvärderingsverksamheten, riktlinjer för inriktningen av utvärderingen, finansieringen av utvärderingen samt för hur utvärderingsplanen ska följas upp. Även om planen har en övergripande karaktär närmar den sig vad som skulle kunna kallas för ett ramverk.

Vår bedömning

Det är slående att de rekommendationer som OECD lämnar är så pass lika mellan länder som i övrigt är förhållandevis olika. Oberoende av var länderna står i utvecklingen av sina uppföljnings- och utvärderingssystem är huvudrekommendationen att de bör ta fram eller utveckla ett ramverk för utvärdering. OECD anser uppenbarligen att ett sådant dokumenterat ramverk saknas i de flesta länder och att det på ett kraftfullt sätt skulle kunna bidra till utveckling av uppföljningen och utvärderingen av skolans prestationer och ytterst av klassrumspraktiken.

Utredningen finner att OECD:s ambitioner när det gäller ramverk för utvärdering på många sätt är vällovliga. Vi menar dock att det finns en tendens att utvärdering i OECD:s tappning ges en så central roll i styrningen av skolan att den nästan av egen kraft förväntas driva utvecklingen av skolverksamheten. Ramverket begränsas därmed inte till att utgöra underlag för politiska beslut och till att fungera som utgångspunkt för professionens utveckling av verksamheten, utan blir ett kraftfullt nationellt styrinstrument i sig. Vi noterar också att OECD:s utvärderingsbegrepp är smalare än det

⁴ Vi har besökt Danmark, Finland, Norge och Skottland. Därutöver har vi genom dokumentstudier tittat särskilt på även Canada, England, Nederländerna, Nya Zeeland och USA.

⁵ Undervisnings- och kulturministeriet (2012:15) *Utvärderingsplan för utbildningen 2012-2015*.

uppföljning- och utvärderingsbegrepp vi använder i detta betänkande, genom att mer avgränsat avse kunskapsresultat.

7.2 Förslag om ett nationellt ramverk för utvärdering av skolan

Utredningens förslag: Ett nationellt ramverk för utvärdering av skolan fastställs av regeringen, t.ex. i form av en skrivelse till riksdagen, och kompletteras med regelbundna planer för utvärdering av skolan. Av dessa planer bör framgå regeringens övergripande prioritering för utvärderingen av skolan för en överblickbar period, t.ex. tre år. Förslaget avser skolväsendet i sin helhet, dvs. även förskola och fritidshem.

Utredningens bedömning: Ett nationellt ramverk för utvärdering av den svenska skolan bör tydliggöra det övergripande syftet med uppföljningen och utvärderingen av skolan samt

- vilka delar systemet består av för att följa upp, utvärdera, granska och beforska skolan,
- vilka syften systemets olika delar har och vad de förväntas ge svar på,
- hur dessa olika delar förhåller sig till varandra, samt
- vilka aktörer som har ansvar för de olika delarna och hur dessa förväntas samverka.

Av ramverket bör också framgå vilken roll utvärderingen har i samband med större reformer. Ett sådant nationellt ramverk bör beslutas av regeringen och utgöra en över tid stabil grund utifrån vilken den svenska skolan långsiktigt kan utvärderas.

Utredningen konstaterar, liksom OECD, att Sverige har ett väl utvecklat system för uppföljning och utvärdering av den svenska skolan, om man med det menar att vi har ett väl utbyggt uppföljningssystem. Vi delar samtidigt bedömningen att det behövs ett tydligare nationellt ramverk för utvärdering av skolan, även om vi inte lägger riktigt samma innebörd i detta begrepp som OECD. Vi föreslår att ett sådant ramverk tas fram.

Ett sådant ramverk bör, som vi ser det, tydliggöra det övergripande syftet med uppföljningen och utvärderingen av skolan. Vidare skulle det tydliggöra dels vilka delar systemet består av för

att följa upp och utvärdera, granska och beforska⁶ skolan, dels vilka syften dessa olika delar av systemet har och vad de förväntas ge svar på och dels hur de förhåller sig till varandra. Utredningens uppdrag avser utvärderingen av grundskolan och motsvarande skolformer. Vi anser dock att ett förslag som detta måste avse skolväsendet i sin helhet. Det innebär att t.ex. även förskolan ska ingå.

Ett nationellt ramverk för utvärdering ska enligt vår uppfattning beskriva förhållanden på nationell nivå. Det ska omfatta den nationellt insamlade statistiken, annan information från myndigheter etc. Det finns dock i många fall en relation mellan den nationella nivån och verksamhetsnivån som bör framgå. Ett sådant exempel är de nationella proven som har syften som verkar på olika nivåer i systemet. Ramverket bör därför även belysa relationen mellan å ena sidan det nationella systemet för uppföljning och utvärdering och å andra sidan den lokala nivån respektive verksamhetsnivån.

Utöver att tydliggöra utvärderingssystemets olika delar och deras syften bör ramverket också tydliggöra vilka myndigheter och andra aktörer som har ansvar för systemets olika delar. Det kan synas självklart var ansvaret ligger när vi har stora centrala skolmyndigheter, som Skolverket och Skolinspektionen. I dag finns det dock flera ytterligare aktörer på den svenska utvärderingsscenen med uppdrag om utvärdering som berör skolan. På nationell nivå handlar det om Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU), Statskontoret, Vetenskapsrådet samt universitet och högskolor. Vidare står Statistiska Centralbyrån (SCB) för insamling av statistik från huvudmän och skolor. SCB publicerar dessutom regelbundet statistik om skolan samt svarar för de s.k. elevpanelerna. Slutligen genomför även andra aktörer, exempelvis Sveriges kommuner och landsting (SKL), viss uppgiftsinsamling och publicerar vissa nyckeltal.

Utredningen menar att det av ett nationellt ramverk bör framgå vilken roll utvärderingen har i samband med större reformer, dvs. såväl inför, under och efter en reform, samt hur sådan utvärdering kan organiseras. Den konkreta utvärderingen av en specifik reform måste dock utformas i förhållande till just den aktuella reformen. Syftet med ramverket är därför att utgöra en infrastruktur för

⁶ Observera att forskningen kan vara oberoende eller genomföras på uppdrag. Att forskningen tas med som en aktivitet i ett ramverk inskränker inte på något vis den oberoende forskningens oberoende.

utformningen av en sådan utvärdering, inte att vara en handbok om hur den ska genomföras i det specifika fallet.

Vi föreslår även att ramverket kompletteras med regelbundet återkommande planer för utvärdering av skolan, t.ex. vart tredje år. Av dessa bör framgå regeringens övergripande prioritering för utvärderingen av skolan för den aktuella perioden. Enligt vår uppfattning skulle sådana planer kunna bidra till att utvärderingen av skolan blir mer långsiktig och i högre grad kommer att baseras på en analys av de verkliga kunskapsbehoven. I en reformfas kommer en sådan plan i stor utsträckning att präglas av reformutvärdering.

Ramverket och planen bör beslutas på politisk nivå

Som redovisats ovan finns bara på statlig nivå många aktörer med ansvar för olika delar i ett tänkt ramverk för utvärdering. Vi föreslår därför att det nationella ramverket för utvärdering ska fastställas av regeringen, förslagsvis i form av en skrivelse till riksdagen. I arbetet med ramverket skulle ansvarsfördelningen mellan myndigheter och andra aktörer tydliggöras och eventuella luckor i systemet identifieras. Ett nationellt ramverk skulle därmed bidra till överblick och ett bättre helhetsgrepp.

Enligt vår uppfattning är det dessutom viktigt att ramverket utarbetas i former som tar hänsyn till den samlade kunskap som olika parter inom skolans område har, dvs. lärare, skolledare, kommuner, myndigheter, forskare, etc. Vi återkommer till hur detta inflytande skulle kunna organiseras i kapitel 11 om organisation och ansvarsfördelning.

Utredningen vill dessutom understryka värdet av långsiktighet när det gäller utvärderingen av skolan. Det är centralt såväl med hänsyn till användbarhet och kvalitet som ur ett demokratiperspektiv. Elever och föräldrar, lärare, rektorer, politiker, media, allmänhet – dvs. alla – ska kunna förvänta sig en kontinuitet när det gäller informationen om den svenska skolan, oavsett om denna beskriver en positiv eller negativ utveckling. När det gäller utformningen av ramverket vore det därför önskvärt med en bred politisk överenskommelse som kan säkerställa att det får fortleva under en längre tid. Ett nationellt ramverk bör på så vis kunna bidra till att etablera en stabilitet och förutsägbarhet över tid när det gäller t.ex. internationella och nationella kunskapsmätningar.

Ett ramverk som beslutas på regeringsnivå måste samtidigt hållas på en övergripande nivå, dvs. handla om principer och de stora dragen, snarare än om detaljer. Annars försvåras berörda myndigheters och andra aktörers arbete med att ge det ett konkret innehåll. Avsikten bör därför vara att ge berörda myndigheter och andra aktörer tydliga förutsättningar för att genomföra sina uppdrag. Ramverket bör även utformas så att det medger viss flexibilitet och anpassning i förhållande till framtida förändringar utan att behöva revideras. Även detta talar för att det inte bör utformas alltför detaljerat.

Arbetet med att utforma ett ramverk kan leda till omprövning

Arbetet med att utforma ett nationellt ramverk för utvärdering av skolan skulle tydliggöra vad Sverige faktiskt har i form av underlag för att utvärdera skolan. Samtidigt skulle utarbetandet av ramverket identifiera vad som saknas, vilka frågor som *inte* kan besvaras med nuvarande system samt vad som är otydligt med avseende på gränsdragning mellan olika aktiviteter och aktörer. Ett sådant arbete skulle med allra största sannolikhet leda till en omprövning av vissa delar av systemets utformning.

Utredningen har inte sett som sin uppgift att lämna förslag till ett färdigt ramverk för utvärdering av den svenska skolan. Vårt arbete kan dock ses som starten på en process mot ett sådant ramverk. Vissa frågor som skulle aktualiseras i samband med ett sådant arbete kommer vi dessutom att ta upp i de följande kapitlen.

8 Försöksverksamhet – för ökad evidens vid reformer

Användningen av försöksverksamhet i samband med reformer eller andra förändringar har nära koppling till utvärdering. Själva poängen med en försöksverksamhet är ju att den ska utvärderas innan beslut fattas om huruvida försöket ska permanentas eller avslutas. Även stegvis införande av en reform kan betraktas som en form av försöksverksamhet.

Det finns flera fördelar med att genomföra försöksverksamhet inför en reform eller större förändring. Genom att pröva i liten skala kan man undvika misstag i full skala. Försök kan därför lämpa sig i de fall utfallet av en åtgärd inte i förväg kan förutses genom resultat från tidigare forskning och utvärdering. I ett försök är det också möjligt att lära av de erfarenheter försöket ger och vid behov göra justeringar innan reformen eventuellt permanentas. Genom att försöksverksamhet ger förutsättningar för att skapa en kontrollgrupp möjliggörs dessutom användningen av utvärderingsdesigner som annars kan vara svåra att få till stånd. Det bör dock understrykas att det finns andra metoder att skapa kontrollgrupper än genom försök (se vidare kapitel 3). Dessutom måste hänsyn tas till effekterna för berörda elever, vilket vi återkommer till i det följande.

8.1 Användningen av försöksverksamhet

Försöksverksamhet är inte någon okänd reform- eller utvecklingsstrategi inom svensk skolpolitik. Användningen av försöksverksamhet är dock en tradition som försvagats i Sverige.¹ Ett skäl till detta är sannolikt att tempot i politiken har ökat över tid. Bered-

¹ Se bilaga 3.

ningstiderna i regeringskansliet är långa, samtidigt som mandatperioderna kan uppfattas som korta.

Utredningen har låtit göra en historisk genomgång av utvärderingens roll i samband med svenska skolreformer (se bilaga 3). Av denna framgår att försöksverksamhet användes i betydande omfattning i samband med parallellskolans avveckling och enhetsskolans införande.² Även inom ramen för SIA-utredningen genomfördes viss försöksverksamhet.³ Mellan åren 1988 och 1993 genomfördes senare de s.k. ÖGY-försöken i regi av Kommittén för Översyn av den Gymnasiala Yrkesutbildningen. Försöken handlade om en förlängning av gymnasieskolans yrkesutbildningar och ett ökat inslag av teoretiska ämnen.⁴ Vidare beslutade riksdagen 1999 om Timplaneförsöket, en från början femårig försöksverksamhet som innebar att grundskolor i högst 80 kommuner, inom ramen för det totalt garanterade antalet timmar för grundskolan, fick bedriva utbildning utan att tillämpa timplanens bestämmelser om antalet timmar per ämne eller ämnesgrupp i grundskolan.⁵

Förutom dessa mer omfattande försöksverksamheter har det även förekommit, och förekommer, försök som rört mer avgränsade frågeställningar, nämligen

- försöken med föräldrastyrelser i grundskolan och elevstyrelser i gymnasieskolan,
- försöksverksamhet med möjlighet för föräldrar att välja att placera barn med utvecklingsstörning i grundskolan i stället för i särskolan,
- försöksverksamhet med lärlingsutbildning,
- försöksverksamhet med undervisning på engelska i grundskolan, samt
- försöksverksamhet med spetsutbildning.

² Se även Marklund, Sixten (1981) *Skolsverige 1950-1975. Del 3 Försöksverksamheten*. Utbildningsförlaget.

³ Bilaga 3. Reformutvärdering i Sverige – Historisk tillbakablick.

⁴ Se t.ex. Hall, Caroline (2009) *Förlängningen av yrkesutbildningarna på gymnasiet: effekter på utbildningsavhopp, utbildningsnivå och inkomster*. IFAU, rapport 2009:7.

⁵ Försöket bedrevs med stöd av förordningen (1999:903) om försöksverksamhet med utbildning utan timplan i grundskolan. Denna kom senare att förlängas i omgångar fram till den nya skollagen (2010:800) trädde i kraft. Timplanedelegationen (dir. 1990:90 och 2003:173) hade uppdraget att förbereda, leda, stödja och utvärdera försöksverksamheten.

Varierande grad av utvärdering

Flera av dessa försöksverksamheter har pågått under lång tid och förlängts genom nya beslut. Det gäller bl.a. Timplaneförsöket som förlängdes tills den nya skollagen (2010:800) trädde i kraft och försöken med föräldrastyrelser i grundskolan respektive elevstyrelser i gymnasieskolan samt försöket att välja skolform för elever med utvecklingsstörning. Försöksverksamheten har i vissa fall varit utformad som ett stegvis genomförande, t.ex. i samband med övergången till enhetsskolan.

Några exempel på utvärderingar av försöksverksamhet utgörs av Skolverkets rapporter från 1999 och 2002 om försöket med möjlighet för föräldrar att välja skolform för elever med utvecklingsstörning⁶, Statskontorets rapport från 2002 av försöket med föräldrastyrelser⁷, Skolverkets utvärdering från 2010 av försöket med undervisning på engelska i grundskolan⁸ och Skolverkets rapport från 2013 om spetsutbildning⁹. Försöket med lärlingsutbildning var organiserat som en offentlig utredning och har rapporterats i tre betänkanden från Lärlingskommittén¹⁰. Skolverket har senare redovisat en rapport om den gymnasiala lärlingsutbildningens tre första år.¹¹ Inom ramen för ÖGY-försöken pågick en viss utvärdering, men i det fallet genomfördes reformeringen av gymnasieskolan innan slutrapporten från försöksverksamheten hann framläggas.

Flertalet av ovan nämnda rapporter har karaktären av kartläggningar genom intervjuer och enkäter. Ibland utmynnar de i bedömningar. Man kan således inte säga att de har utvärderats enligt någon vedertagen definition av begreppet utvärdering, dvs. en systematiskt utförd undersökning av om försöket varit framgångsrikt eller inte (i förhållande till uppsatta mål eller förväntningar om positiv verksamhetsutveckling).

⁶ Skolverket (1999) *Uppföljning och utvärdering av försöksverksamheten med ökat föräldrainflytande över val av skolform för utvecklingsstörda barn – Slutrapport* – (dnr. 96:565) och Skolverket (2002) *I särskola eller grundskola?* (dnr. 103404).

⁷ Statskontoret (2002) *Regeringsuppdrag att göra en utvärdering av försöksverksamheten med lokala styrelser* (dnr. 2001/320-5).

⁸ Skolverkets rapport 351 (2010) *Undervisning på engelska Utvärdering av en försöksverksamhet i grundskolan*.

⁹ Skolverket (2013) *Redovisning av uppdrag enligt förordning (2011:355) avseende omfattning och utvärdering av försöksverksamhet med riksvrekryterande spetsutbildning i grundskolans högre årskurser läsåret 2012/13* (dnr. 69-2013:00527).

¹⁰ SOU 2009:85 *Gymnasial lärlingsutbildning – Hur blev det? Erfarenheter från första försöksåret*, SOU 2010:75 *Gymnasial lärlingsutbildning – Utbildning för jobb Erfarenheter efter två års försök med lärlingsutbildning* och SOU 2011:72 *Gymnasial lärlingsutbildning – med fokus på kvalitet – Hur stärker vi kvaliteten i gymnasial lärlingsutbildning*.

¹¹ Skolverkets rapport 373 (2012) *Gymnasial lärlingsutbildning de tre första åren 2008–2011*.

Timplaneförsöket skiljer ut sig genom det sätt på vilket det till försöket knöts omfattande resurser för både forskning och utvärdering. Det resulterade i sammantaget tio forsknings- och utvärderingsprojekt, varav några använde metodik med kontrollgrupp. Fler-talet av studierna byggde på små urval och kvalitativa studier, i några fall med klassrumsobservationer. Det genomfördes dock även statistiskt säkerställda kvantitativa analyser. Utredningen genomförde dessutom sådana statistiska analyser av t.ex. underlaget till Skolverkets nationella utvärdering av grundskolan (NU 03). Timplanedlegationen sammanfattade resultaten i två betänkanden och publicerade en forsknings- och utvärderingsantologi.¹²

Få metodologiska ställningstaganden i exemplen ovan

Det finns flera viktiga överväganden som behöver göras om en försöksverksamhet ska bli utvärderingsbar, och inte bara framstå som ett successivt införande av karaktären ”trial-and-error”. Framför allt behöver man tydligt formulera vad det är försöket ska pröva – vad är det man vill veta? – och hur det ska utvärderas. Man bör också ta ställning till hur urvalet ska göras av de individer eller verksamheter som ska ingå i försöket och av de som ska utgöra kontrollgrupp. Ska man t.ex. göra ett representativt urval? Eller ska det vara frivilligt att delta i försöket, och vilka konsekvenser får i så fall det för utvärderingen?

För flertalet av de försök som redovisats ovan har inte denna typ av överväganden gjorts, eller gjorts i liten utsträckning, vilket påverkat förutsättningarna för att utvärdera försöken. Typiskt är att deltagandet i försöken byggt på frivillighet. Uppdrag om att utvärdera försöken har i vissa fall lämnats efter att dessa pågått en viss tid. Man kan säga att fokus har legat på vad som skulle prövas, inte på hur det sedan skulle vara möjligt att dra några säkra slutsatser av försöken.

När det gäller Timplaneförsöket fördes en kortfattad diskussion om förutsättningarna för utvärdering i den proposition som föregick försöket. Regeringen resonerade där om att det omfattande urvalet (80 kommuner) skulle göra det möjligt att i ett nästa steg göra lämpliga urval för utvärdering, och att detta skulle skapa goda förutsättningar för utvärdering av god kvalitet. Man förde däremot

¹² SOU 2004:35 *Utan timplan – med målen i sikte*, SOU 2005:101 *Utan timplan – för mål-nriktat lärande* och SOU 2005:102 *Utan timplan – forskning och utvärdering*.

ingen diskussion om konsekvenserna för utvärderingens validitet av att det var frivilligt att delta i försöket samt att det dessutom ställdes kvalitetskrav på kommunerna för att få delta. Regeringens beslut föregicks dock av en departemental utredning¹³.

8.2 Försöksverksamhet som strategi i andra länder

Utredningen konstaterar att det finns länder som har som strategi att använda antingen försöksverksamhet inför att reformer ska genomföras eller stegvis införande av reformer.

I exempelvis England har försöksverksamheter getts en allt viktigare roll. De brittiska regeringar som suttit vid makten under de två senaste decennierna har målmedvetet drivit på en utveckling i riktning mot evidensinformerade arbetsformer i regeringsarbetet. I en rapport från Cabinet Office¹⁴ heter det t.ex. att det fullskaliga införandet av nya reformer, och de mekanismer för implementering som man avser att använda, om möjligt alltid bör föregås av noggrant kontrollerade försöksverksamheter. Man menar att ett införande som sker stegvis inte bara bidrar till att informera de som berörs av implementeringen, utan även gör det möjligt att identifiera och förebygga oavsedda konsekvenser. Ett försök ses som ett viktigt första steg i en reguljär, långsiktig granskning och utvärdering av reformer. Att försöksverksamhet är förenat med kostnader i form av tid och resurser och kan medföra vissa politiska risker balanseras, enligt rapporten, av den större risken att bädda in brister i den slutliga reformen, som med ett försök hade kunnat undvikas.

Samtidigt ska konstateras att utvecklingen tar sig olika vägar i olika länder. Exempelvis har Nederländerna från och med 1980-talet gått från att genomföra utvärdering av försöksverksamheter utförda av olika kommittéer, till att i stället mer kontinuerligt följa utvecklingen genom ett nationellt uppföljningssystem och genom inspektionsverksamhet.¹⁵ Detta ska dock förstås som en följd av att landet bytt reform- och utvecklingsstrategi och gått från strategier som byggde på centralt initierade nationella reformer till en mer stegvis inriktad strategi. Den nuvarande nederländska utvecklings-

¹³ Ds 1999:1 *Utan timplan – med oförändrat uppdrag*.

¹⁴ Cabinet Office (2003) *Trying it out. The Role of 'Pilots' in Policy-Making*. Report of a Review of Government Pilots. Government Chief Social Researcher's Office.

¹⁵ OECD (2012) *Country Background Report for the Netherlands*. OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.

strategin kan därför betecknas som en kontinuerligt pågående, småskalig strategi som har många kännetecken gemensamma med just försöksverksamhet.

Slutligen kan konstateras att behovet av försöksverksamheter med kontrollgrupper varierar mellan olika länder. I en federalt uppbyggd stat som till exempel USA beslutas merparten av skolpolitiken på delstatsnivå. Exakt samma förändringar införs sällan samtidigt i alla delstater. Det innebär att det uppstår vad man skulle kunna kalla naturliga kontrollgrupper. När en reform eller ett särskilt program införs i skolorna i en delstat kan elever i en annan delstat utgöra kontrollgrupp.¹⁶

8.3 Försöksverksamhet stärker förutsättningarna för reformutvärdering

Försöksverksamhet kan bidra till både formativ och summativ utvärdering av en reform. Genom en försöksverksamhet kan föreställningar testas, om på vilket sätt en åtgärd kommer att påverka ett visst problem respektive om den kommer att åstadkomma ett visst förväntat resultat.

Utöver att ge svar på vilka resultat en försöksverksamhet ger upphov till kan den också ge svar på om det föreligger stor variation i hur reformen genomförs. Utvärdering av försöksverksamhet kan vidare ge svar på om reformen ger upphov till oavsedda negativa effekter, bedöma i vilken omfattning de förekommer och vilka konsekvenser de riskerar att få. I bästa fall kan utvärderingen även bidra med förklaringar till varför variationer och negativa effekter uppkommer och ge rekommendationer om vad som kan göras för att förhindra dem.

Utvärdering av en försöksverksamhet kan även leda till att man avstår från att införa en reform. Det gäller t.ex. där det är oklart hur reformen skulle fungera och risken för ett misslyckande bedöms vara överhängande. Att i förväg testa och utvärdera en tilltänkt reform genom ett försök kan vara särskilt viktigt i de fall reformen i någon mening är innovativ, dvs. inte har prövats tidigare, är

¹⁶ Denna skillnad i förutsättningarna för försöksverksamhet i olika länder uppmärksammas även i det brittiska dokumentet *Trying it Out. The Role of 'Pilots' in Policymaking*. Cabinet Office, 2003.

experimentell till sin natur eller av något annat skäl är förenad med en hög grad av osäkerhet.¹⁷

8.4 Etiska överväganden i samband med försöksverksamhet

En förutsättning för att genomföra en försöksverksamhet är att enskilda individer, i vårt fall elever, utsätts för olika behandling. Vissa omfattas av försöket medan andra utgör kontrollgrupp. Om en utvärdering ska ske med experimentella metoder finns det olika tekniker att skapa kontrollgrupper. Även om en utvärdering inte är utformad med kontrollgrupper innebär försöksverksamhet i de flesta fall att vissa elever blir föremål för en intervention av något slag, medan andra elever inte blir det. Det kan gälla även vid stegvisa införanden av förändringar, men behöver inte göra det.

Invändningar av etisk karaktär har i olika sammanhang framförts mot att elever behandlas olika på grund av försöksverksamhet. Om en åtgärd leder till positiva resultat för den grupp som omfattas av åtgärden kan det anses som "orättvist" att övriga elever inte omfattas av åtgärden. Om en åtgärd i stället leder till negativa resultat kan det anses att en orättvisa har begåtts mot dem som omfattats av åtgärden.

Denna typ av etiska överväganden görs dagligen av forskare inom olika discipliner i samband med forskning eller utvärdering som bygger på experimentella ansatser. Det råder därför ingen brist på forskarreferenser om ämnet.

Clarke¹⁸ menar t.ex. att en grundläggande princip vid experimentell forskning är att experiment endast ska vidtas när förhållandena är sådana att det råder osäkerhet kring effekterna av en specifik intervention. I sådana fall uppkommer inte de etiska problemen som annars sammanhänger med skapandet av slumpvisa kontrollgrupper. Där det inte finns någon kunskap om huruvida en åtgärd fungerar utgör det faktum att kontrollgruppen inte berörs av åtgärden inte nödvändigtvis ett etiskt problem. Vid varje enskilt utvärderingstillfälle – och en försöksverksamhet kan här uppfattas

¹⁷ HM Treasury (2011) *The Magenta Book. Guidance for evaluation*.

¹⁸ Clarke, Alan (2007) *Evidence-based evaluation in different professional domains: similarities, differences and challenges* i Shaw, Ian F. m. fl. (red.) *The Sage Handbook of Evaluation*. Sage Publications.

som ett utvärderingstillfälle – ligger ansvaret enligt Clarke på utvärderaren att redogöra för möjligheten att anlägga ett experimentellt förhållningssätt och visa på vilket sätt fördelarna med slumpvisa experiment väger tyngre än att inte göra det.

Andra forskare går längre. Weisburd¹⁹ menar att man snarare bör börja med att fråga sig varför en experimentell design inte ska användas. Det ansvar som bör åligga forskaren är att förklara varför en mindre valid metod skulle läggas till grund för slutsatser om åtgärder och praxis. Weisburd menar att det etiska problemet består i att valet av icke-experimentella metoder kan komma att inkräkta på forskarens grundläggande professionella skyldighet att erbjuda de mest valida svar som forskaren kan få på frågor som han/hon har blivit ombedd att besvara.

Clarke²⁰ menar att undersökningar av hur patienter, sjukvårdspersonal och andra berörda ställer sig till experiment med slumpvis utvalda kontrollgrupper, visar att dessa i mycket mindre omfattning uttrycker motstånd mot detta slags experiment än en del samhällsforskare gör. En slutsats som kan dras från dessa undersökningar är enligt Clarke att man i samband med experiment med slumpvis utvalda kontrollgrupper behöver lägga både tid och engagemang på att klarlägga varför en sådan forskningsdesign behöver användas.

Utredningen konstaterar att etiska överväganden alltid måste göras innan en försöksverksamhet inleds. När det finns starka etiska invändningar ska en försöksverksamhet givetvis inte initieras. All försöksverksamhet, som kan komma ifråga på skolans område, har dock inte samma etiska problematik. I många fall bör denna kunna hanteras på ett ansvarsfullt sätt.

Om exempelvis en försöksverksamhet snabbt visar så stora positiva effekter att det kan uppfattas som oförsvarligt att endast låta försöksgruppen omfattas av dem, är det inget som hindrar att man beslutar att tidigarelägga genomförandet av reformen för samtliga elever. I det fall en försöksverksamhet i stället visar starka oförutsedda negativa effekter kan försöket naturligtvis avbrytas i förtid eller justeras så att de negativa effekterna inte uppstår. Eftersom det gäller ett försök handlar det om ett begränsat och

¹⁹ Weisburd, David (2003) *Ethical practice and evaluation of interventions in crime and justice: the moral imperative for randomized trials*. Evaluation Review, 27(3). Här citerad efter Clarke, Alan (2007) a.a.

²⁰ Clarke, Alan (2007) a.a.

kontrollerbart antal berörda elever. Det bör därför vara möjligt att kompensera de elever som har deltagit i försöket för eventuella negativa effekter.

Det är inte lika enkelt att på detta sätt kompensera elever om det visar sig uppstå negativa effekter vid en intervention i full skala. Vi menar därför att de etiska övervägandena ofta är lika viktiga när en reform ska genomföras i full skala utan föregående försök, som i samband med en försöksverksamhet. I det första fallet drabbas samtliga elever av de negativa effekterna av en inte tillräckligt väl förberedd reform.

8.5 Förutsättningarna för försöksverksamhet i skollagen

Utgör kraven på likvärdighet och rättssäkerhet hinder för försöksverksamhet?

Likvärdighet och rättssäkerhet är två centrala begrepp inom skolan. Av skollagen framgår att alla elever ska ha lika tillgång till utbildning och att utbildningen inom skolväsendet ska vara likvärdig inom varje skolform och inom fritidshemmet oavsett var i landet den anordnas.²¹

Vad som ska avses med likvärdighet inom skolan har diskuterats i många sammanhang.²² Skolverket menar att likvärdighetsbegreppet kan delas upp i tre grundläggande aspekter; lika tillgång till utbildning, lika kvalitet på utbildningen och att utbildningen ska vara kompenserande.²³ Verket hänvisar till formuleringar i den nya skollagen (2010:800) och läroplanen (Lgr 11) men konstaterar att motsvarande formuleringar fanns även tidigare.

Av författningskommentaren till propositionen om den nya skollagen framgår följande:

Begreppet ”likvärdig” innebär inte att utbildningen ska vara likformig i betydelsen likadan utan att kvaliteten i verksamheten ska vara så hög att de fastställda målen kan uppnås oavsett var i landet verksamheten bedrivs. Hur verksamheten ska utformas för att nå de fastställda målen

²¹ 1 kap. 8 och 9 §§ skollagen (2010:800).

²² Se t.ex. Skolverket (2012) *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid* (Rapport 374) och Ingegerd Tallberg Broman, Lena Rubinstein Reich och Jeanette Hägerström (2002) *Likvärdighet i en skola för alla. Historisk bakgrund och kritisk granskning* (Skolverket) samt Skolverket (1995) *Likvärdighet i skolan – en antologi*.

²³ Skolverket (2012) *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid*. Rapport 374.

kan variera beroende på lokala behov och förutsättningar. Det finns olika vägar att nå de fastställda målen. Hänsyn ska tas till barnens och elevernas olika förutsättningar och behov. Det finns utrymme för anpassningar av undervisningen och organisationen av utbildningen till behoven hos olika barn och elever, så länge de tillförsäkras lika tillgång till likvärdig utbildning.²⁴

Skolan ska enligt skollagen även ta hänsyn till olika elevers behov och kompensera för ogynnsamma hemförhållanden och andra olikartade förutsättningar:

I utbildningen ska hänsyn tas till barns och elevers olika behov. Barn och elever ska ges stöd och stimulans så att de utvecklas så långt som möjligt. En strävan ska vara att uppväga skillnader i barnens och elevernas förutsättningar att tillgodogöra sig utbildningen.²⁵

Det likvärdighetsbegrepp som skollagen ger uttryck för ställer krav på verksamhetens kvalitet men ger också utrymme för att verksamheten kan utformas på olika sätt så länge den strävar efter att nå de fastställda målen. Utredningen slutsats är att likvärdighetsbegreppet därigenom inte lägger hinder i vägen för försöksverksamhet.

De rättsäkerhetsfrågor som aktualiseras i samband med försöksverksamhet rör myndighetsutövning mot enskild. Det är t.ex. tveksamt om en försöksverksamhet skulle kunna anordnas när det gäller antagning till gymnasieskolan. Det bör dock vara fullt möjligt att anordna försöksverksamhet vad gäller olika aspekter av betygssystemet i de årskurser där betygen inte ligger till grund för antagning, t.ex. i årskurserna 6 och 7. Det skulle i så fall göra det möjligt att, i det fall frågan är aktuell, göra kontrollerade experiment med olika former av standards för exempelvis antalet betygsteg i betygsskalan, utformningen av kunskapskraven, etc. Man kan också inom detta område tänka sig ett stegvis införande av förändringar, där exempelvis vissa kommuners skolor får utgöra piloter.

²⁴ Prop. 2009/10:165

²⁵ 1 kap. 4 § skollagen (2010:800).

Hur ska skollagens bestämmelse om försöksverksamhet tolkas?

I skollagen (2010:800) bemyndigas regeringen eller den myndighet som regeringen bestämmer att få meddela föreskrifter om försöksverksamhet inom skolväsendet. I sådana föreskrifter får undantag göras från *organisatoriska* bestämmelser i skollagen. Vidare får undantag göras även från andra bestämmelser i skollagen för att möjliggöra en försöksverksamhet med sådan utbildning i gymnasieskolan som inte utgörs av utbildning på nationella program eller introduktionsprogram.²⁶

Vad som avses med organisatoriska bestämmelser är dock inte tydligt. Propositionen *Den nya skollagen – för kunskap, valfrihet och trygghet* (prop. 2009/1:165) ger ingen annan vägledning än att det finns en bestämmelse²⁷ i 1985 års skollag med i stort sett identiskt innehåll. De enda förändringarna som gjorts i den nya skollagen är en anpassning till den nya skollagens begrepp; förändringarna påverkar inte det som diskuteras här.

I prop. 1985/86:10, som föregick 1985 års skollag, motiveras bestämmelsen med att försöksverksamhet spelar en stor roll inom det offentliga skolväsendet. För viss försöksverksamhet föreligger behov av att kunna göra avsteg från föreskrifter som annars gäller i *pedagogiskt och organisatoriskt avseende*. I det flesta fall handlar det om avsteg från föreskrifter i t.ex. skolförordningen eller någon läroplan, men det kan uppkomma behov även av undantag från organisatoriska bestämmelser i skollagen.

I de fall riksdagen bemyndigat regeringen att utfärda en förordning inom ett område, så har regeringen enligt denna tolkning, möjlighet att initiera försöksverksamhet och besluta om undantagsbestämmelser på förordningsnivå, oberoende av om det rör sig om avsteg som är av organisatorisk eller pedagogisk karaktär. Det gäller så länge det inte krävs ett undantag även från bestämmelser i skollagen.

Bestämmelserna i 1985 års skollag bör bedömas mot bakgrund av att denna lag initialt innehöll en mer omfattande reglering av den kommunala organisationen på skolområdet, medan föreskrifter om t.ex. undervisningens innehåll fanns på lägre nivå i hierarkin med bestämmelser.

²⁶ 29 kap. 23 § skollagen (2010:800).

²⁷ 15 kap. 4 § skollagen (1985:1100).

Utredningens slutsats är att med stöd av 1985 års skollag skulle försöksverksamhet ha kunnat initieras vid långt fler tillfällen än vad som faktiskt skett under en lång rad av år.

Effekter av den nya skollagen

2010 års skollag har emellertid påverkat förutsättningarna för att genomföra försöksverksamhet. I skollagen återfinns nu bestämmelser som med 1985 års skollag var placerade på förordningsnivå. Det framgår av prop. 2009/10:165 att det har varit en medveten strategi att binda fler bestämmelser i lagen.

Detta måste enligt vår tolkning innebära att förutsättningarna för regeringen att besluta om försöksverksamhet som inte avser organisatoriska bestämmelser har begränsats. Den omreglering som skett mellan lag och förordning kan alltså ha fått till följd att möjligheterna att bedriva försöksverksamhet har minskat utan att detta varit avsikten.

Oklart med vad som menas med en organisatorisk bestämmelse

Samtidigt som utredningen bedömer att det redan i dag finns ett större utrymme för försöksverksamhet än vad som faktiskt tagits i anspråk, menar vi att innebörden i det i skollagen centrala uttrycket ”organisatorisk bestämmelse” är oklar.

Att uttrycket ”organisatorisk” syftar på något som har med organisation att göra, dvs. med en planlagd indelning av en verksamhet, är uppenbart. Men hur långt kan tolkningen tänjas? Det finns många inslag i en skolverksamhet som *också* har en organisatorisk aspekt. Medger skollagen då att själva sakfrågan kan beröras av undantag – och därmed bli föremål för försöksverksamhet – eller är det endast den organisatoriska aspekten som kan beröras av undantag?

Hur svårt det kan vara att fånga innebörden i begreppet kan illustreras med den förordning som låg till grund för det s.k. timplaneförsöket. Enligt *Förordning om försöksverksamhet med utbildning utan timplan i grundskolan* (1999:903) fick en skola bedriva undervisning inom ramen för det totala garanterade antalet timmar utan att tillämpa föreskrifterna om antalet timmar för ämnen och ämnesgrupper i bilaga 3 till skollagen (1985:1100). Eftersom tim-

planen utgjorde en bilaga även till 1985 års skollag måste alltså timplanen ha betraktats som en organisatorisk fråga. I annat fall hade en försöksverksamhet inte kunnat äga rum inom detta område. Timplanen påverkar dock planeringen av undervisningen, varför tolkningen av begreppet organisatorisk i detta fall måste anses som förhållandevis vid.

Det finns överhuvudtaget många bestämmelser i skollagen som måste anses ha en eller annan organisatorisk aspekt. Så till exempel kan även ämnena i grundskolan anses utgöra en planlagd indelning av verksamheten i skolan. Samma sak gäller skolans indelning i årskurser, läsår och terminer. Frågan är dock om det är denna typ av bestämmelser som är mest meningsfullt att göra undantag från i syfte att bedriva försöksverksamhet.

Utredningen är tveksam till om begreppet ”organisatorisk bestämmelse” är tillräckligt tydligt. I fråga om försöksverksamhet är det å ena sidan viktigt att skollagens bestämmelse skyddar det som verkligen behöver skyddas från experiment, å andra sidan ger utrymme för försök som bidrar till kunskap om och utveckling av skolan. Enligt vår uppfattning handlar skyddsbehovet i första hand om företeelser som innebär att enskilda elevers rättssäkerhet skulle inskränkas eller försök som allvarligt skulle hota likvärdigheten. Det är t.ex. inte rimligt att tänka sig skillnader i myndighetsutövningen gentemot elever i och utanför ett försök som får konsekvenser för eleven, t.ex. betygssättning i årskurs 9.

När det gäller likvärdigheten vill vi samtidigt påminna om att likvärdighet inte nödvändigtvis innebär att allt ska vara likadant för alla elever, däremot att elever ska erbjudas en likvärdig kvalitet. Samtidigt som begreppet ”organisatoriska bestämmelser” möjliggör vida tolkningar, ser vi en risk för att bestämmelsen också begränsar möjligheten till försök som skulle vara till stor nytta, särskilt inför reformer och större förändringar i styrningen av skolan.

8.6 Möjligheterna att bedriva försöksverksamhet bör tillvaratas

Utredningens bedömning: Försöksverksamhet är ett sätt att underlätta utvärdering av reformer som i ökad utsträckning bör tillvaratas. Kunskaper och erfarenheter som vinnas genom försöksverksamhet kan bidra till att lägga en grund för långsiktigt bärkraftiga beslut.

Nuvarande bestämmelse i skollagen om försöksverksamhet är oklar varför det är angeläget att frågan utreds. En fråga är vad som avses med organisatoriska bestämmelser, en annan är vilken typ av bestämmelser i skollagen regeringen faktiskt ska få besluta om undantag för genom att inrätta försök.

Fördelarna med försöksverksamhet framstår på många sätt som uppenbara. Innovationsrådet²⁸ menar t.ex. att verkligheten är så komplex och föränderlig att det är svårt att till fullo förutsäga vad olika förändringar får för konsekvenser och hur människor reagerar på en viss åtgärd eller reform. Genom att förslag provas på riktigt inom ramen för försöksverksamheter ges värdefull information inför beslut om ett permanent och nationellt införande av ett visst förslag. Rådet föreslår därför att instrumentet försöksförfattningar ska användas mer systematiskt för att pröva nya lösningar.

Utredningen menar att den dagliga verksamheten i skolan långsiktigt har mycket att vinna på att skolpolitiska reformer beslutas efter en noggrann prövning och i bred politisk enighet. Brist på kunskap leder till att frågor som till sin karaktär i grunden är empiriska i stället ideologiserats. Kunskaper och erfarenheter som vinnas genom försöksverksamhet kan bidra till att i stället lägga en grund för långsiktigt bärkraftiga beslut. Användning av försöksverksamhet är därför ett sätt att underlätta utvärdering av reformer som bör tillvaratas. Regering och riksdag bör, enligt vår uppfattning, i högre grad utnyttja möjligheten att genomföra försöksverksamheter inför reformer.

Regeringen har möjlighet att fatta beslut om försöksverksamhet när det gäller frågor där riksdagen har bemyndigat regeringen att besluta om bestämmelser. I praktiken innebär det att regeringen kan besluta om försök inom en rad viktiga områden, naturligtvis under förutsättning att försöket inte innebär att enskilda elevers

²⁸ SOU 2013:40 *Att tänka nytt för att göra nytta – om perspektivskiften i offentlig verksamhet*.

rättssäkerhet påverkas negativt eller att försöket allvarligt skulle påverka likvärdigheten. En försöksverksamhet i samband med att grundskolans kunskapskrav utformades inför Lgr11 hade exempelvis kunnat komplettera de utprovningar som genomfördes och gett viktiga erfarenheter som i sin tur kunnat påverka den slutliga utformningen av kraven. Självklart hade inte ett sådant försök kunnat genomföras för årskurs 9, eftersom eleverna ska söka till gymnasieskolan på sina betyg från denna årskurs. Man hade i stället fått välja t.ex. årskurs 7 i ett urval skolor eller ett stegvis införande, med möjlighet till justeringar innan betyg skulle sättas för elever i årskurs 9. Ett sådant förfaringssätt hade dock krävt något mer tid för att nå fram till den slutliga utformningen.

För frågor som regleras i skollagen bör enligt vår uppfattning huvudalternativet vara att det är riksdagen som på regeringens förslag ska besluta om en försöksverksamhet. Vi menar att det är fullt möjligt att tänka sig sådana försök.

Vad utredningen uppfattar som problematiskt med nuvarande regelverk gäller dagens bestämmelse i skollagen om att regeringen eller den myndighet som regeringen bestämmer får meddela föreskrifter om försöksverksamhet inom skolväsendet och att regeringen därvid får besluta om undantag från *organisatoriska* bestämmelser i skollagen. Vi konstaterar att det är högst oklart vad bestämmelsen egentligen syftar till. En fråga är vad som avses med organisatoriska bestämmelser, en annan lika viktig fråga är vilken typ av bestämmelser i skollagen som regeringen faktiskt ska få besluta om undantag för genom att inrätta försök.

Utredningen har inte haft möjlighet att närmare utreda de juridiska och etiska förutsättningarna för detta, men anser att det är angeläget att frågan utreds.

9 Systemet för att följa kunskapsresultat – utvecklingsbehov

I detta kapitel fokuserar vi i första hand på hur förutsättningarna för att följa kunskapsutvecklingen i den svenska skolan kan förbättras, dvs. för att följa trender. Trendmått av god kvalitet utgör ett viktigt underlag, bland andra underlag, för att det ska vara möjligt att följa resultatet av skolreformer. Det bör dock understrykas att trendmått av detta slag inte utgör en förutsättning för exempelvis effektutvärdering. För sådan utvärdering ställs i stället andra kvalitativa krav på data som vi återkommer till i kapitel 10.

9.1 Dagens nationella system för kunskapsbedömning

Att ha tillgång till statistik som belyser utvecklingen av skolans kunskapsresultat, dvs. om skolans prestationer förbättras eller försämras mellan olika år, är central vid uppföljningen av skolpolitiska reformer. Det visar inte minst intresset för resultaten på de internationella studierna. Uppföljning av kunskapsresultat i den svenska skolan sker genom betygssystemet, genom de nationella proven och genom deltagandet i internationella studier.

Utredningen har redan i sitt delbetänkande¹ konstaterat att det är tveksamt att använda betyg eller mått som baseras på betyg för att mäta trender. Det nationella provsystemet har visserligen många syften, men är inte heller det utformat på ett sådant sätt att det möjliggör trendmätning. I dag är det därför endast de internationella studierna som har en sådan utformning och kvalitet att de kan utgöra underlag för att följa kunskapsutvecklingen över tid i den svenska skolan, dvs. att följa trender. Vi menar att detta är ett problem

¹ SOU 2013:30 *Det tar tid – om effekter av skolpolitiska reformer*.

eftersom de internationella studierna endast mäter utvecklingen i vissa ämnen och endast täcker vissa av de mål som återfinns i de svenska styrdokumentet.

9.1.1 Nationella prov och betyg – lämpar sig inte som grund för trendmätning

Här redogör vi inledningsvis för varför vi anser att betyg och nationella prov med nuvarande utformning *inte* lämpar sig som underlag för trendmätning.

De nationella proven

I kapitel 4 redogörs för huvuddragen i det nationella provsystemet² – eller med den benämning Skolverket använder, det nationella systemet för kunskapsbedömning – och konstaterar att detta utformats för att tillgodose en rad olika syften³. Systemet ska

- bidra till ökad måluppfyllelse för eleverna,
- förtydliga målen och visa på elevers starka och svaga sidor,
- konkretisera kursmål och betygskriterier,
- stödja en likvärdig och rättvis bedömning och betygssättning och
- ge underlag för en analys av i vilken utsträckning kunskapsmålen nås på skolnivå, på huvudmannanivå och på nationell nivå.

Skolverket poängterar numera, på sin websida och i kontakt med de institutioner som utvecklar de nationella proven, att de båda sista strecksatserna utgör huvudsyftena för just de nationella proven, men att proven även kan bidra till att de övriga syftena uppfylls. För dessa övriga syften hänvisar man samtidigt till systemets övriga bedömningsstödjande inslag.

I trendmätningssammanhang är det den sista punkten som har relevans, dvs. att de nationella proven ska ge underlag för analys på olika nivåer i systemet. Det bör observeras att det inte ursprungligen har angetts som syfte för provsystemet att det ska bidra till trend-

² Utredningen har även redovisat omfattningen av det nationella provsystemet och effekter av olika slags prov i delbetänkandet *Det tar tid – om effekter av skolpolitiska reformer* (SOU 2013:30).

³ U2004/5293/S, U2005/8381/G, U2006/8951/S, U2008/1045/SV och U2008/2645/S.

analys. Regeringen konstaterar också i november 2011 i ett uppdrag⁴ till Skolverket att de nationella proven inte i nuläget är utformade för att mäta kunskapsutvecklingen över tid. Av uppdraget framgår dock att det är något som nu förväntas. Enligt vår bedömning kräver dock en sådan förväntan mer genomgripande förändringar av systemet än vad regeringens uppdrag ger vid handen, vilket vi återkommer till längre fram.

Frågan om trendmätning är statistiskt komplicerad och har belysts mer på djupet av andra författare än vi kan göra här⁵. Några faktorer bör dock lyftas fram.

Ett grundläggande krav på varje prov måste vara att *proven har både god validitet och god reliabilitet*. Det gäller oberoende av om de ska ligga till grund för trendmätning eller inte. Med validitet avses giltigheten i det mått som används, dvs. hur väl måttet mäter det som det avser att mäta⁶. Med reliabilitet avses tillförlitligheten hos mätningen, dvs. att upprepade mätningar ger samma resultat, oberoende av vem som genomför mätningen.

För att därutöver kunna utgöra grund för just trendmätning krävs bl.a.

- att proven är stabila över tid, dvs. att varje nytt prov innehåller ett antal uppgifter som ingått i tidigare prov, s.k. ankaruppgifter,
- att antalet frågor räcker för att få en tillräcklig täckning av det aktuella ämnets innehåll, samt
- att innehållet i kursplaner och utformningen av betygssystemet i huvudsak är stabila över tid, så att det är samma kunskaper och kunskapsnivåer som jämförs.

Dagens nationella prov svarar inte mot samtliga dessa krav, helt enkelt för att de inte i första hand är utformade med trendmätning för ögonen. Flera av faktorerna ovan står dessutom i konflikt med vissa av det nationella provsystemets nuvarande syften. Bland annat

⁴ Skolverket (dnr. 2011/6543/S) *Uppdrag om nationella prov*.

⁵ Se t.ex. Gustafsson, Jan-Eric (2006) *Ramverk för ett system för uppföljning av kunskapsutvecklingen i grundskolan*, Underlag utarbetat på uppdrag av Skolverket 2006-05-06, Skolverkets rapport 286 (2006) *Rullande stickprovsbaserat system för kunskapsutvärdering av grundskolans ämnen* samt Gustafsson, Jan-Eric, Christina Cliffordson och Gudrun Ericson (2014) *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS Förlag.

⁶ Man talar t.ex. om att det föreligger god innehållsvaliditet om ett prov på ett bra sätt återspeglar den kunskap som eleven ska ha tillägnat sig enligt kursplanen.

Skolverket har vid flera tillfällen⁷ framhållit svårigheterna att möta alla dessa syften med ett enskilt instrument på grund av de inbyggda konflikter som finns mellan dessa olika syften.

Ankaruppgifter kan inte vara offentliga

Den första punkten, att systemen ska innehålla ankaruppgifter som kan återanvändas och därmed bidra till att resultaten blir jämförbara över tid, är kanske det viktigaste. Den står dock i konflikt med andra syften för provsystemet.

Skolverket pekar t.ex. i en rapport till regeringen⁸ på att det finns en konflikt mellan den öppna publicering av provuppgifterna som krävs för att proven ska kunna fungera förebildligt för lärarna och det hemlighållande av provuppgifterna som krävs för att kunna mäta trender. Verket menar att ett bedömningssystem som präglas av stor öppenhet uppmuntrar till dialog och reflektion i undervisningssituationen mellan elever samt mellan lärare och elever. Man menar att fler öppet tillgängliga delar i bedömningssystemet på ett bättre sätt än i dag skulle bidra till att föräldrarna blir delaktiga i sina barns lärande.

Samtidigt måste enligt Skolverket delar av systemet av kostnads-skal vara behörighetsskyddat, dvs. ha begränsad tillgänglighet, för att kunna återanvändas. Till de delar av systemet som därför måste vara behörighetsskyddat hör ”de som har det uttalade syftet att vara summerande eller betygsstödjande, dvs. de nationella proven”. Konsekvensen av detta är att de nationella proven inte kan fylla den förebildliga och konkretiserande roll i förhållande till lärare och elever som sedan länge varit ett av syftena med provsystemet.

De nationella proven har för få frågor

Det svenska provsystemet kännetecknas av att varje prov görs av samtliga elever i den aktuella årskullen. Det kan låta motsägelsefullt att dessa prov inte kan utgöra underlag för trendmätningar, utan att stickprovsundersökningar i stället är att föredra.

För att det ska vara möjligt att använda proven för trendmätning ställs emellertid särskilt stora krav på att proven vid varje tillfälle

⁷ Skolverket 2009-04-27 (dnr. 2009:337) och 2012-10-19 (dnr. 2011:1242).

⁸ Skolverket 2009-04-27 (dnr. 2009:337).

har en god täckning av målen i det aktuella ämnet. Annars riskerar man att resultaten på proven varierar mellan olika år, inte på grund av att elevernas kunskaper varierar, utan på grund av att proven haft olika fokus olika år. I praktiken visar sig detta redan i dag, genom att resultaten på nationella prov varierar mellan olika år, väsentligt mer än betygsresultaten. För att det ska vara möjligt att genomföra trendmätningar krävs därför att proven omfattar ett stort antal frågor som täcker väsentliga delar av kursplanernas centrala innehåll och kunskapskvaliteter. Det handlar om avsevärt fler frågor än en och samma elev kan besvara på ett prov.

Många länder, liksom de internationella studierna, har löst detta genom att utforma sina nationella prov som randomiserade stickprovsundersökningar, där olika elever löser olika uppgifter samtidigt som den samlade bilden blir tillförlitlig. En studie kan då totalt sett omfatta många uppgifter, medan varje elev bara behöver besvara ett mindre antal. I ett sådant stickprovsbaserat system kan sedan resultatet, genom s.k. matrissampling⁹, vägas samman till ett statistiskt representativt mått som kan användas för att följa trender. I ett nationellt provsystem som det svenska, utformat utifrån andra syften, är detta inte möjligt.

Styrdokumentet och bedömningssystemet är inte stabila över tid

Om man ska mäta trender är det vidare betydelsefullt att objektet för mätningen är stabilt, dvs. ”*när man vill mäta förändring, förändra inte måttstocken*”.¹⁰ I Sverige är betyg ett mått på hur väl eleverna nått målen i styrdokumentet. Förändras måttet, eller det som måtten refererar till, bryts jämförbarheten över tid. Det är således en fördel ur ett utvärderingsperspektiv om styrdokumentet, dvs. läroplaner och kursplaner, samt betygssystemet är stabila över tid. Detta är dock förhållanden som inte kan påverkas inom ramen för hur provsystemet utformas.

⁹ Uppgifterna i testet upp i olika versioner av testformuläret på ett sådant sätt att olika elever besvarar delvis olika uppgifter samtidigt som det görs en viss överlappning mellan de uppgifter som erbjuds olika elever. Balanced Incomplete Block (BIB) är en design för detta som beskrivs i rapporten *Ramverk för ett system för uppföljning av kunskapsutvecklingen i grundskolan* (2006) av Jan-Eric Gustafsson.

¹⁰ ”When measuring change, do not change the measure”. Påpekandet formulerades av Beaton och Zwick i artikeln *Disentangling the NAEP 1985-86 reading anomaly* (NAEP Rep. No. 17-TR-21) Princeton NJ: Educational testing service. 1990. Här citerad efter Jan-Eric Gustafsson (2006) *Ramverk för ett system för uppföljning av kunskapsutvecklingen i grundskolan*.

Betyg

Som nämnts ovan är på nationell nivå de genomsnittliga betygsresultaten mer stabila över tid än de genomsnittliga resultaten från de nationella proven. Hur stora variationerna är varierar för olika ämnen, men förhållandet indikerar att de genomsnittliga betygsresultaten sannolikt är mer rättvisande än resultaten från de nationella proven. Orsaken till detta kan bl.a. vara att prov vid ett enstaka tillfälle mäter en mer begränsad del av ett ämnes kunskapsinnehåll än den bedömning som läraren gör. Lärarens bedömning fångar elevens kunskaper bredare och tar dessutom hänsyn till elevens prestationer över tid.

Detta innebär dock inte att betygen fungerar tillräckligt bra för att mäta trender. Ett problem är den betygsinflation som forskare pekar på¹¹. Även Skolverket konstaterade våren 2012 att det empiriska underlaget ger tydliga indikationer på att betygsinflation har förekommit i både grundskolan och gymnasieskolan sedan slutbetyg började utfärdas enligt det målrelaterade betygssystemet i slutet av 1990-talet. I grundskolan var den starkast under de första åren efter 1998, för att sedan minska över tid. Minst var betygsinflationen i ämnen där det förekommer nationella prov.¹²

Det bör observeras att denna analys avser perioden innan den nya betygsskalan infördes hösten 2011. Erfarenheten från 1990-talet pekar på att det nu finns en risk för att lärare initialt är försiktiga med sin betygssättning, för att sedan successivt utnyttja hela betygsskalan. Detta kan nu återigen komma att leda till en kraftigare betygsinflation. Det kan dock även finnas andra orsaker till betygsinflation, exempelvis konkurrensen om elever till följd av det fria skolvalet och allt starkare påtryckningar från elever och föräldrar.

Utredningen bedömer sammantaget att det faktum att betygssystemet inte är stabilt över tid, samt att det av olika orsaker finns risk för betygsinflation, innebär att betyg inte lämpar sig som underlag för att studera trender. Betygsresultat kan dock användas för andra analyser, t.ex. prognosticerar de bättre för kommande studieframgångar än de nationella proven.

Som utredningen pekat på redan i delbetänkandet är det för närvarande särskilt olämpligt att använda betygsresultat för att följa

¹¹ Gustafsson, Jan-Eric, Christina Cliffordson och Gudrun Ericson (2014) *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS Förlag.

¹² Skolverket (2012) *Betygsinflation – betygen och den faktiska kunskapsutvecklingen* (dnr. 2012:387).

trender. Anledningen till detta är att betygssättningen påverkas av de reformer som nyligen genomförts och som gör att såväl kursplaner och kunskapskrav som betygsskalan är nya. Erfarenheten visar att det tar tid innan nya bedömningssystem sätter sig och innan lärare använder hela betygsskalan på det sätt som är avsett.

Sammantaget

Varken de genomsnittliga betygsresultaten eller resultaten från de nationella proven lämpar sig för att ligga till grund för trendmätning. Betygen har visserligen en lägre volatilitet än provresultaten, men är ändå behäftade med problem vad gäller likvärdigheten mellan olika grupper och skolor samt med betygsinflation.

Flera av de syften som är knutna till de nationella proven gör att dessa genomförs av samtliga elever i årskullen. Ett sådant prov, utformat lika för samtliga elever, kan inte omfatta ett tillräckligt stort antal frågor för att ge tillräckligt god validitet för trendmätning. De nationella proven är helt enkelt utformade för andra syften än att ligga till grund för trendmätning och har därför ett begränsat värde i samband med utvärdering av reformer.

Det bör samtidigt understrykas att både betyg och nationella prov har andra viktiga funktioner vid olika former av utvärdering, t.ex. effektutvärdering. Både betyg och nationella provresultat kan t.ex. ligga till grund för olika former av tvärsnittsanalyser och för longitudinella studier (se vidare kapitel 10). Betygen är också den bästa indikatorn vi har på elevers framtida yrkesframgång. Det betyg och nationella provresultat inte duger till är att uttala sig om hur kunskapsnivån utvecklas på nationell nivå över tid.

9.1.2 Internationella studier – de enda instrument som i dag kan mäta kunskapsresultat över tid

Utredningen har i kapitel 4 redovisat Sveriges deltagande i olika internationella studier. Som ett resultat av en ständigt pågående metodutveckling bygger de internationella studierna i dag på en avancerad statistisk modellering. De internationella studierna har ett gott anseende bland experter, bl.a. beroende på att proven hela tiden debatteras och till följd av det utvecklas.

Debatten handlar bland annat om vad de mäter och inte mäter, hur resultaten presenteras och hur de kan förstås. En vanlig invändning är att proven är konstruerade utifrån ett ramverk som ”definierar huvudstrukturen av kunskaper och färdigheter inom ett område, anger vilka typer av uppgifter som skall användas, hur stor andel av det totala antalet uppgifter en viss uppgiftstyp skall utgöra osv.”¹³ Detta ramverk överensstämmer mer eller mindre väl med de deltagande ländernas nationella läro- och kursplaner.¹⁴ Skolverket har dock analyserat överensstämmelsen mellan de internationella studierna och de svenska kursplanerna och funnit att den överlag är god.¹⁵ Där överensstämmelsen inte är perfekt kompletterar dessutom PISA och TIMSS varandra.¹⁶

Ett problem med de internationella studierna är att de omfattar ett begränsat antal ämnen, fokuserar ett begränsat kunskapsinnehåll och genomförs med olika periodicitet och i olika årskurser under grundskoletiden. PIRLS fokuserar på läsförståelse i årskurs 4 och genomförs vart femte år. TIMSS fokuserar matematik och naturvetenskap i årskurs 4 och årskurs 8 och genomförs vart fjärde år. PISA mäter 15-åringars kunskaper (årskurs 9) i läsning, matematik och naturvetenskap och genomförs vart tredje år.

Det finns dessutom risk för både missförstånd och övertolkningar eftersom de internationella studierna i dag bygger på så avancerade statistiska metoder att de inte ens behärskas av alla forskare och att resultaten presenteras i former som kan vara svåra att tillgodogöra sig även för dem som arbetar inom utbildningsområdet.¹⁷

Trots att invändningar av olika slag alltså har rests mot de internationella studierna drar utredningen ändå slutsatsen att de i dag utgör den enda källan för bedömning av hur den svenska skolans kunskapsresultat utvecklas över tid. De ger dock en bild av svenska elevers kunskaper som är begränsad till vissa ämnen och förmågor, vilket aktualiserar frågan om hur ett framtida nationellt system bör utformas.

¹³ Gustafsson, Jan-Eric (2006) a.a.

¹⁴ Skolverket (2004) *Internationella studier under 40 år. Svenska resultat och erfarenheter*.

¹⁵ Skolverket (2006b). Med fokus på läsförståelse – En analys av skillnader och likheter mellan internationella jämförande studier och nationella kursplaner. Skolverkets aktuella analyser 2006. Stockholm: Skolverket och Skolverket (2006c). Med fokus på matematik och naturvetenskap – En analys av skillnader och likheter mellan internationella jämförande studier och nationella kursplaner. Skolverkets aktuella analyser 2006. Stockholm: Skolverket.

¹⁶ Gustafsson, Jan-Eric, Christina Cliffordson och Gudrun Ericson (2014) *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS Förlag.

¹⁷ Gustafsson, Jan-Eric (2006) a.a.

9.2 Användningen av nationella prov i andra länder

Utredningen har besökt Danmark, Finland, Norge och Skottland och diskuterat hur man i dessa länder arbetar med att samla in och bearbeta information om elevers kunskapsresultat som ett led i ländernas utvärdering av skolans resultat. Samtliga dessa länder deltar i någon eller i flera av de internationella studierna, vilket ger viss information om trender. Flera länder har också examensprov, vilka dock knappast ger underlag för att följa kunskapsutvecklingen över tid. Liksom de nationella proven i Sverige kan examensproven inte täcka ett ämnes samtliga mål. Vi har därför uppehållit oss särskilt vid de olika ländernas nationella prov eller tester, vilka kan utgöra komplement till de internationella studierna.

En skillnad mellan länderna gäller om de nationella proven eller testerna liksom i Sverige omfattar samtliga elever i berörda årskurser, eller om testerna baseras på stickprov. I Finland och Skottland baseras de på stickprov. Orsaken är att syftet med dessa prov i första hand är att ge en nationell bild av kunskapsläget. I Finland återrapporteras resultaten från proven till kommunen och skolan, men endast på skolnivå och inte för enskilda elever. Varje skola får en sådan återkoppling om läget i förhållande till riket minst vart femte år. I Skottland görs ingen sådan återkoppling till den enskilda deltagande skolan.

I Danmark och Norge genomförs de nationella proven eller testerna för samtliga elever i berörd årskurs. I båda länderna är det primära syftet med proven att fungera som ett pedagogiskt verktyg för läraren samt att vara en utgångspunkt för skolans kvalitetsarbete. I Danmark sammanställs dock en nationell prestationsprofil som publiceras en gång varje år. 2013 redogörs för utvecklingen från året innan. I Norge publiceras resultat från de nationella proven på nationell nivå, men man understryker samtidigt att proven inte är avsedda för att beskriva utvecklingen över tid. Trendmätande prov har dock nyligen introducerats i matematik och engelska.

Gemensamt för länderna är att proven genomförs för ett strategiskt urval av ämnen samt att provuppgifterna inte publiceras och därmed kan återanvändas. Gemensamt är också att proven inte är av "high-stakeskaraktär", dvs. resultaten har ingen omedelbar betydelse för deltagande skolor, lärare och elever vad gäller t.ex. resursfördelning eller betyg. I flera av länderna finns dock, som nämnts ovan, examensprov som är av "high-stakeskaraktär".

9.3 Tidigare förslag om rullande stickprovsbaserad utvärdering av elevernas kunskapsresultat

Skolverket har redan 2006 i en rapport med anledning av ett regeringsuppdrag redovisat ett förslag till ett framtida nationellt stickprovsbaserat system för kunskapsutvärdering av grundskolans ämnen.¹⁸ Jan-Eric Gustafssons rapport *Ramverk för ett system för uppföljning av kunskapsutvecklingen i grundskolan*¹⁹ utgjorde ett underlag för denna rapport från Skolverket.

Gustafssons rapport är den grundligaste genomgång av förutsättningarna för ett system för mätning av kunskapsutvecklingen över tid (trendmätning) som gjorts i Sverige. Gustafsson redogör bland annat för hur extern kunskapsbedömning och mätning, dvs. bedömning som initieras och genomförs av andra än av den undervisande läraren, vuxit fram med USA som föregångsland och hur de landvinningar som gjorts inom statistisk modellering och psykometri utnyttjats vid utformningen av den återkommande utvärderingen av amerikanska skolbarns kunskaper (National Assessment of Educational Progress, NAEP) och därifrån förts vidare till de olika internationella studier som utvecklats av IEA (TIMSS och PIRLS) och OECD (PISA).

En av Gustafssons slutsatser är att det tycks vara nödvändigt att använda sig av flera olika utvärderingsformer samtidigt vid kunskapsbedömning. För att mäta trender föreslår Gustafsson att ett stickprovsbaserat system för kunskapsbedömning på nationell nivå införs. Stickprovsbaserade prov har, enligt Gustafsson, flera fördelar jämfört med prov där samtliga elever i en årskurs deltar: de kan utformas så att de har hög validitet i förhållande till läroplanens samlade kunskapsmål utan att de samtidigt blir alltför omfattande för varje enskild elev, de minskar risken för allmän provtrötthet och de eliminerar risken för att lärarna börjar anpassa sin undervisning till proven (teach-to-the-test).

I rapporten *Rullande stickprovsbaserat system för kunskapsutvärdering av grundskolans ämnen*²⁰ menade Skolverket att det förslag som Gustafsson lämnat skulle bli alltför kostsamt för ett litet land som Sverige, både i uppbyggnadsskedet och i en tänkt mer regel-

¹⁸ Skolverket (2006) *Rullande stickprovsbaserat system för kunskapsutvärdering i grundskolans ämnen*. Rapport 286.

¹⁹ Denna rapport är publicerad som appendix till Skolverkets rapport *Ett rullande stickprovsbaserat system för kunskapsutvärdering av grundskolans ämnen*. Rapport 286. Skolverket 2006.

²⁰ Skolverket (2006). Rapport 286. Rapporten utgjorde en bilaga till Skolverkets redovisning av ett regeringsuppdrag avseende resultatinformation givet 2004-12-22.

bunden driftssituation. Skolverket föreslog i stället att de internationella kunskapsmätningarna ska integreras i en nationell strategi och helt eller delvis ersätta nationella mätningar inom de kunskaps- eller ämnesområden de mäter. I de fall en internationell studie står i rimlig överensstämmelse med intentioner och ambitioner i de nationella måldokumentet, skulle således de internationella studierna ersätta nationella mätningar och nationella stickprovsundersökningar genomföras endast som komplement till de internationella studierna. Skolverket ansåg att sådana nationella undersökningar borde göras i ett fåtal strategiskt valda ämnen där det nationella kunskapsbehovet inte täcks av internationella studier eller av andra skäl bedöms som särskilt angeläget. En förutsättning för Skolverkets förslag är att Sverige fattar ett långsiktigt beslut om att delta i de internationella studierna och inte som nu förutsättningslöst prövar deltagandet inför varje enskilt tillfälle. Regeringen har inte reagerat på Skolverkets förslag, vilket innebär att något särskilt uppdrag att genomföra denna typ av stickprovsundersökningar inte har lämnats.

Det bör i sammanhanget uppmärksammas att det är möjligt för ett land att i samband med en internationell studie utöka urvalet av elever för att möjliggöra mer fördjupade analyser av resultaten, t.ex. för olika elevgrupper eller olika kommuntyper etc. Det är också möjligt att lägga till egna frågor utöver dem som ingår i den internationella undersökningen för att få en mer nationellt anpassad belysning av ämnet ifråga. I båda fallen är detta förknippat med kostnader. Att tillföra extra frågor kan dessutom påverka elevernas motivation att lösa uppgifterna i negativ riktning. Att proven inte redovisas tillbaka till skolan eller eleverna riskerar att påverka både elevernas och lärarnas engagemang för proven negativt.

9.4 Förslag till utveckling av systemet för att mäta trender

Utredningens förslag: Ett system införs med stickprovsbaserade undersökningar för trendmätning, som kompletterar de internationella studierna. Ett sådant system skulle förbättra möjligheterna att följa kunskapsutvecklingen i den svenska skolan. Systemet bör ha en bredd med avseende på både de ämnen och de kunskapskvaliteter som mäts.

En utredning föreslås med uppdrag att i ett sammanhang utforma ett system för trendmätning genom stickprovsundersökningar och att utreda det framtida systemet för utvärdering av kunskapsresultat i sin helhet. Arbetet bör i den senare delen inriktas på vilka syften som kan och ska knytas till systemet, vilka instrument som bör knytas till respektive syfte och vilka aktörers informationsbehov de olika instrumenten ska svara mot.

I det föregående har olika aspekter av det nuvarande systemet för kunskapsuppföljning beskrivits. Av genomgången ovan framgår att dagens svenska system för kunskapsuppföljning inte förmår att tillhandahålla den slags information som är nödvändig för att följa kunskapsutvecklingen i skolan över tid. Varken betygssystemet eller det nationella provsystemet lämpar sig för trendmätning. Det bör understrykas att det inte är nödvändigt med den typ av absolut kunskapsmätning som trendstudier representerar för att det ska vara möjligt att utvärdera reformer, genom exempelvis effektstudier. Utredningen bedömer dock att det behövs mått av god kvalitet, som belyser hur skolan utvecklas över tid. Systemet bör ha en bredd med avseende på både de ämnen och de kunskapskvaliteter som mäts.

Till det svenska nationella bedömningssystemet har under årens lopp knutits allt fler mer eller mindre förenliga syften, ett förhållande som Sverige inte är ensamt om. Som exempel kan nämnas att Newton²¹ för Englands vidkommande har identifierat inte mindre än arton olika syften kopplade till bedömning (assessment)²². I likhet med Skolverket²³ konstaterar Newton att det är nödvändigt att hålla isär olika typer av syften för bedömning. Eftersom de har uppenbara återverkningar på utformningen av ett system för kun-

²¹ Paul E. Newton (2002) *Clarifying the purpose of educational assessment*. Assessment in education, Vol. 14, No. 2. Newton var vid tidpunkten för artikelns publicering chef för Assessment research, regulation and Standards Division vid Qualifications and Curriculum Authority i England. Han är i dag professor vid utbildningsvetenskapliga institutionen vid University of London.

²² De 18 olika syftena är: social utvärdering, formativ, elevuppföljning, uppflyttning, placering, diagnostik, vägledning, utfärdande av en kvalifikation, urval, licensiering, val av skola, institutionell granskning, fördelning av resurser, grund för intervention i en organisation, utvärdering av en åtgärd, uppföljning på systemnivå, jämförelse och nationellt ansvarsutkrävande.

²³ Se t.ex. Skolverket 2009-04-27(dnr. 2009:337) med anledning av regeringsuppdrag givet i regleringsbrev för 2009.

skapsbedömning behöver vart och ett av syftena enligt Newton tas om hand för sig.

Kraven på att det nationella systemet för kunskapsbedömning ska kunna beskriva kunskapsutvecklingen över tid har stärkts under hand. Som redan nämnts gav t.ex. regeringen 2011 Skolverket i uppdrag²⁴ att kvalitetssäkra ämnes- och kursproven så att de bättre än i dag kan användas för att bedöma kunskapsutvecklingen över tid. Skolverket konstaterade dock i sin redovisning²⁵ att de nuvarande nationella proven är utformade för att stödja en likvärdig och rättvis bedömning och betygssättning, och att det inte är rimligt att använda dem för att också göra årliga trendanalyser av olika kunskapsområden.

För att kunna utgöra grund för trendmätning krävs enligt Skolverket att proven är stabila över tid, dvs. att varje nytt prov innehåller ett antal uppgifter som ingått i tidigare prov. Vidare krävs att innehållet i kursplaner och utformningen av betygssystemet inte förändras. Vi har dessutom i det ovanstående visat att det, för att få en god täckning av ett ämne, även krävs att varje undersökning omfattar ett så stort antal frågor att det inte är rimligt att förvänta sig att en elev vid ett och samma tillfälle ska kunna besvara dem.

Skolverket förordar att de nationella provens nuvarande syften bibehålls och inte utökas till att också omfatta trendmätning. Om kunskapsutvecklingen inom olika kunskapsområden med stor säkerhet ska kunna följas över tid är det enligt verket mer motiverat att undersöka möjligheten att ta fram ett särskilt mätinstrument för just det ändamålet.

Utredningen konstaterar att det nuvarande systemet för kunskapsbedömning har vuxit fram genom successiva förändringar under mer än tjugo års tid. Detta har inte bara medfört att förväntningarna på vad systemet ska bidra till har blivit allt fler utan att det också skapats spänningar mellan de olika förväntningarna. Dessutom har den successiva framväxten lett till att prov i olika ämnen är utformade på olika sätt, vilket försvårar för de lärare som undervisar i flera ämnen för vilka det genomförs prov.

Utredningen delar uppfattningen att det inte är rimligt att förvänta sig att den typ av nationellt provsystem som vi har i Sverige ska kunna användas för att följa trender och föreslår därför att ett system med stickprovsbaserade undersökningar för trendmätning

²⁴ U2011/6543/S.

²⁵ Skolverket 2012-10-19 (dnr. 2011:1242).

som kompletterar de internationella studierna ska införas för detta syfte. Ett sådant system skulle förbättra möjligheterna att följa kunskapsutvecklingen i den svenska skolan och därmed även förbättra förutsättningarna för att utvärdera reformer och andra initiativ på skolans område. Det bör understrykas att behovet att komplettera de internationella undersökningarna inte endast avser andra ämnen och mål i läroplanen än de som ingår i dessa. Även för de ämnen för vilka det genomförs internationella mätningar kan det finnas behov av komplettering.

Utöver de fördelar med ett stickprovsbaserat system som redogjorts för ovan – hög validitet i förhållande till läroplanens samlade kunskapsmål, få frågor per elev, låg risk för provtrötthet, minskad risk för "teaching to the test" – ger ett stickprovsbaserat system även goda möjligheter att analysera variation, t.ex. mellan elever med olika kön eller socioekonomisk bakgrund. Det bör också kunna utformas för att möjliggöra kvalificerade analyser av vilka inslag i olika ämnen som svenska elever lyckas väl med och vilka de lyckas mindre väl med, dvs. kunna ligga till grund för olika nationella åtgärdsprogram.

Erfarenheten från de internationella studierna visar dock att det kan finnas problem med stickprovsundersökningar, som bl.a. har att göra med motivationen, då proven inte har någon avgörande betydelse för eleven eller skolan. Detta bör dock gå att hantera, t.ex. genom att anpassa antalet elever i urvalet så att varje elev inte behöver besvara mer än ett begränsat antal frågor.

Vi bedömer emellertid att det inte är lämpligt att nu endast lägga till ytterligare ett instrument i form av stickprovsundersökningar till ett redan hårt belastat system utan att samtidigt utreda konsekvenserna av detta för övriga delar av systemet. Det bör få konsekvenser för övriga delar av det nationella systemet för kunskapsbedömning om behovet av trendmätning tillgodoses via ett stickprovsbaserat system. Det gäller framför allt för de nationella proven som i dag bl.a. har till syfte att bidra med underlag för en analys av i vilken utsträckning kunskapsmålen nås på nationell nivå.

Den successiva framväxten av systemet för kunskapsbedömning har enligt vår uppfattning nu lett fram till en punkt där systemets olika syften behöver övervägas på nytt. Hela systemet behöver enligt vår uppfattning utredas från grunden för att klargöra dels vilka syften som kan och ska knytas till systemet, dels vilka instrument som bör knytas till respektive syfte och dels vilka aktörers

informationsbehov de olika instrumenten ska svara mot. Det bör därvid observeras att även betyg och resultat från nationella prov utgör viktiga underlag för olika former av utvärdering, t.ex. effektutvärdering. Vi föreslår att en utredning tillsätts med uppdrag att utreda det framtida systemet för kunskapsbedömning med avseende på dessa faktorer. En ambition bör vara att de nationella prov som ska finnas i framtiden utformas på ett mer likartat sätt för olika ämnen. Frågan om att utforma stickprovsundersökningar för att kunna följa kunskapsutvecklingen i den svenska skolan föreslås ingå i uppdraget.

Liknande rekommendationer lämnas i dagarna i en SNS-rapport²⁶. I denna rekommenderas bl.a. en försöksverksamhet med ett rullande, stickprovsbaserat, nationellt kunskapsbedömningssystem för grund- och gymnasieskolan inom ett par ämnen. Vidare redovisas tankar om hur ett sådant nationellt kunskapsbedömningssystem kan kopplas samman med det nationella provsystemet, genom en gemensam uppgiftsbank, i syfte att de nationella proven bättre ska kunna stödja en likvärdig och över tid stabil betygssättning. Detta innebär dock inte att författarna menar att det går att konstruera ett enskilt instrument som svarar mot samtliga olika syften. Utredningen bedömer att rapportens bedömningar kan bidra med värdefulla utgångspunkter för den utredning vi här föreslår.

En sådan utredning bör slutligen ses i perspektivet av vad vi redovisat i föregående avsnitt om behovet av ett utvärderingsramverk. Systemet för kunskapsbedömning utgör en central del av ett sådant ramverk. Förändringar bör göras för att systemets olika delar ska kunna bidra optimalt till helheten samtidigt som varje del kan utformas på ett stabilt och funktionellt sätt.

²⁶ Gustafsson, Jan-Eric, Christina Cliffordson och Gudrun Ericson (2014) *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS Förlag.

10 Data för utvärdering – utvecklingsbehov

Utredningen ska enligt direktiven utreda vilka data som behövs för att i ökad utsträckning kunna utvärdera reformer inom skolan samt, om det behövs, lämna författnings- eller andra förslag för att öka tillgången till sådana data. I det föregående kapitlet fördjupade vi oss specifikt i frågan om hur utvecklingen av kunskapsresultaten i skolan kan följas över tid, dvs. hur trender kan mätas. I detta kapitel redovisar vi huvuddragen i det nuvarande statistiksystemet samt översiktligt den omfattande lagstiftning som finns inom området. Mot denna bakgrund analyseras dels vilka ytterligare data, förutom de som beskriver kunskapsutvecklingen över tid, som behövs för att förbättra möjligheterna för utvärdering av skolpolitiska reformer, dels förutsättningarna för att samla in dessa data.

10.1 Den nationella statistiken om skolan – en infrastruktur för systemförståelse

Liksom för andra samhällsområdet utgör statistiksystemet för skolan en form av infrastruktur för förståelse av skolan och skolans utveckling. Begreppet statistik definieras inte i något regelverk. Av statistiklagen¹ framgår dock att officiell statistik (SOS) ska finnas för allmän information, utredningsverksamhet och forskning.

10.1.1 Kort om skolstatistikens framväxt

Den enda utbildningsstatistik som fanns i Sverige under 1900-talets första årtionden var den statistik som Skolöverstyrelsen (SÖ) förde över folkskolor och läroverk. Ännu så sent som 1970 omfattade

¹ Lag (2001:99) om den officiella statistiken.

grundskolestatistiken endast uppgifter om antalet elever och deras ålder.²

1968 inrättade Statistiska Centralbyrån (SCB) i samarbete med Skolöverstyrelsen ett lärarregister. Registret innehåller uppgifter om alla lärare, även vikarier, som är i tjänst hela eller delar av en viss vecka på hösten. Uppgifterna samlas in från alla skolor, både grundskolan och de frivilliga skolformerna. Lärarregistret ingår numera, liksom elevregistret, i Skolverkets nationella uppföljningssystem och benämns Registret över pedagogisk personal. Registret innehåller identifikationsuppgifter och uppgifter bl.a. om befattning, utbildning och tjänstgöringstidens omfattning.³

1978 gav SCB för första gången ut *Utbildningsstatistisk årsbok*. Denna publikation har därefter utgetts årligen.⁴

1988 började SCB publicera statistik om slutbetyg för grundskolan. Inledningsvis redovisades endast rikssiffror. Så småningom redovisades uppgifterna även på kommunnivå, vilket bidrog till att intresset för statistiken ökade.⁵

Fram till och med budgetåret 1991/92 bestod statistiken över utbildningskostnader endast av en sammanställning av statsbudgetens kostnader. Därefter har kostnadsuppgifter samlats in från kommunerna och de fristående huvudmännen för olika skol- och verksamhetsformer.

Från och med 1998 överfördes ansvaret för statistiken för förskoleverksamheten och skolbarnsomsorgen från Socialstyrelsen till Skolverket, till följd av att bestämmelserna om dessa verksamheter fördes från socialtjänstlagen till skollagen.

Särskilda datainsamlingar utgörs av de longitudinella uppgiftsinsamlingarna som genomförs av SCB respektive projektet UGU vid Göteborgs universitet. SCB svarar för de s.k. *elevpanelerna* som ingår i den officiella statistiken, men också utgör underlag för UGU-projektet. Det senare genomför dessutom ytterligare datainsamling. Elevpanelerna utgör underlag för att beskriva elevströmmar inom skolväsendet. Alltsedan 1980-talet har SCB samlat in och redovisat uppgifter som beskriver elevers vägar genom utbildningssystemet för olika kohorter. SCB ansvarar för urvalsdragningar, driften av panelerna, sammanställning av bakgrundsdata och insamling av elevdata från skolorna medan Pedagogiska institutionen

² Jorner, Ulf (2008) *Summa summarum. SCB:s första 150 år*. Statistiska centralbyrån.

³ Jorner, Ulf (2008) a.a.

⁴ Jorner, Ulf (2008) a.a.

⁵ Jorner, Ulf (2008) a.a.

bearbetar och analyserar uppgifter som samlas in direkt från elever, föräldrar och lärare.⁶

Det framgår att ansvarig myndighet för statistiken om skolan har varierat över tid. Fram till 1960 ansvarade SÖ för skolstatistiken. Från detta år tog SCB över ansvaret.⁷ I samband med en senare statistikreform⁸ blev 1994 Skolverket ansvarig för den officiella statistiken om skoläendet.⁹ SCB kvarstår dock som den myndighet som på Skolverkets uppdrag svarar för själva datainsamlingen och för de s.k. elevpanelerna.

10.1.2 Några grundläggande begrepp

Statistikutredningen 2012¹⁰ konstaterar att begreppet statistik inte kan sägas ha någon helt entydig definition. Det kan enligt statistikutredningen avse en viss form av information, nämligen sifferuppgifter som beskriver en företeelse eller en verksamhet. Statistik kan också avse metoder för att samla in, bearbeta och analysera statistiska material. Statistik är därtill benämningen på en vetenskaplig disciplin. Statistikutredningen 2012 använder begreppet som beteckning för *”numeriska sammanställningar av elementära observationer som kan hänföras till händelser, flöden eller tillstånd”*.

Statistikutredningen 2012 konstaterar vidare att uppbyggnaden av ett statistiskt system kräver tre olika typer av data:

Makrodata: Det som vi kallar ”statistik” i dagligt tal, dvs. aggregerade data om populationer och klasser av objekt, t.ex. summor, frekvenser och medelvärden.

Mikrodata: Data om enskilda objekt. t.ex. företag, individer och händelser, som ligger till grund för statistik, och som kan aggregeras till statistik på ett flexibelt sätt.

Metadata: ”Data om data” eller information om data, dvs. data som definierar och beskriver mikro- och makrodata. Det kan röra

⁶ Att elevpanelerna inte ingår i Skolverkets statistikansvar för den officiella statistiken har sannolikt att göra med att undersökningen är longitudinell och redan pågick när Skolverket bildades. Ett annat skäl kan vara att elevpanelerna innehöll tidsseriedata med personnummer och bedömdes passa bättre inom SCB:s ansvar då Skolverket vid sin tillkomst saknade möjlighet att hantera hopkopplingsbara registerdata.

⁷ Murray, Mac (2011) *Projektet som utredning- och beslutsunderlag* i Allan Svensson (red.) *Utvärdering Genom Uppföljning. Longitudinell individforskning under ett halvsekel*. Gothenburg Studies in Educational Sciences 305. Göteborgs universitet.

⁸ Statistikreformen berörde inte endast skolstatistiken, utan ett stort antal samhällssektorer.

⁹ Förordningen (2001:100) om den officiella statistiken.

¹⁰ SOU 2012:36 *Vad är officiell statistik? En översyn av statistiksystemet och SCB*.

sig om beskrivningar av innehållet i och/eller strukturen för en viss datainsamling.

Med utgångspunkt i denna terminologi handlar detta kapitel i huvudsak om tillgången på mikrodata för olika typer av bearbetning och om möjligheterna att öka tillgången på sådana data.

Distinktionen mellan *registerdata* och *direktinsamlade data* är central för de resonemang som förs i detta kapitel. Med registerdata menar vi data som samlas in via befintliga lokala administrativa system, som i flertalet fall upprättats för andra ändamål än nationella uppföljningsbehov¹¹. Med direktinsamlade data menar vi data som samlas in genom olika enkäter, t.ex. till huvudmän och skolor.

Utvärdering kan slutligen genomföras på många olika sätt och varje utvärderingsdesign har sina särskilda behov av data och krav på hur dessa ska vara tillgängliga. *Longitudinella studier* innebär exempelvis att man följer ett antal individer över tid och gör upprepade mätningar för samma individ. *Tvärsnittsstudier*, där jämförelser görs för en eller flera variabler vid en viss tidpunkt, kräver oftast att det är möjligt att samköra register från olika samhällsområden. I många fall handlar det även här om uppgifter på individnivå.

10.1.3 Skolverkets statistikansvar

Regeringen beslutar om vilka myndigheter som ska vara statistikansvariga för Sveriges officiella statistik (SOS) samt om vilka ämnes- och statistikområden som statistikansvaret ska omfatta. Vilken myndighet som är statistikansvarig inom olika områden framgår av bilagan till statistikförordningen. Den statistikansvariga myndigheten beslutar därefter om statistikens innehåll och omfattning inom sitt statistikområde.¹²

Enligt förordningen om den officiella statistiken är Skolverket ansvarig myndighet för den officiella statistiken inom skolväsendet och barnomsorgen¹³. Vidare har verket enligt skollagen¹⁴ och sin instruktion¹⁵ ansvar för att följa upp och utvärdera skolväsendet samt vissa övriga utbildningar och annan pedagogisk verksamhet.

¹¹ Enligt SOU 2012:36 baseras den officiella statistiken till mer än 95 procent på administrativa data, vilka ursprungligen har samlats in för andra ändamål än statistik.

¹² Förordningen (2001:100) om den officiella statistiken.

¹³ Barnomsorgsbegreppet är borttaget i skollagen (2010:800), men någon följändring i förordningen har ännu inte gjorts.

¹⁴ SFS 2010:800.

¹⁵ Förordning (2011:555) med instruktion för Statens skolverk.

Enligt verkets instruktion ska uppföljningen och utvärderingen öka kunskapen om hur utbildningarna och verksamheterna har utvecklats i förhållande till de nationella målen för utbildningen. Skolverket ges stor frihet att avgöra hur man bäst går till väga för att öka kunskapen om skolan, ett friutrymme som i praktiken begränsas av det stora antal uppdrag som ges i de årliga regleringsbreven och i form av regeringsuppdrag som ges i särskild ordning.

Skolverkets arbete med uppföljningssystemet är inriktat på förvaltning och vidareutveckling av de stora informationssystem som ska ge faktabaserat underlag för diskussion och beslut om förskola, fritidshem, skola och vuxenutbildning. Man kan säga att statistiksystemet erbjuder en infrastruktur för systemförståelse och systemanalys.

Skolverket beslutar i praktiken om innehållet i skolstatistiken

De ämnes- och statistikområden som anges i bilagan till statistikförordningen – för Skolverkets del skolväsende och barnomsorg – är övergripande. Det innebär att det är den statistikansvariga myndigheten som i praktiken definierar variabler och objekt och beslutar vilken eller vilka metoder som ska användas för att samla önskad statistik till önskad kvalitet. Den statistikansvariga myndigheten avgör slutligen vilka ”skattningar av statistiska storheter” som utgör den officiella statistiken, dvs. vilka mått som ska redovisas.¹⁶

Det råd för den officiella statistiken som är inrättat vid SCB har dock beslutat om riktlinjer för beslut om innehåll och omfattning av SOS¹⁷. Av riktlinjerna framgår bland annat att SOS i regel bör vara av löpande karaktär så att man kan följa utvecklingen över tid, men engångsundersökningar och periodiskt återkommande (intermittenta) undersökningar kan också vara officiell statistik. De elevpaneler som SCB svarar för i samarbete med Pedagogiska institutionen vid Göteborgs universitet är exempel på detta.

Den statistikansvariga myndigheten bör enligt rådet anpassa statistikens innehåll och omfattning till samhällsutvecklingen och basera denna anpassning på omvärldsbevakning, inklusive dialog

¹⁶ Rådet för den officiella statistiken (2012) *Riktlinjer för beslut om innehåll och omfattning av den officiella statistiken*.

¹⁷ Rådet för den officiella statistiken (2012) *a.a.*

med användarna. Myndigheten bör även samråda med organisationer som företräder uppgiftslämnarna.

Statistikansvarig myndighet avgör om statistikens kvalitet är tillräckligt god i relation till hur den är avsedd att användas. I kvalitetskravet ingår att beakta aktualitet kontra t.ex. precision. Myndigheten kan även besluta att en undersökning som inte ingår i SOS från en viss tidpunkt ska definieras om till att ingå i systemet.

10.1.4 Dagens statistiksystem

Inom ramen för Skolverkets ansvar för Sveriges officiella statistik (SOS) om förskola, fritidshem, skola, vuxenutbildning och annan pedagogisk verksamhet samlas varje år in uppgifter om barn, elever, personal, kostnader och utbildningsresultat. Uppgifterna sammanställs till statistik per verksamhetsform och skolform. Statistiken presenteras som SOS på riksnivå. Uppgifterna presenteras dessutom av Skolverket på huvudmannanivå eller skolnivå.

Den övervägande delen av Skolverkets uppgifter om elever samlas in på individnivå, vilket möjliggör redovisningar uppdelat på kön, socioekonomisk bakgrund och eventuell utländsk bakgrund. Uppgifter som samlas in aggregerat samlas alltid in uppdelat på kön. Det är SCB som står för den databas där den insamlade informationen lagras. Det Skolverket har tillgång till är utdrag ur databasen där de enskilda individerna avidentifierats. Av detta följer att om en forskare avser att göra en registerbaserad studie och eventuellt samköra olika register, så är det SCB som tillhandahåller data.

Skolverket redovisar en del av statistiken i den årliga publikationen *Beskrivande data*. Den statistik som ingår i SOS presenteras i de tre årliga rapporterna *Utbildningsresultat*, *Barn elever och personal*, samt *Kostnader*. De uppgifter som Skolverket samlar in är även i sin helhet tillgänglig på verkets webbplats, bl.a. i databaserna Jämförelsetal, SIRIS och SALSA. För en mer omfattande presentation av innehållet i det nationella uppföljningssystemet hänvisas till kapitel 4 samt till Skolverkets föreskrifter om uppgiftslämning¹⁸.

¹⁸ Se t.ex. SKOLFS 2012:91 och SKOLFS 2014:1.

Krav på statistiksystemet

Den nationella statistiken ska ge information om resurser, organisation och resultat och utvecklingen ska kunna följas över tid. Det är därför väsentligt att systemet är *långsiktigt* och stabilt över tid. Det är särskilt viktigt för den officiella statistiken som enligt lagen (2001:99) om den officiella statistiken ska finnas för allmän information, utredningsverksamhet och forskning och som ska vara objektiv och allmänt tillgänglig.

För den officiella statistiken anges kvalitetskriterier i lagen om den officiella statistiken¹⁹ avseende följande punkter:

1. *Relevans* (i vilken grad statistiken tillgodoser användarnas nuvarande och potentiella behov)
2. *Noggrannhet* (överensstämmelse mellan skattning och det skattningen avser att mäta)
3. *Aktualitet* (tiden mellan att de uppgifterna finns tillgängliga och den aktuella händelsen/företeelsen som uppgiften avser)
4. *Punktlighet* (den tid det tar för statistikansvarig myndighet att göra uppgifter tillgängliga efter det datum de ska ha lämnats in)
5. *Tillgänglighet och tydlighet* (hur användarna kan få tillgång till, använda och tolka uppgifter)
6. *Jämförbarhet* (för statistik från olika geografiska områden eller sektorsområden eller från olika tidsperioder)
7. *Samstämmighet* (i vilken utsträckning uppgifterna med tillförlitlighet kan kombineras på olika sätt och för olika ändamål)

Generellt gäller att data blir värdefull för utvärdering och forskning om den har samlats in systematiskt under lång tid. Med systematiskt menas här att det handlar om samma data som samlas in på samma sätt. Det handlar bl.a. om

- med vilka intervall som datainsamlingen görs, t.ex. varje eller vart femte år,
- att urvalet görs på samma sätt vid varje mättillfälle, t.ex. samtliga i en ålderskohort eller varje årskurs 3 eller 6,
- att det är samma information som samlas in vid de olika mättillfällena (tidsseriedata),

¹⁹ § 3 a lag (2001:99) om den officiella statistiken.

- att informationen samlas in på ett enhetligt sätt vid de olika mättillfällena, dvs. med samma metoder, begrepp, definitioner, mått, frågor etc.

Datainsamlingen bör således vara regelbunden och systemet måste bygga på enhetliga definitioner, indelningar och tidsperioder som gör det möjligt att aggregera data till nationell nivå och göra jämförelser mellan huvudmän och verksamhetsområden och över tid. Uppföljningen ska spegla verksamhetsformernas särart, samtidigt som det är viktigt att statistiken presenteras på ett likartat sätt.

Den befintliga statistiken om skolan bedöms generellt hålla en hög kvalitet. Det finns dock en kritik vad gäller tillförlitligheten avseende vissa data, bland annat inom ett så central område som likvärdigheten i betygssättning respektive bedömning av de nationella proven. Det handlar dock mer om hur dessa system är utformade än om statistiksystemets funktionssätt.

Långsiktighet och stabilitet påverkas av förändringar

Samtidigt som statistiksystemet måste värna om långsiktighet förväntas det svara mot behovet av snabb styrinformation och ett flertal andra, delvis motsägelsefulla, behov. En svårighet för alla longitudinella databaser är att de är känsliga för förändringar, till exempel av lagar och andra bestämmelser, som kan innebära att informationen inte förblir jämförbar över tid. Politiskt beslutade förändringar, som t.ex. förändringar i betygssystemet, kan därför leda till att tidsserier bryts. Sådana brott är ofrånkomliga – om än beklagliga för statistiker, forskare och andra användare av statistiken. Andra aktuella exempel som påverkat uppgiftsinsamlingen är införandet av betyg i årskurs 6 samt förändringar av i vilken årskurs och i vilka ämnen de nationella proven ges. Även förändringar i omvärlden kan föranleda behov av att utveckla uppgiftsinsamlingen och statistiken, t.ex. tillkomsten av enskilda huvudmän.

För att en förändring av uppgiftsinsamlingen ska kunna träda i kraft måste den beredas och aviseras i god tid för att de administrativa systemen och insamlingsrutinerna hos både uppgiftslämnarna och myndigheten ska kunna ses över.

Registerdata hämtas ur skolornas administrativa system

Skolornas IT-baserade administrativa system är konstruerade för att tillgodose både nationella krav och lokala behov. De nationella kraven består dels av krav på dokumentation som ställs i styrdokumenten, dels av Skolverkets krav på uppgifter till det nationella uppföljningssystemet.

De lokala administrativa systemen är uppbyggda kring eleven, även om vissa uppgifter om ansvarig lärare också ingår. Schema-lägningsmoduler används av många skolor, dock främst i grundskolans högre årskurser och i gymnasieskolan. Lärarna kan även för varje genomförd lektion registrera närvaro och giltig eller oanmäld frånvaro respektive sen ankomst.

Den faktiska användningen av skolornas administrativa system avgörs lokalt. Detta får till följd att uppgifter inte alltid registreras även om funktionaliteten finns. Det är inte heller säkert att uppgifterna om genomförda lektioner avser den lärare som faktiskt undervisade. Om en vikarie undervisar under en kortare tid ändras i allmänhet inte lärarens namn i systemet och lektionen registreras som om den ordinarie läraren hade varit där. Inte heller då en lärare slutar och ersätts med en ny lärare under läsåret är det säkert att skolan ändrar lärarens namn i systemet.

Enligt Skolverket prioriterar de största leverantörerna av skolornas administrativa system registreringen av betyg och att uppgifterna hanteras rättsäkert i systemen med kontroll av behörigheter och versionshantering. Uppgiften om vilken lärare som har satt vilket betyg bedöms därmed ha god kvalitet. Registrering av resultat på nationella prov bedöms däremot ha lägre prioritet på skolorna. Eftersom provens utformning varierar mellan åren ökar risken för felregistrering och systemet blir inte heller lika användbart lokalt då det inte redovisas vilka olika förmågor som provuppgifterna prövar i de olika delproven.

Skolverket konstaterar att de administrativa systemen nu byggs ut för att ge stöd för lärarnas pedagogiska dokumentation. Man menar att då allt fler lärare har tillgång till egen dator, kan de administrativa systemen på sikt komma att få en ökad betydelse för statens möjligheter att implementera reformer men även ge ökade möjligheter att samla in nya uppgifter.

Att införa förändringar i skolornas administrativa system kräver dock enligt Skolverket lång förberedelse, ibland upp till två år. Dels ska systemleverantörerna bygga om systemen och sprida den

nya versionen till sina användare, dels ska skolorna uppdatera till den nya systemversionen och ändra sina arbetsrutiner kring de uppgifter som ska läggas in. Om systemen inte hinner uppgraderas innan en förändring träder i kraft försvåras huvudmännens uppgiftslämnande.

Behov av utveckling och modernisering

Intresset för deltagande i enkätundersökningar minskar. Skolverkets erfarenheter av detta bekräftas av Statistikutredningen 2012²⁰. Det innebär att behovet av registerdata ökar. Vidare ställs allt större krav på statistikens innehåll och systemens användarvänlighet. Att parallellt med den årliga datainsamlingen tekniskt modernisera uppföljningssystemet kommer enligt Skolverkets bedömning att utgöra en stor och resurskrävande utmaning under de närmaste åren. Utredningen återkommer till denna fråga i slutet av kapitlet.

10.1.5 Uppgiftslämnarbördan

Skolverket samlar sedan lång tid tillbaka in en mycket stor mängd uppgifter från skolhuvudmän och verksamheter för den nationella uppföljningen, vilket innebär en belastning för uppgiftslämnarna.

Utredningen har försökt få fram uppgifter om vilka kostnaderna är för uppgiftslämnandet. Detta har visat sig vara svårt eftersom det inte finns någon enhetlig definition för vilka kostnader som ska räknas in. En stor del av de uppgifter som finns i det nuvarande uppföljningssystemet består dock av data som huvudmännen själva behöver för sin administration av verksamheten, dvs. uppgifter som de skulle behöva samla in oavsett om de efterfrågades för den nationella uppföljningen eller inte.

SCB har uppskattat tidsåtgången och kostnaderna för det samlade uppgiftslämnande som krävs för Sveriges officiella statistik (SOS), dvs. den statistik som de statistikansvariga myndigheterna har beslutat ska vara officiell. Kostnaden för uppgiftslämnandet har beräknats till 46 miljoner kronor för kommuner och landsting.²¹ Det framgår inte hur stor andel av detta belopp som avser skolan.

²⁰ SOU 2012:83 *Vad är officiell statistik? En översyn av statistiksystemet och SCB*.

²¹ Statistiska centralbyrån (2013) *Sveriges officiella statistik – årlig rapport för 2012*. Rådet för den officiella statistiken.

Dessa kostnader hänför sig dock endast till själva arbetet med att ta fram de uppgifter som SCB efterfrågar vid datainsamlingen. Det innebär att t.ex. kostnader för att anskaffa och underhålla de administrativa systemen inte är medräknade, inte heller kostnaderna för den lokala uppgiftsinsamlingen. De samlade kostnaderna för uppgiftslämnarbördan är sannolikt betydligt högre än den SCB uppgivit. Skolverket uppger att verkets kostnad för arbetet med den officiella statistiken var 36 miljoner kronor 2012²².

Enligt Skolverket deltar uppskattningsvis 10 000 personer hos huvudmännen i arbetet med att lämna uppgifter till SCB, varav varje år omkring 1 000 bedöms vara nya. Antalet involverade och antalet oerfarna uppgiftslämnare påverkar sannolikt tillförlitligheten i data. Arbetsinsatsen uppskattas till cirka 70 årsarbetskrafter (samtliga huvudmän inom tretton skolformer). Ett annat mått på komplexiteten i uppgiftsinsamlingen är att SCB, enligt uppgift från Skolverket, tar emot elevuppgifter via 33 olika administrativa program som utvecklats av olika it-företag och i vissa fall av enskilda kommuner.

Utredningen har inte haft möjlighet att närmare undersöka vad detta betyder för kvaliteten i de insamlade uppgifterna. Det är dock sannolikt att en mer ändamålsenlig organisation av uppgiftslämnandet dels skulle innebära tidsbesparingar för både de som lämnar uppgifterna som de som tar emot dem, dels förbättra kvaliteten i de data som samlas in.

Skolverket har vid kontakter med utredningen kraftigt understrukt att utrymmet för att öka kraven på skolhuvudmännens uppgiftslämnande är starkt begränsat.

Enligt förordningen om den officiella statistiken²³ ska uppgifter för denna samlas in på ett sådant sätt att uppgiftslämnandet blir så enkelt som möjligt, står i proportion till användarnas behov, och är en rimlig arbetsbörda för uppgiftslämnarna. De statistikansvariga myndigheterna ska vid framställning av officiell statistik i så stor utsträckning som möjligt använda uppgifter ur befintliga register.

Skolverket har dock påpekat att förutsättningarna att hantera uppgiftslämnandet skiljer sig mycket åt mellan olika huvudmän. Det finns fortfarande mindre skolor och huvudmän som inte har något administrativt system alls. Uppgiftsinsamlingen får då ske genom direktinsamling via enkäter. Den administrativa bördan för

²² Statistiska centralbyrån (2013) a.a.

²³ Förordning (2001:100) om den officiella statistiken.

dessas skolor kan bli oproportionerligt tung vid förändringar och då nya data ska samlas in.

Skolverket samråder med Sveriges kommuner och landsting kring nya uppgiftsinsamlingar i enlighet med den samrådsrutin som anges i förordning om statliga myndigheters inhämtande av uppgifter från näringsidkare och kommuner²⁴. Samrådet omfattar bl.a. vilka uppgifter som ingår i undersökningen, tidpunkt för genomförandet, urvalets storlek och kostnader för uppgiftslämnarna.

Även om möjligheten finns att registrera uppgifter i skolornas administrativa system krävs uppgiftsskyldighet för att säkerställa att uppgifterna faktiskt läggs in i systemen och för att systemleverantörerna ska utveckla tekniska lösningar. Vid förändringar är det därför viktigt att beakta helheten av den uppgiftsbörda som riktas mot huvudmännen. Kostnadsaspekter men även krav på kvalitet i data medför begränsningar för vad det är möjligt att samla in inom ramen för det nationella uppföljningssystemet. Eftersom kvaliteten i de lämnade uppgifterna ytterst är beroende av uppgiftslämnarna är det angeläget att upprätthålla förståelse och engagemang hos dessa.

10.2 Regelverket för hantering av data

Det finns en omfattande lagstiftning på statistikområdet. För närvarande finns dessutom flera utredningar som ser över delar av regelverket. Utredningen bedömer att resultatet av dessa utredningar kommer att få betydelse för hur flera av de datafrågor bör hanteras som direktiven till föreliggande utredning reser.

10.2.1 En omfattande lagstiftning

Följande lagar gäller för produktion, lagring och användning av data och statistik:

- *Lagen om den officiella statistiken* (2001:99) anger syftet med den officiella statistiken samt behandlar frågor om krav på denna, t.ex. kvalitetskriterier, uppgiftsskydd, uppgiftsskyldighet, utlämnande av uppgifter i vissa fall, gallring m.m.

²⁴ SFS 1982:668.

- *Förordningen om den officiella statistiken* (2009:100) innehåller bland annat den bilaga där de statistikansvariga myndigheterna anges.
- *Offentlighets- och sekretesslagen* (2009:400) innehåller bland annat bestämmelser om sekretess till skydd för uppgift om enskilda personliga eller ekonomiska förhållanden, samt särskilda bestämmelser om sekretess hos vissa organ.
- *Personuppgiftslagen* (1998:204) innehåller bestämmelser som syftar till att skydda människor mot att deras personliga integritet kränks genom behandling av personuppgifter.
- *Lagen om etikprövning av forskning som avser människor* (2003:460) reglerar forskningens användning av data och statistik.

Även tryckfrihetsförordningen²⁵ är till vissa delar tillämplig. Det gäller kapitel 2 om allmänna handlingars offentlighet. Dessutom finns förordningar om statliga myndigheters inhämtande av uppgifter från näringsidkare och kommuner²⁶, om konsekvensutredning vid regelgivning²⁷, samt om myndigheters inhämtande av yttrande från Regelrådet²⁸.

Vidare finns ett stort antal s.k. registerförfattningar som innehåller särregleringar på olika områden i förhållande till personuppgiftslagen (PUL). Patientdatalagen²⁹, polisdatalagen³⁰ och lagen om behandling av personuppgifter vid Institutet för arbetsmarknads- och utbildningspolitisk utvärdering³¹ är exempel på sådana registerförfattningar.

Även Europeiska unionen har på unionsrättslig nivå reglerat behandlingen av personuppgifter genom det s.k. dataskyddsdirektivet³². Inom ramen för Europarådet har den s.k. dataskyddskonventionen antagits³³. Till det kan läggas europaparlamentets och rådets förordning om framställning och utveckling av statistik om utbildning och livslångt lärande, som reglerar vilken data inom utbild-

²⁵ Tryckfrihetsförordning (1949:105).

²⁶ Förordning (1982:668) om statliga myndigheters inhämtande av uppgifter från näringsidkare och kommuner.

²⁷ Förordning (2007:1244) om konsekvensutredning vid regelgivning.

²⁸ Förordning (2011:118) om myndigheters inhämtande av yttrande från Regelrådet.

²⁹ Patientdatalagen (2008:355).

³⁰ Polisdatalagen (2010:361).

³¹ Lagen (2012:274) om behandling av personuppgifter vid Institutet för arbetsmarknads- och utbildningspolitisk utvärdering.

³² Europaparlamentets och rådet direktiv 95/46/EG.

³³ CETS 108.

ningområdet som medlemsländerna är skyldiga att samla in och leverera till Eurostat³⁴.

Myndighetens samrådsskyldighet och uppgiftslämnarnas skyldighet att medverka vid datainsamling

En statistikansvarig myndighet är enligt förordningen om statliga myndigheters inhämtande av uppgifter från näringsidkare och kommuner skyldig att samråda med organisation eller annan som kan anses företräda de näringsidkare och kommuner som ska lämna uppgifterna³⁵.

När insamlade uppgifter ingår i SOS omfattas uppgiftslämnarna (näringsidkare, kommuner, ideella föreningar etc.) av uppgiftsskyldighet³⁶. Uppgiftsskyldigheten gäller dock inte enskilda individer. Det innebär olika förutsättningar vid insamling av registerdata och andra data som samlas in direkt från enskilda, t.ex. genom enkäter. I det senare fallet behöver den enskilde inte besvara enkäten. Bortfallet riskerar alltså att bli större i samband med direktinsamlad information än vid insamling av administrativa registerdata.

Vid enkätundersökningar är myndigheten i enlighet med statistikförordningen skyldig att uppge att deltagandet i undersökningen är frivilligt. I sådana fall räknas det förhållandet att personen skickar in ett enkätsvar som samtycke, dvs. det krävs inte någon separat undertecknad bekräftelse på samtycket.

Användning av registerdata för forskningsändamål – vad krävs?

Hur data får insamlas och lagras i en databas eller hämtas från olika register och lagras i ett nytt register för forskningsändamål kringgärdas av omsorgsfullt reglerade processer. Några begrepp är centrala för förståelsen av dessa processer.

Sambearbetning av registeruppgifter innebär att uppgifter om en specifik person hämtas från olika register och sammankopplas. I de fall samtliga registeruppgifter finns hos en och samma myndighet, t.ex. SCB, kan sambearbetningen ske inom denna myndighet. Forskarna får då tillgång endast till oidentifierade uppgifter. När regis-

³⁴ EG förordning nr 452/2008 av den 23 april 2008.

³⁵ Förordning (1982:668) om statliga myndigheters inhämtande av uppgifter från näringsidkare och kommuner.

³⁶ 7 § SFS 2001:99.

teruppgifterna finns hos olika myndigheter uppstår en annan situation. För att det ska vara möjligt att koppla uppgifterna kan de då inte vara helt oidentifierbara.³⁷

Ett sätt att lösa detta på är att avidentifiera (avkoda) de data som lämnar SCB genom att personnummer ersätts med löpnummer³⁸. Ingen information kan då hänföras direkt till en enskild individ hos den myndighet till vilken uppgifterna lämnas. SCB har dock kvar en nyckel mellan dessa löpnummer och individernas personnummer. Hela eller delar av registret kan därför på nytt hänföras till en individ och länkas till andra register efter behov från myndigheter, utredare eller forskare.³⁹

I personuppgiftslagen (PUL) görs en distinktion mellan personuppgifter och *känsliga* personuppgifter. Med personuppgift avses all information som direkt eller indirekt kan hänföras till en särskild person som är i livet. Känsliga personuppgifter är sådana som avslöjar ras eller etniskt ursprung, politiska åsikter, religiös eller filosofisk övertygelse eller medlemskap i fackförening samt personuppgifter som rör hälsa och sexualliv. För behandling av sådana känsliga uppgifter gäller särskilda restriktioner. Enligt PUL⁴⁰ får känsliga personuppgifter endast behandlas för forskningsändamål om behandlingen *etikprövats*, dvs. godkänts av en etikprövningsnämnd enligt etikprövningslagen⁴¹.

Om en uppgift kan lämnas ut ska även provas mot offentlighets- och sekretesslagen (OSL)⁴². Prövningen ska avse dels om uppgiften utgör en s.k. *allmän handling*, dels om uppgiften omfattas av *sekretess* eller inte. Någon allmän bestämmelse i OSL om utlämnade av uppgifter för forskningsändamål finns inte.

I PUL föreskrivs vidare att personuppgifter inte får behandlas för något ändamål som är oförenligt med det för vilket uppgifterna ursprungligen samlades in, den s.k. *finalitetsprincipen*. I det fall en tänkt behandling inte omfattas av de ursprungliga ändamålen måste det alltså provas om det nya ändamålet är förenligt med det ursprungliga ändamålet för behandlingen. Av PUL följer vidare att behandlingen av personuppgifter för historiska, statistiska eller vetenskap-

³⁷ Dir. 2013:8 *Förutsättningar för registerbaserad forskning*.

³⁸ 16 § lagen (2001:99) om den offentliga statistiken.

³⁹ Prop. 2011/12:176 *Behandling av personuppgifter vid Institutet för arbetsmarknads- och utbildningspolitisk utvärdering*.

⁴⁰ 19 § personuppgiftslagen (1998:204).

⁴¹ Lag (2003:460) om etikprövning av forskning som avser människor.

⁴² Offentlighets- och sekretesslag (2009:400).

liga ändamål inte ska anses som oförenliga med de ändamål för vilka uppgifterna ursprungligen samlades in.⁴³

Vad gäller för data som inte ingår i den officiella statistiken?

En statistikansvarig myndighet har rent formellt möjlighet att samla in uppgifter som inte ingår i den officiella statistiken (SOS) och presentera dessa på olika sätt, t.ex. i en publikation eller på sin webbplats. En anledning till detta kan vara att undersökningen varken genomförs årligen eller är periodiskt återkommande, eller att den inte fullt ut möter de kvalitetskrav som ställs upp i statistiklagen. En annan kan vara att insamlingen avser nya uppgifter som på sikt, om de har tillräcklig kvalitet, kan komma att ingå i SOS. Detta är ett sätt som Skolverket använder sig av när verket introducerar nya uppgiftsinsamlingar.

Utredningen har inte haft möjlighet att närmare granska i vilken omfattning olika myndigheter samlar in uppgifter och presenterar statistik vid sidan om SOS. En bedömning är dock att detta sker endast i begränsad omfattning. Skolverket har hittills valt att inte samla in registerdata utanför SOS annat än vid utvecklingen av ny uppgiftsinsamling. Trots att Skolverkets insamling av data sker inom ramen för SOS presenteras dock statistiken i vissa fall som SOS, i andra fall vid sidan om SOS. Orsakerna till detta är av mer historisk karaktär.

Det är enligt medarbetare på Skolverket svårt att motivera insamling av registerdata om inte avsikten är att den ska ingå i SOS. Man kan exempelvis inte räkna med att det ska komma in uppgifter i tillräcklig omfattning om insamlingen är frivillig. Om insamlingen inte sker inom ramen för SOS skulle därför uppgiftsskyldigheten behöva regleras på annat sätt, vilket ter sig ganska poänglöst när skyldigheten redan är reglerad inom ramen för SOS.

10.2.2 Pågående utredningar

Regeringen har i propositionen *Ett lyft för forskning och innovation*⁴⁴ tydliggjort sin ambition att stärka förutsättningarna för den forskning som utnyttjar de svenska befolkningsbaserade registren.

⁴³ 9 § personuppgiftslagen (1998:204).

⁴⁴ Prop. 2008/09:50 *Ett lyft för forskning och innovation*.

För närvarande pågår två utredningar som har relevans i sammanhanget. Den ena behandlar frågor om hur registerförfattningar ska utformas i framtiden, den andra frågor om hur de olika lagverken inom området ska bringas i bättre samklang med varandra.

Utredning om registerbaserad forskning

Regeringen har uppmärksammat att i den lagstiftning som finns kan två samhällseliga intressen komma att stå i konflikt med varandra. Det gäller å ena sidan intresset av att bedriva forskning som ger upphov till ny viktig kunskap och å andra sidan intresset av att skydda den enskildes integritet. Mot denna bakgrund har regeringen gett en särskild utredare i uppdrag att undersöka förutsättningarna för att bedriva s.k. registerbaserad forskning. Utredaren ska lämna författningsförslag och andra förslag i syfte att

- registeransvariga myndigheter i större utsträckning ska kunna lämna ut uppgifter för forskningsändamål med hänsyn tagen till skyddet för den enskildes integritet,
- sambearbetning av registeruppgifter för forskningsändamål ska underlättas, och
- göra det möjligt att på ett integritetssäkert sätt samla in personuppgifter till register som förs för särskilda och uttryckligt angivna forskningsändamål samt till longitudinella studier som håller sig inom ramen för sådana ändamål.⁴⁵

Uppdraget ska slutredovisas den 30 juni 2014.

Utredning om registerlagstiftningen m.m.

Känsliga personuppgifter från myndigheter kan lämnas ut om det finns en särskild registerförfattning för det register inom vilket personuppgifterna ingår. Det finns mer än hundra sådana författningar som kompletterar personuppgiftslagen och som innehåller bestämmelser om behandling av personuppgifter, och som därför kan kallas för särskilda registerförfattningar.⁴⁶ Någon fullständig kartläggning och överblick finns inte.⁴⁷

⁴⁵ Dir. 2013:8 *Förutsättningar för registerbaserad forskning*.

⁴⁶ Öhman, Sven (2006) *Särskilda registerförfattningar* i Cecilia Magnusson Sjöberg och Peter Wahlgren (red.) *Festschrift till Peter Seipel*. Norstedts juridik.

⁴⁷ Öhman, Sven (2006) a.a.

Regeringen gav 2011 en särskild utredare i uppdrag att se över registerlagstiftningen och bland annat göra en översiktlig inventering av registerförfattningarna. Utredaren ska närmare analysera hur författningarna fungerar inom tre olika verksamhetsområden och utarbeta en generell modell för reglering av registerfrågor⁴⁸.

Enligt direktiven måste ändringar göras i de aktuella registerförfattningarna nästan varje gång elektronisk informationsöverföring mellan myndigheter ska införas, ändras eller utökas. Vidare måste komplicerade överväganden göras enligt offentlighets- och sekretesslagen (OSL). Om dessutom de grundläggande reglerna i tryckfrihetsförordningen⁴⁹ och OSL samt den principiella uppbyggnaden av registerförfattningarna inte är anpassade till varandra blir det svårt att hitta lösningar som är tillfredsställande från integritets-, effektivitets- och öppenhetssynpunkt i lagstiftningsärenden om elektronisk informationsöverföring mellan myndigheter.

För att kunna effektivisera förvaltningen med hjälp av modern informationsteknik är det därför enligt regeringen viktigt dels att respektive regelverk är väl genomtänkt och så enkelt som möjligt att tillämpa, dels att dessa regelverk är förenliga med varandra. Det är mot denna bakgrund behovet att göra en samlad översyn av dessa regelverk i relevanta delar ska ses.

Uppdraget ska slutredovisas senast den 1 december 2014.

Utredningarna berör även register på skolans område

De båda utredningarnas arbete kan komma att påverka förutsättningarna för registerhanteringen på skolans område. Ett exempel på en sådan särskild registerförfattning som nämns ovan är t.ex. lagen (2012:741) om behandling av personuppgifter vid Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU). Enligt denna får vid IFAU personuppgifter endast behandlas om det behövs för att fullgöra institutets uppdrag att främja, stödja och genom forskning genomföra studier, uppföljningar och utvärderingar. Att det här är fråga om en så bred ändamålsformulering beror på att den ansluter till IFAU:s uppdrag som det formulerats i myndighetens instruktion (2007:911).

⁴⁸ Dir. 2011:86 *Integritet, effektivitet och öppenhet i en modern e-förvaltning*.

⁴⁹ Tryckfrihetsförordningen (1949:105).

10.3 Förslag och bedömningar om behovet av data för utvärdering och utvärdering av reformer

Utredningens förslag: I enlighet med verkets hemställan ges Skolverket möjlighet att samla in individdata för skolan i sin helhet, dvs. även för modersmålsundervisning och skolgång i specialskola, särskola och sameskola.

Utredningens bedömning: Utvärdering av reformer kan baseras på data som samlats in till det löpande uppföljningssystemet och på särskilt insamlad data. Utökade och förbättrade registerdata kan underlätta möjligheterna att utvärdera reformer. För utvärdering av reformer behövs dock ofta riktade uppgiftsinsamlingar, med frågeställningar relevanta för den aktuella reformen.

Det nationella uppföljningssystemet har en potential att utvecklas vad gäller ny uppgiftsinsamling, utökad analys av befintliga data från uppföljningssystemet och de internationella studierna samt det sätt som informationen presenteras på. Vid utvecklingen bör ställning tas till vilka uppgifter som lämpar sig för att samla in i form av registerdata, vilka uppgifter som bör utgöras av direktinsamlade data och ingå i det nationella uppföljningssystemet samt vilken information som inte lämpar sig för att ingå i detta. Detta är bedömningar som bör göras av Skolverket som statistikansvarig myndighet. Det är viktigt att Skolverket inför sina bedömningar samråder med forskare, myndigheter och andra aktörer som använder uppgifterna i systemet.

Det har ingått i utredningens uppdrag att bedöma behovet av data för utvärdering av reformer, varför bedömningar av vissa specifika önskemål om ytterligare data presenteras i det följande.

För att det ska vara möjligt att utveckla och modernisera det nationella uppföljningssystemet behöver denna infrastruktur förstärkas finansiellt.

Med anledning av att det för närvarande pågår två utredningar som arbetar med frågor kopplade till hur förutsättningarna för registerbaserad forskning kan förbättras, lämnas inte här några sådana förslag. Det är emellertid viktigt att regeringen i detta sammanhang bevakar de långsiktiga förutsättningarna för UGU-databasen.

Generellt gäller att olika data behövs för olika syften och olika statistiska metoder kräver olika typer av data. Detta påpekande gränssar till det självklara. Utbildning och utbildningsval påverkar en individs sociala situation, yrkesval, sysselsättning, hälsa etc. Sociala bakgrundsfaktorer har ett starkt samband med hur elever lyckas i skolan. Verkningar av skolpolitiska reformer sträcker sig även utanför själva skolan och kan många gånger avläsas först i vuxen ålder. Det är sammantaget inte enkelt att avgränsa vilken slags information som behövs för att utvärdera skolan eller att utvärdera skolpolitiska reformer.

De behov av data för utvärdering av skolan och av reformer inom skolan som aktualiserats under utredningens arbete, bl.a. av Skolverket, IFAU och olika forskare, fördelar sig på några huvudtyper. Det handlar om behovet av individdata, uppgifter om undervisningens förutsättningar, genomförande och resultat. Det finns även frågeställningar som har att göra med hur data samlas in. Vi återkommer till dessa behov och frågeställningar i det följande.

10.3.1 Några utgångspunkter

Utredningen framhöll redan i kapitel 4 att det löpande nationella uppföljningssystemet har en viktig roll i utvärderingen av skolan. Vi underströk samtidigt att uppföljnings- och utvärderingsystemet knappast på egen hand kan ge en så kvalificerad kunskap att det inte finns behov av att också genomföra särskilda utvärderingar av en reform eller ett reformskede. Ofta behövs också en kompletterande uppgiftsinsamling. I kapitel 6 framhöll vi att det behövs en mångfald av utvärderingsmetoder, både kvantitativa och kvalitativa, för att utvärdera ett komplext system av det slag som skolan utgör.

Utredningen konstaterar vidare att inte minst forskare önskar sig många slags data om skolan, inte endast uppgifter om kunskapsresultat. Till det man särskilt vill ha mer information om hör olika ramfaktorer för undervisningen och möjligheter att på olika sätt koppla samman lärarna med elever och undervisningsformer för att kunna studera det pedagogiska arbetet i skolan och kopplingarna mellan elevresultat och undervisning. Man vill också kunna följa individer från förskolan och in i arbets- och vuxenlivet.

De önskemål om data som olika intressenter uttrycker är i princip utan slut. Ju mer data som är tillgängliga, desto fler möjligheter finns det att genomföra olika slags studier och att samköra

data från olika register. Mot bakgrund av detta faktum är det nödvändigt att peka på ett antal förutsättningar som måste beaktas vid varje diskussion om att utöka uppgiftsinsamlingen inom ramen för det nationella uppföljningssystemet:

- Insamling, lagring, bearbetning, analys och rapportering av data medför kostnader av olika slag, kostnader som fördelas på skolor, huvudmännen och staten. Ökad insamling av data behöver finansieras.
- Uppgiftslämnarbördan är redan stor. När nya data samlas in bör därför alltid övervägas om det samtidigt kan genomföras andra lättnader för uppgiftslämnarna, t.ex. genom att upphöra att samla in andra mindre aktuella data eller på annat sätt underlätta uppgiftslämnandet. Det senare kan dock vara förknippat med investeringar i systemen, som i så fall måste finansieras.
- Såväl i Sverige som utomlands tenderar svarsfrekvenserna i statistiska undersökningar (enkäter) att minska. Detta gäller särskilt när enkäterna är märkta med personnummer eller löpnummer som kan kopplas till personnummer.
- Det finns ett omfattande regelverk att ta hänsyn till såväl när data samlas in som när insamlade data ska användas för olika ändamål, bl.a. forskning.
- Regeringen har på olika sätt markerat att möjligheterna att använda befintliga register för forskningsändamål bör förbättras och den har initierat olika utredningar som bl.a. har till uppgift att belysa hur väl de olika regelverken inom området samspelar med varandra.

10.3.2 Individdata

Fördelarna med en individbaserad uppgiftsinsamling bygger i grunden på att alla uppgifter är likvärdiga och att samma uppgifter finns att tillgå för samtliga individer. Individdata ger stor flexibilitet och många möjligheter i samband med de flesta utvärderingsdesigner. Exempelvis behövs individdata kopplade till personnummer för att man ska kunna göra analyser av skolprestationers betydelse för samsättningsgraden senare i livet. Samma sak gäller om man vill konstruera mått som anger vilken betydelse en viss skola har för elevernas resultatutveckling, s.k. value-added-mått.

Hanteringen av individdata omgärdas av bestämmelser som skyddar individen, vilket framgår ovan. Vid samkörning av register för forskningsändamål finns t.ex. i dag tekniker att avpersonifiera registeruppgifterna, dvs. personnumren tas bort och ersätts av löpnummer. Endast SCB har tillgång till den s.k. kodnyckel som gör det möjligt att koppla samman individerna i de olika registren. Vidare publiceras inte uppgifter för enskilda individer eller för så små grupper att individer indirekt kan identifieras.

Individdata för grundskolan sedan 2008

Skolverket beslutade 2007 att samla in individbaserade uppgifter om grundskolans elever i syfte att öka kvaliteten på den officiella statistiken och minska uppgiftslämnarbördan. Beslutet föregicks av en intern utredning där fördelar och nackdelar med en individbaserad uppgiftsinsamling vägdes mot varandra. Utredningen utmynnade i bedömningen att fördelarna övervägde, med hänsyn till det samhällsintresse som finns för att behandla uppgifterna och det sekretesskydd som uppgifterna omgärdas av.

I samband med utredningen gjorde Skolverket även en utförlig genomgång av hela uppföljningssystemet för att uppmärksamma förekomsten av personuppgifter som i enlighet med personuppgiftslagen kan betraktas som känsliga. De personuppgifter som verket identifierade gällde modersmålsundervisning och skolgång i specialskola, särskola, sameskola samt skolgång i en fristående skola med konfessionell inriktning⁵⁰. Skolverket valde därför att hösten 2008 genomföra den första insamlingen på individnivå i grundskolan med undantag för dessa uppgifter.

Eftersom information om eleverna inom flera skolformer faller inom det som i lagstiftningen benämns som känsliga personuppgifter krävs en förändring i förordningen om den officiella statistiken för att Skolverket ska kunna samla in data om samtliga elevers skolgång. Skolverket hemställde därför hösten 2007 hos regeringen om en ändring i förordningen om den officiella statistiken. Skolverket har på regeringens uppdrag kompletterat hemställan. Regeringen har ännu inte beslutat i frågan.

⁵⁰ Skolinspektionen kategorisering av fristående skolor med konfessionell inriktning har upphört. Denna uppgift har därför inte längre någon relevans.

Individdata för barn i förskola och pedagogisk omsorg från 2014

Det har från forskarhåll länge funnits önskemål om tillgång till data på individnivå avseende barn som deltar i förskola och pedagogisk omsorg, på motsvarande sätt som för elever i grundskolan. Regeringen har gett Skolverket i uppdrag⁵¹ att införa en individbaserad uppgiftsinsamling avseende barn som deltar i förskola och pedagogisk omsorg. Insamling av uppgifter ska påbörjas från och med 2014 och statistiken ska därefter finnas tillgänglig årligen. Individdata finns även för gymnasieskolan – med undantag för gymnasiesärskolan.

Utredningens bedömning och förslag

Det är angeläget att Skolverket, i enlighet med den hemställan som myndigheten gjort hos regeringen, ges möjlighet att samla in individdata för skolan i sin helhet, dvs. även för modersmålsundervisning och skolgång i specialskola, särskola och sameskola.

Möjligheten att göra detta är angelägen bland annat mot bakgrund av att även för specialskolan och sameskolan finns kunskapsmål som motsvarar dem som gäller för grundskolan. Det kan därför ses som en fråga om likvärdighet att resultaten kan följas upp och att utbildningens långsiktiga effekter kan analyseras för dessa skolformer på motsvarande sätt som för grundskolan – trots att det här i flera fall handlar om förhållandevis små elevgrupper. Enligt vår uppfattning ger den lagstiftning som omgärdar uppgiftsinsamling och redovisning av statistik ett nödvändigt och tillräckligt skydd för individen.

10.3.3 Uppgifter om undervisningens förutsättningar och genomförande

Med undervisningens förutsättningar avses här förhållanden som har att göra med undervisningstiden och dess fördelning på olika årskurser, sammansättningen av elevgrupper, förekomsten av särskilt stöd samt lärarnas formella kompetens etc. Med genomförande avses t.ex. skolans arbetssätt och användning av läromedel etc.

⁵¹ Skolverkets regleringsbrev för 2013.

Planerad och faktiskt genomförd undervisningstid

Bland annat från forskarhåll har framförts att det är angeläget att kunna utvärdera hur resultaten påverkas av skolornas fördelning av undervisningstiden. Det finns dock inte någon nationellt insamlad information om detta.

För grundskolan gäller enligt skollagen⁵² att den totala undervisningstiden för varje elev ska vara 6 785 timmar. Därutöver framgår av bilaga till skollagen hur timmarna ska fördelas på olika ämnen eller ämnesgrupper. Både totalt och per ämne eller ämnesgrupp gäller detta för hela grundskolan, dvs. över nio år. Av bilagan framgår också att av det totala antalet får 600 timmar användas för skolans val. Därvid får antalet timmar per ämne eller ämnesgrupp minskas med högst 20 procent. Detta gäller dock inte för svenska eller svenska som andraspråk och för matematik. Enligt skolförordningen⁵³ beslutar huvudmannen efter förslag av rektorn om fördelning av undervisningstiden mellan årskurserna.

Huvudmannen har således en stor frihet att fördela tiden inom ramen för timplanen. Det gäller särskilt de huvudmän som genom beslut enligt en särskild bestämmelse⁵⁴ är undantagna från skyldigheten att följa timplanen. Samtliga skolor och huvudmän måste dock ta hänsyn till ämnenas kursplaner vid planeringen av undervisningen. Med ett par undantag har samtliga ämnen ett centralt innehåll angivet i kursplanen för årskurserna 1–3, 4–6 respektive 7–9. Hem- och konsumentkunskap och moderna språk saknar centralt innehåll för årskurserna 1–3⁵⁵. Det går således inte att helt avstå från att undervisa i ett ämne under de tidigare åren, för att ta igen detta senare.

Bestämmelserna om undervisningstiden infördes i denna form huvudsakligen i och med 1994 års läroplan (Lpo 94). Läroplanskommittén betonade i sitt slutbetänkande vikten av att läroplan, kursplaner och timplan skulle lämna ett friutrymme för professionell lokal utveckling av skolan. Regeringen underströk att timplanens bestämmelser inte skulle få utgöra ett hinder för pedagogiskt nytänkande och förändring. I och med Lpo 94 kom därför målen

⁵² 10 kap. 5 § skollagen (2010:800).

⁵³ 9 kap. 4 § skolförordningen (2011:185).

⁵⁴ 10 kap. 6 § punkt 4 skollagen (2010:800).

⁵⁵ Detta gäller dessutom för teckenspråk för hörande, som kan väljas inom språkvalet i årskurserna 4–9.

för utbildningen att betonas starkare än undervisningstiden i det enskilda ämnet.

Frånvaron av en detaljerad timplan har medfört att det sedan 1994 inte funnits data om fördelningen av undervisningstiden. Det saknas även jämförbar information på huvudmannanivå. Dels samlar inte alla huvudmän (kommunala och fristående skolor) in dessa uppgifter, dels är inte principerna för insamling de samma.

Hösten 2011 införde dock Skolverket en ny uppgiftsinsamling av grundskolornas *planerade undervisningstid* per ämne och årskurs. Skolorna ska uppge den tid som schemaläggs. Det sker däremot ingen insamling av uppgifter om faktiskt genomförda lektioner. Skolverket har främst använt de insamlade uppgifterna i ett regeringsuppdrag om den garanterade undervisningstiden, men ny officiell statistik över den planerade undervisningstiden ska tas fram under 2014. Det kommer enligt Skolverket att göra det möjligt att analysera uppgifterna om skolors planerade undervisningstid i kombination med skolornas resultatuppgifter. Uppgifter om planerad undervisningstid samlas dock in per årskurs och inte på individnivå.

Utrédningen bedömer att det är angeläget att kunna utvärdera hur skolornas fördelning av undervisningstiden påverkar resultaten. Den insamling som Skolverket nu initierat är värdefull och motsvarar en rimlig ambition för en löpande insamling. Rapporteringen i skolornas administrativa system gör att det knappast är möjligt att med någon kvalitet löpande samla in uppgifter om faktiskt genomförda lektioner, vilket annars vore att föredra.

För en mer fördjupad analys räcker dock, enligt vår bedömning, inte uppgifter som utgår från en planerad schemaläggning. Insamling av uppgifter om faktiskt genomförda lektioner bör i stället genomföras i form av särskilda studier som genomförs vid en specifik tidpunkt. Även sådana undersökningar kan genomföras med viss regelbundenhet och därmed också utgöra en del av det nationella uppföljningssystemet. Här måste dock nyttan vägas mot kostnaden för en sådan insamling.

Frågan om hur uppgifter om undervisningstiden kan samlas in är dessutom beroende av vad som händer med bestämmelserna om grundskolans timplan i framtiden.

Undervisningsgruppens sammansättning

Från flera håll har vikten av uppgifter om undervisningsgruppens sammansättning (klasstorlek och vilka klasskamrater som ingår i undervisningsgruppen) betonats då sådana faktorer kan ha betydelse för elevernas studieresultat. För att bättre kunna studera sambanden krävs, menar man, bättre och mer heltäckande information om vilka undervisningsgrupper som varje elev tillhör, dvs. möjligheter att göra en koppling mellan elev och undervisningsgrupp.

Skolverket har tidigare samlat in uppgifter om klasstorlek. Insamlingen upphörde dock när regelverket ändrades på 1990-talet. I den nuvarande insamlingen av uppgifter om elever i grundskolan ska skolorna rapportera klassbeteckning för samtliga elever. Skolorna kan fritt namnge klassbeteckningarna enligt den benämning som används lokalt. Skolverket har hittills bara använt uppgiften som en hjälpvariabel vid återkontakter med skolor om felaktig inrapportering.

På grund av kvalitetsbrister i uppgifterna om klassbeteckning har det enligt Skolverket inte gått att publicera uppgifter om klasstorlek. Verket utreder för närvarande vilket sätt som skulle kunna ge kvalitet i uppgiften om klassbeteckning utan att påtagligt påverka uppgiftslämnarbördan. Förutom att använda klassbeteckningen för att beräkna klasstorlek kan uppgiften även ge information om hur klasser sätts samman. Klassbeteckningen kan däremot inte ge någon helt korrekt bild av förutsättningarna för undervisningen, eftersom de faktiska undervisningsgrupperna varierar över såväl läsår, vecka som dag och mellan ämnena.

Utredningen bedömer att det vore värdefullt om det gick att publicera uppgifter om klasstorlek. Detta trots att klass och undervisningsgrupp inte är samma sak. Med begreppet klass menas i första hand en administrativ indelning av eleverna som gäller för en längre tid.

Vill man studera hur undervisningsgrupperna faktiskt ser ut i skolan krävs dock enligt vår uppfattning andra metoder för uppgiftsinsamlingen, t.ex. olika former av observationsstudier. Detta då en undervisningsgrupp kan skapas av lärarna själva utifrån pedagogiska överväganden, ha varierande varaktighet och bestå av en del av eleverna i en klass eller av elever från flera olika klasser. Att behöva redovisa alla de pedagogiskt motiverade konstellationer

som skapas till och från i verksamheten i register skulle medföra ett betydande merarbete för lärarna.

Särskilt stöd

Omfattningen av det särskilda stöd som skolorna ska ge elever som har svårigheter att nå målen för utbildningen är en fråga som har stort intresse för föräldrar, huvudmän och verksamma i skolan.

Hösten 2011 införde Skolverket en ny uppgiftsinsamling av uppgifter om särskilt stöd i grundskolan på aggregerad nivå per skola, årskurs och kön. Uppgifterna avser förekomst av åtgärdsprogram, särskild undervisningsgrupp, anpassad studiegång och studiehandledning på modersmålet. Däremot samlas inte uppgifter in om orsakerna till det särskilda stödet. Den officiella statistik som publicerades hösten 2013 gav endast en översiktlig bild över det särskilda stödet i grundskolan och vilka tänkbara brister som kan finnas på systemnivå.

Skolverket planerar att samla in uppgifter på individnivå om särskilt stöd. Insamlingen avses ske från hösten 2014 för de elevgrupper där man redan i dag samlar in uppgifter på individnivå för andra variabler.

Det bör vidare noteras att uppgifter om stödundervisning även samlas in inom ramen för de s.k. elevpanelerna. Det gäller dock endast för ett urval elever och omfattar inte lika många typer av uppgifter om stöd som den insamling Skolverket planerar.

Utredningen bedömer att den utveckling som Skolverket för närvarande genomför motsvarar de önskemål som man kan ställa på data över särskilt stöd.

Lärarnas formella kompetens

Det har påpekats att Skolverkets register över pedagogisk personal (lärarregistret) har haft brister, som begränsat registrets användbarhet för att göra mer detaljerade analyser.

Skolverket har under de senaste åren på regeringens uppdrag utvecklat registret över pedagogisk personal (lärarregistret) i syfte att få fram

ett mer detaljerat underlag. Målsättningen är, enligt regeringen, att verket ska kunna följa hur stor del av undervisningen som bedrivs av förskollärare respektive lärare som är behöriga för skolformen, årskursen och ämnet. Genom att kombinera uppgifterna från lärarregistret med Skolverkets register över legitimerade lärare kommer det att vara möjligt att följa upp lärarnas formella ämneskompetens nationellt men även på huvudmanna- och skolnivå.

Utredningen bedömer att den utveckling av lärarregistret som Skolverket genomfört motsvarar de önskemål som framställts i fråga om detta register.

Möjlighet att koppla samman lärare och elev

Ökade möjligheter att statistiskt koppla samman elever med de lärare som undervisat dem har efterfrågats. Värdet av uppgifter som belyser vad som – förutom den formella utbildningen – karaktäriserar en bra lärare har framhållits. Man menar att det skulle innebära nya möjligheter att studera lärarnas betydelse för elevernas resultatutveckling och bidra till att ge en ökad förståelse för vilka läraregenskaper som är särskilt viktiga. För att kunna forska om dessa frågor krävs information om vilka lärare som undervisar vilka elever i vilka ämnen.

Rent tekniskt skulle det sannolikt, enligt Skolverket, vara möjligt att koppla samman uppgifter om elever med uppgifter om lärare. Det finns dock ett antal svårigheter som riskerar få till följd att uppgifterna skulle få låg kvalitet.

För det första försvåras kopplingen mellan lärare och elev av att endast elevens personnummer ska registreras i betygskatalogen, inte lärarens. Det är endast den betygssättande lärarens namn och titel som ska registreras⁵⁶. Namn är aldrig en bra identifierare vid koppling av data.

För det andra kommer inte den undervisande lärarens namn att registreras i betygskatalogen i de fall denne är obehörig. I stället registreras den betygssättande lärarens (eller rektorns) namn⁵⁷.

⁵⁶ SKOLFS 2011:123 8 §.

⁵⁷ Bestämmelserna om att en obehörig lärare inte får sätta betyg trädde i kraft den 1 december 2013.

Detta innebär att det inte är möjligt att analysera betydelsen av om undervisande lärare är behöriga eller inte.

För det tredje finns det fall där det inte räcker med att identifiera de ämnen lärarna undervisar i för att koppla elev och lärare i ett ämne. Det gäller då flera lärare undervisar i samma ämnen på en skola och då en lärare undervisar i ett flertal ämnen. Det behövs även en koppling som talar om vilket ämne just den eleven blivit undervisad i för just den läraren.

För det fjärde möter elever som är i behov av särskilt stöd (ca 14 procent av eleverna i grundskolan) andra personalkategorier än sin lärare, både direkt och indirekt. Eleven kan få direkt stöd, t.ex. av resurspersonal eller speciallärare, men den ansvarige läraren kan också få stöd av t.ex. en specialpedagog för att skapa en undervisningssituation för eleven som utgår från elevens individuella behov.

Utredningen bedömer att den efterfrågade möjligheten att koppla samman lärare och elev på registernivå är intressant då den skulle möjliggöra analyser på nationell nivå av vilken betydelse lärares utbildning har för elevers skolframgång och framtida yrkesliv. Det skulle ge möjlighet att analysera betydelsen av att lärare har en lärarutbildning eller betydelsen av att lärare har utbildning i det ämne de undervisar i.

Mot bakgrund av de svårigheter som redovisats ovan är det dock svårt att motivera en utökad uppgiftsinsamling av det slag som efterfrågas. Det är helt enkelt svårt att åstadkomma en koppling av data mellan lärare och elev i de administrativa registren som svarar mot de kvalitetskrav som bör ställas för att data ska kunna användas vid kvalificerad analys. Det skulle dessutom medföra en kraftigt utökad utgiftsbörda för skolorna samt öka kostnaderna för systemleverantörerna.

Framför allt är det svårt att se den verkliga nyttan, givet de problem som registren är behäftade med. Varje elev i registret har under sin skoltid många olika lärare. Vissa av dessa kommer inte att vara möjliga att koppla till eleven. Eftersom kunskapande är en process så räcker det inte att koppla elev och lärare i exempelvis årskurs 9. Hela skoltiden har betydelse för resultaten, varför man åtminstone behöver uppgifter om undervisande lärare vid ett antal tidpunkter under grundskolan, t.ex. årskurserna 3, 6 och 9, och detta per ämne. Det handlar med andra ord om en ansenlig mängd uppgifter som inte registreras i dag. Att betygssättande lärare inte alltid är lika med undervisande lärare är ytterligare ett problem.

Enligt vår uppfattning är det av olika anledningar knappast realistiskt att kraftigt utöka kravet på registrering i de administrativa systemen. En sådan uppgift skulle antingen belasta lärarna själva, i ett läge när ambitionen är att avlasta dem administrativa uppgifter, eller innebära utökade uppgifter för administrativ personal på skolorna. Dessutom skulle sannolikt den utökade registreringen och samkörningen av många lärare uppfattas som en långtgående kontroll av deras arbete.

Även om vi delar uppfattningen att studier av lärarnas och olika läraregenskapers betydelse för elevernas resultatutveckling är viktiga, anser vi således inte att nyttan av att kunna koppla uppgifter om lärare och elev i de administrativa registren uppväger kostnaderna och den belastning detta skulle innebära för skolorna. Frågor om vad som kännetecknar en bra lärare är dessutom komplexa till sin natur och kan enligt vår uppfattning bättre studeras med andra metoder.

Utredningen konstaterar samtidigt att vissa uppgifter redan finns i skolornas register – även om de i dag inte är sammankopplade i systemen och är exporterbara – som kan samlas in i samband med särskilda studier, företrädesvis för ett representativt urval av skolor.

Skolans arbetssätt

Förutsättningarna för att studera vad som fungerar i klassrummet anses från flera håll i dag vara begränsade. Man noterar att det saknas uppgifter om hur rektorer och lärare använder sin tid i skolan och om vilka pedagogiska arbetssätt som tillämpas.

Forskare har i olika sammanhang efterfrågat kompletterande uppgifter om hur undervisningen bedrivs, gruppindelningar, åtgärder för att individualisera undervisningen, pedagogiska metoder, läxor, prov, samverkan med föräldrarna, etc. Man vill bland annat kunna studera hur detta samspelar t.ex. med lärares kunskaper. Sådan information skulle enligt forskarna kunna samlas in genom exempelvis enkätundersökningar som riktas till ett (roterande) urval av skolor och där enkäterna är utformade så att informationen görs kvantifierbar. Även uppgifter om klassrumsklimat, elevernas motivation, samt hur mycket tid de lägger på sitt skolarbete i skolan och på fritiden, skulle enligt forskarna kunna bidra till ökad kunskap om förutsättningarna för lärandet.

Det nationella uppföljningssystemet kan i dag belysa utbildningens förutsättningar och resultat. Däremot innehåller inte systemet några uppgifter om olika undervisningsprocesser, dvs. om vad som sker i klassrummet, eller hur det upplevs av eleverna. Här aktualiseras i än högre grad än för uppgifter om undervisningens förutsättningar frågan om vilka uppgifter som bör samlas in som registerdata, vad som kan samlas in genom periodiskt återkommande uppföljningar och vad som behöver vara föremål för särskilda studier (ad hoc).

Processerna i klassrummet (frågor om arbetssätt, läromedelsanvändning, klassrumsklimat, elevernas motivation, m.m.) kan belysas med olika metodiska ansatser och tekniker.

Exempelvis har Skolverket vart tredje år sedan 1993 genomfört en enkätundersökning – *Attityder till skolan* – bland elever i årskurs 7–9 samt bland lärare i hela grund- och gymnasieskolan⁵⁸. Vissa år har undersökningen kompletterats med en föräldraenkät. Attitydundersökningen syftar bland annat till att bidra till en fördjupad diskussion om skolan och ge signaler till ytterligare analyser. Undersökningen innehåller uppgifter om lärares kompetens och skicklighet, skolans arbetssätt och miljö för lärande, stöd, motivation, trivsel, trygghet, stress och krav i skolan samt kränkningar.

Utredningen har övervägt om Skolverkets arbete med *Attityder till skolan* kan utvecklas metodologiskt och i fråga om innehållslig bredd för att möta forskarnas behov. Vi konstaterar dock att attitydundersökningen som den hittills varit utformad inte kan utgöra en utgångspunkt för att möta forskarnas efterfrågan på ytterligare data om skolans arbetssätt. Attitydundersökningen görs för det första på ett för litet urval för att det ska vara möjligt att bryta ner data på lägre nivåer, för det andra måste man ange på förhand vilka samkörningar man planerar att göra. Genom att svara på enkäten ger varje individ i urvalet sitt samtycke. Personnummer (löpnummer) finns förvisso för de individer som ingår i urvalet och för de enkätsvar som inkommit kan samkörning göras, men när körningen gjorts avkodas uppgifterna, dvs. personnumret gallras bort. Detta medför att uppgifter i attitydundersökningarna därefter inte kan samköras med uppgifter i andra register, vilket begränsar attitydundersökningens värde för forskningen.

Om Friskolekommitténs förslag om en ”kundundersökning” för skolan realiseras kan detta komma att påverka attitydunder-

⁵⁸ Skolverket (2013) *Attityder till skolan 2012*. Rapport nr. 390.

sökningens utformning. Enligt uppgift från Skolverket kan denna komma att få ett mer avgränsat innehåll jämfört med i dag. Även Skolinspektionens enkätundersökningar till elever, lärare och föräldrar bör vägas in i detta sammanhang. Syfte med dessa är dock att ge underlag inför tillsyn och enkäterna är anonyma, bl.a. för att hålla uppe svarsfrekvensen.

Utöver Skolverkets återkommande attitydundersökning finns också sedan lång tid vissa longitudinella data med koppling till undervisningen i de s.k. elevpanelerna som SCB samlar in i samarbete med bl.a. UGU-projektet vid Göteborgs universitet. Det gäller uppgifter om bl.a. stödundervisning, hemspråksundervisning och språkval. Inom ramen för UGU samlas dessutom uppgifter in om bl.a. inställning till skolan och skolmotivation för elever i årskurs 6 samt genom enkäter till lärare, som undervisar vissa kohorter av elever, personliga uppgifter relaterade till arbetet som lärare, klassstorlek och klassammansättning, undervisningen samt skolans kvalitet.

Utredningen bedömer att det är viktigt att undersöka olika möjligheter för att öka kunskapen om skolans arbetssätt. Det bör vara fullt möjligt att belysa vissa aspekter av skolans arbetsätt genom exempelvis enkätundersökningar som riktas till ett (roterande) urval av skolor och där enkäterna utformats så att svaren är möjliga att omvandla till kvantifierbar information på skolnivå. För att förbättra informationsläget inom detta område krävs att en ny undersökningsdesign utvecklas. För Skolverket som statistikansvarig myndighet är detta en resursfråga. Myndigheten måste också ta hänsyn till skolhuvudmännens uppgiftslämnarbörda.

Vi menar dessutom att det finns anledning att vara försiktig med vilken typ av direktinsamlade uppgifter som registreras i det löpande statistiksystemet. Svar på frågor om hur undervisningen organiseras eller genomförs, kan vara känsliga för "tidens gång". Sättet att tolka en fråga och att besvara den kan påverkas av att kontexten runt frågan och runt den som besvarar den förändras. Dessutom måste det vägas in att det blivit allt svårare att uppnå en hög svarsfrekvens på enkätundersökningar, särskilt om personnummer/löpnnummer registreras.

Vår bedömning är att skolans arbetsätt är ett område där det finns anledning för olika aktörer att successivt utveckla nya instrument eller att komplettera redan befintlig uppgiftsinsamling. Det gäller inte endast Skolverket. Vill man bygga upp longitudinella

register med data av detta slag bör även övervägas vilken roll t.ex. UGU kan spela. Även i detta fall blir det en fråga om finansiering.

Läromedel

Uppgifter om vilka läromedel som används har efterlysts. Uppgifterna skulle kunna användas för att utvärdera effekterna av läromedel och arbetssätt.

Utredningen bedömer att det är av stort värde att kunna följa utvecklingen på läromedelsområdet. Detta bl.a. mot bakgrund av den stora frihet som skolan och läraren har att välja läromedel och besluta om hur dessa ska används i undervisningen. Denna frihet påverkar samtidigt förutsättningarna att studera läromedelsanvändningen i ett statistiksystem. Utredningen konstaterar att Skolverket redan i dag gör en återkommande enkätstudie med it-inriktning. Eventuella ytterligare uppgifter bör ingå i den återkommande enkätundersökningen eller i särskilda riktade undersökningar av läromedelsanvändningen.

10.3.4 Uppgifter om undervisningens resultat

Frågan om systemet för att följa kunskapsresultaten över tid i den svenska skolan har redan behandlats i föregående kapitel. Här kommenteras därför endast några specifika frågor om de nationella proven samt frågan om tidiga bedömningar.

Mer objektiv rättning av nationella prov

Tillgång till goda data för att kunna göra olika slags jämförelser och analyser, t.ex. av orsakssamband, är en fråga som under senare år lyfts fram från flera håll. För att säkra tillgången till goda data krävs, menar man, en mer objektiv rättning av de nationella proven.

Utredningen uppfattar inte att det ligger inom ramen för direktiven att lämna förslag när det gäller frågan om behovet av mer objektiv rättning av nationella prov. Frågeställningen rör inte data som behövs utan handlar mer specifikt om kvaliteten på de data som finns. Vi har dessutom i föregående kapitel föreslagit att en utredning till-

sätts för att se över hela systemet för att följa kunskapsresultaten i den svenska skolan. En sådan utredning bör i så fall även pröva frågan om hur rättningens kvalitet kan förbättras. Några tankar vill vi dock förmedla.

Att prov ska rättas så objektivt som möjligt är självklart, hur detta ska gå till är dock mindre självklart. Förmågan att bedöma är central i lärarens profession. Att ta ifrån läraren uppgiften att bedöma proven skulle därför kunna uppfattas som en deprofessionalisering av lärarkåren. Dessutom skulle provens förebildliga ambition riskera att gå förlorad eller åtminstone minska.

Samtidigt är kravet på likvärdig bedömning vid rättningen av de nationella proven viktig. De omrättningar som Skolinspektionen har genomfört visar att proven i hög grad bedöms på samma sätt av elevernas lärare och Skolinspektionens omrättare. I två delprov, nämligen de längre skrivuppgifterna i engelska och svenska i årkurs 9 och gymnasieskolans kursprov i samma ämnen, finner man dock större avvikelser mellan lärarnas ursprungliga bedömning och omrättarens. Det går dock inte att avgöra vilken rättning som är ”rätt” eller ”fel”, endast att det finns en spridning i bedömningen. Det är också ett välbekant fenomen att variationen i bedömningen ökar vid frågor som kräver längre självproducerade svar. Iakttagelsen ligger i linje med svensk och internationell forskning om bedömarreliabilitet.

Utredningen bedömer, i likhet med Skolverket, att det finns goda möjligheter att utveckla och, om det behövs, reglera formerna för kollegial bedömning i Sverige. Här finns förebilder att tillgå, t.ex. i England som Skottland. Man talar där om *modering* och har skapat olika institutionella arrangemang för detta. En möjlighet är att utforma ett system där prov från en skola lottas på lärare på en eller flera andra skolor för extern medbedömning.

Insamling av provresultat på poängnivå

Forskare har efterfrågat att rapporteringen av resultat på nationella prov genomgående borde ske på poängnivå. För vissa årskurser samlas resultaten i dag endast in på betygsnivå. För att kunna analysera elevernas resultat på nationella prov med den precision som krävs för att olika svårigheter – och mönster bakom dessa svårigheter – ska framträda, vill man kunna fingradera bedömningen mer

än vad som är möjligt i ett antal betygssteg. Provbetyg döljer även alla underliggande resultatskillnader för elever med samma betyg.

De nationella proven består av ett antal delprov. Betyg sätts på varje delprov. De olika betygen på delproven vägs sedan samman till ett sammanfattande provbetyg. Skolverket samlar in betygen per ämne från de nationella proven i årskurs 6 och 9, både delprovsbetygen och det sammanfattande provbetyget. För årskurs 3 samlas uppgifter om delprovspoäng in, liksom om eleven nått kravnivån för godtagbara kunskaper eller inte.

Innan betyg infördes i årskurs 6, samlades uppgifter in om den poäng eleverna uppnått på respektive delprov både i årskurs 3 och årskurs 6. Syftet med insamlingen var att kunna analysera hur väl eleverna klarade de olika delproven (hur väl de hade förvärvat den förmåga som testades på just det delprovet) samt hur fördelningen såg ut, dvs. hur stor andel av elevresultaten som befann sig i gränslandet av de mål som skulle uppnås, inte ens var i närheten eller hade uppnått kravnivån med råge.

Skolverket har inte några data på uppgiftsnivå. De institutioner som gör proven genomför dock egna insamlingar av fullständiga lösningar av provuppgifterna. Dessa insamlingar baseras emellertid enbart på lösningar från elever som är födda på vissa, fördefinierade, datum, vilket möjliggör vissa fördjupade analyser av elevernas provresultat. Vad vi har förstått varierar det sätt som insamlingen sker mellan institutionerna och urvalen kan inte förväntas vara så stora att dessa insamlingar svarar mot forskarnas behov av underlag på poängnivå.

Skolverket anser inte att nyttan med poänginsamling står i paritet med den uppgiftslämnarbörda den skulle orsaka. Man menar att den insamling som sker av betyg i årskurserna 6 och 9 i grundskolan och av provbetyg från årskurserna 3, 6 och 9 är tillräcklig för uppföljningens behov. Man menar vidare att den nya betygsskalan med fler steg ger mer preciserade resultat än den tidigare. Insamlingarna av poäng har enligt verket visat sig vara mycket krävande. Även felrapporteringar är vanligt förekommande när det gäller rapportering av poängnivåer.

För en poängrapportering argumenterar bl.a. Jan-Eric Gustafsson m.fl. i en nyligen utkommen SNS-rapport⁵⁹. Författarna konstateras att information från de nationella proven kan förväntas vara av

⁵⁹ Gustafsson, Jan-Eric, Christina Cliffordson och Gudrun Ericson (2014) *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS Förlag.

intresse för uppföljnings- och forskningsändamål och att den därför bör samlas in. Man menar vidare att det är viktigt att insamlingen av provresultat inte begränsas till det sammanfattande provbetyget, och inte heller till de olika delbetygen. Det är i stället nödvändigt att göra storskaliga insamlingar även av de poängtal som eleverna når på de olika delproven. Som stöd för sina slutsatser hänvisar författarna till Koretz⁶⁰ som visat att prestationsmått som sammanfattas i så grova kategorier som fyra till sex betygssteg knappast är användbara för uppföljningsändamål, eftersom prestationsförändringar i allmänhet är relativt små.

Utredningen bedömer att det är angeläget att resultaten från de nationella proven samlas in på poängnivå, för att möjliggöra analyser med högre precision. Vi inser samtidigt att hindren för en sådan rapportering i dag handlar om såväl de administrativa systemens funktionssätt och hur vissa prov är konstruerade⁶¹ som de rutiner och tidsplaner som institutionerna som utarbetar proven arbetar med. Det bör emellertid undersökas om det med rimliga insatser går att kortsiktigt komma till rätta med dessa problem. Utredningen har samtidigt föreslagit att hela frågan om ett nationellt system för kunskapsbedömning och de nationella proven bör utredas i ett sammanhang. Frågan om insamling av resultat på poängnivå bör ingå i uppdraget till en sådan utredning.

Tidiga bedömningar

Forskare⁶² vill kunna utvärdera hur insatser i de lägre åldrarna påverkar elever med olika förutsättningar. För att kunna bedöma effekterna av olika insatser behövs, menar man, mätningar av elevernas förmågor och färdigheter så tidigt som möjligt, gärna redan innan de börjar i skolan, dvs. i förskolan. Det har konstaterats att

⁶⁰ Koretz, Daniel M. (2008) *Measuring Up. What Educational Testing Really Tells Us*. Harvard Press. Cambridge, Massachusetts, London, England.

⁶¹ Bland annat förekommer det att betyg på delprov inte endast grundas på det samlade antalet poäng på delprovet, utan att eleven också måste ha fått visst antal poäng på delar av delprovet för att betyget ska kunna sättas. Sådana konstruktioner försvårar naturligtvis möjligheterna att låta administrativa system beräkna delprovsbetyget.

⁶² Se Allan Svensson (red.) (2011) *Utvärdering genom Uppföljning. Longitudinell individforskning under ett halvsekel*. Gothenburg Studies in Educational Sciences 305. Göteborgs universitet. Se även Camilo von Greiff, Anna Sjögren och Ing-Marie Wieselgren (2012) *En god start – en ESO-rapport om tidigt stöd i skolan*. Rapport till Expertgruppen för studier i offentlig ekonomi, 2012:2.

en del huvudmän redan i dag kartlägger t.ex. barns språkutveckling genom olika diagnos- och bedömningsmaterial.

Utredningen bedömer att denna fråga aktualiserar den grundläggande frågan om samhällets och forskningens nytta i förhållande till den enskildes integritet.

Frågan om insamling av tidiga bedömningar av kognitiv förmåga skiljer sig på flera sätt från de övriga önskemålen om ytterligare data. Utredningen har i ett tidigare avsnitt redovisat den komplicerade lagstiftning som gäller inom området. När det gäller exempelvis de tester som görs inom ramen för UGU-databasen krävs den enskildes samtycke till att delta. Sannolikheten för stora bortfall är stor om detta skulle göras i nationell skala. Att genom lagstiftning tvinga föräldrarna att låta sina barn delta i tester av barnens kognitiva förmåga förefaller att gå utöver det rimliga. Enligt vår uppfattning lämpar sig denna typ av uppgiftsinsamling bättre för den typ av forskningsdatabaser som exempelvis UGU utgör, och där de deltagande har gett sitt medgivande.

Även den efterfrågan som finns på insamling av resultat från de olika diagnos- och bedömningsmaterial som ibland används i förskolan för att framför allt kartlägga barns språkutveckling anser vi problematisk. Det är en sak vad enskilda förskolor och skolor gör lokalt för att diagnosticera barn och elever, t.ex. i samband med planering och genomförande av undervisning. Det är en helt annan sak vad som är lämpligt att samhället samlar in på nationell nivå – särskilt om detta sker på individnivå för hela populationer. Det går knappast att samla in sådana data utan att bli mer eller mindre föreskrivande om att det ska genomföras denna typ av diagnoser och hur de ska genomföras. Att genomföra särskilda studier som utgår från resultaten från befintliga diagnoser bör däremot vara möjligt under förutsättning att föräldrarna har gett sitt godkännande till det.

10.3.5 Kostnader och elevresultat

Myndigheter och forskare har konstaterat att huvudmännen använder sig av olika system för att hämta in uppgifter om elever och lärare och föreslår att dessa system vidareutvecklas för att möjliggöra nationell insamling av uppgifter om elever och lärare.

Forskare har konstaterat att det finns begränsad kunskap om effekterna av kompensatoriska inslag i huvudmännens resursfördelning och av hur olika typer av resurser påverkar elevers resultat. Enligt forskare vore det önskvärt att kostnadsdata samlas in på skolnivå, gärna uppdelat på olika utgiftsposter. Man önskar också att få tillgång till information om de principer som ligger bakom kommunernas resursfördelningssystem.

Utredningen bedömer att det är svårt att göra något åt att kommunerna använder sig av olika administrativa och ekonomiska system⁶³ med mindre än att staten genom lagstiftning föreskriver ett enhetligt system för samtliga kommuner och landsting, något vi bedömer vara mindre realistiskt. Man har vidare olika system för budgetuppföljning. Det är enligt Skolverket svårt att få in jämförbara uppgifter. Detta är givetvis problematiskt.

Det är dock värt att notera att Norge valt en annan väg. Där inleddes, i samband med övergången till målstyrning under första halvan av 1990-talet, ett utvecklingsarbete av statistiken inom kommunsektorn under rubriken KOSTRA (Kommun-Stat-Rapportering)⁶⁴. Arbetet inriktades på att skapa ett sammanhållet produktionssystem för definitionsmässigt enhetlig ekonomisk statistik och verksamhetsstatistik. Från och med 2002 ingår alla kommuner och fylken i KOSTRA.

Även OECD⁶⁵ har noterat att de svenska kommunerna har olika system för datainsamling, att de inte är standardiserade och inte kan användas för att göra jämförelser mellan kommunerna. Detta gäller även för de fristående skolorna. Varje utbildningsanordnare samlar in sina egna data och det finns inte någon strävan att harmonisera denna datainsamling. Även på skolnivå är datasystemen mycket olika och ojämna och en del skolor saknar system för att hantera data.

OECD konstaterar att det inte finns någon nationell plan för att hantera frågan om standardisering av system för datainsamling i syfte att möjliggöra jämförelser av detta slag. OECD konstaterar slutligen att varken Sveriges kommuner och landsting eller Fristående skolors riksförbund har tagit några initiativ för att standardisera eller aggregera data.

⁶³ När det gäller elevdata arbetar SCB enligt uppgift från Skolverket mot 33 olika systemleverantörer.

⁶⁴ SOU 2009:25 *Samordnad kommunstatistik för styrning och uppföljning*.

⁶⁵ OECD (2011) *Reviews of Evaluation and Assessment in Education. Sweden*.

När det gäller specifikt analyser av hur resursfördelningen påverkar elevers skolresultat är det samtidigt värt att notera att den stora posten, och sannolikt även den viktigaste, utgörs av undervisande personal. Den går att mäta på annat sätt än i utgiftstermer, varför det ändå bör vara möjligt att genomföra meningsfulla analyser.

Frågan om hur kommunernas resursfördelningssystem påverkar elevernas resultat kan med nuvarande redovisningsformer endast studeras genom särskilda studier. Skolverket har i flera rapporter pekat på komplexiteten i att försöka koppla kvantitativa data om ekonomiska resurser till kvaliteten på undervisningen och elevernas studieresultat⁶⁶. Det har också visat sig svårt att identifiera tydliga alternativa resursfördelningssystem. Möjligheterna att belysa detta i ett registerbaserat statistiksystem måste därför anses begränsade.

10.3.6 Utveckling, modernisering och resurser

Vid utvecklingen av uppföljningssystemet måste ställning tas till vilken information som lämpar sig för ett nationellt uppföljningssystem och vilken information som bör hanteras vid sidan av detta. Det innebär dock inte att vi utesluter möjligheten att införliva mer data från enkäter (urval eller totalundersökningar) i systemet, under förutsättning att frågor bedöms ha en sådan karaktär att den information de ger upphov till är stabila över tid. Detta är en fråga som bör bedömas av Skolverket som statistikansvarig myndighet.

Vi uppfattar att det sammantaget finns olika utvecklingsmöjligheter för det nationella uppföljningssystemet. Förutom de förändringar som följer av regeringens reformer finns också ett behov av att modernisera uppföljningssystemet som helhet. Användarna ställer stora krav på såväl statistikens innehåll och aktualitet som användarvänligheten. För att tillgodose användarnas behov behöver Skolverket dels samla ny data och utveckla ny statistik, vilket har diskuterats ovan, dels utveckla en modern, sammanhållen publiceringsdatabas med uppgifter om alla verksamhetsformer på alla befintliga redovisningsnivåer (riket, län, kommun, skola, huvudman).

I praktiken måste dock Skolverket i dag i första hand fokusera på driften av det befintliga systemet. Detta är resurskrävande. Vad vi kan bedöma saknas därför tillräckliga resurser för att utveckla och modernisera systemet. Som redan nämnts måste dessutom

⁶⁶ Se t.ex. Skolverkets rapport 391 (2013) *Kommunernas resursfördelning till grundskolor*.

hänsyn tas till huvudmännens uppgiftslämnarbörda, vilket innebär att ny uppgiftsinsamling bör motsvaras av avlastning i någon annan del av systemet.

Den infrastruktur som uppföljningssystemet utgör behöver därför enligt vår uppfattning en finansiell förstärkning för att det ska vara möjligt att följa utvecklingen inom hela skolväsendet. Det gäller särskilt efter de senaste årens reformer. Förutom den löpande, årliga kostnaden för uppgiftsinsamling och statistikframställning uppstår även initiala kostnader för utveckling och anpassning av de befintliga insamlings- och granskningssystemen samt för att ta fram ny statistik.

Skolverket har i dag inget särskilt anslag för insamling av data för statistikändamål, utan resurserna för detta bestäms genom verkets interna budgetprioriteringar. Sett i perspektivet av myndighetens samlade åtagande kan knappast förväntas några större omfördelningar mellan olika verksamhetsområden till förmån för utveckling av uppföljningssystemet. Vid eventuella förändringar av uppföljningssystemet behöver regeringen därför se till att resurser avsätts inte bara för att samla in nya data utan också för att hantera dem vidare.

10.3.7 Inga förslag om författningsändringar

Enligt direktiven ska utredningen om det behövs lämna författningsförslag för att öka tillgången på data som behövs för att i ökad utsträckning kunna utvärdera reformer.

Insamling, lagring och användning av registerdata som avser individer är omgärdade av ett omfattande, och ibland svårgenomträngligt, regelverk som i första hand syftar till att skydda den enskilde individen. Detta är nödvändigt, men påverkar förutsättningarna för utvärdering och forskning.

Mot bakgrund av att det för närvarande pågår två utredningar som arbetar med frågor kopplade till hur förutsättningarna för registerbaserad forskning kan förbättras, har vi avstått från att lämna förslag om registerförfattningar på skolans område.

Vi bedömer dock att exempelvis den s.k. UGU-databasen vid Göteborgs universitet kan komma att beröras av resultatet av utredningarna. Det är i så fall viktigt att regeringen tar hänsyn till den betydelse som denna internationellt unika databas har för forskning

inom flera vetenskapliga discipliner och att de långsiktiga förutsättningarna för UGU-databasen inte äventyras.

10.3.8 Avslutande diskussion

Utvärdering av reformer kan baseras på data som samlats in till det löpande uppföljningssystemet och på särskilt insamlad data. Vilken metod som bör väljas i samband med en reformutvärdering hänger samman med reformens innehåll, vad det är man vill veta och vad man vill använda utvärderingsresultaten till. Kvantitativt inriktade utvärderingar kräver ofta tillgång till registerdata medan kvalitativt inriktade utvärderingar oftast bygger på direktinsamlade data.

Kännetecknande för de *registerdata* som ingår i statistiksystemet är att de samlas in årligen, hämtas ur de lokala administrativa systemen samt att de består av hela kohorter, dvs. alla elever, alla lärare och alla huvudmän. Skolverket samlar i princip inte in registerdata som inte är avsedda att ingå i den officiella statistiken. *Direktinsamlad data* kännetecknas av att den samlas in via enkäter, observationer eller på annat sätt, att insamlingen sker periodiskt (t.ex. vart tredje år) samt att insamlingen sker genom urval. Direktinsamlad data kan ingå i uppföljningssystemet, som i fallet med Skolverkets attitydundersökning, eller genomföras särskilt i anslutning till utvärdering av en specifik reform.

Ett väl fungerande uppföljningssystem med registerdata som möjliggör olika typer av analyser och framför allt bidrar till en systemkunskap är en viktig grund för reformutvärdering. Utredningen konstaterar samtidigt att det ofta för utvärdering av reformer behövs kompletterande uppgiftsinsamling, som utgår från just de frågeställningar som ska belysas i utvärderingen. Det är dessutom ofta värdefullt att utvärderingar som baseras på registerdata kombineras med utvärderingar som grundas på annan slags empiri. Detta för att åstadkomma en tillräckligt bred och fördjupad utvärdering av reformer och andra mer omfattande åtgärder.

När utredningen behandlat de olika önskemål som framkommit om ytterligare data i det nationella uppföljningssystemet har avvägningar fått göras mellan dels de kunskapsintressen som efterfrågan återspeglar, dels de praktiska och ekonomiska förutsättningarna för att genomföra uppgiftsinsamlingen, nyttan av en sådan insamling samt lämpligheten ur bl.a. ett etiskt perspektiv.

Utredningens slutsats är att långt ifrån all information, t.ex. om skolans arbetsformer och undervisningssätt, lämpar sig för att samlas in regelbundet för att bilda registerdata. För att sådana data ska vara användbara för statistiska analyser över tid krävs att de uppgifter som lagras inte påverkas av yttre förändringar eller över tid.

Vi menar dessutom att hänsyn måste tas till att frågor som ställs vid en regelbunden insamling av data i ett nationellt uppföljningssystem kan ha en indirekt styrande effekt på verksamheten. Den svenska skolan och de svenska lärarna har stor frihet att, inom ramen för läro- och kursplanerna, organisera och genomföra undervisningen, välja läromedel etc. Det vore därför olyckligt om det sätt som uppgiftsinsamlingen utformas leder till en indirekt och oavsedd styrning. Den stora variationen i undervisningsformerna gör dessutom att det är svårt att fånga dessa med mindre än att fördjupade studier genomförs på plats.

I flera fall har önskemålen om ytterligare datainsamling handlat om att öka den information som hämtas ur lokala administrativa register. Det handlar t.ex. om planerad och genomförd undervisning, undervisningsgruppens sammansättning samt om möjligheterna att koppla samman lärare och elev.

Utredningen anser att det vore värdefullt om det gick att hämta fler uppgifter från de administrativa systemen än vad som är möjligt i dag. Flera av de uppgifter som efterfrågas lämpar sig i princip och till sin karaktär för denna typ av registrering, t.ex. uppgifter om planerad och genomförd undervisning respektive undervisningsgruppernas sammansättning. Som vi bedömt ovan blir dock sannolikt kvaliteten i sådana data inte tillräckligt god, helt enkelt för att skolorna inte registrerar de uppgifter man ur ett utvärderingsperspektiv skulle kunna önska att de gör, och inte heller är skyldiga att göra det. Att ålägga skolan, och i praktiken läraren, en skyldighet att utöka registreringen i de administrativa systemen är inte realistiskt i ett läge där ambitionen i stället är att minska den administrativa bördan. Skolverket bör dock undersöka möjligheterna att successivt utöka uppgiftsinsamlingen från de administrativa systemen i de fall uppgifter registreras i tillräcklig omfattning och med acceptabel kvalitet.

I det fall de administrativa systemen inte kan erbjuda de data som efterfrågas med tillräcklig kvalitet bör annan uppgiftsinsamling väljas, genom enkäter eller genom observationer på plats. Behovet av empiri för att studera t.ex. den faktiska organiseringen av under-

visning i undervisningsgrupper löses troligen bäst genom att data samlas in vid ett urval av skolor genom enkäter eller observationsstudier.

Utredningen konstaterar vidare att ändamålsbestämmelserna i PUL tar sikte på bl.a. den vetenskapliga användningen av administrativa data. PUL medger vetenskaplig användning av data som ursprungligen samlats in för andra ändamål. Den uppgiftsinsamling som efterfrågas här handlar dock inte om ett sådant återbruk av data som samlats in för andra, lokala administrativa ändamål, utan om att samla in och sammankoppla data för allmänna forskningsändamål utan att detta har något med de lokala behoven att göra.

Vissa uppföljnings- och utvärderingsbehov förutsätter att man använder sig av direktinsamlingsmetoder, oftast i form av urvalsstudier och ofta genom enkätstudier. Exempel på direktinsamling är regelbundna attitydundersökningar bland elever, lärare och föräldrar som genomförts sedan 1993 samt de longitudinella studier som SCB genomför i samverkan med UGU (elevpanelerna) och den ytterligare insamling som UGU genomför. I båda fallen omfattas uppgifter om bl.a. skolans arbetssätt.

Det finns flera områden som är viktiga för skolans utveckling och där det framförts önskemål om ytterligare data och som kan lämpa sig för direktinsamling. Det handlar t.ex. om undervisningen, dvs. skolans arbetsformer och lärarnas arbetssätt. Data inom detta område är svårt att samla in som registerdata med en rimlig arbetsinsats och det är ofta svårt att få god kvalitet i sådana uppgifter.

Förbättrade underlag av detta slag skulle underlätta möjligheterna att utvärdera reformer. För utvärdering av reformer och andra mer omfattande förändringar kommer dock enligt vår bedömning även att behövas riktade uppgiftsinsamlingar, med frågeställningar relevanta för den aktuella reformen. Detta kan ske genom enkäter eller fallstudier som bygger på observationer av skolans arbetssätt ”på plats”. Sådana fallstudier är förvisso resurskrävande, men kan enligt vår uppfattning bidra med en kunskap av högre kvalitet och leda till djupare insikt än andra former av informationsinsamling.

Det är den statistikansvariga myndigheten, i detta fall Skolverket, som ska ansvara för bedömningen av vilka data som behövs samt göra den nödvändiga avvägningen mellan nytta och kostnader för ny uppgiftsinsamling. I detta ligger också att svara för bedömningen av på vilket sätt nya uppgifter ska samlas in, t.ex. från admi-

nistrativa register eller genom t.ex. enkätstudier. Det är viktigt att Skolverket inför sina bedömningar samråder med forskare, myndigheter och andra aktörer som använder uppgifterna i systemet. Skolverket är dock beroende av de ramar som ges i form av finansiering. Vi menar att Skolverket behöver tillföras resurser för att dels samla ny data och utveckla ny statistik, dels utveckla sättet att redovisa statistiken på.

Det behövs även ytterligare resurser till systemet för att systematiskt analysera den omfattande statistik som samlas in redan i dag. Enligt uppgifter som utredningen fått från Skolverket och forskare finns inom många områden tillgång till mer data än vad man har resurser att bearbeta och analysera. Både den nationella statistiken och underlagen från de internationella studierna PIRLS, PISA och TIMSS skulle kunna bidra med avsevärt mer kunskap. Det är dessutom möjligt att utöka urvalen för de internationella undersökningarna för att möjliggöra en mer fördjupad nationell analys eller att lägga till frågor av nationellt intresse. Detta kostar dock pengar.

11 Organisering för stärkt utvärdering

I kapitel 7 pekar utredningen på behovet av ett nationellt ramverk för utvärdering på skolans område. Av ett sådant ramverk bör framgå hur utvärderingen är organiserad, vilka de centrala aktörerna är samt hur ansvaret är fördelat dem emellan. Det bör också framgå hur och av vem förslag till prioriterade områden för utvärdering tas fram, vem som beslutar om detta och på vilket sätt olika parter inom skolsamhället ges inflytande över detta. Detta är frågor som tas upp i detta kapitel.

11.1 Dagens aktörer i Sverige

I Sverige finns i dag två aktörer med ett direkt ansvar för utvärdering av skolan:

- *Skolverket* ska enligt skollagen (2010:800) på nationell nivå följa upp och utvärdera skolväsendet, övriga utbildningar och verksamheter som står under Statens skolinspektions tillsyn enligt skollagen, och viss annan pedagogisk verksamhet¹. Enligt verkets instruktion² ska myndigheten rapportera till regeringen i de frågor som dessa uppgifter kan ge upphov till. Uppföljningen och utvärderingen ska öka kunskapen om hur utbildningarna och verksamheterna har utvecklats i förhållande till de nationella målen. Myndigheten ska också sammanfatta och publicera resultaten av arbetet med uppföljning och utvärdering. Slutligen ska Skolverket ansvara för officiell statistik³.

¹ 26 kap 24 § skollagen (2010:800).

² Förordning (2011:555) med instruktion för Statens skolverk.

³ Enligt förordningen (2001:100) om den officiella statistiken.

- *Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU)* har sedan februari 2012 enligt sin instruktion⁴ i uppdrag att bland annat utvärdera effekter av olika reformer och åtgärder inom utbildningsväsendet.

Förutom dessa två myndigheter finns som redovisats i kapitel 4 ett antal aktörer som mer marginellt har betydelse för utvärderingen av skolan.

Statskontoret ska enligt sin instruktion⁵ på regeringens uppdrag bistå med utredningar, utvärderingar samt uppföljningar av statlig och statligt finansierad verksamhet samt övergripande frågor om den offentliga förvaltningens funktionssätt. Utredningen har redan i kapitel 4 konstaterat att de ganska fåtal utredningar som Statskontoret under överskådlig tid haft i uppdrag att genomföra kan motiveras med behovet av en i förhållande till den statliga skolförvaltningen oberoende aktör eller på grund av att uppdraget legat inom Statskontorets särskilda kompetensprofil.

Det offentliga utredningsväsendet kan också genomföra utvärdering på regeringens uppdrag.

Skolinspektionen ansvarar enligt skollagen ansvarar för tillsyn och kvalitetsgranskning⁶. Enligt sin instruktion⁷ ska myndigheten dessutom sammanfatta och publicera resultaten av sitt arbete med tillsyn och kvalitetsgranskning. Skolinspektionen har således inte något direkt utvärderingsuppdrag, men arbetar inom kvalitetsgranskningen med en liknande metodik och förväntas bidra till den samlade bilden av den svenska skolan. Skolinspektionen måste därför räknas in i det vi kallat ett ramverk för utvärdering av skolan. Det bör man däremot inte göra med *Riksrevisionen*, trots att även den använder utvärdering som metodik och ibland reviderar den statliga styrningen av skolan. Riksrevisionen har en helt annan roll.

Till dessa aktörer kan läggas *forskarsamhället* som kan genomföra utvärderingar på uppdrag av regeringen, offentliga utredningar eller myndigheter. Motivet för att engagera forskare för att genomföra utvärdering kan vara behovet av att anlita en från den statliga förvaltningen oberoende aktör eller behovet av den kompetens som forskarna besitter i det enskilda fallet. I dessa fall genomför forskarna uppdrag där beställaren styr vad forskaren ska utvärdera.

⁴ Förordning (2012:16) med instruktion för Institutet för arbetsmarknads- och utbildningspolitisk utvärdering.

⁵ Förordning (2007:827) med instruktion för Statskontoret.

⁶ 2 § 26 kap. skollagen (2010:800).

⁷ Förordning (2011:556) med instruktion för Statens skolinspektion.

Man bör inte heller bortse från att den oberoende forskningen bidrar till den samlade kunskapen om skolan. Här spelar *Vetenskapsrådet* en viktig roll. Kommittén för utbildningsvetenskap vid Vetenskapsrådet stödjer efter nationell prioritering den kvalitetsmässigt bästa forskningen och de bästa forskarna inom områden som är relevanta för skolans, förskolans och utbildningsvetenskapens utveckling.

Slutligen bör också nämnas det nya *skolforskningsinstitut* som ska etableras från och med den 1 september 2014. Enligt direktiven⁸ till den organisationsutredning, som för närvarande förbereder för myndighetens inrättande, ska myndigheten ha i uppdrag att systematiskt väga samman och sprida forskningsresultat som bidrar till ökad kunskap om vetenskapligt väl underbyggda och effektiva metoder och arbetssätt i skolväsendet, dvs. att genomföra s.k. systematiska översikter (systematic reviews). Institutets uppdrag är inte att utvärdera. Genom att göra systematiska översikter kommer man dock att bidra med kunskaper om vilka metoder och arbetssätt som fungerar respektive inte fungerar i skolan. På samma sätt som syftet är vid utvärdering handlar det således om ett kunskapande om skolan och skolans verksamhet.

Sammanfattningsvis kan konstateras att det i Sverige finns många aktörer som bidrar till kunskapandet om skolan och som kan ges en roll vid utvärderingen av reformer på skolområdet. Även med en snäv avgränsning av begreppet utvärdering är aktörerna många.

11.2 Internationell utblick

För att få ett perspektiv på det svenska systemet har utredningen valt att belysa utvärderingssystem i några andra länder. Vi har redan tidigare i betänkandet redogjort för hur olika länder byggt upp sina utvärderingssystem och hur de ser på utvärdering av reformer samt på evidensinformerad styrning. Hur de gör det har också konsekvenser för hur länderna har organiserat sin utvärdering av skolan. Framställningen baseras här på litteraturstudier och underlag från bl.a. de besök utredningen gjort i Danmark, Finland, Norge och Skottland samt England, som vi tyvärr inte hann besöka inom de tidsramar vi hade till förfogande, men som är väl dokumenterat. Se vidare bilaga 5 för en noggrannare redogörelse över de olika län-

⁸ Dir. 2014:7 Inrättande av ett skolforskningsinstitut.

dernas utvärdering av skolan och organisering av deras utvärderingssystem.

De beskrivna länderna uppvisar vissa gemensamma drag i hur utvärderingen av skolan organiserats, som skiljer sig från hur organiseringen ser ut i Sverige. I andra avseenden skiljer de sig också tydligt sinsemellan.

Jämfört med Sverige är närheten mellan policyutformning och resurserna för utvärdering och analys stor i flertalet andra länder. Det gäller generellt och för flera av de länder vi besökt, t.ex. Danmark, England och Skottland. Finland och Sverige är mer lika i detta avseende. Norge intar, som vi uppfattar det, en mellanställning med en större närhet mellan ministeriet och den centrala skolmyndigheten än i Sverige. Att Sverige skiljer ut sig så starkt i detta avseende har att göra med den svenska förvaltningsstrukturen, med ett litet regeringskansli och självständiga myndigheter. Detta är en struktur och ansvarsfördelning med en lång historisk tradition.

En annan påtaglig skillnad jämfört med Sverige är att flera av de länder vi besökt har ett helt annat sätt att se på, och organisera för, samverkan med olika berörda parter inom skolsamhället när det gäller utvärdering av skolan. Exempel är *Rådet för utvärdering och kvalitetsutveckling av folkeskolan (Skolerådet)* i Danmark, det *evalueringsråd* som ska svara för den strategiska styrningen av det nybildade *Nationella centret för utbildningsevaluering* i Finland och det *programstyre* som etablerades för utvärderingen av Kunskapslyftet i Norge. I samtliga dessa fall involveras företrädare för såväl huvudmän som fackliga organisationer och i vissa fall företrädare för föräldrar och elever/studerande i styrningen av utvärderingen. I exemplen från Finland och Norge är även forskare starkt företrädare. Även i Skottland förekommer boards med företrädare för olika intressen i styrningen av nationella utvecklings- eller utvärderingsprojekt. Arbetet med att utveckla skolan i Skottland präglas överhuvudtaget av ambitionen om en skolpolitisk konsensus.

Detta sätt att organisera styrningen av utvärderingen av skolan innebär i flera fall att forskarsamhället får en annan och mer uttalad roll. Utredningen har intrycket att det i dessa länder också är vanligare än i Sverige att politiskt initierad utvärdering av reformer på skolans område genomförs av forskare. I Sverige engageras forskare av regering eller myndigheter i första hand i utvärdering av implementering av reformer snarare än av resultatet av desamma.

I Danmark, Norge och England har man, liksom nu är avsikten även i Sverige, skapat organ för att sammanställa och/eller presentera systematiska kunskapsöversikter. Även i Skottland används metodiken. Systematiska översikter utgår från redan genomförd forskning eller kvalificerad utvärdering och kan utgöra viktiga underlag för utvärdering av exempelvis reformer, varför organisationen av dessa insatser nämns här.

När det gäller just utvärdering av reformer skiljer Norge ut sig genom det sätt man organiserat utvärderingen av Kunskapslyftet på – under ledning av ett särskilt programstyre, och med omfattande resurser. Det bör observeras att själva organisationen av detta utformades av den ansvariga myndigheten, efter att denna fått regeringens uppdrag att svara för utvärderingen, och att uppdraget inte påverkades av det regeringsskifte som skedde medan utvärderingsinsatsen pågick. En intressant konsekvens av det sätt som Norge de senaste decennierna utvärderat reformer på, genom omfattande projekt som involverar forskarsamhället, är att det vuxit fram starka forskningsmiljöer inriktade på vad vi skulle kalla för policy-relevant forskningsbaserad utvärdering.

I Danmark, England och Skottland sker utvärdering av reformer mer ad hoc, men det sker. Finland har inte genomfört några mer omfattande reformer på länge, varför det är svårt att ha någon uppfattning om hur en sådan utvärdering skulle utformas.

I Sverige däremot har en stor mängd omfattande reformer genomförts sedan 1990-talets början, utan att det förrän nu tagits initiativ till någon reformutvärdering. När detta nu sker, sker det på initiativ av riksdagen⁹, och handlar om att i efterhand utvärdera effekterna av de reformer som genomfördes i början av 1990-talet. För ingen av de mer omfattande reformer som genomförts sedan 2006¹⁰ har något specifikt utvärderingsuppdrag lämnats från regeringen, varken till Skolverket eller till något oberoende utvärderingsorgan, t.ex. forskare eller utvärderingsinstitut. Sverige saknar sedan länge en tradition att utvärdera reformer på skolområdet, vilket även framgår av kapitel 5 och bilaga 3.

⁹ Riksdagens tillkännagivande om att regeringen borde tillsätta en utredning med uppdrag att analysera och utvärdera de senaste årens stora skolreformer – kommunaliseringen, det fria skolväsendet och fristående skolors etablering (bet. 2010/11:UbU7, rskr. 2010/11:206). Sedan tillkännagivandet har tillsatts en särskild utredare med uppdrag att utvärdera effekterna av kommunaliseringen av skolväsendet (U 2019:09) och givits ett uppdrag till IFAU att utvärdera skolreformerna.

¹⁰ För en redovisning av dessa, se utredningens delbetänkande *Det tar tid – om implementeringen av skolpolitiska reformer* (SOU 2013:30).

Sammantaget uppfattar vi att trenden i vår omvärld är en förstärkt organisering av utvärderingen av skolan samt ett ökat politiskt intresse för utvärdering och reformutvärdering.

11.3 Andra röster om organiseringen av utvärdering

11.3.1 Att nå ut och nå ända fram (SOU 2009:94)

I december 2009 lämnade en utredare betänkandet *Att nå ut och nå ända fram* (SOU 2009:94) till regeringen. Utredningens uppdrag handlade om två delar – policyinriktad utvärdering respektive spridning av forskning med relevans för skolans måluppfyllelse. Utredningen lämnade följande fyra delförslag:

1. Ett *Råd för utbildningsutvärdering* och analys organiserat som en kommitté inom Regeringskansliet. Inriktningen av rådets arbete skulle vara policyinriktade utvärderingar på nationell nivå.
2. En *mäklarfunktion* inom utbildningsområdet med uppdrag att genomföra systematiska översikter av forskningsresultat. Sådana följer vissa standardiserade steg där metoderna är explicita och transparenta och förutsätter en användardialog för att säkerställa relevans och användbarhet etc. Utredaren föreslog ett nordiskt samarbete kring detta, initialt med Danmark som redan har Danish Cleraringhouse.
3. En *funktion för spridning av forskningsresultat*. Utredaren bedömde att Skolverket även fortsatt skulle ha detta ansvar.
4. En *regional organisation* med syfte att se till att information från utvärderingar och forskning sprids och används. Utredaren rekommenderade att de regionala utvecklingscentrumen (RUC) skulle utnyttjas av universitet och högskolor i detta syfte och att deras uppdrag därför skulle utvidgas. Utredaren uppmärksammade att universitetens och högskolornas mer autonoma ställning i förhållande till staten försvårar styrningen av RUC:en.

Något råd för utbildningsutvärdering och analys av det slag som föreslogs av utredaren har inte inrättats. Däremot har IFAU getts det utvidgade uppdrag som redovisats ovan. Det bör dock uppmärksammas att IFAU:s utvärderingar är avgränsade till kvantitativa studier, i första hand effektutvärderingar. Utredarens förslag

innebar ett råd för policyinriktade utvärderingar på nationell nivå som byggde på en bredare syn på utvärdering.

Någon mäklarfunktion i samarbete med andra nordiska länder har inte heller inrättats. Däremot har regeringen tagit initiativ till det nya skolforskningsinstitut som ska etableras från och med den 1 september 2014 och som redovisats ovan. Institutet ska arbeta med att systematiskt väga samman och sprida forskningsresultat som bidrar till ökad kunskap om vetenskapligt väl underbyggda och effektiva metoder och arbetssätt inom skolväsendet, dvs. att bland annat genomföra s.k. systematiska översikter (systematic reviews).

Slutligen bör observeras att Skolverket, helt i enlighet med utredarens tredje delförslag, fortfarande har uppdraget att sammanställa och sprida kunskap om resultat av forskning. Skolverkets uppdrag sammanfaller således i denna del med en av huvuduppgifterna för det skolforskningsinstitut som regeringen ursprungligen aviserades 2012 i forskningspropositionen¹¹. Enligt direktiven¹² till den särskilda utredare som getts i uppdrag att förbereda inrättandet av myndigheten ska myndigheten kunna inleda sin verksamhet den 1 januari 2015.

Det fjärde förslaget, att med stöd av RUC:arna sprida forskningsresultat så att dessa används, har inte lett till något nationellt initiativ. Som en konsekvens av universitetens och högskolornas ökade autonomi har det nationella kravet på att det ska finnas RUC:ar försvunnit. Parterna i de olika regionerna reglerar samarbetet mellan å ena sidan universitet/högskola och å andra sidan skola och förskola. I vilken utsträckning spridning av forskningsresultat ingår i samarbetet varierar. Mer indirekt bygger naturligtvis de kurser och den kompetensutveckling man erbjuder på forskning.

11.3.2 Institutionaliseringsen av utvärdering – ett fenomen i tiden

Utvärdering var ursprungligen en insats som genomfördes vid ett specifikt tillfälle efter beslut av någon instans som hade behov av den. Det kunde vara en myndighet, en kommun, ett företag eller en annan verksamhet. Utvärderingen utfördes av en eller flera indivi-

¹¹ Prop. 2012/13:30 *Forskning och innovation*.

¹² Dir. 2014:7 *Inrättande av ett skolforskningsinstitut*.

der som hade de kompetenser som krävdes för uppgiften. Dessa personer kunde finnas inom eller hämtas utanför organisationen.

Med tiden har utvärdering byggts in som en reguljär del i olika verksamheter. Myndigheter har i sina instruktioner inte sällan utvärdering som en av sina uppgifter. Utvärderingsverksamheten kan därmed sägas ha institutionaliserats. Gröjer¹³ ger i sin avhandling ett exempel på detta från högskoleområdet.

Enligt Gröjer handlade myndigheternas¹⁴ utvärderingsverksamhet inom högskoleområdet först om att ge regeringen beslutsunderlag. Den hade alltså tydliga politiska förtecken. Senare fick den en inriktning mot att ge underlag för förbättringar på högskolorna inom sådana områden där de själva kunde fatta beslut. Utvärderingen kännetecknades alltså av att ge beslutsunderlag på lokal nivå. Denna inriktning har enligt Gröjer förstärkts allt mer. Tillkomsten av Högskoleverket innebar dock enligt Gröjer att utvärderingens syfte att bidra med beslutsunderlag för den politiska nivån återigen tydliggjordes.

Det har också med tiden tillkommit särskilda myndigheter som har utvärdering som sin specifika huvuduppgift. Statskontoret har¹⁵ beskrivit utvecklingen mot att staten organiserar utvärdering i sektorsspecifika myndigheter som är fristående i förhållande till politikens implementering. Myndigheten konstaterar att det är en trend med många olika uttryck. De fristående utvärderingsmyndigheterna har tillkommit under olika tidsperioder, i olika sammanhang och i olika syften.

Enligt Statskontoret presenteras utvärderingar i förarbetena som lösningen på ett antal problem inom skilda områden: effektivitet, styrning, legitimitet, information och kunskap. Med hjälp av utvärderingar

.... hoppas regeringen bl.a. kunna kontrollera myndigheters effektivitet, öka insynen i offentlig verksamhet, få underlag för utveckling av politiken, stöd i styrningen av sina myndigheter och hjälp med att lösa samhällsproblem såsom arbetslöshet och kriminalitet.

Oavsett hur behovet av utvärdering beskrivs är emellertid argumenten för att det behövs en fristående myndighet för att genomföra utvärderingarna i stort sett desamma. Framför allt är det vikten av

¹³ Gröjer, Anette a.a.

¹⁴ De har funnit ett antal myndigheter inom området: Universitetskanslersämbetet (UKÄ), Universitets- och högskoleämbetet (UHÄ), Kanslersämbetet (KÄ) och Högskoleverket. Från och med 2013 heter myndigheten åter Universitetskanslersämbetet.

¹⁵ Statskontoret (2011) a.a.

oberoende som betonas. Här citeras ofta den princip som också Förvaltningskommittén förfäktar och som går ut på att man bör skilja uppgiften att utföra en verksamhet från den att utvärdera resultaten. Ingen annan organisationsform än en fristående utvärderingsmyndighet anses kunna uppfylla kraven på oberoende.¹⁶

Statskontoret menar vidare att även om det i olika förarbeten uttalas att utvärderingsmyndigheterna bör ha utrymme att agera långsiktigt och ta egna initiativ, så är utvärderingsverksamheten hos flera av myndigheterna i stor utsträckning uppdragsstyrd från regeringen. Vissa myndigheter har fått ett större utrymme att agera självständigt. I andra fall har regeringen betonat myndigheternas stabsstödjande roll genom att i myndigheternas instruktioner föreskriva att de löpande ska bistå regeringen med beslutsunderlag och handlingsrekommendationer. Diskussionen om fristående utvärderingsmyndigheter kan därmed, menar Statskontoret, relateras till diskussionen om Regeringskansliets kapacitet för analys, beställning och mottagning av utvärderingar. Ett alternativ till att inrätta fristående utvärderingsmyndigheter med stabsstödjande roller skulle, enligt Statskontoret, kunna vara att lägga motsvarande resurser hos ansvarigt departement. Genom att resurserna förläggs till en myndighet är de emellertid fredade och riskerar inte att ätas upp av den dagliga ärendehantering i Regeringskansliet.

I flera fall har utvärderingsmyndigheterna även andra uppdrag som t.ex. att ansvara för statistikhantering, omvärldsbevakning, analys och metodutveckling. Storleken på utvärderingsmyndigheterna varierar, men Statskontoret bedömer att de generellt måste anses vara små, särskilt i förhållande till de områden som de ska utvärdera.

11.3.3 OECD om organisering och styrning av policyrelevant utvärdering

I en rapport från 2013¹⁷ analyserar OECD sättet att bedriva det centrala regeringsarbetet i tretton länder. Man kommer bl.a. in på frågan om utvärdering av politiska reformer och program. Det bör observeras att iakttagelserna gäller för den statliga förvaltningen generellt, inte endast för skoladministrationen.

¹⁶ Statskontoret (2011) a.a.

¹⁷ OECD (2013) *Value for Money in Government: Sweden*.

De intervjuer som OECD-teamet genomfört i Sverige pekar på två huvudproblem med att lägga ansvaret för utvärdering av policy¹⁸ på självständiga myndigheter på det sätt som sker i Sverige. Man uppmärksammar bl.a. att kopplingen är svag mellan den policy-utvärdering som görs i myndigheterna och det arbete med politik-utveckling och styrning som sker i fackdepartementen. Enligt rapporten indikerade personal i departementen att det är svårt att påverka utvärderingsmyndigheters agenda och att få politiskt relevanta och användbara råd från dem. OECD pekar också på svårigheter att balansera ett ideal om institutionell självständighet och oberoende med kravet att garantera relevanta utvärderingar. Man uppmärksamade också att flera av utvärderingsmyndigheterna inte heller har något tydligt och specifikt fokus på utvärdering utan ofta har en hel rad andra uppgifter.

OECD menar att det primära ansvaret för utvärdering av politiska reformer och program bör ligga på ansvarig minister snarare än på de verkställande myndigheterna, även om de senare naturligtvis bör vara delaktiga under hela processen.

OECD rekommenderar även att Sverige ska överväga att följa det danska exemplet och överföra ansvaret för all reformutvärdering (policy evaluation) till antingen universitet eller till specialiserade forskningsorgan. Man menar att fackdepartementen på detta sätt kan få tillgång till oberoende och relevant forskning om reformer. Man pekar dock på att denna typ av organisering av reformutvärderingen inte är oproblematisk då högskolor och fristående forskningsinstitut principiellt eftersträvar oberoende från regeringens politiska agenda.

11.4 Förslag om ett råd med dubbla funktioner

Utredningens förslag: Ett *Råd för utvärdering av skolan* inrättas med två huvuduppgifter:

1. Löpande bistå regeringen i utvärderingsfrågor avseende skolväsendet. Det innebär t.ex. att
 - utarbeta förslag till ett nationellt ramverk för utvärdering,

¹⁸ Det bör uppmärksammas att det engelska språket gör skillnad på *policy* i bemärkelsen innehåll i en politisk reform eller annan åtgärd och *politik* som avser att besluta om och genomföra en policy, dvs. själva det politiska "spelet". På svenska används uttrycket policy ibland i bemärkelsen riktlinje eller värderingsmässig bas som ligger till grund för ett agerande. Här används ordet policy i den engelskspråkiga betydelsen.

- lämna förslag till nationell plan för utvärdering av skolväsendet, samt
 - inom ramen för en sådan plan vid behov kunna initiera policyrelevant forskningsbaserad utvärdering.
2. Vid större och genomgripande reformer inom skolväsendet bistå regeringen med
- förslag till program för reformutvärdering,
 - att vid behov initiera forskningsbaserad reformutvärdering,
 - att samordna utvärderingsinsatserna i det fall flera aktörer är involverade i reformutvärderingen (forskare, utvärderingsinstitut, myndigheter etc.),
 - att sammanfatta resultat från olika aktörers utvärderingar, samt
 - att sammanfatta identifierade behov av ytterligare stöd till, eller justeringar av, den aktuella reformen.

Rådet organiseras som en kommitté inom Regeringskansliet, med ett begränsat kansli. Rådets ledamöter, dvs. de som ska ha beslutsbefogenhet i rådet, förslås vara forskare som representerar olika forskningsdiscipliner. I övrigt föreslås rådet bestå av sakkunniga som företrädar olika delar av skolsamhället, t.ex. Utbildningsdepartementet, berörda myndigheter och skolvårdsmän, samt företrädare för lärare, rektorer och elever.

Förutsättningarna för ett evidensinformerat politiskt beslutsfattande samt för en kvalificerad utvärdering av politiska reformer kan enligt utredningen förbättras genom en tydligare och mer ändamålsenlig organisering och ansvarsfördelning. Enligt vår uppfattning bör den framtida organiseringen och styrningen bidra till att genomföra regeringens strategier för utvärdering på skolans område. Sådana strategier bör bidra till att det kommer till stånd en kvalificerad utvärdering som bidrar till en förbättrad systemkunskap om skolan samt till att reformer i högre grad än i dag kommer att utvärderas i framtiden.

Ett viktigt instrument för att åstadkomma detta är, som vi ser det, att det utformas ett nationellt ramverk för utvärdering som bl.a. bidrar till en långsiktigt stabil myndighetsstruktur på skolans område där ansvarsfördelningen är tydlig. Ett annat är att det tas fram en övergripande nationell plan för utvärderingen av skolan samt att det vid behov fattas beslut om inriktningen av, och plan

för, specifika utvärderingsstudier. Om ramverk och nationell plan, se vidare kapitel 7. Enligt vår uppfattning bör dessutom olika intressenter på skolans område göras mer delaktiga i de processer som föregår beslut om inriktningen av utvärdering.

Vidare bör den nationella utvärderingen av skolans verksamhet, liksom utvärderingen av reformer på skolans område, kunna bidra till både nationell policyutveckling och lokal skolutveckling. Det innebär att den dels bör förse både den nationella och den lokala nivån med relevant kunskap om resultatet av interventionerna, dels bidra till utvecklingen av skolan utifrån såväl ett nationellt som ett lokalt perspektiv. Att organisera för detta är naturligtvis svårare än om man endast har den centrala nivåns summativa kunskapsbehov för ögonen. Den mer mångfacetterade ambitionen gör det särskilt viktigt att involvera de olika intressenterna på skolans område i arbetet med att utforma utvärderingen av skolan.

Utvärdering har dessutom effekter i sig. Det gäller därför att utforma de olika inslagen i systemet på ett sådant sätt att den lokala verksamheten och utvecklingen inte störs eller att utvärderingen leder till att utvecklingen får en icke avsedd inriktning (oavsedda bieffekter).

Utredningen delar OECD:s slutsats att det är på central politisk nivå ansvaret för att fatta strategiska beslut om utvärderingen av skolan måste ligga. Det är t.ex. regeringen som måste ta ansvar för att reformer utvärderas och att detta sker med den bredd och med de olika perspektiv som skissats ovan. Det kan inte någon av myndigheterna inom området göra.

Skolverket, som annars har i uppdrag att svara för uppföljning och utvärdering av skolan, har t.ex. inte något specifikt uppdrag att utvärdera reformer. Det är svårt att se att myndigheten utan ett sådant uppdrag från regeringen skulle kunna ta initiativ till mer genomgripande reformutvärderingar, då detta skulle innebära att regeringens centrala skolmyndighet oombedd skulle utvärdera regeringens förda politik. Om regeringen efterfrågar och tillför resurser för en sådan utvärdering ställs frågan naturligtvis i en annan dager. Man bör dock beakta att Skolverket är en del av skol-systemet och att det därför i vissa fall kan vara mindre lämpligt att det är just Skolverket som utvärderar en reform. Att myndighetens implementeringsinsatser ska utvärderas av en oberoende aktör är självklart och sker också. Skolverkets roll i styrningen av skolan

gör emellertid att det kan finnas skäl att även i övrigt överväga att lägga vissa former av reformutvärdering på andra aktörer.

Till skillnad från Skolverket har IFAU uppdraget att utvärdera effekter av reformer på utbildningsområdet. Vi har dock redan konstaterat att utvärderingen av effekter av reformer endast utgör en pusselbit i den strategi för utvärdering av skolan och av reformer på skolområdet som behövs.

Bara mängden aktörer som kan ha en roll vid utvärderingen av skolan och av reformer på skolområdet pekar på behovet av en stärkt nationell styrning. Till ovan nämnda myndigheter kan läggas Skolinspektionen, det aviserade Skolforskningsinstitutet, Statskontoret, Vetenskapsrådet samt forskare och andra utvärderingsorganisationer.

Rådet för utvärdering av skolan – med oförändrad myndighetsorganisation i övrigt

Utredningen föreslår att det inrättas ett råd för utvärderingsfrågor inom skolområdet, *Rådet för utvärdering av skolan*. Förslaget är i vissa avseenden inspirerat av det förslag till ett *Råd för utbildningsutvärdering* som föreslogs i betänkandet *Att nå ut och nå ända fram* (SOU 2009:94). Enligt detta tidigare förslag skulle rådet inriktas på policyinriktade utvärderingar på nationell nivå inom utbildningssektorn. Inriktningen av det råd som föreslås här skiljer sig i vissa avseenden från det tidigare förslaget.

För det första är vårt förslag begränsat till utvärderingsfrågor på skolans område, helt enkelt för att vårt uppdrag avser skolan. Vi har mot den bakgrunden inte haft möjlighet att bedöma behoven av ett organ av detta slag för den högre utbildningen. Vi menar dock att rådet bör svara mot hela skolväsendets behov. Det innebär att exempelvis förskolan bör omfattas. För det andra föreslår vi ett råd med två huvuduppgifter – ett grunduppdrag och ett uppdrag kopplat till reformutvärdering.

Grunduppdrag

En av rådets huvuduppgifter bör vara att bistå till regeringen med att utarbeta förslag till det ramverk för utvärdering som föreslås i kapitel 7. Det innebär att rådet kommer att lämna förslag till regeringen om den strategiska inriktningen av den nationella utvär-

deringen av skolan samt föreslå hur ansvarsfördelningen för frågor om utvärdering av skolan bör se ut. Rådet bör ha en god överblick över de olika aktörernas pågående utvärdering av skolan samt, inte minst, över vad som behöver utvärderas.

Rådet föreslås också lämna förslag till den nationella plan för utvärdering som förelås i kapitel 7. Rådet bör inom ramen för den nationella planen, och mot bakgrund av en behovsanalys, ges förutsättningar att initiera policyrelevant forskningsbaserad utvärdering.

Vid reformering av skolan

Rådet föreslås dessutom bistå regeringen i samband med större och genomgripande reformer på skolområdet. Inför genomförandet av en sådan reform bör rådet bistå regeringen med förslag till program för reformutvärderingen. Ett sådant program bör kunna omfatta olika former av forskningsbaserad utvärdering och engagera olika forskningsdiscipliner.

Under en pågående reform bör rådet, med utgångspunkt i det av regeringen beslutade programmet kunna upphandla reformutvärdering av forskare och andra utvärderare samt följa och kvalitets-säkra arbetet inom dessa projekt. Rådet bör också kunna bistå regeringen med att samordna utvärderingsinsatserna i det fall flera aktörer är involverade i reformutvärderingen, t.ex. forskare, utvärderingsinstitut, myndigheter, etc. Målsättningen bör därutöver vara att rådets arbete med reformutvärdering ska kunna resultera i en samlad analys av resultat från de olika aktörernas utvärderingsinsatser som rapporteras till regeringen. En sådan samlad analys kan identifiera behov av att justera reformen eller påverka den nationella strategin för att implementera den.

Detta skulle innebära att rådet i samband med större reformer ges ett uppdrag som i mycket skulle likna det som det norska Programstyret hade vid utvärderingen av Kunskapslyftet i Norge. Om ett mer permanent råd inte inrättas förordar vi att modellen används åtminstone vid utvärderingen av större reformer.

Organisation

Liksom i SOU 2009:94 föreslår vi att rådet ska organiseras som en kommitté inom Regeringskansliet. Vi menar att denna form har flera fördelar. Genom en sådan konstruktion kommer rådet att befinna sig på vad man skulle kunna kalla för ett lagomt avstånd till regeringen – tillräckligt nära för att kunna lämna förslag till regeringens beredning av utvärderingsfrågor, men samtidigt på tillräckligt avstånd för att, liksom en myndighet, vara självständiga i det dagliga arbetet. I samband med genomförandet av de uppgifter som rådet getts av regeringen, t.ex. att samordna och analysera resultaten från reformutvärdering, får det inte råda något tvivel om rådets oberoende. I annat fall äventyras utvärderingens legitimitet och trovärdighet.

En fördel med att organisera rådet i form av en kommitté är att det ger en stor flexibilitet. Under lugnare perioder kan sekretariatet bestå av ett fåtal personer, för att i reformskeden kunna växa för att möta ökade behov av samordning och analys. Vi bedömer dessutom att den flexibilitet som kommittéformen ger kan underlätta rekryteringen av sådana kvalificerade forskare till sekretariatet, som annars kan vara obenägna att lämna den akademiska karriären. I ett sekretariat kan man arbeta under mer begränsad tid eller dela sin tid mellan forsknings- och sekretariatsmiljön.

Rådets sammansättning och medlemmarnas inflytande

Utredningen föreslår att rådet sätts samman av företrädare för olika delar av skolsamhället, dvs. med företrädare för forskare från olika discipliner samt för Utbildningsdepartementet, berörda myndigheter och skolhuvudmän samt för lärare, rektorer och elever. På detta sätt skapas ett forum för de strategiska frågorna om utvärderingen av den svenska skolan. Genom att representera både en verksamhetsmässig bredd och en djup kunskap om utvärdering på skolområdet kan rådet bidra till att utvärderingssystemets olika perspektiv och syften lyfts fram och samordnas. Dessutom kan det bidra till systemets legitimitet.

Ytterligare en fördel med kommittéformen är att den möjliggör att olika företrädare i rådet ges olika typ av inflytande¹⁹. I detta fall

¹⁹ Enligt kommittéförordningen (1998:1474) kan en kommitté bestå av en ordförande och en eller flera ledamöter samt biträdas av sakkunniga, experter och sekreterare. Ledamöter och sakkunniga ska få del av de handlingar som är avsedda för kommittén och har rätt att närvara

förordar vi en ordning där ordförande och ledamöter, dvs. de som fattar kommitténs beslut ska vara välmeriterade forskare med olika inriktning. Vi förordar vidare att övriga representanter bör vara sakkunniga, vilket innebär att de har rätt att ta del av kommitténs underlag, har rätt att delta vid kommitténs möten och överläggningar samt har rätt att lämna ett särskilt yttrande i det fall de inte är eniga med de beslut som fattas. Med en sådan lösning ges myndigheter och företrädare för lärare, skolledare och elever möjlighet att påverka, men gör dem inte direkt ansvariga för de förslag som lämnas till regeringen och de initiativ rådet tar. Vi uppfattar att detta är en ordning som ligger närmare en svensk tradition än om även dessa aktörer skulle göras fullt delaktiga i besluten. Det senare skulle för myndigheternas del innebära en form av rundgång, där de själva var med och i rådet lämnade förslag till sina egna och andra aktörers uppdrag. För organisationerna skulle det leda till att rollfördelningen blev otydlig.

Oförändrade uppgifter för befintliga myndigheter

Det bör observeras att förslaget att inrätta Rådet för utvärdering av skolan innebär oförändrade uppgifter för myndigheterna på skolområdet. De uppgifter som rådet är tänkt att svara för, och de utvärderingar som det kan komma att föreslå eller initiera, genomförs i allt väsentligt inte i dag. Som OECD påpekar har myndigheter som Skolverket många andra uppgifter som tränger ut den mer omfattande och djuplodande utvärderingen. Det innebär att det saknas resurser för den utvärdering som vi anser är nödvändig för att öka kunskapen om den svenska skolan samt för att utvärdera reformer.

Det finns flera anledningar till att vi inte lämnar något förslag som går i en mer långtgående riktning, dvs. föreslår inrättandet av en ny myndighet för utvärderingsfrågor. Vi anser att modellen med ett råd i kommittéform först bör prövas och utvärderas eftersom modellen har stora fördelar. Det behöver dessutom inte vara ett problem att olika typer av utvärdering och kunskapande om

vid kommitténs samträden och att delta i kommitténs överläggningar medan experter ska biträda kommittén i den omfattning kommittén eller ordföranden bestämmer. Ordförande och ledamöter deltar i de beslut som fattas av kommittén. Ledamöter har reservationsrätt. Ledamöter och sakkunniga får skriva särskilda yttranden. Experter slutligen får skriva särskilda yttranden endast i frågor som omfattas av expertuppdraget och om kommittén går med på det.

skolan genomförs av olika aktörer, så länge det är tydligt vem som gör vad och varför. Den ansvarsfördelning som nu utkristalliserar sig bygger på specialisering och innebär att IFAU är den myndighet som har kompetens på i första hand effektutvärdering medan exempelvis det nya institut med inriktning på skolforskning som ska inrättas från januari 2015 kommer att bygga upp kompetens för att systematiskt väga samman och sprida forskningsresultat om metoder och arbetssätt i skolan. Den utredning som ska förbereda inrättandet av myndigheten ska också överväga om denna även bör ha i uppdrag att i begränsad omfattning initiera forskning inom området.

Det föreslagna Rådet för utvärdering av skolan bör dessutom kunna bidra till att Skolverkets roll inom utvärderingen av skolan stärks. Skolverket bör kunna svara för en metodologisk bredd vad gäller utvärderingen av skolan. Myndigheten har dessutom förutsättningar att länka utvärderings- och utvecklingsperspektiven inom samma myndighet. Fördelarna med att bedriva utvärdering i en miljö som även svarar för statistiksystemet, de internationella undersökningarna, de nationella proven etc. är uppenbara.

12 Konsekvenser

Enligt kommittéförordningen¹ ska en utredning redovisa eventuella ekonomiska eller samhällsekonomiska konsekvenser av sina förslag samt föreslå finansiering om förslagen leder till kostnadsökningar eller intäktsminskningar. För nya eller ändrade regler ska kostnads-
mässiga och andra konsekvenser anges i betänkandet². Vidare ska enligt kommittéförordningen eventuella konsekvenser redovisas för det kommunala självstyrelsen redovisas, liksom när ett förslag har betydelse för brottsligheten och det brottsförebyggande arbetet, för sysselsättning och offentlig service i olika delar av landet, för små företags arbetsförutsättningar, konkurrensförmåga eller villkor i övrigt i förhållande till större företags, för jämställdheten mellan kvinnor och män eller för möjligheterna att nå de integrationspolitiska målen. Slutligen framgår av förordningen att regeringen anger närmare i utredningens uppdrag vilka konsekvensbeskrivningar som ska finnas i betänkandet.

I direktiven till föreliggande utredning har regeringen inte angett att det ska göras några sådana konsekvensbeskrivningar. Vi har ändå valt att i det följande diskutera vissa ekonomiska konsekvenser av förslagen. Utredningens karaktär och inriktningen av de förslag vi lämnat gör att vi funnit att det inte varit motiverat att göra konsekvensbeskrivningar för samtliga områden. Utredningen innehåller dessutom ett begränsat antal förslag som föranleder konsekvensbeskrivningar.

¹ 14–16 §§ Kommittéförordningen (1998:1474).

² Enligt 6–7 §§ förordningen (2007:1244) om konsekvensutredning vid regelgivning.

Ekonomiska konsekvenser

De förslag som utredningen lämnat avser

- framtagande av och beslut om ett nationellt ramverk för utvärdering av skolan och framtagande av treåriga planer för utvärdering av skolan
- införande av stickprovsbaserade undersökningar för att följa kunskapsutvecklingen i skolan, samt
- inrättandet av ett råd för utvärdering av skolan.

Rådets kostnader inkluderar kostnader för ramverk och plan

När det gäller ekonomiska konsekvenser hör den första och sista punkten samman, eftersom det är rådet för utvärdering som föreslås ta fram förslag till ramverk respektive plan för utvärdering. Det är således kostnaderna för rådet som behöver beräknas. En grundbemanning bör inte behöva vara större än ett par årsarbetskrafter. Det innebär att en rimlig kostnad rör sig om cirka 2–2,5 miljon kronor per år, inkl. kostnader för ledamöter och sakkunniga.

Kostnaderna för rådet i de fall rådet ges i uppdrag i anslutning till en specifik reform blir naturligtvis avsevärt högre. Exempelvis var de totala kostnaderna för utvärderingen av Kunskapslyftet i Norge cirka 60 miljoner norska kronor. Detta måste dock beräknas särskilt vid varje enskilt reformtillfälle. Vi menar dock att det bör vara självklart att varje reform finansieras, inte endast avseende kostnaderna i verksamheterna, utan också avseende kostnader för implementering och utvärdering.

Kostnaderna för stickprovsbaserade undersökningar måste utredas i ett sammanhang med hela systemet för utvärdering av kunskapsresultaten

Ett införande av stickprovsbaserade undersökningar för att följa kunskapsutvecklingen skulle innebära omfattande kostnader. Skolverket uppskattade i en rapport från 2006³ att det förslag⁴ som samma år lämnats av professor Jan-Eric Gustafsson initialt skulle kosta 100 miljoner kronor och för påföljande cykler 55 miljoner

³ Rapport 286 (2006) *Rullande stickprovsbaserat system för kunskapsutvärdering av grundskolans ämnen*.

⁴ Jan-Eric Gustafsson (2006) *Ramverk för ett system för uppföljning av kunskapsutvecklingen i grundskolan*. Underlag utarbetat på uppdrag av Skolverket.

kronor per tillfälle, dvs. drygt 10 miljoner kronor per år igenomsnitt under antagande att mätningarna i varje ämne upprepades vart femte år. Skolverket bedömde att dessa kostnader var för höga och lämnade ett förslag med vad man menade var en rimligare ambitionsnivå och med rimligare kostnader. Regeringen har ännu inte tagit ställning till förslaget som, mot bakgrund av de senaste årens reformer, nu är delvis inaktuellt. Såväl Lgr11:s nya kursplaner som införandet av betyg i årskurs 6 och nationella prov i flera ämnen och andra årskurser än tidigare har förändrat förutsättningarna.

De gjorda uppskattningarna får ses som räkneexempel. Enligt vår uppfattning är det inte meningsfullt att nu bryta ut frågan om kostnaderna för stickprovsundersökningar ur frågan om utvärderingen av kunskapsresultaten i sin helhet. Den utredning som föreslås få i uppgift att utreda frågan om stickprovsbaserade undersökningar i ett sammanhang med hela systemet för utvärdering av kunskapsresultaten, dvs. även det nationella provsystemet, bör även svara för kostnadsberäkningar och förslag till finansiering av sina förslag.

Behov av att förstärka finansieringen av det nationella uppföljningssystemet

Den infrastruktur som uppföljningssystemet utgör behöver enligt vår uppfattning en finansiell förstärkning för att det ska vara möjligt att följa utvecklingen inom hela skolväsendet. Det gäller särskilt efter de senaste årens reformer. Förutom den löpande, årliga kostnaden för uppgiftsinsamling och statistikframställning uppstår även initiala kostnader för utveckling och anpassning av de befintliga insamlings- och granskningssystemen samt för att ta fram ny statistik. Dessutom saknas tillräckliga resurser för att utveckla och modernisera systemet.

I budgetunderlaget för 2014-2016 menar Skolverket att det behövs en permanent ökning av anslaget med 10 miljoner kronor från och med 2014 för att uppföljningssystemet ska kunna omfatta fler insamlingar än i dag och samtidigt bibehålla den nuvarande kvalitetsnivån. Det handlar om kostnader för utökad uppgiftsinsamling med anledning av beslutade reformer, bl.a. parallella insamlingar under den tid som äldre regelverk fasas ut, fler nationella prov, betyg i årskurs 6 och betyg från teknikprogrammets fjärde år i gymnasieskolan.

Därutöver behövs enligt verket ett tillfälligt tillskott med 4 miljoner kronor för 2014, för att kunna göra justeringar i det befintliga tabellverket.

För att kunna utveckla en ny statistikdatabas, bl.a. mot bakgrund av ökade krav på datatillgänglighet och publicering genom s.k. webbtjänster, bedömer verket att det krävs sammanlagt 20 miljoner kronor under perioden 2014–2016.

Sammantaget handlar detta om dels en permanent förstärkning med 10 miljoner kronor per år, dels tillfälliga förstärkningar med 10 miljoner kronor för 2014, 7 miljoner kronor för 2015 och 7 miljoner kronor för 2016. Utredningen bedömer att detta är lågt räknat.

Skolverket har i dag ingen särskild anslagspost för statistik- och uppföljningssystemet. Medel för förvaltning och utveckling avsätts inom ramen för myndighetens interna budgetprocess. Mot bakgrund av myndighetens samlade uppdrag och det utgiftstryck som vilar på myndighetens anslag är det svårt att tänka sig att medel kan frigöras genom förändringar inom myndighetens nuvarande anslagsramar. En särskild anslagspost för statistik- och uppföljningssystemet skulle tydligare än i dag uttrycka statens ambitioner på området.

Slutligen bedömer utredningen att det är önskvärt att det s.k. UGU-projektet vid Göteborgs universitet ges stabila ekonomiska förutsättningar.

Referenser

- Adamson, Lena (2013) *Kvalitetsutvärdering av högre utbildning – en kritisk granskning av det svenska systemet framtagen på uppdrag av SNS Utbildningskommission. SNS.*
- Albäck, Erik (1996) *Why all this Evaluation? Theoretical notes and empirical observations on the functions and growth of evaluation, with Denmark as an illustrative case.* The Canadian Journal of Program Evaluation, Vol. 11, No. 2.
- Alkin, M. C., & Christie, C. A. (2004). *An evaluation theory tree.* In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists; views and influences.* Thousand Oaks, CA: Sage.
- Alton-Lee, Adrienne (2012) *The Use of Evidence to Improve Education and Serve the Public Good.* Paper prepared for the New Zealand Ministry of Education and the annual meeting of the American Educational Research Association, Vancouver, Canada.
- Amnå, Erik (2013) *Speak truth to Power. Statsvetarna och kommittéväsendet.* Statsvetenskaplig Tidskrift 112 (5).
- Andersson, Christian (2007) *Lärartäthet, lärarkvalitet och arbetsmarknad för lärare.* IFAU, rapport 2007:5.
- Bew, Lord (2011) *Independent Review of Key Stage 2 testing, assessment and accountability.* Final Report. June 2011.
- Biesta, Gert (2007) *Why "What Works" won't work: Evidence-based practice and the democratic deficit in education research.* Educational Theory, Vol. 57, No. 1.
- Black, Paul (1998) *Learning, League Tables and National Assessment: opportunity or hope deferred?* Oxford Review of Education, Vol. 24, No 1.

- Bohlin, Ingemar (2010) *Systematiska översikter, vetenskaplig kumulativitet och evidensbaserad pedagogik*. Pedagogisk forskning i Sverige. Årg. 15, nr. 2/3.
- Buckendale, Chad W., Susan L. Davis, Barbara S. Plake, Stephen G. Sireci, Ronald K. Hambleton, April L. Zenisky, and Craig S. Wells (2009) *Evaluation of the National Assessment of Educational Progress. Final Report*. Center for Educational Assessment. University of Massachusetts-Amherst.
- Breakspear, Simon (2012) *The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance*. OECD Education Working Papers, No. 71, OECD Publishing.
- Brulin, Göran och Lennart Svensson (2011) *Att äga, styra och utvärdera stora projekt*. Studentlitteratur.
- Brunsson, Karin (2000) *Det går inte att styra med resultat*. Statsvetenskaplig Tidskrift, årg. 1203, nr. 3.
- Burns, Tracey and Tom Schuller (ed.) (2007) *Evidence in Education. Linking Research and Policy*. CERI. OECD.
- Böhlmark, Anders och Helena Holmlund (2012) *Lika möjligheter? Familjebakgrund och skolprestationer 1998-2010*. IFAU, rapport 2012:14.
- Böhlmark, Anders och Mikael Lindahl (2012) *Har den växande fri-skolesektorn varit bra för elevernas utbildningsresultat på kort och lång sikt?* IFAU, rapport 2012:17.
- Cabinet Office (1999) *Professional Policy Making for the Twenty First Century*. Report by Strategic Policy Making Team.
- Cabinet Office (2003) *Trying it out. The Role of 'Pilots' in Policy-Making*. Report of a Review of Government Pilots. Government Chief Social Researcher's Office.
- Cable, Vincent (2004) *Evidence and UK Politics*. Article in *Research and Policy Development: Does Evidence Matter?* Meeting Series. Overseas Development Institute.
- Campbell, Carol and Ben Levin (2008) *Using Data to Support Educational Improvement*. Springer Science + Business Media, LLC.

- Carlgren, Ingrid, Ingela Josefsson och Caoline Lidberg (red.) (2005) *Forskning om denna världen II – om teorins roll i praxisnära forskning*. Vetenskapsrådet.
- Cliffordsson, Christina (2004) *De målrelaterade gymnasiebetygens prognosförmåga*. Pedagogisk Forskning i Sverige, 9(2).
- Cliffordsson, Christina (2004) *Betygsinflation i de målrelaterade gymnasiebetygen*. Pedagogisk Forskning i Sverige, 9(1).
- Centrum för skollärandeutveckling (2012) "Att bygga om en rondell i rusningstrafik?" *Uppföljning av skolverkets informationsinsatser vid inåpömlenteringen av skollagen*. Umeå universitet.
- Clarke, Alan (2007) *Evidence-based evaluation in different professional domains: similarities, differences and challenges* i Shaw, Ian F. m. fl. (red.) *The Sage Handbook of Evaluation*. Sage Publications.
- Cochrane, Archibald (1972) *Effectiveness and Efficiency. Random Reflections on Health Services*. London: Nuffield Provincial Hospitals Trust (Reprinted in 1989 in association with the BMJ, Reprinted in 1999 for Nuffield Trust by the Royal Society of Medicine Press, London).
- Commission of the European Union (2007) *Towards more knowledge-based policy and practice in education and training*. Commission staff working document. SEC(2007) 1098.
- Cordingley, Philippa (2004) *Teachers using evidence: Using what we know about teaching and learning to reconceptualize evidence-based practice*. I G. Thomas & R. Pring (red.) *Evidence-based practice in education*. Buckingham: Open University press.
- Coryn, Chris L. S., Lindsay A. Noakes, Carl D. Westine and Daniela C. Schröter (2011) *A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009*. American Journal of Evaluation 32(2) 199–226.
- Curman, Lisa (2011) *Uppföljning och utvärdering av resultat i komplexa situationer. En begrepps- och metodutforskande uppsats med ett utfall från internationellt utvecklingsarbete*. Göteborgs universitet, Förvaltningshögskolan.
- Dahler-Larsen, Peter (2003) *Det politiske i evaluering*. Studies in Educational Policy and Educational Philosophy. E-tidskrift 2003:1.

- Dahler-Larsen, Peter (2005) *Vaikuttavuuden arviointi. (Utvärdering av effektivitet)*. Stakes.
- Dahllöf, Urban (1971) *Svensk utbildningsplanering under 25 år*. Studentlitteratur. Lund.
- Danish Clearinghouse for Educational Research (2007) *Concept Note*, v. 12.
- Davies Philip (2000) *The Relevance of Systematic Reviews to Educational Policy and Practice*. Oxford Review of Education, Vol. 26, Nos. 3 & 4.
- Davies, Philip (2004) *Is Evidence-Based Government Possible?* Jerry Lee Lecture. 4th Annual Campbell Collaboration Colloquium, Washington D.C.
- Davies, Phillip, Kathryn Newcomer and Haluk Soydan (2006) *Government as structural context for evaluation* i Shaw, Ian F., Jennifer C. Greene and Melvin M. Mark (Ed.) *The Sage Handbook of Evaluation. Policies, programs and Practicies*. Sage publications 2007.
- Department for education and skills and Office for standards in education (2004) *A new Relationship with Schools*.
- Department for Education (2010) *The Importance of Teaching. The schools White Paper* 2010.
- Department for Education (2011) *Independent Review of Key Stage 2 testing, assessment and accountability. Government Response*.
- Department for Education (2013) *Analytical Review. The Department (By Steering Group)*.
- Ds 2000:34 *Samhällets grundläggande information – Inventering. Analys. Förslag*. Justitiedepartementet.
- Ds 2005:16 *Att fånga kunlandet om lärande och undervisning. Om villkoren för skollära och lärare att ta del av systematiskt framtagen kunskap om utbildningsverksamhet*. Utbildnings- och kulturdepartementet.
- Education, Audiovisual & Culture Executive Agency (EACEA) (2009) *National Testing of Pupils in Europe: Objectives, Organisation and Use of Results*. Eurydice.

- Ekonomistyrningsverket (2006) *Effektutvärdering. Att välja upplägg*. ESV 2006:8.
- Evans, Jennifer och Benefield, Pauline (2001) *Systematic Reviews of Educational Research: does the medical model fit?* British Educational Research Journal, Vol. 27, No. 5.
- Fazekas, Mihály and Tracey Burns (2012) *Exploring the Complex Interaction Between Governance and Knowledge in Education*. OECD Education Working Paperas, No 67, OECD Publishing.
- Faubert, Violaine (2009) *School Evaluation. Current Practices in OECD Countries and a Literature Review*. OECD Education Working Papers, No. 42, OECD Publishing.
- Finansdepartementet (2013) *Budgetpropositionen för 2014. Utgiftsområde 16. Utbildning och universitetsforskning*.
- Foss Hansen, Hanne och Olaf Rieper (2010) *The Politics of Evidence-Based Policy-Making: The Case of Denmark*. German Policy Studies, Vol. 6, No. 2.
- Foss Hansen, Hanne (2012) *Systematic evaluation governance. New logics in the developement of organisational fields*. Offentlig förvaltning. Scandinaivian.
- Franke-Wikberg, Sigbrit (red.), Urban Dahllöf, Daniel Kallos, Ulf P. Lundgren och Erik Wallin (1989) *Skolan och utvärderingen. Fem professorer tar ordet*. HLS Förlag.
- Fredriksson, Peter, Hessel Oosterbeek och Björn Öckert (2012) *Långsiktiga effekter av mindre klasser*. IFAU, rapport 2012:5.
- Furubo, Jan-Eric (2011) *Effektivitetsrevision – roll, möjligheter och gränser*. Riksrevisionen.
- Glouberman, Sholom and Brenda Zimmerman (2002) *Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like?* Discussion paper no. 8. Commission on the Future of Health Care in Canada.
- Goldacre, Ben (2013) *Building Evidence into Education*. Department for Education, UK.

- Gough, David (2007) *The Evidence for Policy and Practice Information and Co-ordination (EPPI) Centre, United Kingdom* i OECD & CERI (2007) *Knowledge Management. Evidence in Education: Linking Research and Policy*.
- Gough, David, Janice Tripney, Caroline Kenny and Elisabeth Buk-Berge (2011) *Evidence Informed Policymaking in Education in Europe*. EIPEE Final Project Report.
- Gough, David, James Thomas and Sandy Oliver (2012) *Clarifying differences between review designs and methods*. Systematic Reviews 2012, 1:28.
- Gray, J. A. M. (1977) *Evidence-based Healthcare: How to Make Health Policy and management Decisions*. London: Churchill Livingstone.
- von Greiff, Camilo, Anna Sjögren och Ing-Marie Wieselgren (2012) *En god start – ESO-rapport om tidigt stöd i skolan*. Rapport till Expertgruppen för studier i offentlig ekonomi, 2012:2.
- Gröjer, Anette (2004) *Den utvärdera(n)de staten. Utvärderingens institutionalisering på den högre utbildningens område*. Studies in Politics 102. Stockholm University: Department of Political Science.
- Gustafsson, Jan-Eric (2006) *Ramverk för ett system för uppföljning av kunskapsutvecklingen i grundskolan*. Underlag utarbetat på uppdrag av Skolverket.
- Gustafsson, Jan-Eric and Gudrun Erickson (2013) *To trust or not to trust? – Teacher marking versus external marking of national tests*. Educational Assessment, Evaluation and Accountability. Vol. 25, No. 1. Springer.
- Gustafsson, Jan-Eric, Christina Cliffordson och Gudrun Erickson (2014) *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS Förlag.
- Hall, Caroline (2009) *Förlängningen av yrkesutbildningarna på gymnasiet: effekter på utbildningsavhopp, utbildningsnivå och inkomster*. IFAU. Rapport 2009:7.

- Hamilton, Linda (2005) *Assessment as a Policy Tool*. RAND Corporation. Reprint with permission from Review of Research in Education, pp. 25–68, Volume 27, 2003. American Educational Research Association.
- Hanberger, Anders (2001) *What is the Policy Problem? Methodological Challenges in Policy Evaluation*. Evaluation, Vol. 7(1). Sage Publications.
- Hanberger, Anders (2003) *Den dolda utvärderingspolitiken*. Studies in Educational Policy and Educational Policy and Educational Philosophy. E-tidskrift 2003:1.
- Hanberger, Anders (2012) *Framework for exploring the interplay of governance and evaluation*. Offentlig Förvaltning. Scandinavian Journal of Public Administration, 16(3): 9–27.
- Hanushek, Eric and Ludger Woessmann (2010) *The cost of Low Educational Achievement in the European Union*. European Expert Network on Economics of Education. EENEE Analytical Report No. 7.
- Harlen, Wynne (2009) *Reflections on a personal journey in research*. The Association for Science Education.
<http://www.ase.org.uk/resources/scitutors/research/research-wynne-harlen/>
- Harlen, Wynne and Ruth Deakin Crick (2002) *A systematic review of the impact of summative assessment and tests on students' motivation for learning*. Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre).
- Hartman, Laura (red.), Anders Anell, Eva Mörk, Jonas Vlachos, m.fl. (2011) *Konkurrensens konsekvenser – vad händer med svenskt välfärd?* SNS Förlag.
- HM Treasury (2003, updated 2007) *The Magenta Book: guidance notes for policy evaluation and analysis. Magenta Book Background Papers*. Government Social Research Unit.
- HM Treasury (2011) *The Green Book. Appraisal and Evaluation in Central Government*. Treasury Guidance. London:TSO.
- HM Treasury (2011) *The Magenta Book. Guidance for evaluation*.

- HM Treasury (2012) *Quality in policy impact evaluation: understanding the effects of policy from other influences* (supplementary Magenta Book guidance).
- HM Treasury (2012) *Quality in qualitative evaluation: a framework for assessing research evidence* (supplementary Magenta Book guidance).
- Hooge, Edith, Tracey Burns and Harald Wilkoszewski (2012) *Looking Beyond the Numbers: Stakeholders and Multiple School Accountability*. OECD Education Working Papers, No. 85, OECD Publishing.
- Johansson, Thomas (2009) *Etnografi som teori, metod och livsstil*. EDUCARE. Malmö högskola.
- Kahan, Barbara (2008) *Excerpts of Evaluation Frameworks*. Saskatchewan Ministry of Education.
- Karlsson, Ove (1999) *Utvärdering – mer än en metod. En översikt. À jour. En serie kunskapsöversikter från Svenska kommunförbundet*. Nr 3.
- Karlsson, Ove (2003) *Evaluation Politics in Europe: Trends and Tendencies*. Studies in Educational Policy and Educational Philosophy 2003:1.
- Karlsson Vestman, Ove och Inger M. Andersson (2007) *Pedagogisk utvärdering som styrning – En historia från präster till PISA*. Forskning i fokus, Nr. 35. Myndigheten för skolutveckling.
- Koretz, Daniel M. (2008) *Measuring Up. What Educational Testing Really Tells Us*. Harvard Press. Cambridge, Massachusetts, London, England.
- Koretz, Daniel M. (2010) *Implications of Current Policy for Educational Measurement*. Center for K-12 Assessment & Performance Management. Educational Testing Service (ETS).
- Krogstrup, Hanne Katherine (2003) *User Participation in Evaluation – Top-down and Bottom-up Perspectives*. Studies in Educational Policy and Educational Philosophy. E-tidskrift 2003:1.
- Leeuw, Frans L. (2003) *Reconstructing Program Theories: Methods Available and Problems to be Solved*. American Journal of Education, Vol. 24, No. 1.

- Levin, Ben (2004) *Making Research Matter More*. Education Policy Analysis Archives, 12(56).
- Levin, Ben, Creso Sá, Amanda Cooper and Shalini Mascarenhas (2009) *Research Use and Its Impact in Secondary Schools*. CEA/OISE Collaborative Research Project.
- Levin, Ben (2011) *Mobilising research knowledge in education*. London Review of Education, 9:1, 15-26.
- Levinsson, Magnus (2013) *Evidens och existens. Evidensbaserad undervisning i ljuset av lärares erfarenheter*. Göteborgs Studies in Educational Science. Göteborgs universitet.
- Lindgren, Lena (2008) *Utvärderingsmonstret. Kvalitets- och resultatmätning i den offentliga sektorn*. Studentlitteratur.
- Looney, Janet W. (2011) *Alignment in Complex Education Systems: Achieving balance and Coherence*. OECD Education Working Paper, No. 64.
- Jens Ludwig, Jeffrey R. Kling and Sendhil Mullainathan (2011) *Mechanism Experiments and Policy Evaluations*. Journal of Economic Perspectives, Vol. 25, Nr. 3.
- Lundgren (Ulf P. (2002) *Utbildningsforskning och utbildningsreformer*. Pedagogisk Forskning i Sverige, årg. 7, nr. 3.
- Lundgren, Ulf P. (2006) *Political Governing and Curriculum Change – From Active to Reactive Curriculum Reforms. The need for a reorientation of Curriculum Theory*. Studies in Educational Policy and Educational Philosophy. E-tidskrift 2006:1.
- Marklund, Sixten (1981) *Skolsverige 1950-1975. Del 3. Försöksverksamheten*. Utbildningsförlaget.
- Maynard, Alan (2007) *Translating Evidence Into Practice: Why Is It So Difficult?* Public Money & Management.
- Meghir, Costa och Mårten Palme (2004) *Education reform, ability and family background*. IFS Working Papers, W04/10, Institute for Fiscal Studies.
- Mintzberg, Henry (1994) *The Fall and Rise of Strategic Planning*. Harvard business Review.

- Newton, Paul E. (2007) *Clarifying the purposes of educational assessment*. Assessment in Education, Vol. 14, No. 2.
- Niklasson, Laila och Carina Carlhed (2011) *Implementering av Lgr11 – Utvärdering av Skolverkets stöd*. Mälardalens utvärderingsakademi. Mälardalens högskola.
- Nutley, Sandra, Isabel Walter and Huw T. O. Davies (2009) *Promoting Evidence-based Practice: Models and Mechanisms From Cross-Sector Review*. Research on Social Work Practice 2009 19:552. Sage.
- Nutley, Sandra, Sarah Morton, Tobias Jung och Annette Boaz (2010) *Evidence and policy in six European countries: diverse approaches and common challenges*. Evidence & Policy, Vol. 6, No. 2. The Policy Press.
- Oakely, Ann (1998) *Experimentation and social interventions: a forgotten but important history*. British Medical Journal 317(7167).
- O'Day, Jennifer A. (2002) *Complexity, Accountability and School Improvement*. Harvard Educational Review, Vol. 72, Nr. 3.
- OECD (2007) *Evidence in Education. Linking research and policy*.
- OECD (2011) *Country Background Report for Norway*. OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.
- OECD (2011) *Sweden*. OECD Reviews of Evaluation and Assessment in Education.
- OECD (2011) *Denmark*. OECD Review of Evaluation and Assessment in Education.
- OECD (2012) *Country Background Report for the Netherlands*. OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.
- OECD (2013) *Synergies for Better Learning. An international Perspective on Evaluation and Assessment*. OECD Reviews of Evaluation and Assessment in Education.
- OECD (2013) *Value for Money in Government: Sweden*.
- Ofsted (2013) *The framework for the inspection of local authority arrangements for supporting school improvement*.

- Osberg, Deborah and Gert Biesta (2010) *Complexity Theory and the Politics of Education*. Sense Publishers. Rotterdam/Boston/ Taipei.
- Pawson, Ray & Nick Tilley (1997) *Realistic Evaluation*. Sage.
- Plant, Roger (2013) *Analytical Review. Data Systems*. Department for education, UK.
- Prewitt, Kenneth, Thomas A. Schwandt and Miron L. Straf (2012) *Using Science as Evidence in Public Policy*. Committee on the Use of Social Science Knowledge in Public Policy. National research Council of the National Academies. The National Academies Press. Washington, D.C.
- Proposition 1988/89:4 *Skolans utveckling och styrning*.
- Proposition 1990/91:18 *Ansvar för skolan*.
- Proposition 2011/12:176 *Behandling av personuppgifter vid Institutet för arbetsmarknads- och utbildningspolitisk utvärdering*.
- Ramböll (2011) *Country Background Report for Denmark*. OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.
- Ramböll (2011) *Utvärdering av matematiksatningen*. Slutrapportering av ett uppdrag från Skolverket.
- Ramböll (2013) *Evaluering af de nationella test i folkeskolen*.
- Rogers, Patricia (2008) *Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions*. Evaluation 14:29.
- Sahlin-Andersson, Kerstin (2000) *Transnationell reglering och statens omvandling – Granskningssamhällets framväxt*. Föredrag vid konferensen Moderna Tider i Staten den 6 december 2000.
- Sandelowski, Margarete, Corinne I Voils, Jennifer Leeman and Jamie L. Cardell (2012) *Mapping the Mixed Methods-Mixed Research Synthesis Terrain*. NIH Public Access. 6(4): 317–331.
- Ian F. Shaw, Jennifer C. Greene and Melvin M. Mark (2006) *The Sage Handbook of Evaluation*. Sage Publications London, Thousand Oaks, New Delhi.

- Sheppard, S, Lewin, S, Straus S, Clarke M, Eccles MP et al. (2009) *Can We Systematically Review Studies That Evaluate Complex Interventions?* PLoS Med, 6:8.
- Sjögren, Anna (2010) *Betygssatta barn – spelar det någon roll i längden?* IFAU, rapport 2010:8.
- Skolverket (1995) *Likvärdighet i skolan - en antologi.*
- Skolverket (1999) *Att utvärdera skolan.*
- Skolverket (1999) *Uppföljning och utvärdering av försöksverksamheten med ökat föräldrainflytande över val av skolform för utvecklingsstörda barn – Slutrapport. (Dnr. 96:565)*
- Skolverket (2002) *I särskola eller grundskola? Redovisning av ett regeringsuppdrag. (Dnr. 103404)*
- Skolverket (2004) *Internationella studier under 40 år. Svenska resultat och erfarenheter. Aktuella analyser.*
- Skolverket (2005) *De nationella minoriteternas utbildningssituation. Rapport 272.*
- Skolverket (2006) *Rullande stickprovsbaserat system för kunskapsutvärdering av grundskolans ämnen. Rapport 286.*
- Skolverket (2006) *Uppdrag att genomföra och redovisa informationsinsatser om de allmänna råden för individuella utvecklingsplaner. (Dnr. 2005:1450)*
- Skolverket (2007) *Romer i skolan – en fördjupad studie. Rapport 292.*
- Skolverket (2007) *Fem år med maxtaxa. Uppföljning av reformen Maxtaxa och allmän förskola m.m. Rapport 294.*
- Skolverket (2007) *Statsbidrag till personalförstärkningar i skola och fritidshem – slutrapport. Rapport 295.*
- Skolverket (2007) *Hur används individuella utvecklingsplaner? En studie efter införandet av nya bestämmelser. Aktuella analyser.*
- Skolverket (2007) *Hemställan om ändring i förordning om den officiella statistiken. Dnr. 2007:3.*
- Skolverket (2008) *Personalförstärkningar i förskolan. Utvärdering av statsbidrag fördelat under åren 2005 och 2006. Rapport 314.*

- Skolverket (2008) *Redovisning av uppdrag om uppföljning och utvärdering av statsbidraget till personalförstärkningar i förskolan.* (Dnr. 2007:00650)
- Skolverket (2009) *Redovisning av regeringsuppdrag om ämnet juridiska studier i Stockholms kommun. En utvärdering.* (Dnr. 2007:3675)
- Skolverket (2009) *Skolverkets bild av utveckling av kunskapsresultaten i grundskolan och av elevers studiemiljö.* (Dnr. 2008:3010)
- Skolverket (2009) *Redovisning av regeringsuppdrag att ge förslag på hur det nationella provsystemet bör utvecklas och utformas.* (Dnr. 2009:337)
- Skolverket (2010) *Skriftliga omdömen i grundskolans individuella utvecklingsplaner. En uppföljning och utvärdering av skolornas arbete ett år efter reformen.* Rapport 340.
- Skolverket (2010) *Undervisning på engelska. Utvärdering av en försöksverksamhet i grundskolan.* Rapport 351.
- Skolverket (2010) *Redovisning av regeringsuppdrag om yrkestävlingar* (Dnr. 2009:00326)
- Skolverket (2010) *Redovisning av regeringsuppdrag om att genomföra jämställdhetsinsatser i skolan.* (Dnr. 2008:3458)
- Skolverket (2011) *Utvärdering av metoder mot mobbning.* Rapport 353.
- Skolverket (2011) *Kommunalt huvudmannaskap i praktiken. En kvalitativ studie.* Rapport 362.
- Skolverket (2011) *Promemoria om evidens på utbildningsområdet.*
- Skolverket (2011) *Redovisning av uppdrag om att genomföra utvecklingsinsatser i matematik, naturvetenskap och teknik.* (Dnr. 2009:259).
- Skolverket (2012) *Gymnasial lärlingsutbildning de tre första åren 2008–2011.* Rapport 373.
- Skolverket (2012) *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid.* Rapport 374.
- Skolverket (2012) *Högpresterande elever, höga prestationer och undervisningen.* Rapport 379.

- Skolverket (2012) *Beskrivande data 2012. Förskola, skola och vuxenutbildning*. Rapport 383.
- Skolverket (2012) *Redovisning av uppdrag om kvalitetssäkring av nationella prov*. (Dnr. 2011:1242)
- Skolverket (2012) *Redovisning av regeringens uppdrag att följa upp, analysera och redovisa insatser utifrån minoritetspolitikens mål*. (Dnr. 2010:242)
- Skolverket (2013) *Utbildningsresultat. Riksnivå. Sveriges officiella statistik om förskola, skola och vuxenutbildning. Del 1*. Rapport 382.
- Skolverket (2013) *Att fånga undervisningstiden med målen i fokus. Garanterad undervisningstid i grundskolan ur ett huvudmanna- och skolperspektiv*. Rapport 384.
- Skolverket (2013) *Lärarnas yrkesvardag. En nationell kartläggning av grundskolläraarnas tidsanvändning*. Rapport 385.
- Skolverket (2013) *Barn, elever och personal. Riksnivå. Sveriges officiella statistik om förskola, skola och vuxenutbildning. Del 2*. Rapport 388.
- Skolverket (2013) *Attityder till skolan 2012*. Rapport nr. 390.
- Skolverket (2013) *Kommunernas resursfördelning till grundskolor*. Rapport 391.
- Skolverket (2013) *Kostnader. Riksnivå. Sveriges officiella statistik om förskola, skola och vuxenutbildning. Del 3*. Rapport 393.
- Skolverket (2013) *Betydelsen av icke-kognitiva förmågor. Forskning m.m. om individuella faktorer bakom framgång*. Skolverkets aktuella analyser.
- Skolverket (2013) *Redovisning av uppdrag om avvikelser mellan provresultat och betyg i grundskolans årskurs 9*. (Dnr. 2013:00164)
- Skolverket (2013) *Redovisning av uppdrag enligt förordning (2011:355) avseende omfattning och utvärdering av försöksverksamhet med riksrekryterande spetsutbildning i grundskolans högre årskurser läsåret 2012/13*. (Dnr. 2013:00527)
- Skolverket (2013) *Redovisning av uppdrag att följa upp och analysera Skolverkets verksamhet utifrån minoritetspolitikens mål*. (Dnr. 2010:00242)

- Skolverket (2013) *Redovisning av uppdrag om att utvärdera statsbidraget för basfärdigheterna läsa, skriva och räkna.* (Dnr. 2009:571)
- Skolverket (2013) *Redovisning av uppdrag om utvecklingsinsatser inom områdena naturvetenskap och teknik.* (Dnr. 2012:868)
- Skolverket (2013) *Verksamhetsutvecklingsplan för Skolverkets uppföljning/utvärdering.*
- SOU 2004:35 *Utan timplan – med målen i sikte.*
- SOU 2005:101 *Utan timplan – för målinriktat lärande.*
- SOU 2005:102 *Utan timplan – forskning och utvärdering.*
- SOU 2007:28 *Tydliga mål och kunskapskrav i grundskolan. Förslag till nytt mål- och uppföljningssystem.*
- SOU 2007:75 *Att styra staten – regeringens styrning av sin förvaltning.*
- SOU 2009:25 *Samordnad kommunstatistik för styrning och uppföljning.*
- SOU 2009:85 *Gymnasial lärlingsutbildning – Hur blev det? Erfarenheter från första försöksåret.*
- SOU 2009:94 *Att nå ut och nå ända fram. Hur tillgången till policyinriktad utvärdering och forskningsresultat inom utbildningsområdet kan tillgodoses.*
- SOU 2010:75 *Gymnasial lärlingsutbildning – Utbildning för jobb. Erfarenheter efter två års försök med lärlingsutbildning*
- SOU 2011:72 *Gymnasial lärlingsutbildning – med fokus på kvalitet. Hur stärker vi kvaliteten i gymnasial lärlingsutbildning.*
- SOU 2012:36 *Registerdata för forskning. Delbetänkande av Statistikutredningen 2012.*
- SOU 2012:83 *Vad är officiell statistik? En översyn av statistiksystemet och SCB. Slutbetänkande från Statistikutredningen 2012.*
- SOU 2013:30 *Det tar tid – om effekter av skolpolitiska reformer.*
- SOU 2013:40 *Att tänka nytt för att göra nytta – om perspektivskiften i offentlig verksamhet.*
- SOU 2013:56 *Friskolorna i samhället.*

- SOU 2014:5 *Staten får inte abdikera – om kommunaliseringen av den svenska skolan.*
- Spencer, Liz, Jane Ritchie, Jane Lewis and Lucy Dillon (2003) *Quality in Qualitative Evaluation: A framework for assessing research evidence.* National Centre for Social Research. Cabinet Office.
- Stame, Nicoletta (2004) *Theory-Based Evaluation and Types of Complexity.* Evaluation, 10:58.
- Statistiska centralbyrån (2012) *Riktlinjer för beslut om innehåll och omfattning av den officiella statistiken.* Rådet för den officiella statistiken.
- Statistiska centralbyrån (2013) *Sveriges officiella statistik – årlig rapport för 2012.* Rådet för den officiella statistiken.
- Statistiska centralbyrån (2013) *Elevpaneler för longitudinella studier.* Beskrivning av statistiken.
- Statskontoret (2000) *Redovisning av uppdrag att genomföra en kartläggning av det samlade utvärderingsarbetet inom de olika utgiftsområdena.*
- Statskontoret (2001) *Utvärdera för bättre resultat! Att beställa utvärderingar som är till nytta i beslutsfattandet.* Rapport: 2001:22.
- Statskontoret (2002) *Regeringsuppdrag att göra en utvärdering av försöksverksamheten med lokala styrelser.* (Dnr. 2001/320-5)
- Statskontoret (2002) *Utvärderingar – Av vem och till vad. En kartläggning av regeringskansliets utvärderingsresurser.* Rapport 2002:20.
- Statskontoret (2011) *Fristående utvärderingsmyndigheter – en förvaltningspolitisk trend?*
- Stigendal, Lars (2007) *Styrning för ökad effektivitet.* Rapport 122. VERVA.
- Stobart, Gordon (2001) *The validity of National Curriculum Assessment.* British Journal of Educational Studies, Vol. 49, No. 1.
- Sundell, Knut red. (2012) *Att göra effektutvärderingar.* Socialstyrelsen och Gothia Förlag AB.

- Sutcliffe, Sophie and Julius Court (2005) *Evidence-Bases Policy-making: What is it? How does it work? What relevance for developing countries?* Overseas Development Institute.
- Svensson, Allan (red.) (2011) *Utvärdering genom Uppföljning. Långitudinell individforskning under ett halvsekel*. Gothenburg Studies in Educational Sciences 305. Göteborgs universitet.
- Tallberg Broman, Ingegerd, Lena Rubinstein Reich och Jeanette Hägerström (2002) *Likvärdighet i en skola för alla. Historisk bakgrund och kritisk granskning*. Skolverket.
- Thomas, James, Angela Harden, Ann Oakely, Sandy Oliver, Katy Sutcliffe, Rebecca Rees, Ginny Brunton and Josephine Kavanagh (2004) *Integrating qualitative research with trials in systematic reviews*. British Medical Journal, Vol. 328.
- Thomas, James and Angela Harden (2008) *Methods for the thematic synthesis of qualitative research in systematic reviews*. BMC Medical Research Methodology. 2008, 8:45.
- Thompson, Wiliam (ed.) (2010) *Making Reform Happen. Lessons from OECD Countries*. OECD.
- Tierney, Andrew (ed.) (2008) *Measuring Improvements in Learning Outcomes. Best Practices to Assess the Value-Added of Schools*. OECD.
- Torgerson, Douglas (1986) *Between Knowledge and Politics: Three Faces of Policy Analysis*. Policy sciences, Vol. 19. No.1.
- United Kingdom Treasury (2011) *The Green Book. Appraisal and Evaluation in Central Government*. The Stationery Office, London.
- Undervisnings- och kulturministeriet (2012) *Utbildning och forskning 2011-2016. Utvecklingsplan*. 2012:2.
- Undervisnings- och kulturministeriet (2012) *Utvärderingsplan för utbildningen 2012-2015*. 2012:15.
- Vedung, Evert (2009) *Utvärdering i politik och förvaltning*. 3:e upplagan. Studentlitteratur.
- Vedung, Evert (2005) *Anette Gröjers avhandling om den utvärderade(n)de staten*. Fakultetsopposition den 15 oktober 2004. Statsvetenskaplig Tidskrift. Årg. 107, nr. 4.

- Weisburd, David (2003) *Ethical practice and evaluation of interventions in crime and justice: the moral imperative for randomized trials*. Evaluation Review, 27(3).
- White, Howard (2010) *A Contribution to Current Debates in Impact Evaluation*. Evaluation 16:153.
- Weiss, Carol H. (1993) *Where Politics and Evaluation Research Meet*. Evaluation Practice, Vol. 14, No. 1, JAI Press Inc.
- Visscher, Adrie J., and Robert Coe () *School performance feedback systems: conceptualisation, analysis and reflection*. School effectiveness and school improvement, 14 (3). Pp. 321-349.
- Voils, Corinne I., Margarete Sandelowski, Julie Barroso and Victor Hasselblad (2008) *Making Sense of Qualitative and Quantitative Findings in Mixed Research Synthesis Studies*. NIH Public Access. 20(1):3-25.
- Öman, Sören (2006) *Särskilda registerförfattningar i Cecilia Magnusson Sjöberg och Peter Wahlgren (red.) Festskrift till Peter Seipel*. Norstedts juridik.

Kommittédirektiv 2012:53

Förbättrade resultat i grundskolan

Beslut vid regeringssammanträde den 24 maj 2012

Sammanfattning

En särskild utredare ska bedöma när de åtgärder som har vidtagits från 2007 och framåt för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan kan förväntas få avsedd effekt. Utredaren ska vidare utreda vilka relevanta data som behövs för att i ökad utsträckning kunna utvärdera de reformer som genomförs på skolområdet och vid behov lämna författnings- eller andra förslag för att öka tillgången till sådana data. Dessutom ska utredaren belysa hur reformer på skolområdet bör genomföras för att underlätta utvärdering av reformerna.

Utredaren ska senast den 30 april 2013 redovisa den del av uppdraget som avser att bedöma när åtgärderna för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan kan förväntas få avsedd effekt. Uppdraget ska slutredovisas senast den 29 november 2013.

Åtgärder som vidtagits för ökad måluppfyllelse

För att Sverige ska vara en framstående kunskapsnation krävs utbildning och forskning i världsklass och ett utbildningssystem som har kapacitet att se potentialen hos varje elev och student. Regeringen har därför under de senaste åren påbörjat arbetet med att lägga om svensk skolpolitik. De reformer och insatser som regeringen har aviserat och tagit initiativ till syftar till att höja kvaliteten i utbildningen och öka elevernas kunskaper. Genom ett

omfattande reformarbete på skolområdet har därmed förutsättningarna för skolan att genomföra sitt huvuduppdrag stärkts.

Barns lärande börjar tidigt. Det är viktigt att alla barn får stöd och stimuleras i sin utveckling redan under förskoleåren, oavsett om barnet är i hemmet, i den pedagogiska omsorgen eller i förskolan. Regeringen har, efter beslut av riksdagen, genomfört ett reformarbete för att stärka förskolans pedagogiska uppdrag, bl.a. genom att förtydliga läroplanens mål om språkutveckling, matematik, naturvetenskap och teknik. När det gäller grundskolan och motsvarande skolformer kan som exempel på stora och viktiga reformer nämnas införandet av nya läroplaner med tydligare och mer konkreta kursplaner och en ny betygsskala med fler steg.

Från och med höstterminen 2011 tillämpas en helt ny skollag. Den innebär bl.a. att elevens rätt till särskilt stöd har stärkts, lärares och rektorers befogenhet för att skapa studiero har förtydligats och så långt som möjligt har lika rättigheter och skyldigheter för kommunala och fristående skolor införts.

Elevernas kunskaper måste kontinuerligt följas upp och utvärderas för att skolan tidigt ska kunna stödja eleverna så att fler kan nå skolans kunskapsmål. Det är särskilt viktigt för elever som av någon anledning upplever svårigheter i skolarbetet, t.ex. på grund av funktionsnedsättningar. För detta arbete måste skolan ha förbättrade redskap för uppföljning och utvärdering, både av elevens kunskapsutveckling och den egna verksamheten. För att stärka uppföljningen och utvärderingen har regeringen därför infört nationella prov i fler årskurser och ämnen i grundskolan. Som ytterligare stöd i detta arbete har regeringen också infört skriftliga omdömen från årskurs 1 i grundskolan. Vidare har riksdagen beslutat att införa betyg från årskurs 6 i grundskolan och motsvarande skolformer från och med höstterminen 2012.

Även gymnasieskolan har reformerats i syfte att höja kvaliteten i utbildningen.

Lärares kompetens och undervisningens utformning är avgörande för elevernas studieresultat och möjlighet att tillgodogöra sig utbildningen. En ny lärarutbildning har därför införts från och med hösten 2011. Vidare har regeringen genomfört särskilda satsningar på lärares kompetens. Genom lärarlyftet gavs lärare möjlighet till fortbildning för att förstärka sin ämnesteoretiska och ämnesdidaktiska kompetens. Genom lärarlyftet II ges nu anställda lärare med lärarexamen möjlighet att komplettera sin behörighet i ämnen som de undervisar i men där de saknar behörighet. Ytterligare insatser

inom området är bl.a. en skärpning av behörighetsreglerna, att speciallärarutbildningen har återinförts samt införandet av lärarlegitimation. Det kan dock nämnas att vissa krav på legitimation föreslås börja tillämpas först den 1 december 2013.

I ett så decentraliserat skolväsende som det svenska blir tillsyn och kvalitetsgranskning viktiga inslag för att säkerställa en likvärdig utbildning med hög kvalitet. Regeringen inrättade därför 2008 Statens skolinspektion och riksdagen förstärkte resurserna för tillsyn och kvalitetsgranskning. I och med den nya skollagen har Skolinspektionen försetts med förstärkta sanktionsmöjligheter.

Regeringen har också tagit initiativ till ett antal stödåtgärder för att stimulera utvecklingen i skolan. Exempel på det är det särskilda stimulansbidraget för att stärka basfärdigheterna läsa, skriva och räkna och bidrag till stärkt undervisning i matematik, naturvetenskap och teknik.

Regeringens bedömning är att de reformer som har genomförts eller planeras att genomföras kommer att ge eleverna bättre förutsättningar att nå kunskapsmålen i skolan.

Uppdraget att bedöma när åtgärderna får effekt

Utbildningsområdet är centralt för att upprätthålla och stärka svensk tillväxt, konkurrenskraft och sysselsättning. Dessutom är utbildning ofta avgörande för individen. Enligt skollagen (2010:800) ska alla, oberoende av geografisk hemvist och sociala och ekonomiska förhållanden, ha lika tillgång till utbildning i skolväsendet. Barn och elever ska ges stöd och stimulans så att de utifrån sina förutsättningar utvecklas så långt som möjligt enligt utbildningens mål.

Under de senaste årtiondena har kunskapsresultaten sjunkit i den svenska skolan. Detta framgår bl.a. av den årliga betygsstatistiken, av Statens skolverks nationella utvärderingar av grundskolan 1992 och 2003 samt av resultat från de internationella studierna Trends in International Mathematics and Science Study (TIMSS), Programme for International Student Assessment (PISA) och Progress in Reading Literacy Study (PIRLS). Som en följd av detta har också andelen elever som är behöriga till gymnasieskolans nationella program sjunkit. Därför har regeringen sedan 2007 tagit initiativ till flera åtgärder för att höja kunskapsnivån och öka

måluppfyllelsen såväl i grundskolan som i övriga utbildningsformer inom skolväsendet.

Att förändra skolväsendet är en långsiktig process som innefattar många olika delmoment. Det kan därför ta tid dels innan verksamheten i landets skolor har förändrats i den riktning som regering och riksdag har avsett, dels innan resultaten av reformerna och insatserna kan utläsas t.ex. i form av andelen elever som får godkända betyg från grundskolan. För att säkerställa att de åtgärder som har vidtagits från 2007 och framåt för att höja kunskapsnivån och öka måluppfyllelsen ger avsedd effekt behövs en bedömning av när man kan förvänta sig denna effekt.

Uppdraget omfattar enbart grundskolan och motsvarande skolformer, dvs. grundsärskolan, specialskolan och sameskolan. Studier visar att personer som misslyckas redan i grundskolan löper större risk för problem som arbetslöshet, psykosociala problem och kriminalitet senare i livet. Att tidigt stimulera elever och vid behov ge stöd är därför ofta avgörande för hur deras fortsatta utbildning gestaltar sig.

Utredaren ska

- bedöma när de åtgärder som har vidtagits från 2007 och framåt för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan och motsvarande skolformer kan förväntas få avsedd effekt.

Uppdraget att analysera förutsättningarna för utvärdering

Både i den nationella och internationella debatten om offentlig förvaltning utgör utvärdering av reformer ett aktuellt tema. Styrutredningen konstaterade 2007 i betänkandet Att styra staten – regeringens styrning av sin förvaltning (SOU 2007:75) att regeringen, utöver den löpande information som myndigheterna tar fram, behöver information för mer eller mindre omfattande omprövningar av politiken. Inom EU, OECD och flera enskilda länder pågår också en diskussion om hur användningen av utvärderingar kan öka och hur utformningen av offentlig verksamhet i högre grad ska kunna baseras på kunskap.

Det omfattande reformarbetet när det gäller framför allt den svenska barn- och ungdomsutbildningen innebär att behovet av utvärderingar inom utbildningsområdet har förstärkts. Utbildning har stor betydelse för enskilda individer och för deras ställning på

arbetsmarknaden. I ett vidare perspektiv är utbildningsväsendet också viktigt för tillväxt, konkurrenskraft och sysselsättning. Områdets politiska och ekonomiska vikt medför att behovet av beslutsunderlag av hög kvalitet ökar.

Inom utbildningsområdet har emellertid utvärderingen av reformer och politiska beslut under lång tid spelat en relativt undanskymd roll. Regeringen har därför från och med 2012 gett det tidigare Institutet för arbetsmarknadspolitisk utvärdering (IFAU) ett utvidgat uppdrag att även utvärdera effekter av olika reformer och åtgärder inom utbildningsväsendet. Myndigheten har därmed också bytt namn till Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU).

Ett antal faktorer kan påverka möjligheten att utvärdera reformer, t.ex. sättet att genomföra reformerna. Utvärderingen av de reformer som genomförs underlättas dessutom om det finns lätt tillgängliga och tillförlitliga data. Skolverket är statistikansvarig myndighet för skolväsendet och samlar in en mängd information om barn och elever, lärare, resultat på nationella prov och betyg liksom kostnadsstatistik. I databasen SIRIS, där information kan fås uppdelat på skolnivå, finns utöver statistisk information, också kvalitetsredovisningar, information om statsbidrag samt resultat av Skolinspektionens kvalitetsgranskning och tillsyn och anmälningar.

Utredaren ska

- utreda vilka relevanta data som behövs för att i ökad utsträckning kunna utvärdera reformer inom de obligatoriska skolformerna, dvs. grundskolan, grundsärskolan, specialskolan och sameskolan,
- om det behövs, lämna författnings- eller andra förslag för att öka tillgången till sådana data, och
- belysa hur reformer inom de obligatoriska skolformerna bör genomföras för att underlätta utvärdering av reformerna.

Samråd och redovisning av uppdraget

Utredaren ska samråda med Statens skolverk, IFAU och övriga berörda myndigheter inom skolområdet. Utredaren ska också samråda med andra relevanta myndigheter och organisationer.

Utredaren ska senast den 30 april 2013 redovisa den del som avser att bedöma när åtgärderna för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan kan förväntas få avsedd effekt. Uppdraget ska slutredovisas senast den 29 november 2013.

(Utbildningsdepartementet)

Kommittédirektiv 2013:73

Tilläggsdirektiv till Utredningen om förbättrade resultat i grundskolan (U 2012:04)

Beslut vid regeringssammanträde den 27 juni 2013

Förlängd tid för uppdraget

Regeringen beslutade den 24 maj 2012 kommittédirektiv om att en särskild utredare ska bedöma när de åtgärder som har vidtagits från 2007 och framåt för att höja kunskapsnivån och öka måluppfyllelsen i grundskolan kan förväntas få avsedd effekt (dir. 2012:53). Utredaren ska också utreda vilka relevanta data som behövs för att i ökad utsträckning kunna utvärdera de reformer som genomförs på skolområdet och vid behov lämna författnings- eller andra förslag för att öka tillgången till sådana data. Dessutom ska utredaren belysa hur reformer på skolområdet bör genomföras för att underlätta utvärdering av reformerna. En delredovisning lämnades till regeringen den 2 maj 2013. Enligt direktiven ska uppdraget slutredovisas senast den 29 november 2013.

Utredningstiden förlängs nu och uppdraget ska i stället slutredovisas senast den 28 februari 2014.

(Utbildningsdepartementet)

Kunskap och skolreform

En studie om underlag för och utvärdering av
reformer på skolområdet

Oktober 2013
Mats Björnsson

På skolans liksom på socialvårdens och militärväsendets område, är de slutliga avvägandena självfallet av politisk art och bör så vara. Men dessa avväganden kan bli annorlunda om de baseras på en bredare grund av systematiskt gjorda erfarenheter och inte enbart på rent tyckande från de berörda parternas sida.

(Torsten Husén, *Skola för 60-talet*, s. 21 f.)

1. Inledning

Den här genomgången handlar om svenska reformer rörande grundskolan. Närmare bestämt är det en genomgång av

- vilket slags underlag som har funnits för de större reformerna;
- om reformerna utvärderats och i så fall av vem och om eventuella effekter, samt
- specifikt hur det som benämnts Nationell utvärdering av grundskolan har gestaltat sig och vilken ev. inverkan den haft.

Tidshorizonten sträcker sig från formandet av en grundskola under 1940- och 1950-talet fram till i dag. I syfte att ringa in ”större skolreformer” har jag funnit det ändamålsenligt att utgå från relevanta ”reformkluster” under denna tidsrymd, snarare än från enstaka politiska (reform)beslut.

Det finns knappast några helt entydiga svar på de frågor som uppdraget bygger på. Det är relativt lätt att konstatera att viss forskning eller vissa publicerade studier fanns vid tillfället för ett propositionsarbete, men det är ofta svårt att säga vilken roll denna empiri faktiskt spelade. Motsvarande gäller för utvärderingar.

För uppdraget har jag tagit del av propositioner, utredningsbetänkanden, rapporter, debatt- och faktaböcker m.m. Jag har också genomfört intervjuer med några personer som haft en central ställning i de senaste decenniernas skol-Sverige och som överblickar en längre period av skolpolitik och skolforskning. Dessa personer förtecknas i referenslistan. Det har inte varit möjligt att ta del av allt relevant material, men jag anser att underlaget är tillräckligt för de allmänna bedömningar som görs.

2. Den stora grundskolreformen: perioden 1940-tal –1960-tal

Stellan Arvidssons *Enhetsskolan*, skriven 1950, inleds med orden

Genom riksdagsbeslutet den 26 maj 1950 inleddes en ny tid för den svenska skolan. Skolreformen av 1950 är visserligen bara den sista länken – hittills – i en lång kedja av skolreformer. Men den skiljer sig från sina föregångare på två sätt. Först och främst är den mer genomgripande än några av dessa. Men dessutom avser den att inleda en period av experimenterande, som lämnar fältet fritt för utveckling i

skilda riktningar. Beslutet i riksdagen avser att vara upptakten till en *ständigt fortskridande skolreform*.

Det Arvidsson säger ger en tidsmarkör men illustrerar också två för denna rapport centrala förhållanden. Han pekar på att det är fråga om en genomgripande skolreform. Han pekar också på att reformen inte bygger på en fix och färdig ritning. Genom försök som följs ska i stället erfarenheter dras, och läggas till grund för vidare beslut.

1940 års skolutredning skulle utreda och lägga fram förslag om bl.a. hur man skulle få en ”..större enhetlighet inom den på olika skolformer splittrade svenska skolorganisationen” (Arvidsson, a.a s. 5). Den kom att 1943–47 framlägga ett digert utredningsmaterial, inte mindre än 20 SOU-publikationer. I materialet ingick bl.a. särskilda bilagor med empiriskt forskningsmaterial (se referenslistan).

Skolkommissionen, som var parlamentariskt sammansatt, tillsattes 1946 och skulle på grundval av vad 1940 års skolutredning kommit fram till, finna politiska lösningar. Huvudbetänkandet lades fram 1948 (SOU 1948:27). Inte mindre än 12 delegationer och ett 80-tal experter knöts till kommissionen.

Omfattande utredningsmaterial och studier

Totalt omfattade utredningsmaterialet som låg till grund för 1950 års riksdagsbeslut omkring femtusen sidor (Marklund, 1980 s.2). Torsten Husén, Nils-Erik Svensson och Kjell Härnqvist hörde till dem som gjorde studier under 1940- och 1950-talen, bl.a. kring vad som benämndes begåvningsreserven. Dessa fick betydelse för förslagen till 1962 års beslut. Som exempel på det underlag som togs fram kan nämnas Kjell Härnqvists studie *Reserverna för högre utbildning. Beräkningar och metoddiskussioner*. SOU 1958:11¹. En kärnfråga som studerades var den s.k. differentieringsfrågan, dvs. om eller från vilken ålder folkskoleelever och realskoleelever skulle kunna undervisas tillsammans.

1950 års beslut om försök med en nioårig enhetsskola var en politisk kompromiss och detta faktum var också huvudskälet till att

¹ Sin titel till trots handlar denna i hög grad om hur stor andel av eleverna i folkskola och realskola som kunde anses ha förutsättningar för senare högre studier.

man beslöt om en omfattande försöksverksamhet. Föredragande statsrådet Weijne anförde i en tidstypisk formulering:

Mot slutet av 1950-talet kan tiden vara inne för en sammanfattande utredning i differentieringsfrågan, varvid praktiska och vetenskapliga erfarenheter böra kunna sammanfattas till en hållbar grund för den organisation, som sedan successivt skall genomföras.

(prop. 1950:70, s. 192)

Uppföljningen hamnade hos SÖ

Skolöverstyrelsen fick i riksdagsbeslutet ansvaret för att leda och följa upp försöksverksamheten, genom inrättande av en särskild avdelning. I beslutet ingick också en utbyggnad av statens psykologisk-pedagogiska institut. Rapporter från SÖ utkom årligen och en sammanfattande rapport för försöksåren överlämnades 1959 till 1957 års skolberedning. Försöken blev med tiden rätt omfattande. Läsåret 1957/58 deltog dryga 143 000 elever och 5 700 klasser.

Under processen fanns motsättningar i synen på hur man skulle gå fram. Inom alla partier insåg man att något behövde göras med det splittrade systemet. Men försöksverksamheten kritiserades för att inte vara tillräckligt bred och det kom också kritik mot att erfarenheterna inte pekade mot den sammanhållna enhetsskola (ofta benämnd 8 + 1) som regeringen förespråkade. Högerpartiets representanter i riksdagen efterfrågade från tid till annan under 1950-talet var de vetenskapliga och "objektiva" beläggen för enhetsskolelösningen fanns.

Även 1957 års skolberedning innebar ett omfattande arbete. Den tillsattes för att dra de politiska slutsatserna av försöksverksamheten. Huvudbetänkandet (Grundskolan, SOU 1961:30) var på nära 900 sidor, vartill fogades ytterligare betänkanden, rapporter och specialbetänkanden. Experter och ledamöter i olika arbetsgrupper m.m. uppgick till långt över 100 personer (Se Marklund, 1982, s. 424 ff). Urban Dahllöfs *Kursplaneundersökningar i matematik och modersmålet* (SOU 1960:15) är exempel på sådana empiriska studier som ingick i beredningens framlagda material.

Reformbesluten kring grundskolan skulle följas upp och det behövdes i den 'rullande reformen' inte en ny stor utredning. Det blev i stället SÖ som fick en central roll. Ett särskilt utvecklingsanslag kom till och SÖ knöt kontakter med vetenskapliga forskningsinstitut. FoU-medlen växte från 2 milj.kr 1963 till 30 milj.kr.

1975². Insatserna rörde sådant som ämnesmetodik och jämförande-utvärderande projekt. Bland andra det relativt nybildade International Association for the Evaluation of Educational Achievement (IEA) fick stöd. S.k. kommunblock för utveckling inrättades 1964. SÖ:s s.k. särskilda pedagogiska arbetsgrupper (SPA) fick betydelse för den 'rullande' grundskolreformen. (Marklund, 1982).

Ytterligare ett steg mot en gemensam 9-årig skola

Under 1960-talet fortsatte diskussionen om differentieringen eftersom 1962 års grundskola fortfarande innebar val av olika studievägar. I praktiken tvingades eleverna i huvudsak välja väg redan i årskurs 7 och årskurs 9 var uppdelad på olika utbildningsvägar. Motsättningarna var inte mindre nu, snarare hade den relativa fred som beslutet 1962 innebar, tagit en ända. Enligt Marklund (1982) så var det dock inte debatten utan statistiken som fick en avgörande betydelse för nästa steg, dvs. uppgifter om hur eleverna faktiskt valde. De praktiska alternativen (flertalet s.k. tvåbokstaviga linjer i årskurs 9) valdes bort i hög utsträckning, då dessa uppfattades som utbildningsmässiga återvändsgränder³. G-linjen dominerade stort. Regeringen valde här att dra konsekvenserna av denna utveckling och minska på uppdelningen efter elevval.⁴ Den proposition som Olof Palme som ansvarigt statsråd lade fram (prop. 1968:129) om en ny läroplan för grundskolan innebar att sista steget i utvecklingen mot en i allt väsentligt gemensam 9-årig grundutbildning för alla barn i Sverige hade tagits. Man kan därmed säga att det tog cirka 30 år av utredande, försök, värdering – och politiska processer – för att nå dit.

Drygt 40 år efter beslutet om grundskolans införande lade Costa Meghir och Mårten Palme fram en studie som visar att grundskolereformen medförde att de unga som omfattades av förändringen, jämfört med en kontrollgrupp, kom att nå en högre utbildningsnivå och att reformen hade en klassutjämnande effekt.⁵

² Ett belopp som i dagens penningvärde motsvarar drygt 150 milj. kr.

³ Det är inte helt långsökt att här jämföra med situationen 2013 för de yrkesinriktade studievägarna i gymnasiet.

⁴ Kontroverser förelåg. Svenska Arbetsgivareföreningen föredrog en tydligare uppdelning efter fallenhet, medan fr.a. TCO aktivt agerade för att göra hela den 9-åringa skolgången i princip gemensam. SÖs ledning var till en början splittrad i denna fråga, men överdirektör Orring fick senare vika sig för generaldirektör Lövbeers linje, som anslöt till TCO:s syn.

⁵ *Educational reform, ability and family background*, 2004.

Utvärderingens roll under reformarbetet

Det är uppenbart att de tre stora utredningar som tillsattes gjorde omfattande utredningar och kartläggningar, och aktivt nyttjade kunskapen hos det fåtal professorer i pedagogik, som då fanns att tillgå. Det är också uppenbart att de bidrag dessa lämnade till stor del var av empiriskt slag; ofta undersökningar på ett större antal elever med gängse tillhörande statistiska beräkningar och analyser. Det är knappast någon tvekan om att det var den tidens mest avancerade pedagogiska forskning som kom att nyttjas integrerat i reformarbetet.

Den sammantagna utredningstiden för de tre stora utredningarna var 10–12 år, en närmast ofantlig tidsresurs sett med dagens ögon. Genom att ha tid på sig skulle man också hitta lösningar som var hållbara och partiöverskridande⁶. Tempot i beslutsfattandet var annorlunda. Förankring genom brett remissförfarande sågs som självklart och viktigt.

Försöken följdes av en särskilt inrättad avdelning inom SÖ. I så måtto skedde en både planmässig och, kan man säga, formativ uppföljning, under en längre följd av år. Hur viktig den blev för den fortsatta processen är en annan fråga. Det rådde, som har påpekats, kontrovers både kring huruvida försöken var tillräckligt spridda och om korrekta slutsatser drogs i samband med grundskolbeslutet. SÖ:s uppföljning bidrog sannolikt till att driva processen framåt, att hålla den på spår. Det var heller inte utan betydelse med den utökning av FoU-verksamheten som skedde under 1960-talet och den kraftigt förstärkta finansieringen av den pedagogiska forskningen.

Sett som en nationell omfattande verksamhet hade uppföljningen knappast den typ av systematik som man i dag skulle förknippa med en nationell utvärdering. Dess syfte var i huvudsak att söka besvara frågor kring vad som var möjligt, vad som fungerade och inte. Om 1950 års beslut ses som en reform så var det en reform som syftade till att pröva för att få underlag och råd(rum) inför kommande beslut.

⁶ Det bör infogas att detta ofta också var den retorik som man utövade i samband med reformerna. Självfallet kan detta också ha haft rent politiska implikationer och syften, behoven av att få råd(rum), söka vinna tid för att bättre säkra sin linje i ett senare skede, o.d.

3. Skolans inre arbete, styrfrågor: perioden 1970–1988

Den svenska grundskolan var organisatoriskt fullt utvecklad 1970 till sina yttre former; vad det därefter gällde var mer att utveckla dess inre former och innehåll.

Erfarenheten visar att det är en sak att besluta om och genomföra en organisatorisk yttre skolreform och en annan, ofta betydligt svårare, att förändra de vanor och traditioner som finns i skolinstitutionen och hos dess bärare: framför allt lärarna.

SIA-utredningen

SIA-utredningen (Skolans inre arbete) var 1970-talets stora utredning. Den tillsattes under senhösten 1970 i form av en sakkunnigutredning. Bakgrunden var en debatt om oreda, bristande motivation och skolsvårigheter och utredningen kom därmed att i vid mening fokusera på skolan som arbetsmiljö. Urbaniseringen och utbyggnaden av grundskolan hade medfört en tillväxt av stora skolenheter. Andelen barn som sågs i behov av särskilda stödinsatser visade en galopperande utveckling; utredningen fann i sina undersökningar att så mycket som 40 procent av eleverna 1972 kom i kontakt med specialundervisning (SOU 1974:53 s. 137). Utredningens uppgift blev att söka lösningar på en rad av dessa problem.

Till utredningen knöts liksom tidigare ett större antal experter och grupper. Utredningen producerade under 1974, före överlämnandet av huvudbetänkandet, ett antal forskningsrapporter fördelade på tre departementsskrivelser. Dessa rörde skolsvårigheter, skolan som arbetsplats respektive lokalresurser. Huvudbetänkandet (SOU 1974:53) som lades fram i juli 1974 omfattade tillsammans med en separat bilagedel ca 1 200 sidor. Bilagedelen (SOU 1974:58) innefattade 18 bilagor, se referenslistan. Utredningen genomförde också i egen regi riksgiltiga undersökningar.

SIA-utredningens svar

Utredningens svar på de debatterade problemen blev att dessa inte kunde lösas med statliga generella insatser. I stället borde kommuner och skolor ges större möjligheter att själva forma en verk-

samhet som svarar mot situationen och behoven. Utredningen pekade i en separat rapport på att statsbidragssystemet i sig drev en expansion av stödundervisningen. Försök med friare resursanvändning genomfördes i ett antal kommuner. Försöken utvärderades vid Göteborgs universitet (Kihlborn & Lundgren, 1974).

I betänkandet föreslogs att skoldagen borde göras mer sammanhållen (s.k. samlad skoldag), att skolan borde utveckla mer av samverkan med det omgivande lokalsamhället, att resursanvändningen borde bli friare, och att arbetslag och arbetsenheter inom skolorna borde tillskapas för att utifrån lokala behov hantera t.ex. resurser för stödet till elever, m.m.

Det restes tidigt krav på oberoende och vetenskaplig utvärdering av skolan, krav som skulle komma att upprepas under många år framöver. 1971 inrättades Pedagogiska nämnden vid SÖ, sammansatt av bl.a. politiska företrädare. Nämnden skulle ha en mer självständig ställning än tidigare rådgivande organ och ha FoU- och utvärderingsfrågor som uppgift. Marklund (1982) menade att nämnden inte hade resurser att skapa något stort och samlat utvärderingsprojekt, men man utvärderade avgränsade delar och kommenterade utvärderingar som bl.a. SÖ bedrev.

SIA-besluten innebar att ett omfattande läroplansarbete startade. SÖ fick uppgiften att arbeta fram ett förslag till läroplan och lade 1978 fram den s.k. regnbågsrapporten med förslag till läroplan. Det slutliga förslaget, vars grunddrag lades fram i propositionen (1978/79:180), hade dock bearbetats inom Utbildningsdepartementet.

Annan utvärdering

I takt med att forskningsresurserna växte, liksom skolans omfattning och betydelse, kom också med tiden fler forskarstudier att genomföras. Några av dessa hade fokus på att, direkt eller indirekt, värdera beslutade skolpolitiska reformer. Nämnas kan t.ex. *Vad trycks om grundskolan – ett försök till lägesbeskrivning av grundskolan, Skolreformer och lokalt gensvar, Utvärdering av 35 grundskolor genom upprepade lägesbedömningar 1980–1985 samt Från lag till bruk. Hemspråks-reformens genomförande*⁷.

⁷ Se här referenslistan.

En summering

SIA-utredningen bildade grund för en rad förändringar i grundskolan genom olika riksdagsbeslut. Utredningen, som arbetade i fyra år, tog fram omfattande utredningsmaterial, också i form av forskningsrapporter och empiriska studier (se referenslistan). Även här genomfördes försök i utredningens regi, försök som följdes upp och till del också utvärderades. Vad gäller försöken med friare statsbidragsanvändning talar dock tillgängliga uppgifter om att utredningen inte kom att bry sig om utvärderingens huvudresultat⁸.

Det empiriska underlaget skiljer sig delvis från det från den föregående perioden som studerats. Mer av kvalitativa studier har kommit in. 1970-talets reformer byggde på kartläggningar och undersökningar av förhållanden i skolan. Lgr 69 och Lgr 80 kan ses som justeringar av en grundskola som formades i praktiken ute på skolorna, och den senare av de två bär tydliga spår av decentraliseringssträvanden. Man kan se en del av SIA-utredningen som en utvärdering av den dåvarande grundskolan. Någon samlad reformutvärdering skedde dock inte. Pedagogiska nämnden vid SÖ utgjorde ett slags embryo, men dess resurser var för små.

4. Den andra stora systemomformningen: 1990-talet

Under 1980-talet kom trenden mot decentralisering att fortsätta. Utvärderingens betydelse för att från statens sida kunna följa utfallet mot uppsatta mål, kom att aktualiseras alltmer. Strax efter mitten av 80-talet tillsattes en styrberedning som lade fram sitt betänkande (SOU 1988:20) *En förändrad ansvarsfördelning och styrning på skolområdet*. En tidigare tillkallad expertgrupp lade fram rapporten *Ansvarsfördelning och styrning på skolområdet* (Ds U 1987:1). Denna expertgenomgång skulle förutom att redovisa hur det dåvarande styrsystemet hade vuxit fram och tankarna bakom, också ”bedöma och redovisa i vilken mån det finns forskningsresultat och annat dokumenterat material som är relevant för diskussionerna om skolans styrning”⁹.

Mål- och resultatstyrningen kom genom ett riksdagsbeslut 1988 att etableras som styrprincip för staten, delvis som svar på proble-

⁸ Intervju med Ulf P. Lundgren.

⁹ PM upprättad inom Utbildningsdepartementet 1986-04-16, cit. i DsU 1997:1.

met med ständigt växande offentliga utgifter. Men några stora skolreformer kom inte till under detta decennium, annat än alldeles i slutet, då genom beslutet om att föra över huvudmannaskapet för skolledare och lärare till kommunerna (prop. 1989/90:41). SÖ bedrev ett arbete med att utveckla ett system för utvärdering, vilket beskrivs i avsnitt 6.

Expertutredningen och styrberedningen utgjorde till någon del underlag för styrningsreformer. Men att ta tag i huvudmannaskapets frågan, dvs. det statliga arbetsgivaransvaret för lärare och skolledare, var något som av allt att döma krävde en avsevärd politisk insats (Haldén, 1994). Kommunaliseringsbeslutet 1989 var knappast av den arten att man kan säga att det fanns empiriskt underlag för det. Det följdes upp av en rad beslut; om ansvaret för skolan, förändringar i skollagen, direktiv till en betygsberedning och en läroplanskommitté, utformning av ett generaliserat statsbidragssystem samt nedläggning av SÖ och länskolnämnderna och inrättande av det nya Skolverket. Under detta reformarbete blev det hösten 1991 regeringsbyte. Det ledde till nya direktiv till Läroplanskommittén men inte till någon egentlig kursändring. Läroplansförslagen, som till stor del byggde på mål- och resultatstyrningens principer, kunde efter bred remissbehandling antas av en enig riksdag, där även beslut om ett nytt målrelaterat betygssystem ingick (prop. 1992/93:220)¹⁰. Det som däremot mer tydligt bar den nya regeringens signum var den proposition om valfrihet och fristående skolor (prop. 1991/92:95) som lades fram våren 1992. Till det kom den nästkommande år framlagda propositionen (1992/93:230) Valfrihet i skolan.

Tre propositioner

Följande tre propositioner får ses som centrala under denna period:

Den s.k. *kommunaliseringspropositionen* (prop. 1989/90:41) är volymmässigt påfallande tunn, endast 14 sidor. Den redogör i korthet för diskussioner som förts under en längre tid, bl.a. i nämnda decentraliserings- och styrutredningar sedan slutet av 1970-talet. Det konstateras därvid att man tidvis hållit fram ett odelat kommunalt driftansvar som en önskvärd möjlighet, men att man dithills inte lagt några förslag om detta. I övrigt är propositionen en argumentering för varför det nu var tid att göra denna överföring av det

¹⁰ Detta var första gången som förslag till nya läroplaner remitterades.

statliga arbetsgivaransvaret till kommunerna och för hur likvärdighet skulle garanteras.

Den s.k. *ansvarspropositionen* (prop. 1990/91:18) är bredare till sitt innehåll och syfte. Även här tas utgångspunkt i decentraliseringssträvandena och hänvisas i korthet till den expertgrupp, styrberedning och styrproposition som berörts ovan, tillika det förslag till nytt statsbidragssystem för skolan som arbetats fram i Utbildningsdepartementet (Ds 1990:32) och som även remitterats.

Avsnittet ”Principiella utgångspunkter för förslagen” inleds med följande text:

Medborgarna ställer allt större krav på den offentliga verksamheten. Kraven gäller bättre tillgång till och kvalitet på service eller tjänster. Kraven omfattar också inflytande över vad som presteras samt möjligheter att välja olika alternativ.

(a.a. s.17)

Det var inte längre möjligt, menade man, att klara krav och behov genom att tillföra mer resurser till offentlig sektor. I stället måste man få ut mer för insatta resurser och ett sätt att åstadkomma det är att effektivisera genom att föra ut besluten längre ut, till dem som arbetar i verksamheten, och till brukare, kunder, föräldrar, etc. På det sättet presenterades redan här, i sak, den ideologiska grunden för de förslag om fristående skolor, som strax efter valet 1991 lades fram av den borgerliga regeringen. I så måtto kanske man kan säga att de olika systemreformerna hängde samman.

Propositionen om valfrihet och fristående skolor (prop. 1991/92:95), som lades fram för riksdagen i mars 1992, är liksom kommunaliseringpropositionen volymmässigt liten (19 sidor). Föredragande statsrådet Ask hänvisade inledningsvis till nyligen gjorda riksdagsuttalanden i frågor om att kunna välja skola, och informerade om de utredningar som pågick vad gällde bl.a. avgifter samt effekter av redan beslutade åtgärder. I övrigt redovisades förslagen om valfrihet och fristående skolor, och överväganden i anslutning till dem.

Tunt underlag

Propositionerna ger uttryck för en politisk vilja att genomföra reformer som innebar tämligen radikala förändringar av styrningen av den offentliga skolan. De allmänna politiska utgångspunkterna

var, liksom under de föregående cirka 15 åren, decentralisering, minskad statlig styrning, ökat inflytande för kommuner, skolpersonal och föräldrar (elever) samt ökad valfrihet. Men de reformer som nu lades fram var skarpare och mer långtgående. Något empiriskt underlag i form av undersökningar, t.ex. av hur den dittillsvarande decentraliseringen fallit ut, anfördes inte. Något omfattande underlag i form av utredningar redovisas inte heller; det rör i sådana fall mera tekniska frågor om hur bidragssystem ska utformas, o.d.¹¹

Utvärdering?

Såvitt kunnat utläsas lades under de 20 år som följde på kommunaliseringsbeslutet inget uppdrag från regeringen om utvärdering av de beslutade reformerna. Det nya Skolverket har sedan sin tillkomst ett generellt och centralt uppdrag att ansvara för utvärdering inom skolektorn. Skolverket genomförde också redan 1992 en bred utvärdering av grundskolan, omfattande samtliga grundskolans ämnen. Den utgör emellertid inte en utvärdering av de då nyligen tagna besluten. En ny mer begränsad utvärdering av grundskolan genomfördes 1995, men inte heller den avsåg att kasta ljus över styrreformer. Det gör däremot en del andra utvärderingar som verket gör under 1990-talet. I Bilden av skolan 1996, där verket summerar en lång rad egna undersökningar, konstaterar man att kommunernas ansvarstagande för skolan lämnar en hel del i övrigt att önska. Denna slutsats kommer igen i många av verkets utvärderingar under lång tid framåt.

Även om Skolverket som sektorsmyndighet upprepade gånger pekade på att det ansvarstagande som decentraliseringen bygger på inte fungerade som tänkt, ställdes knappast frågan om huruvida 'reformerna fungerar'. Inriktningen av Skolverkets uppdrag utgick från den beslutade decentraliseringen och mål- och resultatstyrningen. Verket hade därför svårt att på allvar utvärdera det nya systemet. Verkets uppgift var, och blev framför allt alltmer, att verka för att reformerna *skulle* realiserars¹². De stora utvärderingar som genomfördes syftade till att se hur grundskolan fungerade snarare än hur reformerna fungerade, även om dessa frågeställ-

¹¹ Däremot var nya läroplaner, kursplaner, timplaner och betygssystem föremål för utredning.

¹² Se Björnsson, 2003. *Elva år med Skolverket. En myndighet under påverkan*.

ningar givetvis tangerar varandra. Detsamma torde gälla den s.k. friskolereformen. Många studier – framför allt efter millennieskiftet – kom att ge bilder av vad det växande antalet fristående skolor och fritt skolval betytt. Det gäller studier från såväl Skolverket som från forskare och andra aktörer. Skolverket har (senare) fört fram att reformerna tycks ha bidragit till ökad segregation. Det torde dock vara först genom t.ex. SNS forskarrapport 2011 som det läggs fram en relativt bred genomgång och analys av de studier som finns inom området¹³. Detta pekar mot att det för en viss typ av reformer är först efter en ganska lång tid som det går att dra mer säkra slutsatser om följderna av dessa. Samtidigt innebär ett längre tidsspänn att många faktorer då har kommit att påverka skeendet; det blir allt svårare att isolera effekten av en specifik reform.

Har de utvärderingar som gjorts haft någon effekt?

Allmänt sett noterades självfallet från politiskt håll Skolverkets ofta framförda bedömning att det nya ansvaret för kommunerna inte satt sig, eller sköts dåligt. Det gav till och från argument för regering och riksdag att sätta in åtgärder. Främst handlade det om att införa kompletterande verktyg i styrningen, alltifrån tydligare krav på lokal utvärdering och kvalitetsredovisningar till mer krav på elevuppföljning och införande av skolinspektion, etc. Efter reformerna i början av 90-talet ändrade man – under fortsatt retorik om mål- och resultatstyrning – successivt karaktären av den statliga styrningen ganska väsentligt.

Våren 2011 beslutar Riksdagen om ett tillkännagivande att en utvärdering av de senaste årens skolreformer borde göras. I utskottets betänkande sägs att ”de skolreformer som ska utvärderas är kommunaliseringen, det fria skolvalet och frågan om de fristående skolornas etablering”. Regeringen kom därefter med hänvisning till riksdagens tillkännagivande att dels tillsätta en enmansutredning för att se på kausala effekter av kommunaliseringen, dels ge IFAU¹⁴ i uppdrag att utvärdera nämnda reformer. Redovisning av dessa uppdrag har när denna rapport skrivs ännu ej skett.

¹³ Hartman, L., m.fl., 2011 *Konkurrensens konsekvenser. Vad händer med svensk välfärd?*

¹⁴ Institutet för arbetsmarknads- och utbildningspolitisk utvärdering.

5. De senaste årens skolpolitiska beslut: perioden 2008–2012

En tid efter regeringsskiftet 2006 fattades en lång rad beslut om kursplane- och läroplansförändringar, utökade betygs- och provsystem, läsa-skriva-räknasatsning, m.m. De frågor som dessa beslut handlar om har delvis genomlysts i termer av existerande forskning, m.m.¹⁵ Här ska i stället redogöras för karaktären på det beslutsunderlag som utgjort en del av den officiella beredningsprocess som legat till grund för propositionerna ifråga. De propositioner som valts ut är sådana som – i enlighet med den bedömning som gjorts i betänkandet "Det tar tid" – kan antas få synliga konsekvenser på skolenivån, för lärare och elever. Det är dels "Tydligare mål och kunskapskrav – nya läroplaner för skolan" (prop. 2008/09:87), dels "Betyg från årskurs 6 i grundskolan" (prop. 2009/10:219).

Målen: brist på utvärdering av systemet

Den förstnämnda av de två propositionerna syftar, som titeln anger, till att formulera tydligare mål för ämnesundervisningen. Den vilar på ett utredningsbetänkande, SOU 2008:28, "Tydliga mål och kunskapskrav i grundskolan". Betänkandet är omfattande (drygt 360 sidor exkl. bilagor). I inledningen av betänkandet sägs att det, vid genomgången av "varför inte målsystemet genomförts", ofta saknats information om grunderna för olika beslut, och att därför en mängd intervjuer genomförts. Den bild som dessa gett anges som i stort samstämmig, vilket anses ha skapat en stabil grund för de förslag som läggs (s. 55).

Intrycket som utredaren hade av att det saknas dokumentation kring de frågor som utredningen skulle lägga fram förslag om, kommer igen senare i betänkandet.

Hur styrdokument som läroplaner och kursplaner uppfattas och används i skolan finns det än så länge begränsade kunskaper om. Enligt uppgifter till utredningen togs det inom Skolverket under senare delen av 1990-talet initiativ till att styrdokumentet skulle utvärderas, något som dock aldrig kom att genomföras. En bidragande orsak kan ha varit att det uppfattades som känsligt och möjligen olämpligt att Skolverket skulle utvärdera sitt eget arbete. I verkets årsredovisningar finns under

¹⁵ *Det tar tid. Om effekter av skolpolitiska reformer*, SOU 2013:30.

senare delen av 1990-talet en del uttalanden om hur lärare uppfattar kursplanerna, men redovisningarna är mycket kortfattade.

(a.a. s. 279)

Det empiriska underlaget för förslagen om förändringar av främst kursplanerna består följaktligen främst av de samtal med personer i och utanför Skolverket som utredningen genomfört. Det saknas inte dokument från Skolverket som tar upp problematiken med gällande mål och målstruktur, inte minst avseende det faktum att mål att uppnå ofta kommit att uppfattas som ”grundkurs” och som planeringsmål, i strid med hur dessa var tänkta. Detta förhållande framgår också av propositionen ifråga. Det har däremot inte gjorts några systematiska undersökningar härav.

Tidigare betyg: ett rent politiskt ställningstagande?

Underlaget för regeringens förslag om betyg från årskurs 6, i stället för som då i årskurs 8, finns i Ds 2010:15. En arbetsgrupp hade här fått i uppdrag att ”föreslå de författningsändringar som är nödvändiga för att betyg från och med årskurs 6 i grundskolan ska kunna införas.” Det fanns således inget uppdrag att utreda om eller varför betyg skulle ges tidigare. Motiven anges endast översiktligt i skrivelsen.

En tungt vägande orsak till utvecklingen inom grundskolan är bl.a. bristen på en tidig uppföljning av varje elevs kunskapsutveckling samt att de elever som behöver det inte får det stöd de har rätt till. Därför finns det starka skäl att vidta ytterligare åtgärder för förbättrad uppföljning av varje elevs kunskapsinhämtande i syfte att fler elever når en högre grad av måluppfyllelse.

(s. 31)

Ställningstagandet att betyg ska börja sättas fr.o.m. årskurs 6 hanteras som ett uteslutande politiskt avgörande. Den (negativa) utvecklingen i grundskolan är utgångspunkt, att tidigare betygsättning ska leda till en förbättring, är något som förutsätts.

Kortare utredningstider

Det har i flera sammanhang konstaterats (se t.ex. SOU 2007:75) att de offentliga utredningar som tillsatts under senare tid getts betydligt mindre tidsutrymme än tidigare. De utredningar som belysts här utgör inget undantag. Utredningen om tydligare mål arbetade i drygt ett år, underlaget vad gäller betyg från årskurs 6 i cirka 5 månader. Från utredningar om tre–fyra år eller mer, till utredningar om drygt ett år i genomsnitt. Såvitt gäller de underlag och utredningar som omfattats av denna genomgång så sker detta skifte successivt och någonstans mellan 1970- och 1980-talet.

Utvärdering av senare års beslut om skolan

Vilken utvärdering har regeringen initierat med anledning av den omfattande mängden politiska beslut om förändringar i skolsystemet?¹⁶ Ett sätt att se detta är genom regleringsbreven till den myndighet som har ansvaret för nationell utvärdering, Skolverket. Regleringsbreven summerar år för år samtliga rapporteringskrav och uppdrag som ligger på myndigheten, oavsett i vilket sammanhang dessa har getts. Regleringsbreven avseende 2011, 2012 och 2013 har gått igenom. De omfattar vardera ett 90-tal uppdrag inkl. krav på återrapportering/redovisning.¹⁷

I samtliga regleringsbrev anges att Skolverket regelmässigt ska rapportera om vilka uppföljningsstudier och utvärderingar som genomförts under året. Sett till de *uppdrag om utvärdering* som riktas till verket rörande grundskolan anges följande objekt under de tre åren:

- för satsningen på basfärdigheter (fr. år 2008) ska en extern aktör anlitas för utvärdering,
- insatser ifråga om ämnet idrott och hälsa,
- insatser ifråga om näringsriktig skolmat,
- metoder mot mobbning,
- matematiksatsningar och satsning på naturvetenskap och teknik,

¹⁶ De har beskrivits av företrädare för alliansen som de största reformerna på skolområdet sedan 1800-talet. Se t.ex. www.unt.se/debatt/storsta-skolreformerna-sedan-1842-2461189.aspx

¹⁷ Vissa av dessa upprepas från ett regleringsbrev till ett annat.

- insatser ifråga om nyanlända, bl.a. ska en extern aktör för utvärdering anlitas.

Till det kommer att verket i några fall ska samråda eller samverka med annan myndighet i samband med utvärdering, t.ex. med IFAU om Sfi-bonus, med Socialstyrelsen om ett australiskt självmords-förebyggande utbildningsprogram.

Vissa av satsningarna är relativt omfattande, t.ex. matematiksatsningarna, medan andra är av specifik och begränsad art. Det är antingen specifika frågor eller statliga resursinsatser som ska utvärderas. Något explicit bredare utvärderingsuppdrag som siktar på de omfattande reformerna läggs inte på Skolverket. Skolverket har också vid upprepade tillfällen informerat om att nästan all verksamhet är uppdragsrelaterad, dvs. att verkets resursmässiga utrymme för egna initiativ är mycket begränsat.

Av Skolverkets verksamhetsplan för 2013 framgår dock att verket gör vissa utvärderingsinsatser kopplade till de senaste årens förändringar. Så t.ex. genomförs studier kring hur reformerna i grundskolan landar, framför allt hos lärarna. Främst sker detta genom intervjuer. Skolverket talar inte om detta som utvärdering; snarare ses det som en tidig implementeringsstudie. Det pågår också nationell utvärdering av de praktiskt-estetiska ämnena och en planering för utvärdering av samhällsorienterande ämnen. Resursinsatsen som krävs för godkänt deltagande i de internationella kunskapsmätningarna bestäms i liten grad av Skolverket. Det medför att de två nu nämnda undersökningar som genomförs på eget initiativ resursmässigt har ett relativt marginellt utrymme.

En lite förenklad bild utifrån genomgången av de senaste årens många skolpolitiska beslut blir att det empiriska underlaget är begränsat, om något alls, och att det inte getts något bredare uppdrag om utvärdering av förändringarna, vid sidan av de internationella kunskapsmätningarna som successivt genomförs. De av riksdagen 2011 begärda utvärderingarna utgör delvis ett undantag.

6. NU – en lång och krokig väg

Som redan framgått fick SÖ uppgiften att följa upp den stora försöksverksamheten på 1950-talet. Utvärdering kom under de två följande decennierna att diskuteras och i viss mån genomföras, såväl av SÖ som inom utredningar och av andra aktörer. Politiska

krav restes, till och från, på vad man kallade "oberoende" och "vetenskaplig" utvärdering, då man ansåg att en myndighet inte bör utvärdera det man själv varit med om att forma (se t.ex. Richardsson, a.a. s. 155). När frågorna om styrning och statens roll alltmer aktualiseras under 1980-talet, växte också frågor om utvärdering, och främst då på nationell nivå. I ett flertal budgetpropositioner från mitten av decenniet och framåt, togs nationell utvärdering upp.

En väsentlig fråga i detta sammanhang är hur man skall kunna kontrollera hur de statligt angivna målen och riktlinjerna förverkligas. Jag finner här det som SÖ anför i sin anslagsframställning om ett nationellt utvärderingsprogram intressant. I likhet med SÖ menar jag att ett centralt uppföljnings- och utvärderingsprogram bör innefatta precisering och konkretisering av vad de vanligtvis mer allmänt formulerade målen kan avse. Programmet bör också innefatta bedömning och, där det så är möjligt, mätning av i vad mån målen uppnåts. Ett arbete med en sådan inriktning mot utvärdering och kontroll av att statens intentioner uppfylls bli enligt min mening en av den statliga skoladministrationens huvuduppgifter i ett decentraliserat och målstyrt system.
(1985 års budgetproposition)

Inom SÖ pågick ett arbete, i stort sett ända sedan 1982 års reform som resultat av SAK-kommittén¹⁸, med att forma en nationell skolutvärdering. Erland Ringborg kom som ny generaldirektör 1986 och hade som grundinställning att SÖ inte "skulle veta hur andra skulle göra utan skulle veta bäst om hur det stod till". Det var lite oklart exakt varifrån initiativen till nationell utvärdering av skolan kom, men man kan säga att det låg i tiden.¹⁹

Arbetet med att ta fram instrument för en utvärdering av grundskolan skedde i samarbete med forskare vid pedagogiska institutioner i landet. Det synes inte ha varit ett konfliktfritt samarbete, vare sig mellan forskare sinsemellan eller mellan dessa och SÖ²⁰. 1989 genomfördes den första utvärderingen som omfattade elever i årskurs 2 och 5, med avsikt att årskurs 9-elever skulle ingå i nästa omgång. SÖ kom att lansera resultaten av studien en tid före nedläggningen av myndigheten den 1 juli 1991. Samma år som studien publicerades kom också en kritik mot SÖ:s utvärdering i en skrift där fem pedagogikprofessorer medverkade²¹. Bland annat kritiserades metoderna för att vara alltför snävt orienterade och för

¹⁸ Skoladministrativa kommittén.

¹⁹ Intervju med Erland Ringborg.

²⁰ Intervjuer med Erland Ringborg och Mats Ekholm.

²¹ Skolan och utvärderingen. Fem professorer tar ordet.

ensidigt byggda på differentialpsykologiska och kvantitativa synsätt.

Skolverket tar över

Skolverket bildades som en helt ny myndighet i stället för det nedlagda SÖ sommaren 1991. Vissa förberedelser hade gjorts inom det s.k. Skolprojektet som pågick inom departementet under 1990/91. Att nationell utvärdering i förbindelse med lokal utvärdering, skulle spela en central roll i myndighetens arbete stod klart, men inte hur.²² I detta skede kom den modell för nationell utvärdering av grundskolan som SÖ initierat, att få en betydelse. Den representerade en existerande plan och ett konkret projekt som utan större dröjsmål kunde sösättas och bära den nationella utvärderingen ett steg framåt.²³

En hel del förändringar gjordes i förhållande till det kritiserade upplägget 1989. En del av de tidigare uppgifterna utmönstrades, instrumenten breddades mot att få fram mer kvalitativa data, och ett problemlösande prov togs fram²⁴. Som nämnts ingick också årskurs 9 i utvärderingen. En bred samverkan med forskare föregick själva genomförandet, och dessa var också aktiva i analysen av resultaten. Huvudresultaten av NU 1992 offentliggjordes i samband med att Skolverket lade fram rapporten Bilden av skolan 1993.

Utvärderingen som en styrmodell

Rapporten Bilden av skolan 1993 gavs bred spridning i media och diskussionen fokuserades på bl.a. arbetsformerna i grundskolans avslutande år. Utvärderingen granskade en skolverksamhet som styrdes av 1980 års läroplan.

Genom sin inriktning på även ämnesvisa spørsmål kan man säga att NU i hög grad kom att rikta sig inte bara till statsmakterna utan också till kommuner, skolor och lärare. Såväl de ämnesvisa som de summerande rapporterna spreds brett. Detta var en del av Skolverkets strategi; genom den nationella utvärderingen av lokala förhållanden skulle de lokala enheterna få ett underlag att värdera sin

²² Se t.ex. Halldén, 1997.

²³ Intervjuer med Staffan Lundh respektive Oscar Öqvist.

²⁴ Av någon betydelse torde det ha varit att Skolverkets generaldirektör var en av dem som offentligt kritiserat SÖ:s utvärderingstänkande.

egen verksamhet mot och också stimuleras att själva utvärdera. Förutsättningen för modellen, som i viss mån kan ses som en modell för såväl styrning som utvärdering och utveckling, var dock att det fanns en form av ekologi; att det växte fram lokala utvärderingar att bygga de nationella på, något som knappast infriades.

En tid efter det att Skolverket lanserat NU 92 kom åter en forskarkritik. Den vände sig nu mot Skolverket och framför allt mot den styrmodell som Skolverket synes företräda med sin utvärdering. Organisationsforskarna Kerstin Sahlin-Andersson och Bengt Jacobsson kritiserade i *Skolan och det nya verket* idén om att man med utvärdering ska kunna styra och utveckla skolan. De ansåg inte att modellen fungerade på det sättet i förhållande till skolor och lärare, och därutöver menade de att en tydligt normerande myndighet behövs i en organisation av den karaktär som skolan utgör.

Kritiken verkade inte avsätta några tydliga spår inom Skolverket. Däremot konstaterades det att NU var en stor och dyr ansats med sin forskarstödda utvärdering av samtliga grundskolans ämnen. 1995 gjordes en uppföljning av NU 92 i vissa delar – under benämningen UG 95 – som visade på en viss nedgång i kunskaperna.²⁵ Ytterligare tre år senare, 1998, gjordes en utvärdering mot läroplanens övergripande mål. ”Det var en mycket bra utvärdering, men den fick inte mycket genomslag. Den hade inga bjärta budskap av typen att något höll på att gå snett eller så.”²⁶

Först 2003 kom NU tillbaka, med en replik på 1992 års studier.²⁷ Vissa s.k. ankaruppgifter m.m. behölls från proven som ingick i NU 92 medan andra var nya. Resultaten offentliggjordes hösten 2004 ungefär samtidigt som resultaten från de internationella kunskapsmätningarna PISA 2003 och TIMSS 2003. Samtliga mätningar kunde visa på en betydande tillbakagång jämfört med respektive föregående mätning, resultat som fick stor uppmärksamhet.

Även efter NU 2003 konstaterades att denna ansats är dyr och tung att genomföra.

Vid den här tiden, alltså runt 2004, så hade också de internationella kunskapsmätningarna som PISA och TIMSS etablerat sig på det sättet att de behövde tänkas in i ett långsiktigt nationellt utvärderingsperspektiv. Verket sökte hitta en stabil väg genom förslag och propåer till

²⁵ Se vidare längre fram om effekterna av denna.

²⁶ Intervju med Staffan Lundh.

²⁷ Att NU återkom i början av 2000-talet hade samband med intresset hos generaldirektören vid den tiden.

departementet. Det gällde senare också frågor om långsiktig finansiering av dem, där vi sedan några år har fått en särskild anslagspost.²⁸

Post NU?

Vid ungefär samma tidpunkt pågick en diskussion om de nationella proven och vilken roll dessa kunde och borde spela. Regeringen gav Skolverket ett omfattande uppdrag kring resultatinformation. Uppdrag E rörde systemet med nationella prov. Skolverket lämnade en redovisning på denna punkt i oktober 2003 där man hävdade att proven av flera skäl inte kunde tjäna som en löpande utvärdering på nationell nivå. Skolverket menade att ”.. ett rullande stickprovsbaserat system för kunskapsutvärdering utarbetas i ämnen som inte har nationella prov. Exempelvis kan tre ämnen utvärderas år 1, tre andra år 2, ytterligare tre andra år 3 osv. Efter t.ex. 5–6 år upprepas cykeln.” (Skolverkets redovisning, s. 16). I bilagan till redovisningen (s. 136) skrev man att ett sådant rullande system ”bör utredas”.

Verket fick ganska snart härfter (december 2004) ett nytt uppdrag att utreda förutsättningarna för att införa just ett mer systematiskt, rullande stickprovsbaserat system för kunskapsutvärdering för samtliga ämnen i grundskolan. Skolverket lämnade ett omfattande underlag som svar på uppdraget i september 2006. I processen med förslaget lät man Jan-Eric Gustafsson vid Göteborgs universitet lämna ett förslag, som verket fann alltför ambitiöst och dyrbart. Man lämnade därför en reviderad modell som sitt förslag (Skolverkets rapport 286).

Skolverket har till dags dato inte fått någon reaktion på sina förslag. Vartefter har man, i den mån resurserna medgett, beslutat söka genomföra en del av idéerna i sitt förslag. Under 2013 pågår utvärdering av praktiskt-estetiska ämnen, samt planering vad gäller samhällsorienterande ämnen.

²⁸ Intervju med Staffan Lundh.

7. PISA, TIMSS och PIRLS

Några år inpå det nya millenniet kan konstateras att de internationella kunskapsmätningarna har blivit fler och återkommande, och dessutom politiskt alltmer intressanta. I praktiken är det knappast Skolverket som beslutar om deltagande²⁹. Från att ha haft en prövande inställning till de internationella studier som successivt kom att föreslås, så byter nu Skolverket fot och inriktar sig på att

- i princip delta i alla kunskapsmätningar,
- se dessa som en del av den nationella utvärderingen av skolan, samt att
- söka se till att förutsättningarna planeringsmässigt för detta säkras.

Det är den andra punkten som här är intressant. Sverige deltar numera i såväl PISA och TIMSS som PIRLS, vilket drar såväl ekonomiska resurser som resurser i form av expertis. Stickproven i studierna är tillräckligt stora för generaliserbarhet och såväl validitet som reliabilitet i testerna anses hög. Dessutom erbjuds möjlighet till jämförelse med andra länder, något som visar sig alltmer intressant, politiskt och medialt (se t.ex. Pettersson, 2008). Med tanke på att det efter 2004 inte har offentliggjorts någon bred och samlad utvärdering av den svenska grundskolan har de internationella mätningarna kommit att helt dominera vad gäller redovisning av skolans resultat. Dominansen gäller såväl i sak, som ifråga om den bild som den politiska skoldebatten, och politiken i sig, haft att grunda sig på.

Skolverket konstaterar i en rapport till regeringen:

Samtliga Skolverkets nationella utvärderingar från NU-92 och framåt har varit egeninitierade och inte regeringsuppdrag (och har som sådana finansierats inom förvaltningsanslaget). Besluten att genomföra utvärderingarna har varje gång fattats (relativt ad-hoc-mässigt) internt inom Skolverket, och det har inte funnits någon långsiktig plan – varken från Skolverkets eller statsmakternas sida – att genomföra denna typ av nationella studier med vissa intervall.³⁰

Den nationella utvärderingen har i och med PISA m.fl. mätningar stärkts i så måtto att den fått andra instrument för att löpande följa resultaten i form av uppmätta kunskaper hos eleverna och genom

²⁹ Skolverket meddelade 2004 sin bedömning att Sverige inte skulle delta i PIRLS 2006, men efter en fråga i riksdagen gav statsrådet Östros till svar att Sverige skulle delta.

³⁰ Skolverket, Rapport 2006:286.

jämförelserna med andra länder. Den har sannolikt också stärkts i det att mätningarna, såvitt det går att bedöma, utgör tillförlitliga trendmått, dvs. visar förändring över tid. Det är tveksamt om Skolverket hade kunnat få fram så pass tillförlitliga trendmätningar på egen hand. Den nationella utvärderingen får dock anses ha försvagats – sett över de tio år som förflutit sedan NU 2003 – genom att Skolverket inte kunnat kompensera för de begränsningar som trots allt ligger i de internationella studierna: den ensidiga fokuseringen på kunskapsprov, mätningen mot internationell och inte nationell standard, den snäva fokuseringen på en del av läroplanen och bristen på analyser av processer och förutsättningar. Skolverket har sökt möta dessa uppenbara begränsningar genom att lämna förslag till regeringen om rullande stickprovsundersökningar och genom att nu, 2013, genomföra viss utvärdering mot nationella mål i praktiskt-estetiska ämnen. PISA och dylika studier bör kanske betraktas som en form av kvalificerad uppföljning av kunskapsresultat, resultat som givetvis är en central *del* av en nationell utvärdering.

8. Hur har utvärderingsresultat använts och påverkat?

I princip skulle det vara möjligt att göra en genomgång av alla de större skolutredningarna och propositionerna under de senaste 25 åren för att spåra inflytande från reformutvärderingar. Det finns dock en del problem förknippade med det. Det första är helt enkelt bristen på studier som tydligt kan betecknas som reformutvärderingar avseende grundskolan. Det finns studier som har inslag av utvärdering av en reform men knappast mer³¹. Det andra problemet är att följa effekten av framlagda utvärderingsresultat. De offentliga utredningarna på skolområdet fångar i regel upp den empiri som t.ex. skolmyndighetens offentliggjorda studier utgör. Men att avgöra om denna empiri ligger till grund för de förslag som utredningen lägger fram, kan redan det visa sig svårt. Att gå vidare för att se om och i så fall hur denna empiri får betydelse i regeringens

³¹ Det kan observeras, även om distinktionerna inte är glasklara, att NU presenterades som en *systemutvärdering*, där resultaten och processerna bakom skulle fångas för att ställas mot gällande mål/läroplan. En *reformutvärdering* får förutsättas fokusera på om utfallet svarar mot just reformens bärande delar eller idéer. Detta kan vara föremål för en analys i sig, men som inte tas upp här. Läggas till bör också att såväl NU 92 som UG 95 avsåg skolverksamhet som till större del styrts av Lgr 80 och inte av Lpo 94.

beslut är än svårare. Många propositioner ger inte den motivbakgrunden utan pläderar i en politisk kontext för vilka problem man vill lösa och vad det är man vill uppnå.

Allmänt sett har Skolverkets underlag kommit att användas frekvent inom Regeringskansliet och även omnämnts relativt ofta i t.ex. budgetpropositionerna. Regeringen har också direkt uttalat detta, dvs. pekat på att man med den myndighetsstruktur och målstyrning som gäller i statsförvaltningen kommit att vara starkt beroende av det som myndigheten lägger fram och att man i hög grad vilar på detta underlag. Det är däremot svårt att bedöma hur underlaget påverkat.

Ett måhända specifikt exempel får belysa något av detta. Skolverket publicerade 1996 resultaten av UG 95 (se ovan), den studie som till del jämförde kunskaperna 1995 i årskurs 9 med dem 1992. Studien visade en tendens till fallande resultat. En resurssatsning som sedan presenterades i regeringens vårbudget 1997 kunde därmed hämta argument ur en utvärdering med empiriska data, som kunde motivera eller legitimera satsningen. Det är samtidigt uppenbart att satsningen just då snarare kom till stånd till följd av främst två förhållanden: den under en period kraftiga minskningen av lärartätheten i kommunerna (det var nu politiskt önskvärt att öka den), och att statsfinanserna åter möjliggjorde insatser (det var nu politiskt möjligt att vidta åtgärder).

Svårgenomskådat

Det går att peka på ett antal exempel på att en utvärdering gett anledning till ett politiskt initiativ. Det kan handla om dåliga resultat i en internationell kunskapsmätning i matematik som leder till en nationell matematiksatsning eller studier av brister i hur nya styrdokument implementeras, som leder till justeringar av de fortsatta insatserna. PISA har t.ex. vid ett flertal tillfällen refererats till i propositioner, etc. Men det är svårt att se mönstret i vad som får betydelse och när, om man inte samtidigt kan iaktta de konkreta ekonomiska och politiska förhållanden som i hög grad styr beredningsprocesserna hos regering och riksdag. Det kan ses som "politiska fönster" som är mer eller mindre öppna. Till det kommer den mediala logiken "good news is no news" som innebär att t.ex. studier som visar på sjunkande kunskapsresultat ger stort eko.

9. Avslutande reflektioner

I korthet ger genomgången följande bild. Från att ha tillsatt stora utredningar med ofta omfattande utredningsexpertis och lång utredningstid, tillsätts senare under perioden utredningar med betydligt kortare utredningstider, ibland med snävare uppdrag. Från de tidiga utredningarna, vilka grundade sig på och presenterade empiriska och för ändamålet relevanta undersökningar, kommer senare sådant underlag att i stor utsträckning lysa med sin frånvaro. De kortare utredningstiderna gör i sig att utrymmet för att göra egna undersökningar inom ramen för utredningarna begränsas starkt. Men man kan inte heller se tydliga spår av att redan existerande omfattande empiri i stället byggs in i betänkandena.

Det är vanskligt att göra enkla jämförelser över loppet av mer än ett halvt sekel. Uppenbart gällde delvis andra normer för beslutsfattande och delvis andra politiska ideal på 1950-talet jämfört med i dag. Den samhälleliga kontexten för reformer förändras. Värdet av stora kompromisser och "gediget underlag" ses inte som lika högt i dag. Många har påpekat att de politiska processerna nu går snabbare och att det ofta inte är realistiskt att vänta flera år på en utredning.

När det gäller utvärdering av reformerna, så är bilden delvis en annan än om vi talar om utredningsunderlag som föregått reformbeslut. Från 1960-talet växer utvärdering fram i termer av resursramar, forskningsstudier och mängden utvärderingsrapporter. Trots det allmänna talet om vikten av utvärdering, går det inte att se någon tydlig utveckling över decennierna när det gäller gedigen utvärdering av skolreformer. Genomgången visar en ojämn, ryckig resa; det kommer inte till stånd någon stabil, bred, förankrad och systematisk utvärdering av skolan. Det säkras inte en tillräcklig massa av empiriska studier. Det förs under decennierna diskussioner om vem som borde ansvara för en sådan: den statliga förvaltningen, fristående aktörer eller ansvariga myndigheter. Men dessa diskussioner landar sällan eller aldrig annat än i tillfälliga initiativ. Pedagogiska nämnden på 1960-talet är ett sådant initiativ, resurser till IFAU 2011 ett annat. Myndigheter som SÖ och Skolverket söker bygga upp system för utvärdering av grundskolan men dessa bär inte långt p.g.a. nedläggningar, omgörningar av myndigheter, brist på resurser, internationella standarder, etc. Skälen kan vara många till detta sakernas tillstånd, men det mest iögonfallande är ändå det faktum att något brett uppdrag till myndigheter om

utvärdering av beslutade reformer inte lämnats efter 1990. Ett undantag utgörs delvis av den av riksdagen nyligen begärda utvärderingen av 1990-talets reformer.

Särskilt kan detta vara förvånande med tanke på att de idéer som slår rot i Sverige i slutet av 1980-talet lägger så stor vikt vid just utvärdering som del av styrningen.³² Om sättet att ta fram underlag för reformer tidigare kunde ses som 'framtungt' genom det gedigna beredningsarbetet före reformbeslut, kunde man kanske nu vänta sig mer av en 'baktung' modell med reformutvärdering, dvs. *ex post*, men så ser inte ut att ha blivit fallet.

Ett skäl till att någon bred, stabil och förankrad utvärdering av reformer avseende grundskolan inte kommit till stånd, är sannolikt framväxten av de återkommande internationella kunskapsmätningarna, främst tillkomsten av PISA 2000. PISA som mätning mot en form av överenskommen *global curriculum* är bara ett verktyg av flera för att öka kunskapen om skolans resultat. Slående är att dessa mätningar ger lite av förståelse för förutsättningar och processer i svensk skola och får ses som både hårt och snävt resultatinriktade. I en högt politiserad debatt om skolan i Sverige ser PISA, TIMSS och PIRLS ut att få räcka till som besked om den svenska skolans status och utfall och därmed också räcka som bränsle för debatten³³. Betydande empiriskt grundade analyser och utbildningsforskning, såväl kvantitativt som kvalitativt, med sikte på att belysa utvecklingen i svenska skolor skulle dock behövas för att komplettera dessa resultatmätningar och utgöra en nationell utvärdering värd namnet.

Ett annat spår i diskussionen om underlag och utvärdering av skolreformer rör utbildningsforskningens vägar. OECD konstaterar i sin *Evidence in education. Linking research and policy*, att den pedagogiska forskningen ofta kommit att ägna sig åt annat än sådana studier som är möjliga att nyttja i nationell utvärdering och policyskapande. Konflikterna kring kvantitativa respektive kvalitativa synsätt har säkerligen varit till förfång på det sättet att de skymt det ganska uppenbara, nämligen att detta inte är en fråga om antingen-eller utom om både-och. Många metoder behövs i modern utbildningsforskning och analys som ska kunna ge en information av sådant djup att de duger att grunda väsentliga samhällsbeslut på.

Det är lätt att säga att det behövs mer forskning och utvärdering om skolan. Vad som synes behövas är inte minst ett uthålligt poli-

³² Idégods som mål- och resultatstyrning hör till "idéer som reser" enligt Røvik, 1999.

³³ Pettersson, 2008.

tiskt intresse för att noggrant ta vara på samlade erfarenheter inför beslut om viktiga skolreformer.

Referenser och källor

Litteratur:

- Agrell, Jan, *Den nya skolan och försöksverksamheten: en principiell granskning*, J. A. Lindblad, Uppsala, 1953.
- Andersson, D. 2012. *Tillsammans – i lust eller blomstring*. Hjalmarson Högberg Förlag.
- Arrhén, E. 1956. *Enhetsskolan – vardagen och atomtidsåldern*. Bokförlaget Medborgarskolan, Uppsala.
- Arvidson, S. 1950. *Enhetsskolan: riksdagsbeslutet, kommunal planering, försöksverksamheten*, Gleerup, Lund.
- Arvidson, S., Stenholm, B. & Blomberg, I. 1961. *Enhetsskolan blir grundskolan*, Liber, Stockholm.
- Björnsson, M. 2003. *Elva år med Skolverket. En myndighet under påverkan*. Magisteruppsats, Försvarshögskolan.
- Carle, T & Helén, G. 1961. *Den nya skolan*. Forum, Stockholm.
- Dahllöf, U. 1967. *Skoldifferentiering och undervisningsförlopp*.
- du Rietz, L., Andrén, G. Lundgren, U.P. 1987. *Ansvarsfördelning och styrning på skolområdet. Ett beredningsunderlag utarbetat inom utbildningsdepartementet*. Ds U 1987:1. Stockholm: Utbildningsförlaget.
- Ekholm, M., Fransson, A. och Lander, R. 1987. *Skolreformer och lokalt gensvar. Utvärdering av 35 grundskolor genom upprepade lägesbedömningar 1980–1985*. Göteborgs universitet.
- Franke-Wikberg, S. (red), Dahllöf U., Kallós, D., Lundgren, U.P., Wallin, E. 1989. *Skolan och utvärderingen. Fem professorer tar ordet*. Stockholms universitets förlag.
- Haldén, E., 1997. *Den föreställda förvaltningen*. Statsvetenskapliga institutionen, Stockholms universitet.
- Husén, T. 1963 *Skola för 60-talet*, Stockholm: Almqvist & Wiksell.
- Härnqvist, K. 1960. *Individuella differenser och skoldifferentiering*. 1957 års skolberedning (Utgiven som SOU 1960:13).

- Larsson, H-A. 2011. *Mot bättre vetande En svensk skolhistoria*. SNS förlag.
- Lundgren, U.P. 2002. *Utbildningsforskning och utbildningsreformer*. I Pedagogisk forskning i Sverige, årg. 7, nr.3, s. 233–243.
- Lundgren, U.P. 1999. *Ramfaktorteori och praktisk utbildningsplanering*. I Pedagogisk forskning i Sverige, årg. 4, nr 1, s. 31–41.
- Lundgren, K. 2013. *Kan landet lagom vara bäst?* SNS förlag.
- Läroverkslärarnas riksförbund (red.). 1956. *Hur skall det gå för enhetsskolan?: en skrift om skolreformen i program och försök*, Norstedt, Stockholm.
- Marklund, S. 1960. *Skolreformen och lärarutbildningen: en undersökning utförd på uppdrag av 1957 års skolberedning*, Stockholm.
- Marklund, S. 1959. *Enhetsskolan under tio år: kort redogörelse för försöksverksamheten läsåren 1949/50–1958/59*. Skolöverstyrelsen.
- Marklund, S. 1980. *Skolsverige 1950–1975.1. 1950 års skolbeslut*. Liber Utbildningsförlaget.
- Marklund, S. 1982. *Skolsverige 1950–1975. 2. Från Visbykompromissen till SIA* Liber Utbildningsförlaget.
- Meghir, C., Palme M. 2004. *Educational reform, ability and family background*. IFS Working Papers W04/10, Institute for Fiscal Studies.
- Molander, P., m.fl.. 2002. *Vem styr? Relationen mellan regeringskansliet och myndigheterna*, SNS Förlag.
- OECD. 2007. *Evidence in education. Linking research and policy*. Paris: OECD.
- Pettersson, D. 2008. *Internationell kunskapsbedömning som inslag i nationell styrning av skolan*. Uppsala universitet.
- Richardsson, G. 2010. *Svensk utbildningshistoria. Skola och samhälle förr och nu*. Studentlitteratur, Lund.
- Román, H. 2013. *Från folkskoleinspektion till utbildningsförvaltning: Stockholms skolförvaltning efter 1958 – svensk skola i ett kommunalpolitiskt perspektiv*. (Manuskript)
- Rothstein, B. 2003. *Sociala fällor och tillitens problem*. Studentlitteratur.
- Sandström, B., Ekholm, M. 1984. *Stabilitet och förändring i skolan*. (SÖ. Utbildningsforskning FoU Rapport 50). Stockholm: Liber Utbildning.

- Sjöstrand, W. 1954. *Skolreformen i stöpsleven: svensk skolpolitik i vetenskaplig belysning: ett diskussionsinlägg*, Stockholm: Natur och kultur.
- Svensson, N-E. 1962. *Ability grouping and scholastic achievement*. Stockholm: Almqvist & Wiksell.
- Vedung, E. 2009 *Utvärdering i politik och förvaltning*. Lund: Studentlitteratur.
- Vestman Karlsson, O., Andersson, I.M. 2007. *Pedagogisk utvärdering som styrning – En historia från präster till PISA*. Forskning i fokus 35. Skolverket.

Elektronisk resurs:

www.unt.se/debatt/storsta-skolreformerna-sedan-1842-2461189.aspx
www.dn.se/debatt/vardens-kontrollsystem-bättre-for-andra-omraden/

Utredningsbetänkanden, m.m.

- SOU 1944:20 Skolan i samhällets tjänst. (1940 års skolutredning). Bilagor (urval):
1. Hygieniska förutsättningar för skolarbetet (Hjärne).
 2. Den psykologiska forskningens nuvarande ståndpunkt om den psykiska utvecklingen hos barn och ungdom m.m. (Anderberg, Elmgren Katz & Landquist).
-
4. Lärjungeurvalet till studielinjer med den nuvarande realskolans mål (Dahr).
 5. Skolungdomens vägledning till utbildning och yrke (Neymark).
- SOU 1948:27 Riktlinjer för det svenska skolväsendets utveckling. (1946 års skolkommission).
- SOU 1960:13 Individuella differenser och skoldifferentiering (Härnqvist).
- SOU 1960:15 Kursplaneundersökningar i matematik och modersmålet (Dahllöf).
- SOU 1961:30 Grundskolan. (1957 års skolberedning).
- SOU 1974:36 Staten, skolan och kommunerna. (SSK).

- DsU 1974:1 Skolans som arbetsplats (SIA-utredningen).
- DsU 1974:2 Problem och svårigheter i skolan (SIA-utredningen).
- DsU 1974:3 Grundskolans lokalresurser (SIA-utredningen).
- SOU 1974:53 Skolans arbetsmiljö (SIA-utredningen).
- SOU 1974:58 Bilagor till Skolans arbetsmiljö (urval):
2. Förstärkningsresursens omfattning i grundskolan.
 3. Förstärkningsresursens ekonomiska värde i grundskolan.
 4. Kostnader för s.k. klassföreståndartimmar på grundskolans högstadium och i gymnasieskolan.
 5. Kostnader för att öka ledningsvolymen i rektorsområden.
 6. Förändrat antal lärartjänster till följd av ändringar i specialundervisningens struktur.
 7. Beräkning av antal nya arbetstillfällen enligt SIA:s förslag.
 8. Undervisningsgrupperna i grundskolan. En studie av deras storlek och den tid eleverna tillbringar i grupper av olika storlek.
-
16. Kumulativa relativa frekvenser över klasstorleken i obligatoriska skolan läsåren 1961/62 och 1972/73.
 17. Kostnadsberäkningar för skolledarutbildningen.
 18. Skolledarutbildningens innehåll.
- SOU 1978:65 Skolan, en ändrad ansvarsfördelning (SSK-utredningen).
- DsU 1987:1 Ansvarsfördelning och styrning på skolområdet. Ett beredningsunderlag utarbetat inom utbildningsdepartementet (Expertgrupp).
- SOU 1988:20 En förändrad ansvarsfördelning och styrning på skolområdet (Styrberedningen).
- SOU 1992:94 Skola för bildning (1991 års Läroplanskommitté).
- SOU 2007:28 Tydliga mål och kunskapskrav i grundskolan.
- Ds U 2010:15 Betyg från årskurs 6.

Propositioner, utskottsbetänkanden

- Kungl. Maj:ts proposition 1950:70 angående riktlinjer för det svenska skolväsendets utveckling.
- Kungl. Maj:ts proposition 1962:54 angående reformering av den obligatoriska skolan m.m.

Kungl. Maj:ts proposition 1968:129 angående revidering av läroplan för grundskolan.

Proposition 1975/76:39 Skolans inre arbete.

Proposition 1976/77:85 Statsbidrag för grundskolan.

Proposition 1978/79:180 Läroplan för grundskolan.

Proposition 1980/81:97 Skolforskning och personalutveckling.

Proposition 1988/89:4 Skolans utveckling och styrning.

Proposition 1989/90:41 Kommunalt huvudmannaskap för lärare och skolledare.

Proposition 1990/91:18 Ansvar för skolan.

Proposition 1991/92:95 Valfrihet och fristående skolor.

Proposition 1992/93:220 Ny läroplan för grundskolan och nytt betygssystem för grundskolan m.fl. skolor.

Proposition 1992/93:230 Valfrihet i skolan.

Proposition 2008/09:87 Tydligare mål och kunskapskrav – nya läroplaner för skolan.

Proposition 2009/10:219 Betyg från årskurs 6 i grundskolan.

Utbildningsutskottets betänkande 2010/11 UbU:7 Grundskolan.

Intervjuer:

Mats Ekholm, prof. em., f.d. generaldirektör i Skolverket (1999–2003) och i Myndigheten för skolutveckling (2003–2004).

Ingrid Lindskog, senior rådgivare, Skolverket, f.d. kansliråd i Utbildningsdepartementet, f.d. chef för grundskol- och förskolenheten, Skolverket.

Ulf P. Lundgren, prof. em., f.d. generaldirektör i Skolverket (1991–1999).

Staffan Lundh, f.d. utredare i Utbildningsdepartementet, avdelningschef, Utvärderingsavdelningen, Skolverket (1991 –

Erland Ringborg, f.d. generaldirektör i Skolöverstyrelsen (1986–1991).

Oscar Öqvist, f.d. ansvarig för NU vid Skolöverstyrelsen resp. Skolverket.

Exempel på uppföljningar och utvärderingar genomförda av Skolverket på regeringens uppdrag åren 2002 – 2013

Här redovisas ett urval av utvärderingar från senare år som utförts av Skolverket på regeringens uppdrag. Urvalet har gjorts med utgångspunkt i om utvärderingen reser principiella frågor eller frågor som av andra skäl är särskilt intressanta att belysa ur ett utvärderingsperspektiv. Fokus ligger således inte på utvärderingarnas sakinhåll utan på hur utvärderingsuppdraget har formulerats och genomförts.

Utvärdering av försöksverksamhet med möjlighet för föräldrar att välja skolform för sina barn med utvecklingsstörning (2002)

Regeringen har vid flera tillfällen gett Skolverket olika uppdrag om att utvärdera försöksverksamheten. Rapporten *I särskola eller i grundskola?* (2002) utgör en fortsättning på två tidigare rapporter *Hur särskild får man vara? En analys av elevökningen i särskolan om värderingar* (2000) och *Kvaliteten i särskolan – en fråga om värderingar* (2001).

Denna rapport bygger på olika enkäter som skickats till sarsko-leansvariga i samtliga landets kommuner. Lärare och rektorer har även lämnat kompletterande synpunkter genom intervjuer. Den huvudsakliga empirin i rapporten bygger dock på observationer gjorda i nio fallstudierkommuner.

Rapporten avslutas med ett kapitel benämnt Slutdiskussion. Verket menar t.ex. att ”studien väcker några grundläggande frågor”, men kapitlet innehåller inga bedömningar av det slag som man kan förvänta sig av en utvärdering.

Kartläggning av de nationella minoriteternas utbildningssituation (2005)

Skolverket betecknar, i linje med regeringens uppdrag, rapporten *De nationella minoriteternas utbildningssituation* (Rapport 272, 2005) som en kartläggning och inte som en utvärdering. Empirin i rapporten utgörs av data hämtade från Skolverkets uppföljningssystem, en förstudie i form av telefonintervjuer med ansvarig för modersmålsundervisning på grundskolenivå (hur många som intervjuades framgår ej av rapporten). Vidare genomfördes fallstudier i fem kommuner och en enkät till samtliga kommuner (svarsfrekvens 86 %).

Trots att rapporten betecknas som en kartläggning och i huvudsak är deskriptiv innehåller den såväl bedömningar som förslag till åtgärder. Empirin och metodologin vid denna kartläggning skiljer sig inte från dem verket brukar använda i samband med utvärdering.

Uppföljning av maxtaxan (2007)

I rapporten *Fem år med maxtaxa. Uppföljning av reformen Maxtaxa och allmän förskola m.m.* (Rapport 294, 2007) anger verket följande syfte med uppföljningen:

Syftet med slutrapporten är att redovisa regeringsuppdraget samt att *diskutera* (vår kursivering) huruvida de mål och syften som anges i propositionen Maxtaxa och allmän förskola m.m. har uppnåtts, om det finns andra resultat eller effekter av reformen än de tänkta och hur dessa i sådant fall förhåller sig till de tänkta effekterna.¹

Skolverket betecknar rapporten som en uppföljning trots att de syften som anges snarare förknippas med utvärdering än med uppföljning. Rapporten bygger på empiri som består av offentlig statistik, föräldraenkäter och kvantitativa beräkningar av effekter av maxtaxan på vistelsetider och kommunernas ekonomi.² Rapporten avslutas med ett kapitel benämnt Diskussion och slutsatser. Kapitlet innehåller allmänt hållna bedömningar som inte fullt ut är kopplade till reformens fyra politiska mål.

¹ Skolverket (2007) a.a.

² Wikström, Magnus (2007) *Kommunalekonomiska effekter av maxtaxa och ökad tillgänglighet inom förskola och skolbarnomsorg*. Umeå Economic Studies no. 704.

Rapporten skulle lika gärna ha kunnat betecknas som utvärdering. De är dessutom oklart varför myndigheten väljer att spela ner betydelsen av resultaten genom att i syftet använda ordet ”diskutera”.

Statsbidraget till personalförstärkningar i skola och fritidshem (2007)

Skolverkets rapport *Statsbidrag till personalförstärkningar i skola och fritidshem* – slutrapport (Rapport 295, 2007) betecknas som en utvärdering. Verket använder sig av en mix av metoder: beskrivande statistik samt kvalitativa data i form av fallstudier och en genomgång av kommunernas uppföljningar av bidraget. Man har även låtit en forskare göra tre studier³ som med utgångspunkt i olika registerdata och med stöd olika statistiska metoder som t.ex. regressionsanalys bl.a. analyserar hur förändrade resurser i skolan i form av lärartäthet samvarierar med förändringar i elevernas studieresultat.

Skolverket konstaterar i slutkapitlet Avslutande diskussion och slutsatser att det sammantaget är svårt att belysa långsiktiga effekter i form av förändringar i elevers resultat eller grad av måluppfyllelse som en följd av statsbidragets införande. Verket bedömer dock att det har skett personalförstärkningar som motsvarar de uppsatta målen under perioden.

Romer i skolan – en fördjupad studie (2007)

I Skolverkets rapport *Romer i skolan – en fördjupad studie* (Rapport 292, 2007) redovisar verket ett regeringsuppdrag om att göra en fördjupad studie av situationen för romer i den svenska skolan. Orden ”värdera” eller ”bedöma” förekommer inte i uppdraget, men det verket enligt uppdraget ska göra ligger ändå mycket nära utvärdering: kartlägga kommunernas insatser, belysa situationen för romska barn, förslå åtgärder som kan öka kommunernas förutsättningar att informera om de rättigheter som romer och andra nationella minoriteter har.

³ Andersson, Christian (2007) *Lärartäthet, lärarqualitet och arbetsmarknaden för lärare*. IFAU. Rapport 2007:5.

Studien har genomförts som fallstudier (strukturerade intervjuer) vid arton skolor. Även Skolverkets nationella statistik har utnyttjats. Till studien har knutits sju representanter för romska organisationer som deltagit som experter och medverkat i hela studien. Det rör sig alltså i allt väsentligt om en kvalitativt inriktad studie.

Trots att studien inte är rubricerad som en utvärdering innehåller den ett antal förslag till åtgärder. Verkets förslag har till viss del grund i vad regeringens uppdrag, men går utöver den fråga där regeringen efterfrågar åtgärdsförslag.

Utvärdering av statsbidraget till personalförstärkningar i förskolan (2008)

Den empiriska grunden för Skolverkets rapport *Redovisning av uppdrag att utvärdera statsbidraget för basfärdigheterna läsa, skriva och räkna* (2013-03-25) utgörs av data som SCB årligen samlar in på Skolverkets uppdrag. Verket har även skickat ut en enkät till kommunerna där dessa ombetts att besvara fyra huvudfrågor: Vilka åtgärder som kommunerna använt statsbidraget till, hur bidraget fördelats, vilka resultat som uppnåtts med bidraget och övriga kommentarer.

En fallstudie har genomförts i åtta kommuner i syfte att få en bild av hur kommunerna fördelat statsbidraget och vilka personalförstärkningar som gjorts. Empirin har kompletterats med telefonintervjuer dels med de 17 kommuner som fått beslut om innehållande av bidrag eller återkrav, dels med de kommuner som varit undantagna från kravet på att öka personaltätheten.

För att belysa hur kommunerna har fördelat statsbidraget jämfört med ordinarie resursfördelningssystem har en annan studie av Skolverket använts (det s.k. RORIS-projektet).

För att besvara frågan om hur kommunerna har ökat personaltätheten har verket genom en multipel regressionsanalys undersökt vilka faktorer som har inverkan på en kommuns förutsättningar att genomföra personaltäthetsförändringar.

Skolverket analys utmynnar i konstaterandet att statsbidraget har haft avsedd effekt. Detta får anses vara en samlad bedömning från verkets sida.

Utvärdering av ämnet judiska studier i Stockholms kommun (2009)

Rapporten *Redovisning av regeringsuppdrag om ämnet judiska studier i Stockholms kommun* (2009-10-01) bygger på intervjuer, enkäter, dokumentanalys, observationer och statistisk analys av registerdata om antal elever och elevers betyg. Verket har även gjort sökningar i två lokala arkiv.

Apropå hur studien genomförts konstaterar Skolverket följande:

När det gäller hur själva undervisningen genomförts har en studie av detta slag emellertid klara begränsningar. Den empiri som samlats in om genomförandet av undervisningen utgår främst från elever och lärares uppfattningar om denna.

Utvärderingen utmynnar i ett antal slutsatser. Slutsatserna utgörs av en förteckning över brister och problem som verket funnit och kan därmed jämföras med en bedömning av utvärderingsobjektet.

Utvärdering av undervisning på engelska (2010)

Skolverkets rapport *Undervisning på engelska. Utvärdering av en försöksverksamhet i grundskolan* (Rapport 3351, 2010) bygger på data från verkets uppföljningssystem, en kommunenkät, en skolenkät och på halvstrukturerade gruppintervjuer på sex skolor med engelskspråkig undervisning.

Utvärderingen har en i huvudsak kvalitativ inriktning och en deskriptiv karaktär. Under rubriken Avslutande diskussion görs ett antal bedömningar, t.ex. om behovet av att stärka lärarnas kompetens.

Utvärdering av försöksverksamhet med lärlingsutbildning (2011)

För denna försöksverksamhet tillsatte regeringen den s.k. Nationella lärlingskommittén. Denna har haft till uppgift att följa, bedöma samt dra erfarenheter från den pågående försöksverksamheten. Kommittén har levererat tre betänkanden. Dessa har dock inte varit upplagda som utvärderingar.

Skolverket har förutom att redovisa omfattningen av försöksverksamheten och hur statsbidraget använts publicerat en rapport

(*Gymnasial lärlingsutbildning de tre första åren 2008–2011*, rapport 373, 2012) som handlar om deltagande, elevernas erfarenheter och studieresultat. Rapporten omnämns både som redovisning och som utvärdering (s. 16). Rapportens empiri är huvudsakligen kvantitativ och bygger främst på registeruppgifter men också på en enkät som skickades till 2 683 elever vid 163 skolor (svarsfrekvensen var 51,2 procent).

Till sin karaktär är rapporten deskriptiv och kartläggande och den innehåller inga bedömningar av det slag man förväntar sig av en utvärdering. Däremot avslutas rapporten med att ett antal s.k. förbättringsområden pekas ut.

Utvärdering av försöksverksamhet med riksrekryterande spetsutbildning i grundskolans högre årskurser (2011)

Skolverkets rapport *Redovisning av uppdrag enligt förordning (2011:355) avseende omfattning och utvärdering av försöksverksamhet med riksrekryterande spetsutbildning i grundskolans högre årskurser läsåret 2012/13*. Att Skolverket ska utvärdera försöksverksamheten framgår av förordningen. Verkets utvärdering av de åtta spetsutbildningar som påbörjats 2012/13 grundar sig på semistrukturerade intervjuer (dvs. har utgått från framtagna intervjuguiden) med rektorer, lärare och föräldrar som gjorts på respektive skola. Totalt har elva rektorer, arton lärare och trettiofyra föräldrar intervjuats. Vidare har verket intervjuat 53 elever i s.k. fokusgrupper om mellan sex till åtta elever på varje skola. Slutligen har en enkät gjorts bland elever som deltar i försöksverksamheten.

Rapporten innehåller bedömningar.

Utvärdering av mobbning (2011)

Skolverkets rapport *Utvärdering av metoder mot mobbning*. (Rapport 353, 2011) utgör en av de större utvärderingar som Skolverket genomfört. För utvärderingen anlätade myndigheten ett antal forskare. Som utgångspunkt för utvärderingen valdes en longitudinell, kvasiexperimentell design. Skälen var enligt rapporten att vederbörande praxis angående utvärdering av evidensbaserad programimplementering mer eller mindre uteslutande vilar på den randomiserade kontrollmodellen. Denna modell är dock, enligt rapporten,

inte möjlig för en statlig myndighet i Sverige ”som varken har rätt att ingripa i skolornas verksamhet eller av etiska skäl skulle kunna påbjuda skolor att arbeta med ett antimobbingsprogram vars effekter är osäkra”. Uttalandet görs förvisso i en enskild rapport men har en tydlig karaktär av principiellt ställningstagande.

Rapporten utmynnar i den övergripande bedömningen att inget specifikt program bör rekommenderas svenska grundskolor.

Utvärdering av matematiksatsningen (2011)

Skolverket fick 2009 ett uppdrag av regeringen att genomföra en satsning på matematikämnet i syfte att stärka skolornas eget utvecklingsarbete med att höja kvaliteten i matematikundervisningen, för att på sikt öka måluppfyllelsen. Skolverkets huvudsakliga aktivitet har varit att fördela medel till olika lokala projekt. Skolverket uppdrog 2011 åt Ramböll att utvärdera satsningen så långt (*Utvärdering av matematiksatsningen*. Ramböll 2011).

Huvudfrågeställningar har bland annat varit i vilken utsträckning satsningen bidragit till ökad kvalitet i undervisningen, om den påverkat kunskapsresultaten och den bidragit till metodutveckling. Empirin har erhållits genom enkäter (två omgångar till 2 473 projekt och 375 kontaktpersoner), fallstudier (7 st.), intervjuer med företrädare på Skolverket och Utbildningsdepartementet samt dokumentstudier.

Så långt följer Rambölls utvärdering den standard som utvecklats inom Skolverket. Vad som skiljer denna utvärdering från Skolverkets är att rapporten ger utrymme för en mer omfattande metodologisk diskussion. Denna diskussion handlar dels om svårigheter igenomförandet av utvärderingen – egentligen svårigheten att mäta måluppfyllelse. Fokus skiftade under arbetet med utvärderingen från effekter för eleverna till effekterna på undervisningen.

Utvärderingen bygger också på en programteoretisk ansats. Den talar om logiken hos matematiksatsningen som intervention i termer av insatser – resultat – effekt. Analysen görs i förhållande till en modell för programlogik.

Rapporten innehåller såväl bedömningar som slutsatser och rekommendationer.

Garanterad undervisningstid (2013)

Skolverkets rapport *Att fånga undervisningstiden med målen i fokus. Garanterad undervisningstid i grundskolan ur ett huvudmannas- och skolperspektiv* (Rapport 384, 2013) grundar sig på en empiri bestående av en insamling av data om skolors planerade undervisningstid, en enkät riktad till kommunala och enskilda huvudmän samt en intervjustudie. En statistisk analys är enligt Skolverket omöjlig att genomföra eftersom det i dagsläget inte är möjligt att koppla elevers resultat till planerad undervisningstid.

Enligt regeringens uppdrag ska verket genom fältstudier undersöka hur den garanterade undervisningstiden efterlevs och säkerställs samt analysera förhållandet mellan elevernas resultat och den garanterade undervisningstiden.

Rapporten avslutas med ett kapitel benämnt Sammanfattande diskussion. Av denna framgår att Skolverket inte ”på ett enkelt sätt” kan åskådliggöra ett statistiskt samband mellan elevernas resultat och undervisningstid. Verket finner det vidare problematiskt att de undersökta huvudmännen och skolorna inte följer bestämmelserna om den garanterade undervisningstiden.

Uppdraget om garanterad undervisningstid illustrerar för det första att regeringen beställer analyser som man inte försäkrat sig om att de är möjliga att genomföra. För det andra att regeringen direkt anger med vilken metod ett uppdrag ska genomföras. För det tredje att man ger ett uppdrag som har tydlig utvärderingskaraktär utan att man anger detta.

Kommunernas resursfördelning till grundskolor (2013)

Skolverkets rapport *Kommunernas resursfördelning till grundskolor* (Rapport 391, 2013) betecknas i enlighet med regeringens uppdrag som en fördjupad studie. Studien omfattar landets samtliga 290 kommuner, men i flera av analyserna ingår endast 50 s.k. urvalskommuner. Dessa kommuner – ”vilka kan räknas till landets mest bostadssegregerade kommuner” – har ingått i tidigare studier av Skolverket. Det empiriska underlaget för studien utgörs dels av en enkät ställd till cirka 350 grundskolerektorer i dessa 50 kommuner, dels av de kvalitetsgranskningar som Skolinspektionen genomfört i 30 av urvalskommunerna under 2012.

I rapporten studeras inte sambandet mellan kommunernas resursfördelning och skolors eller elevers resultat. Skälet är att verket anser att det finns metodologiska svårigheter med att genomföra den typen av analyser. Svårigheterna sammanhänger med "den komplexitet som finns när det gäller att försöka koppla kvantitativa data på ekonomiska resurser till kvaliteten på undervisningen och till elevernas studieresultat ... Effekterna på elevernas resultat beror i hög grad på hur resurserna används och vad som händer i klassrummet".

Skolverket menar vidare att data inte alltid tillåter "att kommunerna kan jämföras med varandra. Det beror bland annat på att kommunerna redovisar skolornas resultat på olika sätt och att inte alla uppgifter kan fås på skolnivå." Alla kommuner kan inte alltid koppla resurserna och kostnaderna till enskilda skolor.

Skolverket har i första hand använt så kallad regressionsanalys för att undersöka sambandet mellan kommuners kostnader och olika egenskaper hos kommunerna.

Rapporten avslutas med ett kapitel benämnt Diskussion och slutsatser. Kapitlet innehåller ett avsnitt med den tydliga rubriken Skolverkets bedömning. Skolverket bedömer bland annat att det knappast är möjligt att på empirisk väg identifiera det för alla kommuner effektivaste resursfördelningssystemet.

Utvärdering av statsbidraget för basfärdigheterna läsa, skriva och räkna (2013)

Av *Redovisning av uppdrag om uppföljning och utvärdering av statsbidraget till personalförstärkningar i förskolan* (Skolverket 2008-04-02) framgår att verket i enlighet med regeringens uppdrag anlitar en extern aktör för att utvärdera satsningen under åren 2008–2012. Uppdraget har bestått i att besvara hur statsbidraget använts och vilka effekter satsningen haft. Av uppdraget framgår inte vilka slags effekter man efterfrågar. Den externa utvärderaren har genomfört uppdraget i form av en enkätundersökning (20 kommuner och 20 fristående skolor) och genom intervjuer med företrädare för huvudmän som beviljat statsbidrag (40 skolhuvudmän, 30 rektorer samt 70 lärare).

Av förordningen framgår att statsbidraget får användas till att stärka arbetet med basfärdigheterna läsa, skriva och räkna för att

öka elevernas måluppfyllelse, särskilt för de elever som riskerar att inte nå målen.

Skolverket redovisar att medlen använts på följande sätt:

- Personalförstärkningar (47 procent)
- Läromedel (23 procent)
- Kompetenshöjande insatser (18 procent)
- Kompetenshöjande insatser m.a.a. den individuella utvecklingsplanen med skriftliga omdömen (7 procent)
- Övriga insatser (5 procent)

Utvärderingen följer de två målsättningarna att stärka arbetet med basfärdigheterna, respektive att öka måluppfyllelsen.

Av enkätsvaren framgår att nio av tio av de tillfrågade anser att statsbidraget har förstärkt pågående arbetet med basfärdigheterna. Två tredjedelar uppger att statsbidraget har bidragit till ändrade metoder i arbetet med basfärdigheterna.

Även för måluppfyllelse baseras bedömningen på uppfattningar redovisade i enkätsvar. Enkätundersökningen visar att så många som tre fjärdedelar av de svarande uppskattar att de elever i årskurs 1–3 som berörts av satsningen har nått en ökad måluppfyllelse i svenska. I matematik är det ungefär hälften av de svarande som uppger att elever i årskurs 1–3 har nått en ökad måluppfyllelse.

I princip alla intervjuade uppger att det är svårt att härleda observerade effekter till just denna satsning inte minst på grund av att statsbidraget i tid sammanfaller med den s.k. Matematiksatsningen.

Utvärderingen avslutas med ett avsnitt benämnt Slutsatser som innehåller allmänt hållna bedömningar: "Satsningen kan sägas ha bidragit till att stärka arbetet med basfärdigheter ..."

Internationell utblick - med fokus på organiseringen av nationell utvärdering

För att få ett perspektiv på det svenska systemet har utredningen valt att här belysa utvärderingssystem i andra länder, med fokus på hur länderna har organiserat sin utvärdering av skolan. Tyvärr visade det sig inte genomförbart att arrangera ett besök i England inom de tidsramar vi förfogat över. Systemet i England är dock väl dokumenterat, och intressant inte minst i relation till Skottland, varför vi tagit med även denna del av Storbritannien.

Danmark

Utvärderingen av skolan har stärkts i Danmark under senare år, bl.a. till följd av reformer som beslutades 2006. Varje elev ska numera ha en *individuell studieplan* och eleverna ska genomgå obligatoriska *nationella prov*. Vidare har *kvalitetsrapporter* och *självvärdering* tillkommit på skolnivå.

Syftet med det danska utvärderingssystemet för skolan är, enligt Undervisningsministeriet, att få kunskap om huruvida kvaliteten i folkeskolen förbättrats. Den löpande utvärderingen är, menar man, en del av den enskilda folkeskolens kvalitetsutveckling.

Det danska systemet karaktäriseras, som vi uppfattar det, av en tydlig strävan att utvärderingsresultaten i första hand ska bidra till den lokala skolutvecklingen. Det gäller även det sätt på vilket tillsynen genomförs.

På övergripande nationell nivå (systemnivå) används i första hand resultat från internationella undersökningar, nationella resultat från de nationella proven samt betygsgenomsnitt från de avslutande proven i grundskolan för att följa utvecklingen. Därutöver genomförs externa utvärderingar. Att resultaten från de nationella proven kan användas för att följa utvecklingen på nationell nivå

beror på att frågorna inte är offentliga och således kan återanvändas. De kan därför inte bidra till lärarnas gemensamma tolkning av målen, dvs. vara förebildliga, men ger kommuner och skolor en signal om kunskapsnivån. Samtliga elever deltar i proven som genomförs digitalt. Man har avstått från urvalsundersökningar eftersom man menar att sådana inte ger skolorna något tillbaka.

OECD pekar vidare på att Danmark överfört de flesta (nio av tretton) forskningsinstitutioner som arbetar med reformutvärdering från staten till universitetssektorn. Erfarenheterna från Danmark visar enligt OECD att fackdepartementen kan få tillgång till oberoende och relevant forskning utan att denna genomförs inom staten. OECD noterar att en sådan överföring inte är problemfri. Kärnfrågan handlar – föga förvånande – om vilket inflytande den politiska nivån kan ha på beslut om kommande forskningsinsatser. OECD rekommenderar dock att Sverige följer det danska exemplet och överför ansvaret för all statligt beställd reformutvärdering till antingen universitet eller till specialiserade forskningsorgan.

Aktörer på nationell nivå

Styrelsen för IT och lärande (UNI-C), som är en del av Undervisningsministeriet, ansvarar bl.a. för ministeriets statistik och analys och bidrar till ministeriets politikutveckling och uppföljning med datainsamling, kvalitetssäkring av data, analyser, modellarbete samt förmedling av statistik och analyser.

2006 inrättades en särskild myndighet med ansvar för uppgifter kopplade till granskning och utvärdering av den danska skolan. Myndigheten omorganiserades 2011 och fick i samband med namnet *Kvalitets- och tillsynsstyrelsen*. Styrelsen, som även den utgör en del av Utbildningsministeriet, har följande fyra huvuduppgifter:

- *Kvalitetsutveckling*, bl.a. genomförande av de internationella studier som Danmark deltar i samt framtagande av stödmaterial till skolorna om lokal utvärdering
- *Prov och examen*, dvs. utveckling och genomförande av de nationella proven samt de avslutande examensproven
- *Tillsyn*
- *Bidrag och kontrollverksamhet*

Även inrättandet av *Rådet för utvärdering och kvalitetsutveckling i folkeskolen (Skolerådet)* tillhör de initiativ som togs 2006. Rådet är ett oberoende organ vars uppgift är att följa, värdera och ge råd till undervisningsministern om kvaliteten i folkeskolen och ungdomsskolen. Rådet består av ett ordförandeskap och 21 övriga medlemmar som representerar en rad centrala intresseorganisationer på folkeskole- och ungdomsskoleområdet. Skolerådet leds av ett ordförandeskap som består av 3–5 personer med särskild insikt i folkeskolen.

Det är Skolerådets uppgift att besluta om vilka externa utvärderingar som ska genomföras på skolans område, av reformer eller av annan karaktär. Sådana uppdrag kan läggas direkt på *Danmarks Evalueringsinstitut (EVA)* eller upphandlas i konkurrens. I det senare fallet kan EVA delta som en anbudsgivare bland andra. Andra anbudsgivare kan vara konsulter inom utvärderingsområdet eller forskare inom högskolan.

Danmarks Evalueringsinstitut (EVA) arbetar med utvärdering, analyser och redskap som bidrar till att utveckla kvalitet. EVA arbetar med utvärdering och kvalitetsutveckling från förskola till högskola. Den kunskap som produceras används på alla nivåer i utbildningssystemet – från verksamhetsnivå till kommuner och ministerier.

EVA inrättades som ett oberoende statligt institut under Undervisningsministeriet 1999. En föregångare var Evalueringscentret och som existerade under 1992–1999. Sedan 1999 har EVA styrts av en egen lag, som har setts över med jämna mellanrum. Fram till den större översyn som gjordes 2005 beslutade EVA:s egen styrelse om allt som har med EVA:s utvärderingsverksamhet att göra. Från 2005 beslutas det som har att göra med utvärdering av folkeskolen av Skolerådet. Eva tar dock fram förslag till utvärderingsprojekt,

EVA har genomfört flera utvärderingar av 2006 års skolreformer. För den nu pågående reformeringen av folkeskolen finns än så länge ingen plan för utvärdering. När gymnasieskolan reformerades 2005 fanns en sådan plan och EVA fick flera utvärderingsuppdrag.

Evidens har blivit ett allt viktigare inslag i dansk skolpolitik. EVA föredrar dock att tala om en vetenskaps- eller kunskapsbaserad verksamhet och har en vidare syn på vad som bör betraktas som effektutvärdering än vad företrädare för den s.k. evidensrörelsen har. Man lägger stor vikt vid att ge verktyg till de operativt verksamma så att de kan utveckla sin praxis och menar att det är viktigt att lärarna förstår betydelsen av utvärdering.

Utöver att genomföra nationella utvärderingar erbjuder EVA även rådgivning och sparring till kunder, som t.ex. avser att genomföra projekt eller starta ett utvecklingsarbete. Man kan hjälpa till med att formulera klara mål för projektet och säkra ett effektivt samspel med andra parallella projekt samt verkar för att brukarna ska bli ägare av projektet och dess resultat.

Lokala huvudmän kan även få hjälp av EVA med att genomföra utvärderingar och utredningar inom en rad olika teman på skolområdet. Det kan handla om att forma utvärderingens eller utredningens design och att genomföra utvärderingen eller utredningen.

The Danish Clearinghouse for Educational Research är en enhet vid universitetet i Århus och grundades 2006¹. Tillkomsten hade sin bakgrund i den kritik som OECD vid den tidpunkten riktat mot Danmark för bristande koppling mellan dansk utbildningsvetenskaplig forskning och hur forskningsresultat användes vid politiskt beslutsfattande².

Danish Clearinghouse ska bidra till att politiker och praktiker har tillgång till tillförlitlig och kvalificerad kunskap om fostran, undervisning och utbildning som kan användas i den pedagogiska praktiken och vid politikutformning. Danish Clearinghouse samlar in, analyserar och förmedlar kunskap från antingen forskning om utbildning eller från försöks- och utvecklingsarbete på danska skolor. Man genomför ingen egen primärforskning eller egna utvärderingar. Danish Clearinghouse hjälper även forskningsmiljöer att få en större och säkrare överblick av befintlig forskning så att de lättare kan identifiera var det finns behov av mer forskning.

Danish Clearinghouse använder metodiken systematic reviews. Man menar att varje steg i den mycket bestämda processen är en möjlig självständig produkt. I vissa fall kan det t.ex. vara tillfyllest att i ett bibliotek tillgängliggöra en samlad och systematisk överblick över tillgänglig information inom ett område utan att fullfölja processen till en fullständig genomförd systemtic review. Man har vidare anslutit till den grupp av organisationer som menar att systematic reviews kan baseras på såväl kvantitativa som kvalitativa studier.

¹ www.eippee.eu/cms/Default.aspx?tabid=3240

² The Danish Clearinghouse for Educational Research. Concept note, v.1.2. June 2007.

England

Olika *nationella prov* utgör en central del av utvärderingsramverket i England. Dessa prov kallas i dag för *National Curriculum Tests* (NCT:s).³

Vid 5 års ålder bedömer elevernas lärare deras utveckling i förhållande till the Foundation Stage profile (från och med 2008). Därefter genomförs nationella prov vid 7 års ålder, vid 11 års ålder samt vid 14–16 års ålder i årskurserna 7, 8 och 9. Inledningsvis gör eleverna nationella prov i engelska och matematik, från 11 års ålder även i de naturvetenskapliga ämnena (science). Merparten av proven bedöms av lärarna. Provet i matematik vid 11 års ålder bedöms dock externt och skolornas resultat publiceras. Det finns även alternativa standardiserade prov som stöd för lärarnas bedömning i engelska och matematik för åk 7 och 8 och i naturvetenskap för åk 9.

Den brittiska regeringen uppdrog 2010 till en kommitté att se över testsystemet, sedan omfattande kritik riktats av bland annat lärare och skolledare mot främst de tester som genomförs vid 11 års ålder. Provsystemet har även utsatts för kritik från vetenskapligt håll. Till exempel har Paul Black⁴ pekat på att tilltron till ökad extern bedömning genom nationella tester innebär att betydelsen av lärares formativa bedömning för skolresultaten underskattas samt att provsystemet har brister i fråga om reliabilitet och validitet.⁵ Invändningar avseende provsystemets tillförlitlighet har även rests av andra forskare.⁶

Kommittén⁷ rekommenderade bl.a. ökat förtroende för lärarna och en rimlig balans mellan lärares egna bedömningar och olika test samt ökad betoning på elevernas kunskapsutveckling. Det skriftliga provet borde ersättas med lärarbedömning, medan matematik även fortsättningsvis borde utvärderas externt. Regeringen godtog kommitténs förslag som grund för det fortsatta reformarbetet.⁸

³ Proven benämndes tidigare Standard Attainment Tests (SAT:s).

⁴ Paul Black var ordförande i den arbetsgrupp (TGAT) som i slutet av 1980-talet hade i uppdrag att utarbeta ett förslag till hur ett nytt system för utvärdering av elevernas kunskapsresultat skulle utformas.

⁵ Black, Paul (1998) *Learning, League Tables and National Assessment: opportunity or hope deferred?* Oxford Review of Education, Vol. 24, No 1.

⁶ Se t.ex. Stobart, Gordon (2001) *The validity of National Curriculum Assessment*. British Journal of Educational Studies, Vol. 49, No. 1 och Harlen, Wynne and Ruth Deakin Crick (2002) *A systematic review of the impact of summative assessment and tests on students' motivation for learning*. Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre).

⁷ Bew, Lord (2011) *Independent Review of Key Stage 2 testing, assessment and accountability*. Final Report. June 2011.

⁸ Department for Education (2011) *Independent Review of Key Stage 2 testing, assessment and accountability. Government Response*.

Skolinspektion har länge utgjort ett tungt inslag i det engelska utvärderingsramverket. Inspektionen genomförs av *The Office for Standards in Education, Children's Services and Skills* (Ofsted) som är en oberoende myndighet som etablerades vid mitten av 1990-talet. Från att ha haft ett uppdrag som syftade till att inspektera och ställa krav på skolornas prestationer har myndighetens uppdrag under senare år mer inriktats mot att stödja skolornas arbete med självvärdering.

Aktörer på nationell nivå

Arbetsfördelningen mellan regeringen och olika delar av den centrala statsförvaltningen är i England en annan än i Sverige. *Det engelska utbildningsministeriet* (*Department of Education, DfE*) är större än sin svenska motsvarighet och rymmer funktioner som i Sverige ligger på självständiga myndigheter. Det innebär att ministeriet självt har en stor betydelse inte bara som utformare av utvärderingssystemet, utan också för driften av delar av detta system.

I Storbritannien finns ett omfattande datainsamlingssystem (*National Pupil Database, NPD*) som i allt väsentligt motsvarar det svenska uppföljningssystemet. Medan det i Sverige är en självständig myndighet som ansvarar för insamlingen och sammanställningen av provdata och andra data (socio-ekonomiska, demografiska m.m.) sker detta i England inom en särskild avdelning (*Data and Statistics Division, DSD*) inom utbildningsministeriet.

Standards and Testing Agency (STA) är en s.k. verkställande myndighet (*executive agency*⁹) inom DfE. STA ansvarar för utveckling och genomförande av alla lagstadgade kunskapsprov från och med de tidiga skolåren till och med slutet av key stage 3 (årskurs 9).

The Economic and Social research Council (ESRC) är Storbritanniens största organisation för att finansiera forskning om ekonomiska och sociala frågor. ESRC inrättades 1965 och är ett icke-offentligt organ (NDPB) som får huvuddelen av sin finansiering genom Department for Business, Innovation and Skills (BIS). ESRC stödjer oberoende, högkvalitativ forskning som har en inverkan på

⁹ De övriga verkställande organen inom DfE är Education Funding Agency och National College for Teaching and Leadership. Dessa organ betraktas som delar av departementet, men behandlas lednings- och budgetmässigt separat och har till uppgift att utföra någon del av de exekutiva funktionerna i Förenade kungarikets regering.

näringslivet, den offentliga sektorn och den tredje sektorn¹⁰. ESRC stödjer omkring 4 000 forskare och doktorander inom akademiska institutioner och fristående forskningsinstitut.

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) inledde sin verksamhet 1993. Centret genomför systematiska kunskapsöversikter och utvecklar metoder för att ta fram forskningsöversikter inom det samhällsvetenskapliga området. Syftet är att göra tillförlitliga forskningsresultat tillgängliga ”för de människor som behöver dem, oavsett om de handlar om politiskt beslutsfattande, yrkesutövning eller personliga beslut”. Arbetet bedrivs huvudsakligen inom utbildnings- och hälsovårdsområdet samt inom det sociala området. Centret samverkar med en lång rad forskningsinstitutioner och arbetet finansieras av det Ekonomiska och sociala forskningsrådet (ESRC), flera myndigheter och välgörenhetsorganisationer samt nationella och internationella partners.

The Centre for the Economics of Education (CEE) tillkom 2000 och fick fram till 2010 ett grundbidrag från utbildningsministeriet och ministeriet för näringsliv, innovation och kompetens (the Department for Business, Innovation & Skills). CEE är ett multidisciplinärt centrum med tre parter: The Centre for Economic Performance at LSE, the Institute for Fiscal Studies och the Institute of Education. De tre parterna har en framstående meritlista i fråga om arbetsmarknadsfrågor och utbildningsekonomi inklusive yrkesutbildning. CEE:s forskning är starkt orienterad mot empiriska studier. Man använder en mångfald metoder som t.ex. utvärderingsmetodik, statistisk analys av mikrodata och detaljerade fallstudier.

Finland

Det finska skolsystemet har inte på länge varit föremål för några stora reformer eller förändringar. I juni 2012 beslutades emellertid om att inleda arbetet med en översyn av läroplansgrunderna bl.a. för den grundläggande utbildningen. Arbetet leds av Utbildningsstyrelsen och ska ske i nära samarbete med intressegrupperna och i interaktion med utbildningsanordnaren och arbetet ska vara avslutat i april 2014. De nya läroplanerna ska tillämpas i undervisningen från och med hösten 2016. Våren 2012 beslutades även om en

¹⁰ Med tredje sektorn brukar avses t.ex. idéburna organisationer, ekonomiska föreningar och kooperativa företag, vissa stiftelser samt det civila samhället.

revidering av timplanerna. Vidare görs fortsatt satsning på lärare och på att behålla läraryrkets attraktionskraft.

2012 antogs vidare nya kvalitetskriterier för bl.a. den grundläggande utbildningen. Dessa syftar till att göra det möjligt att på ett praktiskt plan utvärdera hur de beslut som omfattar skolväsendet i ett kort respektive långt perspektiv inverkar på skolornas verksamhet.

Utvärdering av utbildningen görs på lokal, regional och riksnivå. Utbildningsanordnarna är skyldiga att utvärdera sin utbildning, främst i förhållande till målen för utbildningen, samt delta i externa utvärderingar som baserar sig på utbildningslagstiftningen. Regionalt görs utvärdering och uppföljning av skolan av länsstyrelserna, t.ex. avseende service och hur efterfrågan på utbildning uppfylls. Det riksomfattande systemet för utvärdering av utbildningen bygger på

- ett system för utvärdering av inlärningsresultaten (stickprov)
- framtagande av indikatorer
- utvärderingsprojekt med varierande teman (läges- eller temautvärderingar).

Dessutom deltar Finland i internationella utvärderingar.

I lagstiftningen anges att utvecklingen av utvärderingssystemet bör inriktas på oavhängighet, träffsäkerhet, effektivitet och ekonomi samt effekter. Med effekter avses, vad vi förstår, snarast de effekter utvärderingarna har, inte effektstudier. Vikten av att skapa ett system som är direkt och betjänar läroanstalterna, eleverna och vårdnadshavarna betonas.

Ett generellt intryck är att det finska skolsystemet bygger på stark tillit från den nationella nivån till kommunerna och ytterst till professionen. Det finns ingen nationell tillsyn och t.ex. står inte frågan om en likvärdig och "rättvis" betygssättning i fokus för debatten på det sätt som den gör i Sverige.

Aktörer på nationell nivå t.o.m. 2013

På central nivå har det fram till nyligen funnits främst två aktörer vad gäller nationell utvärdering på skolområdet – Rådet för utbildningsutvärdering och Utbildningsstyrelsen.

Den externa utvärderingen av skolan har under en följd av år organiserats av *Rådet för utbildningsutvärdering*, ett självständigt organ som sorterat under undervisnings- och kulturministeriet. Rådet har bestått av högst 14 medlemmar, som utnämnts av undervisnings- och kulturministeriet. Medlemmarna är företrädare för såväl akademien som för kommunsektorn och för de fackliga organisationerna. Tillsammans med undervisnings- och kulturministeriet har rådet ansvarat för utvärdering, och utveckling av utvärdering, i grundskola och gymnasium, inom yrkesutbildningen och den yrkesinriktade vuxenutbildningen samt inom det fria bildningsarbetet.¹¹ Sekretariatet för Rådet för utbildningsutvärdering verkar som en separat institution vid Jyväskylä universitet.

Utvärderingsrådet har utvecklat utvärderingsmetoder för olika aktörers behov. Inom nationella utvärderingar tar man tillvara utbildningsanordnarnas och läroanstalternas självutvärderingar och lyfter fram god praxis. För att utvärderingen av utbildningen ska bli så transparent och använd som möjligt har publiceras resultaten i utvärderingsrådets publikationsserie. Skolor rangordnas inte i rapporterna.

*Utbildningsstyrelsen*¹² svarar för utveckling av utbildningen, från förskoleundervisningen till vuxenutbildningen och det fria bildningsarbetet samt den grundläggande konstundervisningen. Det innebär att en central uppgift för Utbildningsstyrelsen är att arbeta med läroplansgrunderna för utbildningarna, grunderna för yrkesexamina och fristående examina samt att utveckla utbildningen genom olika projekt. Bland övriga uppgifter ingår att producera utbildningsindikatorer och prognoser, samla uppgifter om utbildningens finansiering, publicera utbildningsguider, ordna och finansiera fortbildning för lärare och annan personal inom utbildningssektorn samt utveckla och producera läromedel.

Fram till nu har Utbildningsstyrelsen även utvärderat inlärningsresultat inom sitt verksamhetsområde, d.v.s. från förskoleundervisning till vuxenutbildning. Syftet med styrelsens utvärdering av utbildningen har varit att utreda i vilken mån de uppställda målen för undervisning och utbildning uppnås.

De viktigaste målen för Utbildningsstyrelsens utvärdering av utbildning har varit att

- utveckla utbildningen och stödja lärandet

¹¹ Rådet för utbildningsutvärdering (2004) a.a.

¹² Informationen i detta avsnitt är hämtad från Utbildningsstyrelsens webbplats (www.oph.fi) den 12 juni 2013.

- säkra kvaliteten på utbildningen
- producera information som behövs i utvecklingsarbete och beslutsfattande på lokal, regional och riksomfattande nivå samt för internationella jämförelser.

Utbildningsstyrelsen har svarat för de sampelbaserade utvärderingar av inlärningsresultat inom den grundläggande utbildningen som genomförts sedan 1998.

Aktörer på nationell nivå fr.o.m. 2014

Undervisningsministeriet har nyligen beslutat om en utvärderingsplan för utbildningen 2012–2015. Av denna framgår utgångspunkter för utvärderingen, organisering av utvärderingsverksamheten, riktlinjer för utvärderingen, dvs. vad den externa utvärderingen ska fokusera på, finansiering samt uppföljning av utvärderingsplanen. Man kan säga att detta på en övergripande nivå utgör ett slags ramverk för utvärdering.

Av utvärderingsplanen framgår att utvärderingsfunktionerna vid Rådet för utbildningsutvärdering, Rådet för utvärdering av höskolorna och vid Utbildningsstyrelsen från och med början av 2014 ska koncentreras till ett nationellt *center för utbildningsevaluering*. Syftet är att samla utbildningsutvärderingens viktigaste uppgifter och kompetenser till en tydligare, effektivare och mer slagkraftig helhet och att stärka utvärderingsverksamhetens oberoende och självständighet. Därtill, menar man, uppnås funktionella och ekonomiska synergier samt stärks den finska utbildningsutvärderingens nationella och internationella roll. Förändringen innebär att t.ex. de sampelbaserade utvärderingarna av inlärningsresultat som Utbildningsstyrelsen svarat för blir ett ansvar för det nya centret.

Nationella centret för utbildningsevaluering föreslås av den arbetsgrupp som förberett inrättandet, vara ett ämbetsverk inom undervisnings- och kulturministeriets förvaltningsområde. Centret föreslås bli ett expertorgan som utför externa utvärderingar av utbildningar och för att ta fram kunskap som underlag för utbildningspolitiska beslut och utveckling av utbildningen. Centret föreslås av en särskild utredare placeras i Helsingfors med en del av sin verksamhet fortsatt i Jyväskylä.¹³

¹³ Beskrivningen av förslaget är hämtad från Undervisnings- och kulturministeriets hemsida.

Beslutsfattandet föreslås bli delat mellan direktören och ett evalueringsråd med ansvar för den strategiska styrningen. Evalueringsrådets medlemmar föreslås företräda sakkunskap på området, t.ex. expertis inom utvärdering av utbildning. Evalueringsrådet föreslås vidare bestå av personer som företräder olika utbildningssektorer, lärarutbildningen, forskningen, arbetslivet och de studerande.

Enligt förslaget ska direktören ha beslutanderätten för den operativa ledningen medan evalueringsrådet ska fatta beslut om centrets strategiska riktlinjer. Rådet ska ha en central roll vid utarbetandet av förslag till den utvärderingsplan som fastställs av undervisnings- och kulturministeriet. Planen är det viktigaste redskapet för den strategiska styrningen av utvärderingsverksamheten. Rådet ska även organisera särskilda utvärderingar samt besluta om resultatet av en granskning godkänts eller inte. Enligt förslaget utnämner statsrådet (regeringen) direktören för evalueringscentret och tillsätter evalueringsrådet.

Norge

Norge har under flera år arbetat intensivt med att förstärka den nationella uppföljningen och utvärderingen inom utbildningsområdet och det finns numera en stark utvärderingstradition i norsk utbildningspolitik¹⁴. Detta präglar såväl det löpande uppföljnings- och utvärderingssystemet som det sätt som reformer utvärderas på.

Utdanningsdirektoratet använder ett nationellt, stabilt ramverk för att utvärdera och rapportera om det norska skolsystemet som innehåller fem huvudområden: läranderesultat, lärandemiljön, andel elever som fullföljer studierna i gymnasieskolan, resurser och fakta om skolorna. Norge genomför dessutom återkommande granskningar av prioriterade politikområden. Ett aktuellt exempel är det särskilda indikatorssystemet för att följa kvaliteten i lärarutbildningen och i yrket.

På systemnivå har Norge utvecklat ett system för kontinuerlig uppföljning, *Nasjonalt kvalitetsutvurderingssystem* (NKVS). Systemet, snarast att betrakta som ett ramverk för utvärdering, försörjer bland annat olika forskningsprojekt med data. Datainflödet sker

¹⁴ Avsnittet om Norge bygger på information som lämnats till utredningen i samband med dess kontakter med företrädare för det norska Kunskapsdepartementet och Utdanningsdirektoratet i Norge.

genom Norges deltagande i internationella studier, statistikinsamling, resultat från nationella prov, brukarundersökningar och enkäter.

Den insamlade informationen publiceras på flera olika sätt: särskilda rapporter om specifika teman, årliga analytiska rapporter, utvärderingsrapporter i nationellt prioriterade frågor, internetpublicering och en årlig summerande rapport baserad på utvärdering av skolor.

Aktörer på nationell nivå

Utdanningsdirektoratet, som lyder under Kunnskapsdepartementet, har ett brett arbetsfält som spänner från läroplaner och ramverk för motsvarande förskola, grundskola och gymnasieskola, till examen och analyser samt regelverk och tillsyn. Direktoratet ansvarar för merparten av de olika typer av uppgiftsinsamling som igår i NKVS.

Utdanningsdirektoratet har t.ex. ansvar för att utarbeta nationella prov och andra prov för grundskolan och gymnasieskolan samt att utforma villkoren för examen, dvs. den avslutande bedömningen av en elevs kompetens. Direktoratet har också ansvar för tillsynen över landets skolhuvudmän, dvs. kommuner, fylkeskommuner och privata skolor. Syftet är att säkra att barn och unga får en likvärdig utbildning av hög kvalitet.

Direktoratet har dessutom ansvar för utvecklingsprojekt, forskning och statistik om grundskolan och gymnasieskolan. Brukarundersökningar genomförs för elever, lärlingar, lärare, instruktörer och föräldrar, vilket förser skolor, skolhuvudmän och myndigheter med viktig kunskap. Resultaten och analyserna, som publiceras på Utdanningsdirektoratets webbsidor, är tänkta att bidra till förbättring av norsk skola. Utdanningsdirektoratet har vidare ansvar för att koordinera de många internationella studier som Norge deltar i.

Utdanningsdirektoratet har upprättat ett *nationellt forskningsråd* som ger råd i frågor om utveckling av forskningspolicy och om finansiering av forskning samt skapar mötesplatser. De flesta uppdragen anförtros olika forskningsinstitut, universitet och högskolor under Utbildnings- och forskningsministeriets och Utdanningsdirektoratets överinseende.

I Norge finns dessutom ett *Kunnskapssenter for udanning* knutet till Norges forskningsråd. Kunskapscentrets uppgift är bl.a. att presentera systematiska kunskapsöversikter i en öppen och tillgänglig

databas som ska kunna användas av alla som arbetar i förskola, skola, högre utbildning och vuxenutbildning samt med politikutformning. Centret ska sammanställa både nationell och internationell forskning och förmedla den bästa tillgängliga kunskapen inom utbildningsområdet.

Organisering av reformutvärderingar

Norge lägger allt större vikt vid kunskapsunderlag inför beslut om politiska reformer. Evidensinformerade inslag i samband med politikformulering kommer bland annat till uttryck i att litteraturlistor redovisas i de senaste årens stortingsmeldingar (propositioner). Till större reformer som genomförts under de senaste decennierna har dessutom knutits omfattande utvärderingsresurser. Sättet att organisera utvärderingarna har skilt sig åt mellan de olika reformerna, t.ex. med avseende på hur samarbetet mellan politiken och forskningen har utformats.

I fallet med *Reform -94*, som avsåg gymnasieskolan, beslutade Stortinget om en löpande utvärdering som skulle vara forskningsbaserad, dvs. grunda sig på forskningsmässig metodik och kompetens. Departementets kontroll av arbetet med utvärderingen var tydlig. Det följdes nära eftersom det fanns ett uttalat önskemål om att kunna ta nya beslut under vägen. För att säkra att utvärderingen blev policyrelevant skrevs ettåriga kontrakt mellan forskare och departementet. Forskarna hade krav på sig att lämna rapporter vid vissa tidpunkter (varje halvår och år). För att täcka behovet av kompetens genomfördes utvärderingen av flera forskningsmiljöer i samverkan.

För att utvärderingen skulle bli intressant och leva upp till forskningsmässiga krav krävdes samtidigt att forskarna hade ett självständigt handlingsutrymme. För att skapa ett visst avstånd mellan uppdragsgivare och uppdragstagare skapades en uppföljningsgrupp bestående av nio medlemmar med olika kompetenser. Ordföranden i uppföljningsgruppen utnämndes av Norges forskningsråd. Uppföljningsgruppen hade den överordnade kontrollen över projektet och kommenterade utkast till rapporter från forskargrupperna. Uppföljningsgruppen hade vidare till uppgift att ge råd till departementet om styrningen och vid kontraktsförhandlingar. Uppföljningsgruppens sekretariat utgjordes av tjänstemän i Kunskapsdepartementet.

Utöver uppföljningsgruppen tillsattes även en referensgrupp med representanter för flertalet av dem som berördes av reformen. En av forskningsmiljöerna påtog sig också rollen som koordinerande forskningsinstitution.

Arrangemangen runt utvärderingen ledde till en upprörd debatt om att regeringen direktstyrde forskningen och om det olämpliga i att departementet var uppdragsgivare till utvärderingen av en reform som det självt var ansvarigt för. Samarbetet mellan parterna överlevde dock debaclet och ledde efterhand till en ökad vana vid dialog parterna emellan. Dessutom ägde en viktig kompetensuppbyggnad rum inom forskningsmiljöerna, som har kommit till användning vid senare reformutvärderingar.

I samband med *Reform -97*, som avsåg grundskolan, var skolministerns utgångspunkt att den kritikstorm som utbröt i samband med den föregående reformutvärderingen inte skulle upprepas. Denna gång medgav därför upplägget ett stort utrymme för forskningen att själv bestämma inriktning och rapporteringstillfällen.

Utvärderingarna av de reformer som genomfördes 1994 och 1997 kan således sägas representera två extremer på en skala mellan kraftig styrning och stor självständighet. I utformningen av utvärderingen av *Kunskapslyftet 2006*, som avsåg grundskolan och gymnasieskolan, sökte man dra nytta av erfarenheterna från de båda föregående utvärderingarna.

Utdanningsdirektoratet tillsatte en *programstyrelse* som leddes av en välmeriterad och erkänd forskare. Även om Programstyrelsen fick sitt uppdrag av Utdanningsdirektoratet var det Programstyrelsen som utlyste uppdragen och slöt kontrakt med de olika forskningsmiljöerna och därefter höll i kontakten med dessa. Programstyrelsen har bestått av företrädare för olika intressenter på skolområdet, t.ex. brukare (elever och föräldrar), fackliga organisationer, forskare och personer på policynivå.

Programstyrelsen svarade även för återkopplingar till forskarna. Dessa återkopplingar innebar även en diskussion om kvaliteten i utkasterna till forskarrapporterna, dvs. om teori, metoder, empiriskt underlag, konsistens mellan underlag och slutsatser, osv. Man samlade dessutom in frågeställningar som inom olika delar av utbildningssektorn ansågs viktiga att granska och sorterade fram dem som bedömdes vara av generellt intresse för sektorn och som dessutom var knutna till reformen. De utvalda frågeställningarna infogades därefter i utvärderingsuppdragen till de olika forskargrupperna.

Tjänstemännen i departementet höll den politiska ledningen informerad om resultaten från utvärderingarna samt lyfte in dem i olika aktuella politiska processer.

Skottland

Skottland¹⁵ har sedan början på 2000-talet arbetat med att utveckla och introducera en ny läroplan i barn- och ungdomsskolan – *Curriculum for Excellence*. Läroplanen introducerades i de lägre årskurserna under läsåret 2010/11 och under läsåret 2012/13 för elever i secondary schools, dvs. för elever från ca 11 till 18 års ålder. Arbetet med att implementera läroplanen har dock varit en successiv process som fortfarande pågår. Arbetet leds av The Curriculum for Excellence Management Board som består av representerer för såväl den statliga förvaltningen som företrädare för en lång rad organisationer som företräder t.ex. skolhuvudmän, lärare och föräldrar.

Den nya läroplanen syftar till att säkerställa att alla barn och ungdomar i Skottland utvecklar kunskaper, färdigheter och olika personliga förmågor. Enligt läroplanen ska eleverna ges möjlighet att uppleva framgång som elev, känna självförtroende som individ, utvecklas till ansvarstagande medborgare och kunna bidra effektivt till det gemensamma samhällslivet.¹⁶

The Scottish Survey of Literacy and Numeracy (SSLN) är en årlig urvalsundersökning som följer de nationella resultaten i läsning och räkning (växelsvis vartannat år) vid tre åldrar under motsvarande grundskolan. SSLN utvecklades för att stödja bedömningsformer i anslutning till Curriculum for Excellence. SSLN är ett samarbete mellan den skotska regeringen, Education Scotland och Scottish Qualifications Authority (SQA). Skottland deltar även i de internationella PISA-undersökningarna.

Det skotska systemet är annars i hög grad inriktat på att stödja den lokala nivån i dess arbete med att utveckla och utvärdera verksamheten. *The National Assessment Resource (NAR)* är t.ex. ett verktyg online för att stödja lärares bedömning i anslutning till Curriculum for Excellence.

¹⁵ Skottland är en självstyrande riksdel (eller constituent country) inom Storbritannien. Skottland har numera ett omfattande självstyre, bl.a. på skolpolitikens område, varför Skottland i detta sammanhang jämföras med ett land.

¹⁶ Fri översättning av de grundläggande förmågorna Successful learners, Confident individuals, Responsible citizens och Effective contributors.

Den skotska inspektionen har en viktig roll. Varje år inspekteras ett urval skolor och kommuner. Inriktningen är att bidra till förbättring. Inspektörerna fokuserar på elevernas lärande och prestationer. Skolornas självvärderingar är viktiga utgångspunkter för inspektionen. Inspektionen förfogar inte över några sanktionsinstrument.

Generellt präglas arbetet med att utveckla skolan i Skottland av en skolpolitisk konsensus. Denna har dock inte vilat på något formellt kontrakt, utan bygger mer på en tradition där inte minst dialogen mellan olika parter inom skolsamhället ses som central.

Aktörer på nationell nivå

Organisatoriskt har Skottland genomgått en motsatt utveckling jämfört med Sverige. 2011 bildades *Education Scotland* bl.a. genom en sammanslagning av organisationen Learning and Teaching Scotland (LTS) och den tidigare inspektionsmyndigheten Her Majesty's Inspectorate of Education in Scotland (HMIe). LTS huvuduppgift var att utveckla och stödja den skotska läroplanen och att utveckla kvaliteten på lärande och undervisning, bl.a. genom att erbjuda råd, stöd, kompetensutveckling, etc.

Education Scotland har i dag ett brett ansvarsområde. Man ansvarar för råd och stöd i syfte att utveckla undervisningen och elevernas lärande, bl.a. genom arbetet med Curriculum för Excellence och genom IT. Man svarar också för inspektionen och för rapporteringen av resultat från denna. Vidare stödjer man de lokala huvudmännen och verksamheterna i arbetet med att främja positiva relationer, sociala och känslomässig välfärd och gott uppförande, etc.

När det gäller det som står i fokus för denna utredning, dvs. kunskapandet om skolan och skolans utveckling, hämtar Education Scotland underlag från bl.a. statistik och egna utvärderingar, inspektioner och PISA. *The Scottish Survey of Literacy and Numeracy (SSLN)* är också en viktig källa. Man intresserar sig också för framgångsrika skolsystem i andra länder.

Ansvaret för SSLN är delat mellan flera aktörer. *The Scottish Government Education Analytical Services Division* har det centrala ansvaret och svarar för projektledning, design, analys, genomförande och rapportering om SSLN. *Education Scotland* har tillsammans med *The Scottish Qualifications Authority (SQA)* ansvar för utvecklingen av kriterier för bedömning av SSLN. Man ansvarar

för olika nivåer. Education Scotland ansvarar också för att tillgängliggöra resultat från SSLN för lärare, skolor och huvudmän i syfte att stödja och informera om lärande och arbetsformer i klassrummet medan SQA har ansvaret för alla operationella aspekter av SSLN, dvs. distribution, administration på skolorna, datafångst etc. Till projektet har knutits en *SSLN Project Management Board*.

Inom Education Scotland arbetar man utifrån en strategi för det man benämner *knowledge management*. Denna består av tre teman – att bygga professionella relationer, att bygga en kunskapsbas och att bygga värden utifrån kunskap. Man arbetar också med att utveckla en forskningsstrategi. Denna går ut på att stärka kapaciteten inom forskning och utvärdering, att bygga upp en forskningsgrund och att skapa länkar mellan forskning, politik och praktik (*knowledge brokering*). Education Scotland identifierar följande utvecklingsområden för den egna verksamheten:

- Tätare dialog med forskarsamhället inom utbildningsområdet.
- Mer strategiskt angreppssätt när man lägger ut forskningsuppdrag.
- Mer proaktiv roll när man lämnar råd i samband med politikutformning.
- Göra tillförlitlig kunskap tillgänglig för de verksamma (lärarprofessionen).
- Stödja och sprida resultat från professionella undersökningar.

Av detta framgår att Education Scotland utöver att svara för stöd, statistik, utvärdering, inspektion, etc, också har en roll i relation till skolforskningen, både som beställare och som kunskapsmäklare. Man understryker att det, för att verka som kunskapsmäklare, krävs oberoende, trovärdighet, transparens, tillgång till en robust kunskapsbas och klara kopplingar till politikutvecklingsprocessen.

Statens offentliga utredningar 2014

Kronologisk förteckning

1. Vissa bostadsbeskattningsfrågor. Fi.
2. Framtidens valfrihetssystem
– inom socialtjänsten. S.
3. Boende utanför det egna hemmet
– placeringsformer för barn och unga. S.
4. Det måste gå att lita på konsument-
skyddet. Ju.
5. Staten får inte abdikera
– om kommunaliseringen av den svenska
skolan. U.
6. Män och jämställdhet. U.
7. Skärpta straff för vapenbrott. Ju.
8. Översyn av statsskuldspolitiken. Fi.
9. Förändrad assistansersättning
– en översyn av ersättningssystemet. S.
10. Ett steg vidare – nya regler och åtgärder
för att främja vidareutnyttjande
av handlingar. S.
11. Kunskapsläget på kärnavfallsområdet
2014. Forskningsdebatt, alternativ och
beslutsfattande. M.
12. Utvärdera för utveckling – om utvärdering
av skolpolitiska reformer. U.

Statens offentliga utredningar 2014

Systematisk förteckning

Justitiedepartementet

Det måste gå att lita på konsumentskyddet. [4]
Skärpta straff för vapenbrott. [7]

Socialdepartementet

Framtidens valfrihetssystem
– inom socialtjänsten. [2]
Boende utanför det egna hemmet
– placeringsformer för barn och unga. [3]
Förändrad assistansersättning
– en översyn av ersättningssystemet. [9]
Ett steg vidare – nya regler och åtgärder för
att främja vidareutnyttjande av handlingar.
[10]

Finansdepartementet

Vissa bostadsbeskattningsfrågor. [1]
Översyn av statsskuldspolitiken. [8]

Utbildningsdepartementet

Staten får inte abdikera
– om kommunaliseringen av den svenska
skolan. [5]
Män och jämställdhet. [6]
Utvärdera för utveckling – om utvärdering av
skolpolitiska reformer. [12]

Miljödepartementet

Kunskapsläget på kärnavfallsområdet
2014. Forskningsdebatt, alternativ och
beslutsfattande. [11].